2

ALGORITHMS FOR A CONSTRAINED OPTIMIZATION PROBLEM WITH APPLICATIONS IN STATISTICS AND OPTIMUM DESIGN

by

BERNARD TORSNEY

A dissertation submitted to the

UNIVERSITY OF GLASGOW

for the degree of

Doctor of Philosophy

1981

ACKNOWLEDGEMENTS

I wish to acknowledge my gratitude to several people;

to David Silvey for his generous supervision, his unfailing willingness to discuss an issue of concern and also for introducing me to the problem;

to my colleagues, past and present, in the Statistics department for many helpful discussions and in particular Mike Titterington for several useful suggestions;

to my wife Kathleen for her support and in particular for checking the manuscript and typescript;

to Mary Richmond for her prompt, efficient and patient typing.

I am also indebted to the University of Glasgow for waiving tuition fees in view of the fact that I was a member of the University staff while carrying out this research.

SUMMARY

The concern of this thesis is algorithms for solving the following constrained optimisation problems:

(P1) "Maximise (concave) $\phi(p)$ over $\mathcal{P} = \{p = (p_1,\dots,p_J): p_j \geq 0, \sum_{j=1}^{J} p_j = 1\}$".

(P2) "Maximise (concave) $\psi(x)$ over the polygon

$$\mathcal{P}(\mathcal{U}) = \{x = x(p) = \sum_{j=1}^{J} p_j u_j \;:\; p = (p_1,\dots,p_J) \in \mathcal{P}\}"$$

(The set $\mathcal{U} = \{u_1,\dots,u_J\}$ contains the vertices of the polygon and might be a discretization of a continuous space.)

(P3) "Maximise (concave) $\Phi(\theta)$ over $\{\theta = (\theta_1,\dots,\theta_t) : \theta_j \geq 0, C\theta = b\}$, where C is a matrix of order $s \times t$, rank(C) = s".

Problem (P2) is a special form of (P1). Problem (P3) is a generalisation of (P1), but can be an example of (P2).

Chapter 1 opens with a list of examples but is mainly devoted to an outline of the optimal linear regression design problem. This can be viewed as an example of (P2), taking $u_j = v_j v_j'$, $x(p) = M(p) = \sum_{j=1}^{J} p_j v_j v_j'$, where v_j is a vector of length k, M(p) a matrix of order $k \times k$. This dictates that the criteria principally studied are functions of the matrix $D = AM^{-}(p)A'$, where A is of order $s \times k$, rank(A) = s, and the null space of M(p) is contained in that of A. The most general criterion considered is $\phi(p) = \phi_t(p) = -\mathrm{tr}(LD^t)$, where $t > 0$ and L is a nonnegative definite matrix of order $s \times s$. When $t \neq 1$, L is always taken to be the identity matrix.

Adopting the terminology of the design context, we refer to p as a design and define the support of p, denoted by Sup(p), as that subset of $\mathcal{U}$ to which p allocates positive weight.

Chapters 2 for (P2) and 3 for the design problem are devoted to a derivation of optimality conditions. The emphasis is on a differential calculus approach in contrast to a lagrangian one. An

important tool is the directional derivative $F_\psi(x,y)$ of $\psi(\cdot)$ at x in the direction of y, and also a normalised directional derivative. Properties of $F_\psi(x,y)$ are enumerated, differentiability is expressed in terms of it, a concept of constrained stationarity is defined and optimality theorems for (P2) are derived in chapter 2.

In chapter 3, $F_\psi(\cdot,\cdot)$ is derived for standard design criteria. These can be nondifferentiable but, in general they satisfy a concept defined as support differentiability. Optimality theorems are derived for both the differentiable and the nondifferentiable cases. Lagrangian approaches are also reviewed. Examples yielding explicit solutions are examined.

The remaining chapters are devoted to the main topic of algorithms with chapter 4 settling some preliminaries.

Algorithms of various types are considered for (P1), (P2). Some new classes are proposed, others are reviewed, minor improvements occasionally suggested. With the exception of some cutting plane algorithms which are examined at the end of chapter 4, these algorithms aim to identify for (P2) an optimising p*, as opposed to an optimising x*. It is not the intention to make rigorous comparisons between these algorithms although some empirical results are reported at the end of chapter 7. In contrast a three-stage Composite Algorithm is a proposal made in the light of a key discussion in the concluding sections of chapter 6.

This proposal forms one of two main outputs of the thesis, the other being to report in chapters 8 and 9 some results concerning a conjectured monotonicity of a particular algorithm.

An iteration of the first stage algorithms, studied in chapter 5 and called vertex direction algorithms, takes a step towards or away from a vertex. These can identify a small subset of containing Sup(p*).

The third stage algorithms are of the constrained steepest ascent and Newton-Raphson type and are mainly appropriate when Sup(p*) has been identified. These are studied in chapter 6, the former type of iteration being derived for (P3) in particular.

The second stage of the Composite Algorithm recommends a technique which can cope with a $\mathcal{U}$ containing a small number of vertices not in Sup(p*). One recommendation in chapter 7 is an iteration of the form

$$p_j^{(r+1)} \propto p_j^{(r)} h(d_j^{(r)}, \delta_r)$$

where $d_j^{(r)} = \partial\phi/\partial p_j^{(r)}$, δ_r is a free parameter and $h(\cdot,\cdot)$ is a function enjoying particular properties, including that of being positive.

It is for the case $h(d,\delta) = d^{\delta}$ and for special values of δ, that theoretical and empirical results concerning monotonicity are reported, in particular for functions enjoying two properties of $\phi_t(p)$, namely positive derivatives and homogeneity of degree $(-t)$. The theoretical results are aided by establishing links with the EM algorithm and by proving a moment inequality.

Various matrix results are also derived in several chapters.

A final chapter (10) considers generalisations in various directions and contains further results relating to monotonicity.

TABLE OF CONTENTS

CHAPTER 1

A CONSTRAINED OPTIMIZATION PROBLEM; EXAMPLES

§1.1 A Problem

This thesis is primarily concerned with how to solve the following problem:

(P1) "Maximise a function $\phi(p)$ on the probability simplex

$$\mathcal{P} = \{p = (p_1, p_2, \cdots\cdots, p_J) : p_j \geq 0, \sum_{j=1}^{J} p_j = 1\}$$

The equality constraint $\Sigma p_j = 1$ renders the problem a nondegenerate constrained optimisation problem, the full constraint region being a closed bounded convex set.

Many examples of (P1) arise in the field of statistics. We list a few to start with.

Ex 1.1.1 Possibly the simplest example is that of finding the maximum likelihood estimators of the probability parameters of a multinomial likelihood. The likelihood is of the form

$$\phi(p) = c(o) p_1^{o_1} p_2^{o_2} \cdots p_J^{o_J}.$$

The optimum of course is $p_j^* = o_j/n$, $n = \Sigma o_j$

Ex 1.1.2 A second example which has received a lot of attention recently is that of estimating the mixing parameters (probabilities) of a mixture distribution given data $\underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_n$. The simplest example of this would arise when the component probability models $f_j(\underset{\sim}{y})$ of the mixture are themselves free of any unknown parameters. Then the likelihood is

$$\phi(p) = \prod_{i=1}^{n} \{\sum_{j=1}^{J} p_j f_j(\underset{\sim}{y}_i)\}$$

The recent literature includes Smith and Makov (1978), Murray and Titterington (1978), Dempster, Laird and Rubin (1977).

Ex 1.1.3 Another example arises in the field of paired comparisons. Suppose J treatments $T_1, \ldots, T_J$ are compared on a pairwise basis, n_{ij} comparisons being made on treatments T_i and T_j, $i < j$. Assume that in any single comparison of T_i and T_j there is a probability π_{ij} that T_i

will be preferred to T_j ($i \neq j$), the same for all such pairwise comparisons, with $\pi_{ij} + \pi_{ji} = 1$. Let o_{ij} denote the observed number of times that T_i is preferred to T_j ($i \neq j$) and assume that there is no ties so that, for $i<j$, $o_{ij} + o_{ji} = n_{ij}$. Assuming also independence between each pairwise comparison, the likelihood for the data would be

$$L_o(\pi) = \prod_{i<j} (\pi_{ij})^{o_{ij}} (\pi_{ji})^{o_{ji}}$$

Many models suggest that π_{ij} is of the form $\pi_{ij} = \{p_i/(p_i + p_j)\}$, $p_i > 0$. See Bradley and Terry (1952), Davidson (1969), Bradley (1965). This relationship however only defines the p_i's relative to each other, for it would follow that $\pi_{ij} = \{cp_i/(cp_i + cp_j)\}$. In order to find a particular set of p_i's corresponding to the maximum likelihood estimator of π_{ij}, a restriction must be imposed on Σp_i, and $\Sigma p_i = 1$ is a natural choice. Finding the corresponding estimates of the p_i's requires solution of (P1) with

$$\phi(p) = \left(\prod_{i=1}^{J} p_i^{o_i}\right) \Big/ \prod_{i<j} (p_i + p_j)^{n_{ij}} \qquad , \quad o_i = \sum_{\substack{j=1 \\ j \neq i}}^{J} o_{ij}$$

<u>Ex 1.1.4</u> A final example is contained in Morgan and Titterington (1977). They seek to solve (P1) for the case

$$\phi(p) = \prod_{\substack{i=1 \\ i \neq j}}^{J} \prod_{j=1}^{J} \{p_j/(1-p_i)\}^{n_{ij}}$$

The solution yields maximum likelihood estimates under a particular case of quasi-independence in a contingency table whose diagonal entries are either missing or excluded from consideration. Quasi-independence states that only for some i and j can the cell probabilities $p_{ij}(\Sigma\Sigma p_{ij} = 1)$ be factorised into the form $p_{ij} = a_i b_j$, in contrast to full independence in which such factorisation holds for all i and j.

The above example arises when the mover-stayer model of Blumen, Kogan and McCarthy (1955) is postulated for the transition probabilities of a J-state Markov chain. This implies that the conditional probabilities of state change are

$$p_{j|i} = p_j/(1 - p_i) \qquad , \qquad i \neq j$$

This therefore proposes quasi-independence for all $i \neq j$, namely, that $p_{ij} = \{q_j p_i/(1 - p_i)\}$, $i \neq j$, for some probability vector $q_1, q_2, \ldots, q_J$.

Note that making partial use of the constraint $\Sigma p_j = 1$, the function $\phi(p)$ can be rewritten as

1.1.1 $$\phi(p) = \prod_{\substack{i=1 \\ i \neq j}}^{J} \prod_{j=1}^{J} \{p_j / \ell_i(p)\}^{n_{ij}} \quad , \quad \ell_i(p) = \sum_{\substack{t=1 \\ t \neq i}}^{J} p_t \ .$$

It is possibly not surprising that problem (P1) crops up in various forms in the statistical literature given that probabilities are not infrequently parameters of probability models. This of course is particularly so in the case of likelihoods for categorical data.

Other examples arise in the form of the following more general problem.

(P2) "Solve (P1) for $\phi(p) = \psi\{x(p)\}$ for some given function $\psi(\cdot)$ where $x(p) = \sum_{j=1}^{J} p_j u_j$, $u_1, \ldots, u_J$ being a given set of points. Equivalently maximise $\psi(x)$ subject to x belonging to the convex polygon $\mathcal{P}(\mathcal{U})$ with the finite set of vertices $\mathcal{U} = \{u_1, u_2, \cdots u_J\}$, i.e., $\mathcal{P}(\mathcal{U}) = \{x : x = x(p) = \sum_{j=1}^{J} p_j u_j \ , \ p \in \mathcal{P}\}$."

Note that we could alternatively state that $x(p) = E_p(u)$, where u is a random variable assuming the value u_j with probability p_j.

While (P2) clearly yields a particular type of example of (P1) we can conversely view (P1) as a particular case of (P2), namely that which takes u_j to be the unit vertex e_j.

Another example of (P2) can be the following clear generalisation of (P1).

(P3) "Maximise a function $\Phi(\theta)$, $\theta = (\theta_1, \cdots, \theta_t)$, subject to (i) $\theta_j \geq 0$, (ii) $C\theta = b$ where C is an $s \times t$ matrix of rank s."

Clearly (P1) is a particular case of (P3) and viewing (P3) in this light, that is, as a generalisation of (P1), it might be appropriate to replace $\Phi(\cdot)$ by $\phi(\cdot)$. This is done in chapter six.

However the feasible region may define a finite polygon $\mathcal{P}(\mathcal{U})$ for which vertices will be some of the following intersections; namely, the intersections of the region $C\theta = b$ with $(t-s)$ of the regions $\theta_j = 0$. If this is so, we then have in (P3) a particular case of (P2) and it is natural then to replace $\Phi(\cdot)$ by $\psi(\cdot)$. This is done in chapter ten.

One such occurrence of (P3) would arise when testing linear hypotheses about the parameters in multinomial models for categorical data. These parameters are of course probabilities so that the constraint $C\theta = b$ must either include as a component that $\underline{1}'\theta = 1$, where $\underline{1}$ is a vector whose components take the value 1, or state that various subsets of the components of θ should sum to unity. We will consider an example of such a linear hypothesis and discuss (P3) in the last chapter. In a strange way we now can effectively have (P3) as a particular case of (P1).

Yet another example of (P2) will be seen to be a general optimal linear regression design problem. This we will study in the ensuing sections of this chapter.

We thus have a wide range of examples of (P1) and (P2), a justification for our study; and the need for a study of how to solve these problems is that typically they do not possess explicit solutions. Numerical techniques must be employed. It is the remit of this thesis to study algorithms which have been formulated for finding an optimising p^* particularly in the case of (P2) and also to propose a further class of algorithm.

Why new algorithms? One general argument is that there is a dearth of numerical thecniques for the solution of constrained optimisation problems. An argument that (P2) in particular requires special treatment, is the following one.

Problem (P2) differs in a number of ways from other examples of (P1), such as examples 1.1.1,...,1.1.4.

(i) One may only be interested in an optimising x^* as opposed to an optimising p^*, $x^* = x(p^*)$.

(ii) While there may be a unique optimising x^* there could be many optimising p^*'s.

(iii) Frequently an optimising p^* may put $p_j^* = 0$, i.e. the optimum lies on the boundary of $\mathcal{P}$.

In contrast there is almost certainly a unique optimising p^* in the case of examples 1.1.1 to 1.1.4 (certainly in the case of ex.1.1.1), otherwise the parameters would be inestimable. Furthermore p^* certainly does not lie on the boundary of $\mathcal{P}$ in the case of examples 1.1.1,1.1.2,1.1.4 and it is unlikely to do so in example 1.1.3 assuming $n>J$.

In such a case we effectively have a simpler constrained optimisation problem, a problem having one active constraint, the simple linear equality $\sum_{j=1}^{J} p_j = 1$. It would seem that it should be a simple matter to devise a simple neat modification to standard unconstrained optimisation techniques to take account of this.

Techniques as we shall see can be similarly devised for finding an optimising x^* in the case of (P2).

However it would seem that standard numerical techniques cannot be so neatly modified to cope with optima which explicitly lie on boundaries of constraint regions.

It is for this reason that algorithms have been formulated for finding an optimising p^* in the case of (P2), in particular for the optimal linear regression design problem. It is this problem in fact, which has motivated our study of (P2), and this is why we consider the design problem in detail for the remainder of this chapter; while examination of algorithms will not begin until chapter four.

This section closes by observing some of the properties of examples 1.1.1,..,1.1.4 that will be seen to be relevant later.

(a) All four functions are homogeneous including example 1.1.4 when taken in the form 1.1.1. With hindsight this is not all that surprising since independence is a common assumption in the formulation of probability models.

Of interest to note is that the equality $\sum p_j = 1$ is an informative constraint to impose on a function satisfying the homogeneity condition that $\phi(cp) = c^t \phi(p)$. Study $\phi(p)$ subject to $\sum p_j = 1$ and one has an informed picture of the general behaviour of $\phi(\cdot)$ on the positive quadrant at least.

(b) With the exception of example 1.1.3 the functions have positive derivatives as is evident from the following respective expressions for $\partial\phi/\partial p_j$:

(i) $\phi(p)\{o_r/p_r\}$

(ii) $\phi(p)\left\{\sum_i\{f_r(y_i)/\sum_j p_j f_j(y_i)\}\right\}$

(iii) $\phi(p)\left\{o_r/p_r - \sum_{s\neq r} n_{rs}/(p_r+p_s)\right\}$

(iv) $\phi(p)\left\{(1/p_r)\sum_{i\neq r} n_{ir} + (1-p_r)^{-1}\sum_{j\neq r} n_{rj}\right\}$

In the case of example 1.1.3 there will typically be both positive and negative derivatives when p is in the positive quadrant, because $\sum p_r\,\partial\phi/\partial p_r = 0$, a consequence of the fact that $\phi(p)$ is a homogeneous function of degree zero.

(c) In some instances the functions are concave.

The latter property is nice in that it guarantees the existence of a unique maximum while the first two properties, not important in themselves, prove basic ingredients in the formulation of an algorithm.

The design criteria that will be considered and also the particular example of problem (P3) to which we referred will also be seen to enjoy (a) and (b).

§1.2 Optimum Linear Regression Design

The concept of an optimum regression design arises when an observable univariate variable y has the probability model $p(y \mid v, \theta, \tau)$ where in particular

$$\theta = (\theta_1, \theta_2, \cdots, \theta_k)'$$
$$v = \{f_1(x), f_2(x), \cdots\cdots, f_m(x)\}' .$$

The k components of θ are unknown parameters of interest, while τ is a set of nuisance parameters.

The quantity x, possibly a vector, is a regressor variable whose value is restricted to a closed bounded space $\mathcal{X}$, called the design space, which will typically be continuous but can be discrete.

The functions $f_j(x)$, $j = 1,\ldots,m$ are of known bounded form.

The regression is linear for the case $m = k$ if y depends on v and θ only through the linear mean.

1.2.1 $$E(y) = v'\theta \quad .$$

In order to obtain an observation on y, a value for x must first be selected from $\mathcal{X}$. It is assumed that x can be set to any chosen value in $\mathcal{X}$ without error.

Given this control over the selection of x, a natural question to consider is at what values of x should observations, say n, on y be taken in order to obtain a 'best' inference or as reliable an inference as possible for all or some of the parameters θ.

Such a 'best' selection of x values or allocation of the n observations to the elements of $\mathcal{X}$ is termed an optimal design or optimal regression design.

The mode of inference must first be decided upon. For the moment let us suppose that it is point estimation. It will be seen that the solution proposed for this case will hold good for other modes of inference too.

It is desired then to choose n values $(x_1, x_2, \ldots, x_n) = \underline{x}$ to yield 'best' point estimates $\hat{\theta}(\underline{x})$ of some or all of the parameters θ.

Clearly this poses an optimisation problem. As we shall now see the standard mathematical formulation is to approximate it by a particular case of (P2). Rigorous referencing will be deferred to a summary on the literature of optimum regression design in section 1.5.

Suppose by some method of point estimation the estimator $\hat{\theta}(\underline{x})$ of θ is obtained. Typically the components $\hat{\theta}_j(\underline{x})$ will be correlated. Arguably then the k×k matrix $D\{\hat{\theta}(\underline{x})\} = E\{[\hat{\theta}(\underline{x})-\theta][\hat{\theta}(\underline{x})-\theta]'\}$ the dispersion matrix of $\hat{\theta}(\underline{x})$ about θ, contains information about the accuracy of $\hat{\theta}(\underline{x})$ not only in its diagonal elements, which of course measure the mean squared deviation of $\hat{\theta}_j(\underline{x})$ from θ_j, but also in its off-diagonal cross product deviation terms. Generally speaking the "smaller" $D\{\hat{\theta}(\underline{x})\}$ the better the accuracy of $\hat{\theta}(\underline{x})$.

A best $\underline{x}$ must in some sense make $D\{\hat{\theta}(\underline{x})\}$ small. However this matrix will typically not only depend on $\underline{x}$ but also on θ and τ so that a best $\underline{x}$ would depend in particular on the very parameter vector θ for whose estimation an optimal design is sought. This is particularly so in the case of nonlinear models, that is, models in which the means are not linear in the unknown parameter θ.

One exception though is given by least squares estimation in the case of the linear model 1.2.1 with the addition of some assumptions justifying the use of this method of estimation.

Assume the model 1.2.1 and let y_i denote the observation obtained at x_i so that $E(y_i) = v_i'\theta$, $v_i = \{f_1(x_i), f_2(x_i), \ldots, f_k(x_i)\}'$, $i=1,\ldots,n$. It is of note that typically there will be several equalities between the x_i's, more than one observation being taken at the same x value. Suppose also that $y_1,\ldots,y_n$ are independent random variables with common variance σ^2. The y_i's then satisfy the standard linear model

1.2.2 $$E(Y) = \{L(\underline{x})\}\theta \quad , \quad D(Y) = \sigma^2 I_n \ ,$$

where $Y = (y_1,y_2,\ldots,y_n)'$, $x = (x_1,x_2,\ldots,x_n)'$, $L(\underline{x})$ is the $(n\times k)$ matrix whose (i,j)th element is $f_j(x_i)$ and $D(Y)$ denotes the dispersion matrix of Y.

Least squares estimators are a conventional choice for this model having the optimality of being best linear unbiased. They are solutions of

1.2.3 $$[L'(\underline{x})L(\underline{x})]\hat{\theta}(\underline{x}) = L'(\underline{x})Y$$

Consider the case in which it is desired to estimate all k parameters. Then $\underline{x}$ must at least have been chosen to ensure that $[L'(\underline{x})L(\underline{x})]$ is nonsingular, in which case there is only one solution to 1.2.3, namely

1.2.4 $$\hat{\theta}(\underline{x}) = [L'(\underline{x})L(\underline{x})]^{-1}L'(\underline{x})Y.$$

Given that $E\{\hat{\theta}(\underline{x})\}=\theta$ the matrix $D\{\hat{\theta}(\underline{x})\}$ is in fact the dispersion matrix of $\hat{\theta}(\underline{x})$ and has the familiar form

1.2.5 $$D\{\hat{\theta}(\underline{x})\} = \sigma^2[L'(\underline{x})L(\underline{x})]^{-1}$$

It therefore does not depend on θ and depends only on the additional parameter σ^2, in that each of its elements is proportional to σ^2. It is therefore possible to determine a priori, an $\underline{x}$ which makes $D\{\hat{\theta}(\underline{x})\}$ "small", namely an $\underline{x}$ which makes the $k \times k$ matrix $[L'(\underline{x})L(\underline{x})]$ large in some sense.

This has given definition in principle to the linear regression design problem. In the next section the problem is examined in more detail with a view to simplification and streamlining, a process which results in expressing it in the form (P2). Examples of criteria which 'maximise' $[L'(\underline{x})L(\underline{x})]$ will be given in section 1.4.

§1.3 Simplifying The Design Problem

§1.3.1 A first simplification is that it is unnecessary to continue reference to the regressor variable x. The basic model of formula 1.2.1 states that $E(y) = v'\theta$ where $v = \{f_1(x),\dots,f_k(x)\}'$ for some $x \in \mathcal{X}$. Clearly this is equivalent to stating that $E(y) = v'\theta$ where $v \in \mathcal{U}$ for some closed bounded k-dimensional space $\mathcal{U}$. In relation to x, $\mathcal{U} = \{v : v = \{f_1(x),\dots,f_k(x)\}' , x \in \mathcal{X}\}$. That is $\mathcal{U}$ is the image under the vector function $\underline{f} = (f_1,\dots,f_k)'$ of $\mathcal{X}$. From now on $\mathcal{U}$ will be referred to as the design space.

Typically $\mathcal{U}$ will be continuous but a second simplification is to assume that $\mathcal{U}$ is discrete. A 'justification' for this will be given later on in this section.

Suppose then that a discrete $\mathcal{U}$ consists of J distinct vectors $v_1,\dots,v_J$. Then the basic model is

$$E(y) = v'\theta \quad , \quad v \in \mathcal{U} = \{v_1, v_2, \cdots\cdots, v_J\} .$$

In order then to obtain an observation on y, a value for v must first be chosen from the J elements of $\mathcal{U}$ to be the point at which to take this observation. That $\mathcal{U}$ is taken to be discrete suggests that this can be done without error.

The design problem can now be expressed more concisely. At which of the points v_j should observations be taken and, if n observations in total are allowed, how many observations should be taken at these points in order to obtain 'best' least squares estimators of θ ?

Suppose we take n_j observations at v_j, $n_j \geq 0$, $\sum_{j=1}^{J} n_j = n$. Note that the possibility that $n_j = 0$ is included. What values of the n_j produce the best least squares estimators of θ ?

As we have seen the n_j must make $[L'(\underline{x})L(\underline{x})]$, or, in an obvious notational change, $[L'(\underline{n})L(\underline{n})]$, 'big' in some sense, $\underline{n} = (n_1, n_2, \cdots, n_J)$.

It is now possible to obtain neater expressions for this matrix $[L'(\underline{n})L(\underline{n})]$.

Suppose the vector Y, in the linear model $E(Y) = [L(\underline{n})]\theta$, is such that the first n_1 components are the observations taken at v_1, the next n_2 are those observations taken at v_2 and so on. Then $L(\underline{n})$ is such that $L'(\underline{n}) = [L_1' : L_2' : \cdots\cdots : L_J']$, where L_j is the $(n_j \times k)$ matrix each of whose rows is v_j'; that is, $L_j' = [v_j\, v_j \ldots v_j]$ (n_j columns) and hence

$$L'(\underline{n})L(\underline{n}) = \sum_{j=1}^{J} L_j' L_j$$

Two expressions can be obtained for the $(k \times k)$ matrix $L_j' L_j$ which yield two corresponding expressions for $L'(\underline{n})L(\underline{n})$. Since

$$L_j' L_j = \underbrace{[v_j\ v_j \cdots\cdots v_j]}_{n_j \text{ cols.}} \left.\begin{bmatrix} v_j' \\ v_j' \\ \vdots \\ v_j' \end{bmatrix}\right\} n_j \text{ rows}$$

we have

1.3.1 $$L_j' L_j = n_j v_j v_j'$$

and proceeding further

$$L_j' L_j = v_j n_j v_j' = V_j D_j V_j'$$

where D_j is a $(J \times J)$ matrix all of whose entries are zero except the j^{th} diagonal element which takes the value n_j and V_j can be any $(k \times J)$ matrix whose j^{th} column is v_j.

In particular

1.3.2 $$L_j' L_j = V D_j V'$$

where V is the matrix $[v_1\ v_2 \ldots v_J]$.

Let the matrix $M(\underline{n})$ denote $L'(\underline{n})L(\underline{n})$. Then the following two expressions are obtained for $M(\underline{n})$.

1.3.3 $$M(\underline{n}) = \sum_{j=1}^{J} n_j v_j v_j'$$

1.3.4 $$M(\underline{n}) = VNV'$$

$\left.\right\}$ $n_j \geq 0$, $\sum n_j = n$, n_j integer

where $N = \text{diag}\{n_1, n_2, \cdots\cdots, n_J\}$.

Further $M(\underline{n}) = nM(p)$ where

1.3.5 $\quad M(p) = \sum_{j=1}^{J} p_j v_j v_j'$

or $\qquad\qquad\qquad\qquad\qquad\qquad p_j \geqslant 0\ ,\ \sum p_j = 1$

1.3.6 $\quad M(p) = VPV' \qquad\qquad\qquad np_j$ integer ,

where $p_j = n_j/n$ and so is the proportion of observations taken at v_j , and, $P = \text{diag}\{p_1, p_2, \cdots\cdots, p_J\}$

The matrix $M(p)$ is one that we will deal with often mainly in the form of equation 1.3.5, occasionally in the form of equation 1.3.6. Clearly in the former form it is a particular case of the $x(p)$ of (P2), namely $M(p) = \Sigma p_j u_j$, where $u_j = v_j v_j'$. As with $x(p)$, so also with $M(p)$, we can regard the proportions p_j defining a probability distribution on $\mathcal{V}$. Then formula 1.3.5 defines $M(p)$ as an expectation, namely

1.3.7 $\quad M(p) = E_p\{vv'\}$

where $P(v = v_j) = p_j$.

Of course $M(p)$ is a symmetric matrix.

Returning again to the design problem it is clear that it can be regarded as either choosing $\underline{n}$ to make $M(\underline{n})$ large or as choosing p to make $M(p)$ large since, for given n, $M(p)$ is proportional to $M(\underline{n})$.

More prominent now are the constraints on the problem. The n_j must be nonnegative integers summing to n. The p_j must be rational proportions or fractions summing to 1, with the proviso that np_j be an integer.

The problem in fact is a constrained integer programming type problem and in the design context is described as an exact design problem. Typically integer programming problems are difficult or at least laborious to solve even without additional constraints, mainly because the theory of calculus cannot be used to define the existence of or to identify optimal solutions. Furthermore a solution would have to be worked out separately for different values of n. By the nature of the problem then, no formula for an optimal exact design could be devised that would express it as a 'function' of n. Nevertheless one could not avoid having to solve such a problem if, for given n, one doggedly chose to seek optimal n_j's directly.

However if the approach first seeks optimal p_j's then some mild relaxation of the rigidities of the above problem can be employed. This is to relax the restriction that the p_j be rational or that np_j be integer. Require only that p be a probability vector and seek a p which makes $M(p)$ 'big'.

This is a simpler or more flexible problem to solve and yet one that is not much visibly different from the original. It is clearly another example of (P2) as the following formal statement illustrates.

(P4) "Solve (P1) for $\phi(p) = \psi\{M(p)\}$ where $M(p) = \sum_{j=1}^{J} p_j v_j v_j'$ for some given function $\psi(\cdot)$ with a matrix argument. Equivalently maximise $\psi(M)$ over $\mathcal{M} = \{M : M = M(p), p \in \mathcal{P}\}$".

In the notation of (P2) $x = M$, $u_j = v_j v_j'$, both $(k\times k)$ symmetric matrices.

In the design context (P4) is known as an approximate discrete optimal design problem. Examples of suitable functions $\psi(M)$ will be given in the next section.

Note that it is an optimising p^* as opposed to an optimising M^* that we wish to discover.

Such a solution p^*, is referred to as a ϕ- optimal design.

Naturally enough an approximate solution that would be proferred to the original exact design problem would be np^*,rounded to a 'nearest' exact design. Hopefully this would be near a fully optimum exact design.

From now on we consider only the approximate discrete optimal design problem and think of a design as defined by a set of weights or probabilities p_j, p_j being assigned to $v_j \in \mathcal{V}$. It should be noted that some authors record a design in the following more complete manner

$$\begin{Bmatrix} v_1 & v_2 & & v_J \\ & & & \\ p_1 & p_2 & & p_J \end{Bmatrix} .$$

Such a design may put weight $p_j = 0$. In particular as we have observed for (P2) we may have $p_j^* = 0$.

<u>Definition 1.3.1 Support of a design</u>

The support of a design p is defined to be those vertices in $\mathcal{V}$ enjoying non-zero weight under p. The support of p* is the support of the optimum. The term will also be used in reference to the general problem (P2).

We define some further terminology which will be maintained throughout this thesis.

(i) $\mathrm{Sup}(p)$ denotes the support of p. Hence $\mathrm{Sup}(p) \subseteq \mathcal{V}$ $(\mathrm{Sup}(p) \subseteq \mathcal{U}$ in the general case$)$.

(ii) $n\{\mathrm{Sup}(p)\}$ denotes the number of vertices in Sup(p).

(iii) V_p denotes a matrix whose columns are the vectors in Sup(p), while their corresponding non-zero weights are the entries of the diagonal matrix P_p.

We note the following further formulae for M(p).

Suppose $n\{\mathrm{Sup}(p)\} = t$ so that

$$\mathrm{Sup}(p) = \{v_{i_1}, v_{i_2}, \cdots\cdots\cdots, v_{i_t}\} \subseteq \mathcal{V},$$

$$V_p = [v_{i_1} \; v_{i_2} \qquad\qquad v_{i_t}],$$

$$P_p = \mathrm{diag}\{p_{i_1}, p_{i_2}, \cdots\cdots\cdots, p_{i_t}\},$$

where $p_{i_j} > 0$, V_p is $(k \times t)$, P_p is $(t \times t)$, $\mathrm{rank}\{(P_p)\} = t$.

In general $\mathrm{rank}\,(P_p) = n\{\mathrm{Sup}(p)\}$

We have of course

1.3.8 $\quad M(p) = \sum_{j=1}^{t} p_{i_j} v_{i_j} v_{i_j}'$

but appealing to formula 1.3.6 we also have

1.3.9 $\quad M(p) = V_p P_p V_p'$

Further since P_p is a square matrix of full rank it follows that

1.3.10 $\quad \mathrm{rank}\{M(p)\} = \mathrm{rank}\{(V_p)\}.$

Finally in concluding this section we recall formulae 1.3.7 for $M(p)$ and note that (P4) could be restated as seeking a probability distribution p^* on the set $\mathcal{V}$ which maximises $\phi(p) = \psi\{M(p)\}$ where $M(p) = E_p\{vv'\}$. We will now make use of this observation. A similar version of (P2) could of course be invoked.

§1.3.2 Consider now the case of $\mathcal{V}$ continuous. What will be an optimal solution in this case? Clearly the practical problem will still demand an exact design as a solution. Note that an exact design would confer a discrete probability distribution on $\mathcal{V}$. It would allocate weights (rational weights) say $\tilde{p}_1, \ldots, \tilde{p}_J$ to a finite set of points $\mathcal{V}_D = \{v_1, \ldots, v_J\}$ in $\mathcal{V}$ and zero weight to all other points in $\mathcal{V}$.

However it will be no less difficult to discover an optimum exact design in this case than in the case of $\mathcal{V}$ discrete. Again an approximating problem comes to the rescue. It is called the continuous optimal design problem and is simply the extension of the approximate discrete optimal design problem taken in the form mentioned at the end of the last section. The problem is

(P5) "Find a probability measure $p^*(\cdot)$ on the continuous space $\mathcal{V}$ which maximises $\phi(p) = \psi\{M(p)\}$ where

$$M(p) = E_p(vv') = \int_{\mathcal{V}} vv' dp(v)"$$

A continuous analogue to (P2) could be similarly defined.

One might have thought that the solution to this approximating problem would be of no practical value, that an optimising $p^*(\cdot)$ would be a continuous probability measure on $\mathcal{V}$ and that (P5) would thereby be a seemingly more difficult problem than (P4) to solve.

However, as we shall see, Caratheodery's theorem guarantees that at least one solution to (P5) is always a discrete solution.

It is time to study some properties of the matrices $M(p)$ where p is any probability measure.

(i) Let $\mathcal{ND}$ denote the set of nonnegative definite symmetric matrices. Then $M(p) \in \mathcal{ND}$. This is to be expected since the original matrix $M(\underline{n})$ is a dispersion matrix. The symmetry is obvious, while the simple argument,

$$x'M(p)x = x'E_p\{vv'\}x = E_p\{x'vv'x\} = E_p\{(x'v)^2\} ,$$

neatly establishes the nonnegative definite label.

That $M(p)$ is of order $k \times k$ but symmetric means that it can be represented by a point in $k(k+1)/2$ dimensions as opposed to k^2 dimensions.

(ii) Let $\mathcal{M} = \{M(p) : p \text{ is any probability measure on } \mathcal{V}\}$. Let p_v be the probability measure that puts unit weight at the point $v \in \mathcal{V}$. Hence $M(p_v) = vv'$. Clearly $vv' \in \mathcal{M}$ and in fact $\mathcal{M}$ is the convex hull of the set $\{vv' : v \in \mathcal{V}\}$.

We can now appeal to Carathéodery's theorem.

<u>Carathéodery's Theorem</u>

Each point M of the convex hull $\mathcal{M}^*$ of any subset $\mathcal{U}$ of n-dimensional space can be represented in the form

$$M = \sum_{j=1}^{n+1} \alpha_j u_j \quad , \quad \alpha_j \geq 0 \; , \; \sum_{j=1}^{n+1} \alpha_j = 1 \quad , \quad u_j \in \mathcal{U} \; .$$

If M is a boundary point of the set $\mathcal{M}^*$ then $\alpha_{(n+1)}$ can be put to zero.

Such representations are not unique. $\square$

It follows that each $M \in \mathcal{M}$ has at least one representation of the form

$$M = \sum_{j=1}^{J} p_j v_j v_j' \quad ,$$

where $J \leq \{[k(k+1)/2]+1\}$, unless M is a boundary point of $\mathcal{M}$ in which case $J \leq k(k+1)/2$. It will be seen, for the design criteria which we will consider, that an optimum M^* must lie on the boundary of $\mathcal{M}$.

Typically J will be smaller than the above limits. If $J < k$ then M will be singular, since then the rank of M can be at most J.

Thus we have that any 'continuous' design measure and in particular any 'continuous' optimal design measure can be replaced by at least one finite discrete probability distribution, and so we have a justification for having initially assumed $\mathcal{V}$ discrete, for such an optimal design will have a discrete finite support. This optimal support we could regard as the design space. However, typically $\text{Sup}(p^*)$ will not be known except in some instances when it, or some finite discrete subset of $\mathcal{V}$ containing it, can be identified using intuition or, geometrical or symmetry arguments.

In general problem (P5) is that bit more difficult than (P4) in that, typically, $\text{Sup}(p^*)$ must in a sense be computed, possibly only approximately, as a prelude to determining p^* and this is essentially done by some of the algorithms which we will consider.

Given that we need to use algorithms to solve (P4) and hence (P5) it is arguable that there is no need to justify assuming $\mathcal{V}$ to be discrete. Any programmed numerical technique must discretise a continuous space; if solutions are not discrete, numerical techniques will only produce discrete approximations. Effectively design algorithms will work with discrete $\mathcal{V}_D$'s, containing approximations, as indicated below, to the optimal support $\text{Sup}(p^*)$.

An ideal discretisation would seem intuitively to be some form of "uniform grid" on a continuous $\mathcal{V}$, but typically this is difficult to determine when $\mathcal{V}$ is an image under some $\underline{f}$ of some $\mathcal{X}$. In practice the discretisation that is used is the image under $\underline{f}$ of a uniform grid on $\mathcal{X}$.

Whatever discretisation $\mathcal{V}_D$ of a continuous $\mathcal{V}$ that is used however, it will inevitably be the typical case that $\mathcal{V}_D$ will exclude points of a true optimal support, $\mathcal{V}_D^*$. The optimal design for $\mathcal{V}_D$, which is what we would have to compute unless the initial $\mathcal{V}_D$ is altered, will then not be that for $\mathcal{V}$. Clearly though if $\mathcal{V}_D$ is a fine enough grid on $\mathcal{V}$ then the former should be a good approximation to the latter. Recipes exist however for modifying a $\mathcal{U}_D$ to, as it were, a better approximation to Sup(p*) with an accompanying modification to a current design on $\mathcal{U}_D$.

From now on we will almost always assume that a design space $\mathcal{U}$ is discrete.

§1.4 Choice Of Design Criteria; Their Properties

§1.4.1 The standard design criteria are now examined. In this section we consider the case when interest is in inference about all of the parameters Θ of the linear model 1.2.1. The matrix $M(p)$ must therefore be nonsingular and hence positive definite.

Three main maximising criteria have been formulated for this case, these being D-optimality, A-optimality and E-optimality. The corresponding ψ-functions requiring maximisation over $\mathcal{M}$ are

D-optimality: $\psi_D(M) = \ln\{\det(M)\} = -\ln\{\det(M^{-1})\}$.

A-optimality: $\psi_A(M) = -\text{trace}(M^{-1}) = -\text{tr}(M^{-1})$

E-optimality: $\psi_E(M) = -\lambda_{max}(M^{-1})$,

where $\lambda_{max}(A)$ denotes the largest eigenvalue of A.

A further criterion which will be discussed later is

L-optimality: $\psi_L(M) = -\text{tr}(LM^{-1})$.

where L is $k \times k$ and nonnegative definite.

Clearly to maximise any of these functions over $\mathcal{M}$ is to look for an $M(p)$ that is 'big'. The first three criteria have the following more specific inference improving motivations.

The A-optimal criterion has the simplest interpretation. It seeks to minimise the sum of the variances of the least squares estimators or their average variance (A for average).

That of E-optimality is seen in the light of the result that since non-singular $M(p)$ and hence its inverse is a positive definite symmetric matrix then

$$\lambda_{max}(M^{-1}) = \max_{\underline{b}'\underline{b}=1}\{\underline{b}'M^{-1}(p)\underline{b}\}$$

In fact any normalised eigenvector corresponding to $\lambda_{max}(M^{-1})$ is a maximising $\underline{b}$.

Hence, since $\sigma^2(\underline{b}'M^{-1}(p)\underline{b})$ is proportional to the variance of $\underline{b}'\hat{\underline{\theta}}$, the E-optimal criterion seeks to best estimate that linear combination $\underline{b}'\underline{\theta}$ of the θ_j's which is worst estimated among comparable linear combinations satisfying the normalisation constraint that the coefficient vector $\underline{b}$ have unit length.

Various motivations for D-optimality exist. These extend beyond the idea of point estimation and all fall into the realm of explicit joint inference. Possibly the most practical and most obvious one is when the mode of inference simply seeks to identify a subset of the parameter space that is most strongly suggested by the data to contain the true value of the parameters $\underline{\theta}$. Such a region would, for the model $E(\underline{Y}) \propto [M(p)]\underline{\theta}$, $D(\underline{Y}) = \sigma^2 I_n$ take on, either exactly or approximately, the shape of an ellipsoid of the form $\{\underline{\theta} : (\underline{\theta}-\hat{\underline{\theta}})'[M(p)](\underline{\theta}-\hat{\underline{\theta}}) \leq k(\underline{Y})\}$ for some critical value $k(\underline{Y})$. This would be the case whether one used the methods of classical or likelihood inference with the additional assumption that the vector $\underline{Y}$ has a multivariate normal distribution or whether one used the principle of Strong Evidence that is yet to appear in the literature!

The smaller such a region of inference the more informative is the data. The D-optimal criterion chooses $M(p)$ to make the volume of the above ellipsoid smallest because it is the case that this volume is proportional to $\{\det[M(p)]\}^{-1/2}$.

Other motivations for D-optimality lie in hypothesis testing under a normal linear model, though these would be equivalent to taking the ellipsoid above to be a classical confidence ellipsoid.

One other criterion, which appears in the early design literature, falls into this category when interest is in all the parameters and is known as G-optimality (G for generalised variance). It is not though a matrix motivated function of $M(p)$ but, as shall be seen in chapter 2, it has been shown to be equivalent to D-optimality. It seeks to minimise the maximum over $\mathcal{V}$ of $\underline{v}'M^{-1}(p)\underline{v}$. Since $\sigma^2\underline{v}'M^{-1}(p)\underline{v}$ is the variance of $\underline{v}'\hat{\underline{\theta}}$, which is the estimate of the mean of y at $\underline{v}$, and hence $\sigma^2[\underline{v}'M^{-1}(p)\underline{v} + 1]$ is the variance of a predicted value of y at $\underline{v}$, it seeks to predict as well as possible the worst predicted value of y.

We have then three alternative criteria excluding G-optimality. Separate treatment of them however proves unnecessary for they, or 1 to 1 mappings of them, are special cases of

$$\psi_t(M) = -\{(1/k)\mathrm{tr}(M^{-t})\}^{1/t}.$$

That D and E-optimality emerge as particular cases is due to the respective facts that for positive definite M

(i) $\underset{t\to 0}{\mathrm{Lt}}\{\psi_t(M)\} = -\{\det(M^{-1})\}^{1/k} = -\{\det(M)\}^{-1/k}$

and

(ii) $\underset{t\to\infty}{\mathrm{Lt}}\{\psi_t(M)\} = -\lambda_{max}(M^{-1})$,

while A-optimality is clearly equivalent to the case t = 1.

Results (i) and (ii) can be proved directly in a number of ways. In particular they can be established by a proof analogous to that which would prove the two corresponding moment results, below, of which, interestingly, (i) and (ii) are particular cases.

Suppose x is a discrete positive valued random variable with probability distribution given by $P(x = x_i) = q_i$, $i = i,\ldots,k$, where $x_i > 0$, $q_i > 0$, $\sum q_i = 1$. Let $f(t) = \{E(x^t)\}^{1/t}$. Then

(i) $\underset{t\to 0}{\mathrm{Lt}}\{f(t)\} = \prod_{i=1}^{k} x_i^{q_i}$,

(ii) $\underset{t\to\infty}{\mathrm{Lt}}\{f(t)\} = \max_{1\le i\le k}\{x_i\}$.

For a proof see Bechenback and Bellman (1961, p.16).

Since the eigenvalues $\lambda_1,\ldots,\lambda_k$ of M^{-1} are positive, M being positive definite, and $\mathrm{tr}(M^{-t}) = \sum_{i=1}^{k}\lambda_i$, the above matrix results are corollaries which arise in the case $q_i = 1/k, x_i = \lambda_i$.

It was Kiefer (1974) who observed the above generalisation which of course has the advantage of making possible a unified treatment of D, A and E-optimality. The function

$$\phi(p) = -\mathrm{tr}\{M^{-t}(p)\}$$

will appear frequently in the ensuing chapters.

§1.4.2 Suppose now that we are interested in inference about only a subset of size s of the parameters in the vector θ of the linear model 1.2.1, or more generally about the s linearly independent combinations $\alpha = A\theta$ where A is an $(s \times k)$ matrix of rank s.

Again for improving inferences about α comparable to those considered for θ a good design will be one which makes the variance-covariance matrix $D(\hat{\alpha})$ of least squares estimators $\hat{\alpha}$ of α small. Only designs p under which α is estimable can, of course, be contemplated. However it is possible that the complete vector θ may not be estimable under such a design.

Consider that the least squares equations for θ are from 1.2.3

1.4.1 $$nM(p)\theta = L'(n)Y = g$$

The vector θ then will be inestimable if the matrix $M(p)$ is singular for then there will be multiple solutions $\hat{\theta}$ to 1.4.1. Of course, it is possible for this to be the case and that a unique solution is suggested for α by 1.4.1 as the mathematics below will bear out.

However, we first explain from a practical design point of view how a singular $M(p)$ can arise under a design suitable for α. Recall formula 1.3.9 that $M(p) = V_p P_p V_p'$ and let $n\{Sup(p)\} = t$ so that V_p is $k \times t$. Then from 1.3.10 we have

$$\text{rank}\{M(p)\} = \text{rank}\{(V_p)\} \leq \min(k,t)$$

Now if θ were to be estimable then $M(p)$ must be non-singular as stated before. A necessary condition then for this is that $t > k$; i.e., the design p must put weight at a minimum of k points in $\mathcal{U}$ and in fact at a set containing k linearly independent points in order to have $\text{rank}\{(V_p)\} = \text{rank}\{M(p)\} = k$.

At the other extreme suppose we were only interested in estimating the one linear combination $\alpha = c'\theta$ for c a vector (i.e. $A = c'$, $s = 1$) and suppose that $c \in \mathcal{U}$. Then a design which we might consider is the design p_c which puts all weight at c and so has $V_p = c$, $t = 1$. Thus $M(p_c)$ is singular. In some instances p_c will be an optimal design. Similarly if we wish to estimate s linear combinations $\alpha = A\theta$, it can be that a suitable design p need only take observations at s linearly independent points in $\mathcal{U}$, in which case t = s. Certainly t must be at least s but could be smaller than k in which case M(p) is again singular, but the t points must contain s linearly independent points in order that $\text{rank}\{M(p)\} \geq s$.

We turn now to the mathematical detail. According to Graybill (1969, Theorem 7.3.1, p.142) or Searle (1971, Theorem 8, p.28) the set of all solutions to 1.4.1, if solutions exist, is given by

1.4.2 $$\hat{\theta} = (1/n)M^-(p)g + [I - M^-(p)M(p)]h$$

for any h and for any generalised inverse $M^-(p)$ of M(p).

Note that a generalised inverse of a matrix M is any matrix M^- satisfying $MM^-M = M$. This does not define M^- uniquely if M is singular. One particular example, of which we will make particular use, is that unique matrix M^+, known as the Moore-Penrose generalised inverse, which not only satisfies $MM^+M = M$, but also $M^+MM^+ = M^+$ and symmetry of (MM^+) and (M^+M).

From 1.4.2 it follows that a solution to $\hat{\alpha}$ is given by

1.4.3 $$\hat{\alpha} = (1/n)AM^-(p)g + [A - AM^-(p)M(p)]h$$

Now, if there is to be a unique $\hat{\alpha}$, then it must be the case that

1.4.4 $$AM^-(p)M(p) = A$$

in which case

$$\hat{\alpha} = (1/n)AM^-(p)g = (1/n)AM^-(p)L'(\underline{n})Y$$

Hence

$$\begin{aligned} E(\hat{\alpha}) &= (1/n)AM^-(p)L'(\underline{n})L(\underline{n})\theta \\ &= (1/n)AM^-(p)\{nM(p)\}\theta = A\theta = \alpha, \end{aligned}$$

and

$$\begin{aligned} D(\hat{\alpha}) &= (\sigma^2/n)AM^-(p)L'(\underline{n})L(\underline{n})M^-(p)A'/n \\ &= \sigma^2 AM^-(p)\{nM(p)\}M^-(p)A'/n^2 \\ &= \sigma^2\{AM^-(p)M(p)\}M^-(p)A'/n = \sigma^2 AM^-(p)A'/n. \end{aligned}$$

Hence

1.4.5 $$D(\hat{\alpha}) \propto AM^-(p)A'.$$

Some further consequences including properties that we would expect of $D(\mathcal{Q})$ can be derived from the following implication of 1.4.4.

Equation 1.4.4 may be rewritten as

1.4.6 $\quad A = Y'M(p)$

where $Y' = AM^{-}(p)$ for any generalised inverse.

Hence, for estimability of α, $M(p)$ must belong to $\mathcal{M}_A = \{M: M \in \mathcal{N}\mathcal{D}$ and $A = Y'M$ for some $Y\}$ and more generally if some matrix M is to be the matrix of a design under which α is estimable then it must be that $M \in \mathcal{M} \cap \mathcal{M}_A$.

Consequences of 1.4.6 to which reference will be made are the following:

<u>Consequence (i)</u> $\{M(p)\}\underline{z} = 0 \Rightarrow A\underline{z} = 0$, i.e. $\mathcal{N}\{M(p)\} \subseteq \mathcal{N}(A)$ where $\mathcal{N}(B))$ denotes the null space of B.

<u>Consequence (ii)</u> Equivalently any vector $\underline{u}$ which is a linear combination of the columns of A' is also a linear combination of the columns of $M(p)$, i.e. $\mathcal{R}\{M(p)\} \supseteq \mathcal{R}(A')$ where $\mathcal{R}(B)$ denotes the range space of B . In set terminology we have that

$$\{\underline{u} : \underline{u} = A'\underline{b},\ \underline{b} \in E_s\} \subseteq \{\underline{w} : \underline{w} = \{M(p)\}\underline{d},\ \underline{d} \in E_k\}$$

where E_t denotes t-dimensional Euclidean space. The result follows from the symmetry of $M(p)$ using the argument

$\underline{u} = A'\underline{b} \implies \underline{u} = \{M(p)\}\underline{d}$ where $\underline{d} = Y\underline{b}$.

<u>Consequence (iii)</u> A consequence of (ii) is that

$\text{rank}\{M(p)\} \geq \text{rank}\{(A)\} = s$, as we have already observed.

<u>Consequence (iv)</u> The matrix $AM^{-}(p)A'$ is nonsingular since

$\text{rank}\{AM^{-}(p)A'\} = \min\{\text{rank}\{(A)\}, \text{rank}\{M^{-}(p)\}\} = \text{rank}(A) = s$,

by dint of the facts that $\text{rank}\{M^{-}(p)\} \geq \text{rank}\{M(p)\}$ (See Graybill (1969), Theorem 6.6.8) and that, from (iii), $\text{rank}\{M(p)\} \geq \text{rank}(A)$.

<u>Consequence (v)</u> For any $M \in \mathcal{M}_A$ the matrix $AM^{-}A'$ is the same for any generalised inverse M^{-}. This follows since $AM^{-}A' = Y'MM^{-}MY = Y'MY$. In particular $AM^{+}A' = AM^{-}A'$.

Consequence (vi) For any $M \in \mathcal{M}_A$ and for any v in the range space $\mathcal{R}(M)$ of M the vector AM^-v is the same for any generalised inverse M^-. Since $v \in \mathcal{R}(M)$, there exists w such that $Mw = v$. Hence $AM^-v = AM^-Mw = Aw$ by an appeal to equation 1.4.4.

Consequence (vii) We finally note equivalent statements to (i) and (ii). Recall from 1.3.9 and 1.3.10 that $M(p) = V_pP_pV'_p$ where $\text{rank}\{M(p)\} = \text{rank}\{(V_p)\}$. It follows from Corollary 5.4.4.1 of Graybill (1969) that $\mathcal{R}\{M(p)\} = \mathcal{R}(V_p)$ and equivalently there is equality between their orthogonal complements, namely $\mathcal{N}\{M(p)\} = \mathcal{N}(V'_p)$. Further since the columns of V_p make up Sup(p) then $\mathcal{R}(V_p) = L\{\text{Sup}(p)\}$, the linear subspace spanned by Sup(p). Hence to offer estimability of $\alpha = A\theta$ a design p must be such that

$$\mathcal{R}(A') \subseteq L\{\text{Sup}(p)\}$$

This states that the coefficient vectors (the columns of A') of θ in the linear combinations $\alpha = A\theta$ must be linear combinations of the support vectors of p. Intuitively this is sensible, for then each component of α is linear in the terms $v'\theta$ for $v \in \text{Sup}(p)$, and, since estimability of the latter is guaranteed, then so is that of α.

Equivalently the orthogonal compliment of $L\{\text{Sup}(p)\}$ should be contained in $\mathcal{N}(A)$ i.e. $\mathcal{N}(V'_p) \subseteq \mathcal{N}(A)$.

§1.4.3 We finally turn to optimal design considerations. Clearly the claim must be that 'best' least squares estimators of $\alpha = A\theta$ will be provided by designs which make $AM^-(p)A'$ small among M(p) satisfying 1.4.4.

Specific criteria which have been proposed include designs which maximise $\phi(p) = \psi_i\{M(p), A\}$ for i = D, A, E, L where

$$\begin{aligned}
\psi_D\{M,A\} &= -\ln \det\{AM^-A'\} \\
\psi_A\{M,A\} &= -\text{tr}\{AM^-A'\} \\
\psi_E\{M,A\} &= -\lambda_{max}\{AM^-A'\} \\
\psi_L\{M,A\} &= -\text{tr}\{L(AM^-A')\}
\end{aligned}$$

where L is s×s and nonnegative definite.

Clearly these are generalisations of $\psi_i(M)$ for i = D,A,E,L and enjoy corresponding generalised statistical motivations. In another analogy one to one mappings of the first three are special cases of

$$\psi_t\{M,A\} = -\{(1/s)\,\mathrm{tr}(AM^-A')^t\}^{1/t}$$

The definition of the fourth criterion is possibly unnecessary since the constraint that L be nonnegative definite, needed to ensure that it is sensible to minimise $\psi_L\{M,A\}$, means that $L = B'B$ where B is $t \times s$, $\mathrm{rank}\{(B)\} = t$. Thus

$$\psi_L\{M,A\} = \psi_A\{M,AB\}$$

However equally

$$\psi_A\{M,A\} = \psi_L\{M,A\} \text{ for } L = I_s$$

Also

$$\psi_A\{M,A\} = \psi_L(M) = -\,\mathrm{tr}\{(LM^-)\} \text{ for } L = A'A.$$

The criterion in fact is linear in M^- and the class of such criteria is discussed in detail by Fedorov (1972), Tsay (1976a).

A particular case of $\psi_L(M)$ is given by $L = M(q) = \sum q_j v_j v_j'$, $q \in \mathcal{P}$, which could be regarded as a prior distribution on the design space. If $M^- = M^{-1}$ then

$$\psi_L(M) = \sum q_j v_j' M^{-1} v_j$$

which for $M = M(p)$ is proportional to a weighted average of the variances of the least squares of the parameters $\alpha_j = v_j'\theta$.

The observation that $\psi_L\{M,A\} \doteq \psi_A\{M,AB\}$ suggests, of course, that the criterion $\psi_L\{M,A\}$ is concerned about inferences relating to $\alpha = AB\theta$. This suggests further restrictions concerning L, namely that we should have $\mathcal{N}(M) \subseteq \mathcal{N}(AB)$ and more generally that $\mathcal{N}(M) \subseteq \mathcal{N}(A'LA)$ for general L.

The criteria $\psi_i\{M,A\}$ for i = D,L,E will be respectively referred to as D_A- optimality, L_A- optimality, E_A- optimality.

We now consider two particular cases of the matrix A and their implications for $\psi_t\{M,A\}$.

(i) The variable t in $\psi_t\{M,A\}$ only comes into play if $s \geq 2$, for if A is $(1\times k)$, as in $A = c'$ where c is $(k\times 1)$, then

$$\psi_t\{M,c'\} \quad = \quad - c'M^- c \quad \text{for any } t.$$

This criterion is known in the literature as c-optimality. To maximise $\phi(p) = \psi_t\{M(p),c'\}$ is to minimise $c'M^-(p)c$ which for a given p is proportional to the variance of the least squares estimator $c'\hat{\theta}$ of $c'\theta$.

Recently Pukelsheim (1979) reviews this criterion extensively.

(ii) If we are interested in estimating only the first s parameters $\theta_1,\ldots,\theta_s$ of θ then the relevant value of A is $[I_s : 0]$ where I_s is the $s\times s$ unit matrix and 0 is the $s\times(k-s)$ zero matrix. It follows that

$$AM^-A' = (M^-)_{11}$$

the leading $s\times s$ matrix in the partition

$$M^- = \begin{bmatrix} (M^-)_{11} & (M^-)_{12} \\ (M^-)'_{12} & (M^-)_{22} \end{bmatrix}$$

Let the corresponding partition of M(p) be

$$M(p) = \begin{bmatrix} M_{11} & M_{12} \\ M'_{12} & M_{22} \end{bmatrix}$$

We now illustrate that if M(p) satisfies the conditions 1.4.4 or 1.4.6 guaranteeing estimability of $\theta_1,\ldots,\theta_s$ then the $s\times s$ matrix M_{11} is nonsingular and singularity of M(p) will only arise if M_{22} is singular. We do this by showing that

$$\text{rank}\{(M_{11})\} = s,$$

$$\text{rank}\{(M)\} = s + \text{rank}\{(M_{22})\} \qquad , \quad M = M(p).$$

Let $M_1 = \begin{bmatrix} M_{11} & M_{12} \end{bmatrix}$, $M_2 = \begin{bmatrix} M'_{12} & M_{22} \end{bmatrix}$ so that

$$M = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}$$

Now the condition that $MY = A' = [I_s : 0]'$ for some Y implies that $M_1Y = I_s$. Hence $\mathcal{R}(M_1) \supset \mathcal{R}(I_s)$ and thus $\text{rank}\{(M_1)\} \geq s$. Since M_1 is $s \times k$ and $s \leq k$ we conclude that $\text{rank}\{(M_1)\} = s$.

Now appeal to corollary 12.2.20.1 of Graybill (1969). This states that $\mathcal{R}(M_{12}) \subseteq \mathcal{R}(M_{11})$ for any nongegative definite matrix M. It follows that $\mathcal{R}(M_{11}) = \mathcal{R}(M_1)$ and hence that $\text{rank}\{(M_{11})\} = \text{rank}\{(M_1)\} = s$.

To establish the second condition above we first show that there exists a matrix C of order $s \times (k-s)$ which is such that the $s \times s$ matrix $[M_{11} - CM_{22}C']$ is nonsingular.

By theorem 12.2.20 of Graybill (1969) there exists for any nonnegative definite matrix M a matrix C which is such that $CM_{22} = -M_{12}$. In fact $C = -X_1X_2'$ where $M = XX'$ and $X' = [X_1' : X_2']$ relates to an appropriate partitioning of X.

Now from the equation $MY = A'$ we have, for an appropriate partition $[Y_1' : Y_2']'$ of Y, that

$$M_{11}Y_1 + M_{12}Y_2 = I_s$$

$$M_{12}'Y_1 + M_{22}Y_2 = 0$$

It follows that $M_{12}Y_2 = -CM_{22}C'Y_1$ by the argument $M_{12}Y_2 = -CM_{22}Y_2 = CM_{12}'Y_1 = -CM_{22}C'Y_1$. Hence we have that

$$(M_{11} - CM_{22}C')Y_1 = I_s$$

thus establishing nonsingularity of $(M_{11} - CM_{22}C')$ in view of the fact that it is of order $s \times s$. We can therefore conclude that $\text{rank}\,(M_{11} - CM_{22}C') = s$.

To establish the result $\text{rank}\{(M)\} = s + \text{rank}\{(M_{22})\}$ consider that

$$\begin{bmatrix} I_s & C \\ 0 & I_{k-s} \end{bmatrix} \times \begin{bmatrix} M_{11} & M_{12} \\ M'_{12} & M_{22} \end{bmatrix} \times \begin{bmatrix} I_s & 0 \\ C' & I_{k-s} \end{bmatrix} = \begin{bmatrix} M_{11} - CM_{22}C' & 0 \\ 0 & M_{22} \end{bmatrix}$$

Denote the product on the left by N. Clearly

$$\text{rank}\{(N)\} = \text{rank}\{(M_{11} - CM_{22}C')\} + \text{rank}\{(M_{22})\} = s + \text{rank}\{(M_{22})\}$$

Now the matrix $\begin{bmatrix} I_s & C \\ 0 & I_{k-s} \end{bmatrix}$ is an upper triangular matrix with non-zero diagonal entries and so is nonsingular with a consequent rank of k. Thus $\text{rank}\{(N)\} \geqslant \text{rank}\{(M)\}$.

Finally

$$M = \begin{bmatrix} I_s & C \\ 0 & I_{k-s} \end{bmatrix}^{-1} \times \begin{bmatrix} M_{11} - CM_{22}C' & 0 \\ 0 & M_{22} \end{bmatrix} \times \begin{bmatrix} I_s & 0 \\ C' & I_{k-s} \end{bmatrix}^{-1}$$

and so by a similar argument

$$\text{rank}\{(M)\} > \text{rank}\{(M_{11} - CM_{22}C')\} + \text{rank}\{(M_{22})\}$$

Hence the desired result is established.

A distinguishing feature then of condition 1.4.6 in this context is that it implies nonsingularity of M_{11}. Only then is estimability of $\Theta_1, \ldots, \Theta_s$ and nonsingularity of the dispersion matrix of their least squares estimators guaranteed, a matrix which in this case is $(M^-)_{11}$. Nonsingularity of the matrix $(M^+)_{11}$ can in fact be established in the following way.

Let P, orthogonal, denote the eigenmatrix of M. Then

$$M = PDP' = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \times \begin{bmatrix} T & 0 \\ 0 & S \end{bmatrix} \times \begin{bmatrix} P'_{11} & P'_{21} \\ P'_{12} & P'_{22} \end{bmatrix}$$

where $D = \begin{bmatrix} T & 0 \\ 0 & S \end{bmatrix}$ is the diagonal matrix containing the eigenvalues of M, and T is nonsingular, and partitions are such that leading matrices are of order s×s. Then

$$M^+ = PD^+P' = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \times \begin{bmatrix} T^{-1} & 0 \\ 0 & S^+ \end{bmatrix} \times \begin{bmatrix} P'_{11} & P'_{21} \\ P'_{12} & P'_{22} \end{bmatrix}$$

Hence while $M_{11} = P_{11}TP'_{11}$, $(M^+)_{11} = P_{11}T^{-1}P'_{11}$. Since P is orthogonal, the matrix P_{11} is nonsingular and hence so is $(M^+)_{11}$.

We now derive a simplified expression for the matrix AM^-A' when $A = \left[I_s : 0\right]$.

Rhode (1965) derived the following counterpart partition of a generalised inverse of a partitioned symmetric matrix. If

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M'_{12} & M_{22} \end{bmatrix}$$

then

$$M^- = \begin{bmatrix} R^- & -(R^-M'_{12}M^-_{11}) \\ -(M^-_{11}M_{12}R^-) & M^-_{22} + M^-_{22}M'_{12}R^-M_{12}M^-_{22} \end{bmatrix}$$

where

$$R = M_{11} - M_{12}M^-_{22}M'_{12}$$

$$= M_{11} - M_{12}M^-_{22}M_{21}$$

This, of course, is similar to the partition of the inverse of a nonsingular matrix. Whatever formula is employed for R^-, M^-_{11}, M^-_{22} the above is always a generalised inverse of M. However employing R^+, M^+_{11}, M^+_{22} need not render $M^- = M^+$. An exception to this occurs if R is nonsingular, this being sufficient to guarantee $M^-MM^- = M^-$ and symmetry of MM^-, M^-M. In turn a sufficient condition for nonsingularity of R, established by Rhode, is that M_{11} be of full rank and $\text{rank}\{(M)\} = \text{rank}\{(M_{11})\} + \text{rank}\{(M_{22})\}$. We have already shown that this is satisfied by $M = M(p)$ if $M(p)$ satisfies condition 1.4.6 for $A = \left[I_s : 0\right]$. We thus obtain the simplification

1.4.7 $AM^{-}A' = (M^{-})_{11} = R^{-} = R^{-1} = \left\{M_{11} - M_{12}M^{-}{}_{22}M_{21}\right\}^{-1}$

$$\left\{M_{11} - M_{12}M^{+}{}_{22}M_{21}\right\}^{-1}$$

The latter step follows since $AM^{-}A' = AM^{+}A'$ under 1.4.6.

The particular case of maximising $\psi_D(M,A)$ is therefore equivalent to maximising $\log_e\det\{M_{11}-M_{12}M^{+}{}_{22}M_{21}\}$, a criterion which is traditionally known as D_s-optimality.

A further particular case of $\psi_t\{M,A\}$ clearly arises when $A \doteq I$, the function $\psi_t\{M,A\}$ becoming $\psi_t(M)$. All the particular criteria considered above, or one-to-one mappings of them, are therefore particular cases of $\psi_t\{M,A\}$.

Kiefer (1974) contemplates a wider class of criteria still in

$$\psi_+\{M,A,L,m,r\} \quad = \quad -\left\{\mathrm{tr}\left\{L\times(AM^{+}A')^{m}\right\}\right\}^{r}$$

and in

$$\psi\{M,A,L,m,r\} \quad = \quad -\left\{\mathrm{tr}\left\{L\times(AMA')^{m}\right\}\right\}^{r}$$

These though, in particular the latter, have less obvious statistical motivation.

§1.4.4 Some properties of $\psi_t\{M,A\}$ are now examined. The functions $\psi\{M,A,L,m,r\}$ and $\psi_+\{M,A,L,m,r\}$ also possess them in certain circumstances.

(a) The function $\psi_t\{M,A\}$ is

(i) concave on the set $\mathcal{PD}$ of positive definite matrices, when of course M is nonsingular and $M^{-} = M^{-1}$.

(ii) increasing on $\mathcal{PD}$; i.e., if $M_1 \succ M_2$ $\left(\text{i.e.} \{M_1 - M_2\} \in \mathcal{PD}\right)$ then $\psi_t\{M_1,A\} \geq \psi_t\{M_2,A\}$.

(iii) such that $\phi(p) = \psi_t\{M(p),A\}$ defines a function which is homogeneous of degree -1.

(b) Also the D-optimal criterion is invariant under a linear transformation of the design space.

Kiefer (1974) painstakingly establishes (a)(i) and (a)(ii), in particular concavity of $\psi\{M,A,L,m,r\}$ and $\psi_+\{M,A,L,m,r\}$ under a range of restrictions on L,m,r. These cover the particular case $\psi_t\{M,A\}$. Also other authors have established concavity for specific criteria; namely, Silvey (1974) for D_s-optimality, Sibson (1974a) for D_A-optimality, Ford (1976) for D-optimality, A-optimality, E-optimality and for other criteria as well.

It therefore would seem sufficient to briefly outline Kiefer's proof of concavity for $\psi_t\{M,A\}$, or equivalently convexity of

$$-\psi_t\{M,A\} = \{(1/s)\,\mathrm{tr}(D^t)\}^{1/t} \qquad , \qquad D = AM^-A'$$

His general approach is to find an increasing function $G(\cdot)$ such that $G[-\psi_t\{M,A\}]$ is convex. Three main results play a part.

(a) The matrix $D = AM^{-1}A'$ is convex in $M \in \mathcal{PD}$. This is a consequence of the result that, for $M_1, M_2 \in \mathcal{PD}$, $\{\alpha M_1^{-1} + (1-\alpha)M_2^{-1}\} \succ \{\alpha M_1 + (1-\alpha)M_2\}^{-1}$. See Ford (1976), Fedorov (1972).

(b) Let $\lambda_i(M)$ denote the i^{th} largest eigenvalue of $M \in \mathcal{PD}$.

Then $\sum_{i=1}^{m}\lambda_i(M)$ is convex in M; that is, if

$$a_i = \alpha\lambda_i(M_1) + (1-\alpha)\lambda_i(M_2),$$

$$b_i = \lambda_i\{\alpha M_1 + (1-\alpha)M_2\},$$

then, for $M_1, M_2 \in \mathcal{PD}$, it is the case that

$$\sum_{i=1}^{m} a_i \geq \sum_{i=1}^{m} b_i \quad , \qquad m = 1,2,\ldots,k,$$

with equality in the case m = k. See Fan (1959) and Beckenbach. and Bellman (1961, p.75).

Hence $\sum_{i=1}^{m}\gamma(a_i) \geq \sum_{i=1}^{m}\gamma(b_i)$ where $\gamma(\cdot)$ is a convex function.

(c) The function $\gamma(x) = \begin{cases} x^t, & t \geq 1, \; t < 0 \\ -x^t & 0 \leq t \leq 1 \end{cases}$

is convex on the positive real line.

A function $G\{-\psi_t(M,A)\}$, convex in M over $\mathcal{PD}$, is generated by a hierarchical application of these results. We note that for t = 1 convexity follows immediately from the result involving matrices M_1 and M_2 in (a), while concavity of D-optimality readily follows from a parallel result to that in (b); namely, for $M_1, M_2 \in \mathcal{PD}$,

$$\det\left\{\left[\alpha M_1 + (1-\alpha)M_2\right]_{mm}\right\} \geq \left[\det\{(M_1)_{mm}\}\right]^{\alpha} \times \left[\det\{(M_2)_{mm}\}\right]^{1-\alpha}$$

for $m = 1,..,k$, where M_{mm} is the leading $(m\times m)$ submatrix of M. See Bechenback and Bellman (1961, p.74).

The other properties are more easily dealt with.

The "increasing" property (ii) of $\psi_t(M,A)$ would seem to readily follow from the result that $\lambda_i(M_1) \geq \lambda_i(M_2)$ if $M_1 \succ M_2$, $\lambda_i(M)$ being the i^{th} largest eigenvalue of M. See Beckenback and Bellman (1961,p.72). It is this property which guarantees that an optimal M^* must lie on the boundary of $\mathcal{M}$.

The homogeneity property is basically a consequence of the linearity of M(p) in p.

The invariance property of D-optimality can be easily seen to follow from formula 1.3.6 for M(p). Suppose $\mathcal{V} = \{v_1,..,v_J\}$ is transformed to $\mathcal{W} = \{w_1,..,w_J\}$ under the linear transformation $w_j = Hv_j$, H is $(k\times k)$. Then a design assigning weight p_j to w_j has design matrix.

$$\begin{aligned} M(p) &= WPW' \\ &= HVPV'H', \end{aligned}$$

where V,W are respectively $k\times J$ matrices whose j^{th} column is v_j, w_j.

Hence

$$\det\{M(p)\} = \det(VPV') \times \left[\det(H)\right]^2$$

Some comments on these properties are in order.

Concavity on $\mathcal{PD}$ of course means that those M^*'s maximising $\psi_t\{M,A\}$ over $\mathcal{M} \subset \mathcal{PD}$ form a convex set. There may though be several p^* such that $M^* = M(p^*)$ even if there is a unique maximising M^*, an observation we have made with respect to the general problem (P2). While this is a desirable property, the "increasing" property and the homogeneity prove more useful in the formulation of an algorithm and with hindsight are properties to be expected. The former is really a manifestation of the fact that we selected criteria with a view to making $M(p)$ "big". An optimal $M(p^*)$ must therefore be "bigger" or at least as "big" as any other matrix $M(p)$. It would seem reasonable to claim that M_1 is at least as big as M_2 if $M_1 \succcurlyeq M_2$. This relationship is certainly satisfied by $M_i = n_i M(p)$, $i = 1,2$, when $n_1 \geqslant n_2$ for a given design p. The matrix $\sigma^2 M_i^{-1}$ is the covariance matrix of the least squares estimator of the parameter θ of the linear model 1.2.1, obtained from a sample of size n_i, $i = 1,2$, each sample allocating the same weight to v_j (assuming $n_i p_j$ integral). Clearly the larger sample corresponding to the matrix M_1 is more informative about θ. It follows that this must also be true in the case $M_i = M(p_i)$ if $M(p_1) \succcurlyeq M(p_2)$.

A particular consequence or reflection of this "increasing" property is that the function $\phi(p) = \psi_t\{M(p),A\}$ has positive derivatives, as is evident from the formulae at the end of this section. Examples 1.1.1, 1.1.2, 1.1.4 have already been seen to enjoy this property.

These same functions also enjoy the homogeneity property. In this design context it is a property which really must hold because there is a certain arbitrariness in changing from the real constraint

$\sum_{i=1}^{J} n_i = n$ of the exact design problem, where n is the available number of observations, to the constraint $\sum_{j=1}^{J} p_j = 1$,

of the approximate design problem. Why not

$\sum_{j=1}^{J} p_j = c$? Clearly the "optimum" n_i, suggested by an optimal p for a given n, should be independent of the choice of c. The homogeneity property implies this. If p^* is optimal for $c = 1$ then bp^* is optimal for $c = b$.

To conclude we note that in the ensuing chapters we shall take the following clearly equivalent criterion to $\psi_t\{M(p),A\}$ as our general design criterion; namely

$$\phi(p) = -\mathrm{tr}\{AM^-(p)A'\}^t \quad , \quad t > 0.$$

This is a homogeneous function of degree (-t) and, as shall be seen, has the positive derivatives.

$$\partial\phi/\partial p_j = \begin{cases} t v_j' M^{-1}(p)A'\{AM^{-1}(p)A'\}^{t-1} AM^{-1}(p)v_j \ , \text{ if } M(p) \text{ is nonsingular} \\ t v_j' M^{+}(p)A'\{AM^{+}(p)A'\}^{t-1} AM^{+}(p)v_j \ , \text{ if } v_j \in L\{Sup(p)\}, \end{cases}$$

where $L\{Sup(p)\}$ is the linear subspace spanned by Sup(p), while

$$\partial\phi/\partial p_j^+ = 0 \quad , \text{ if } v_j \notin L\{Sup(p)\}$$

where

$$\partial\phi/\partial p_j^+ = \lim_{\varepsilon\downarrow 0}\{[\phi(p+\varepsilon e_j) - \phi(p)]/\varepsilon\} \qquad (e_j = j^{th} \text{ unit vector})$$

Note that $\sum p_j \partial\phi/\partial p_j = t \times \mathrm{tr}(AM^+(p)A')^t = -t\phi(p)$, a consequence of the homogeneity of $\phi(p)$.

In particular if $A = I$ so that $\phi(p) = -\mathrm{tr}\{M^{-t}(p)\}$, then

$$\partial\phi/\partial p_j = t v_j' M^{-(t+1)}(p) v_j$$

As we have said these two features of $\phi(p)$, homogeneity of negative degree and positive derivative, will rise to play an important role in the formulation of an algorithm.

§1.5 Optimal Regression Design - A Review

To conclude this chapter we give a brief description of the chronological development of optimal regression design.

The first recognisable contribution in the field seems to be that of Smith (1918) in which she calculated optimal designs for polynomial regression models with what was in effect the G-optimal criterion in mind. She did not have the benefit of matrices, only tedious algebraic tools.

There then appears to have been no further work in the field until a cluster of papers in the 1940's; these being Wald (1943), Hotelling (1944), Rao (1946), Mood (1946). Wald's paper is a founding one.

The subject finally seems to take seed in the 1950's, contibutions including Box and Wilson (1951), Elving (1952), Chernoff (1953), de la Garza (1954), Elving (1955) Ehrenfield (1955), de la Garza et al (1955), de la Garza (1956), Guest (1958), Kiefer (1958, 1959), Elving (1959), Raghavaro (1959), Kiefer and Wolfowitz (1959).

These early papers were inevitably very specific. They have a particular criterion in mind and also a particular regression model, i.e. particular forms for the regression functions $f_j(\underline{x})$. The regression models were often polynomial regressions on the interval (0,1) or (-1,1) or were determined by weighing experiments. Finally the problem as a whole had to be such that analytic solutions could be derived. They were therefore free to consider the appropriate continuous design problem when $\mathcal{X}$ was continuous, and it was not natural to express the problem explicitly in terms of the transformed design space $\mathcal{V} = \underline{f}(\mathcal{X})$.

Kiefer is by far the foremost contributor in the field. His 1958 paper proved to be the prelude to over twenty publications some joint with others, the last as recently as Kiefer (1978). Not surprisingly he made the first major breakthrough in the field. Kiefer and Wolfowitz (1960) proved the equivalence of D-optimality and G-optimality but more importantly as a result they derived a necessary and sufficient condition for D-optimality. We will see the details in the next chapter.

This left the way open for the study of more ambitious problems and more importantly made possible the formulation of algorithms for the computation of D-optimum designs. However this was not immediately realised. Problems considered were still fairly specific including polynomial regression but possibly in several variables and sometimes with extrapolation in mind. See Hoel (1961a, 1961b, 1965a, 1965b), Hoel and Levine (1964), Clark (1965), Atwood (1969). See also Studden and Van Armin (1969) for optimum designs in the case of a spline regression model. On a more general level an almost successful attempt though was made to extend the equivalence theorem to the D_s-optimal criterion by Karlin and Studden (1966).

The challenge to formulate an algorithm was not taken up until Fedorov (1969) and Wynn (1970) devised almost identical iterative schemes for D-optimality. From this point on the subject has developed rapidly mainly on two fronts; namely, the formulation of further algorithms and the extension of the results of D-optimality to more general criteria. Authors who have formulated iterative schemes include Fedorov, Wynn, Atwood, Silvey, Titterington, Pazman, Sibson, Tsay, Wu. Details will be given in later chapters.

Whittle (1973), Kiefer (1974) were, apart from Karlin and Studden (1966), the first to extend the results of D-optimality. They both generalised the Equivalence Theorem. Other authors include Silvey and Titterington (1974), Sibson (1974a) , Silvey (1974), Fellman (1974), Fedorov and Malyutov (1972). Also more recently Pukelsheim (1980) has solved the design problem in the case of the general criterion $\psi_t\{M,A\} = -\mathrm{tr}\{AM^{-}A'\}^t$.

This essentially completes our review. It is only a skeleton and it must be emphasised that we only had in mind an optimum linear regression design problem. There are other topics in the field of design. There is that of optimum design in the case of nonlinear regression models. As we have said, a crucial change from the linear to the nonlinear design problem is that the matrix $D\{\hat{\theta}(\underline{x})\}$ then depends in a non-simple way on all unknown parameters. It is not possible therefore to determine an optimal design a priori. A sequential approach is needed. See Ford (1976), White (1973, 1975).

Another problem that has been studied in regression design is that of designing to discriminate between models. See Atkinson and Cox (1974), Fedorov and Atkinson (1975), Fedorov (1975, 1978), Hill (1976). Studies of this type to some extent attempt a response to a valid criticism of optimum regression design, namely that one can rarely if at all, fully know the true model.

We have not made reference to the subject of classical or factorial designs. These of course formulate linear models. One would not naturally view these as linear regression models, but one can of course do so. It is natural to consider whether or not these are optimum regression designs with respect to some of the criteria of section 1.4. See Ford (1976), Kiefer (1975a).

Finally optimum regression design has been considered in a Bayesian framework. See Brooks (1972, 1974, 1977).

The above is a very brief summary of the field of optimum design and by no means has the literature been exhausted. Extensive reference lists are contained in St. John and Draper (1975), Fedorov and Malyutov (1972), Fedorov (1972), Fellman (1974), Ford (1976), Hill (1976), Ash and Hedayat (1978), Federer and Ballam (1972).

Fedorov (1972) has been until recently the only English text on the subject of optimal regression design.

Ash and Hedayat (1978) is a review paper introducing a special issue of Communications in Statistics on Optimal Design Theory. The contents were contributed by Cheng; Fedorov; Kiefer; Kurotschka; Silvey; Titterington and Torsney; Studden; Wu. These therefore are among the most recent publications.

We conclude by concurring with Ash and Hedayat that, while Fedorov's book is important, there has clearly been a need for a more contemporary text on optimal regression design. We therefore welcome the appearance of the monograph, Silvey (1980).

CHAPTER 2

CONSTRAINED OPTIMALITY, GENERAL

§2.1 Introduction

The aim of this chapter is to determine conditions under which a p^* will be optimal for problem (P2).

First we note that there are two approaches which we could adopt in solving (P2). We could seek out an optimising p^* directly or first determine an x^* maximising $\psi(x)$ over $\mathcal{P}(\mathcal{U})$ and then find a p^* such that $x(p^*) = x^*$. The former approach, which in the main we will adopt, would require conditions explicitly defining an optimising p^*.

Of course, whatever approach we use, conditions defining the optimum of a constrained optimisation problem need to be defined. The conventional approach to such an exercise seems to be to employ lagrangian techniques, an approach which is developed in Whittle's text "Optimization under Constraints" (1971) and which is used in Rockafeller's book "Convex Analysis" (1970). Often this may well be the only feasible approach. However, the following back to basics pictorial scenario can lead to simple p^*-defining optimality conditions in the case of problem (P2).

Consider the problem of the constrained maximisation of a function $\psi(x)$ subject to $x \in S$, S a closed bounded continuous subset of n-dimensional space, of which (P2) is an example.

The function $\psi(\cdot)$ will sketch out a ψ-surface over S, a single peak mountain if $\psi(x)$ is strictly concave on S, a valley if $\psi(x)$ is convex on S, if such can be the case. Clearly one must be at a highest point on the ψ-surface in S if one is at a maximising x^*. Such an x^* may or may not lie on the boundary of S.

Suppose S had the form of the following two dimensional region

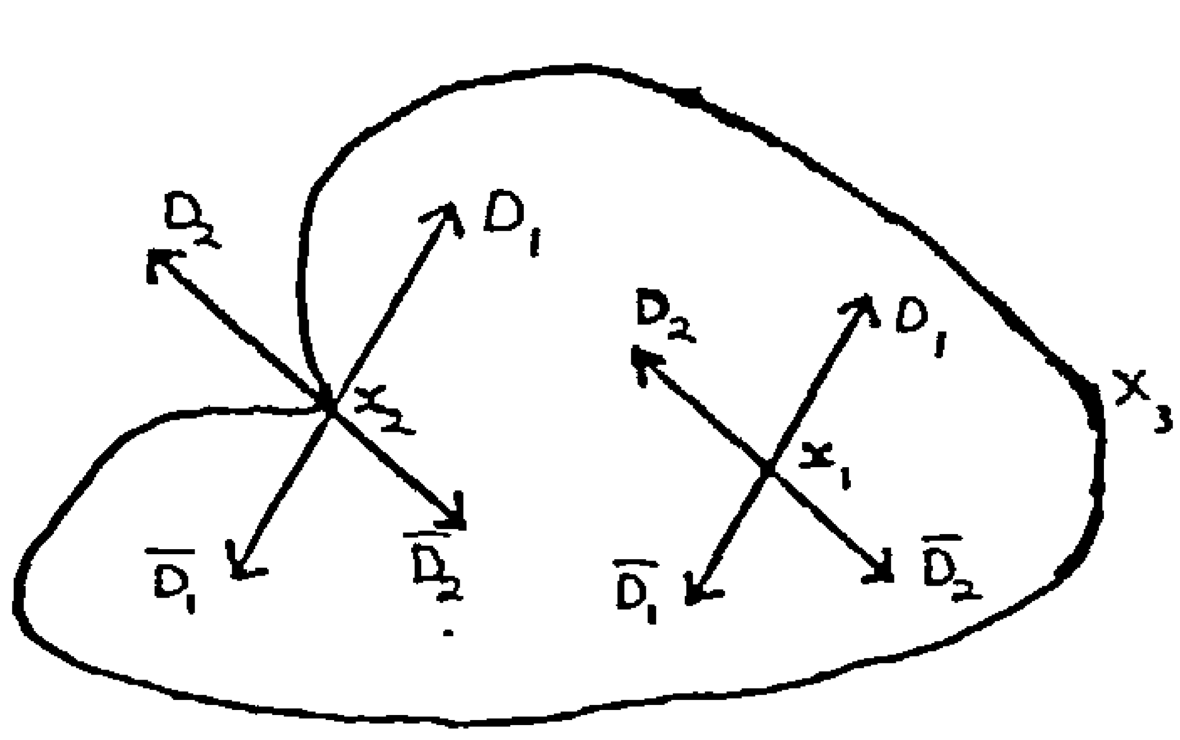


Figure 2.1

Consider the case when $x^* = x_1$ and hence lies in the interior of S and imagine a walk over the surface beginning at x^*. Then one cannot move uphill whatever direction one takes. In particular for a given direction D_i one cannot move uphill whether one moves off in that direction or about turns and moves off in the opposite direction $\bar{D}_i$.

Suppose now $x^* = x_2$. One now may walk uphill from x^* if the direction of the walk takes one immediately out of S as in the case of D_2. However, one still cannot do so if the direction does not immediately lead out of S as in the case of $\bar{D}_2$ and of D_1 and $\bar{D}_1$. Note that the pair D_1, $\bar{D}_1$ illustrate that it can be the case in n dimensions that one immediately remains in a set S both when one moves off in a particular direction D_1 or in the opposite direction $\bar{D}_1$ from a point x_2, even though x_2 lies on the boundary of S. This of course will not be true for all directions and need not hold for any direction as in the case of the boundary point x_3.

A converse picture can obviously be drawn if x^* is to be a minimising point in S.

For the case when $\psi(x)$ sketches a continuous surface this picture can be given familiar expression using the tools of calculus. In particular the phrase 'cannot move uphill from x^* in a given direction' will mean that the 'rate of change' or derivative of $\psi(\cdot)$ at x^* must be nonpositive in that direction. Hence we would have that the derivative of $\psi(\cdot)$ at a maximising x^* must be nonpositive in any direction which leads immediately into S from x^* and, in particular, in all directions from an interior point x^* of S.

Typically this nonpositivity will mean strict negativity in the case of a direction into S from a boundary point x^*, when the direction, as in the case of $\bar{D}_2$ in figure 2.1, is such that the opposite direction leads immediately out of S. Typically the derivative in the opposite direction will be positive.

Derivatives at a maximising x^* may also be strictly negative in directions leading into S whether x^* is a boundary point or not, if $\psi(\cdot)$ is not smoothly changing at x^* (not 'differentiable' at x^*).

In particular derivatives at x^* may be strictly negative in a pair of opposite directions which both lead immediately into S, such as D_1, $\bar{D}_1$ in the case of $x^* = x_1$ and of $x^* = x_2$. The derivative at $x^* = x_2$ may well also be negative in the direction D_2.

However, if $\psi(\cdot)$ is smoothly changing at x^* in all directions then a distinguishing feature is that, in contrast to the above, the derivative of $\psi(\cdot)$ at x^* must be zero in a pair of opposite directions which both lead immediately into S. A function cannot change smoothly and have negative derivatives (or positive derivatives) in both of two opposite directions. A maximum x^* must enjoy some features of the familiar property of stationarity.

This completes the introductory argument. We have used some ideas or terms loosely, in particular "derivative in a direction". We define a directional derivative in the next section and formalise the above picture on optimality in later sections of this chapter.

First, we note that several approaches in the spirit of the above, using a directional derivative, exist and various authors have brought them to bear on the design problem in particular; namely, Kiefer (1974), Silvey (1974), Whittle (1973).

Both Whittle and Rockafeller also include the definition of a directional derivative in their afore-mentioned texts, Whittle (1971) and Rockafeller (1970). Yet neither author makes much use of the concept in determining optimality conditions and moreover there is little evidence of it in other literature on constrained optimisation. As we have said earlier the emphasis is on the use of a lagrangian approach.

At first sight it seems surprising that the concept of a directional derivative has not been more fully developed and has not been applied to optimisation problems, constrained optimisation in particular. However, it may be that its application will only lead to simple optimality conditions in a particular type of problem such as (P2).

§2.2 A Directional Derivative F(x,y): General Properties

§2.2.1 In defining a directional derivative we adopt the convention employed by Whittle (1973). Then he had a general design criterion in mind.

Definition 2.2.1 Directional Derivative

The directional derivative of a function $\psi(\cdot)$ at a point x in the direction from x to the point y is defined to be

$$F(x,y) = \underset{\varepsilon \downarrow 0}{\mathrm{Lt}} \left\{ \left[\psi\{(1-\varepsilon)x + \varepsilon y\} - \psi(x) \right] / \varepsilon \right\}$$

assuming that the limit exists.

The choice of notation is due to a convention which regards this derivative as specifically a Fréchet derivative. Occasionally we will use the notation $F_\psi(x,y)$ when there is a need to emphasise which function is under consideration. In another context concerning influence curves the term F(x,y) has been referred to by Andrews et al (1972,p.30), as a Von Mises derivative. They refer to Von Mises (1947). See also Hampel (1968, 1971), Eplett (1980).

Whittle (1971) also uses the following alternative but equivalent definition, as does Rockafeller (1970).

Definition 2.2.2

The directional derivative of a function $\psi(\cdot)$ at x in the direction of the vector $\underline{m}$ is defined to be

$$\mathfrak{F}(x,\underline{m}) = \underset{\varepsilon \downarrow 0}{\mathrm{Lt}} \left\{ \left[\psi\{x + \varepsilon \underline{m}\} - \psi(x) \right] / \varepsilon \right\} .$$

Clearly $F(x,y) = \mathfrak{F}(x,\underline{m})$ where $\underline{m} = y-x$, while $\mathfrak{F}(x,\underline{m}) = F(x,x+\underline{m})$.

It will be seen that definition 2.2.1, which allows the direction $\underline{m}$, of interest, to be determined by a point y as above, is the more useful and indeed leads to a generalisation of some standard calculus.

We could at this stage state an optimality theorem, but this will be postponed until later.

§2.2.2 A number of properties of F(x,y) are collected together in this and the next section.

(G1) Sensibly the definition confines attention to that 'curve' sketched out by the ψ-surface over the 'line' joining x and y, namely the values of $f(\varepsilon)$ for $0 \le \varepsilon \le 1$ where $f(\varepsilon) = \psi\{(1-\varepsilon)x + \varepsilon y\}$.
If it exists, $F(x,y) = f'_+(0)$, the right hand derivative of $f(\varepsilon)$ at zero. Provided the value of $+\infty$ is allowed, (see Whittle (1973), Silvey (1974)), $F(x,y)$ does exist if $\psi(\cdot)$ is concave (so that $f''(0)<0$ for all x,y) and if $\psi(\cdot)$ is finite. This is so because the latter conditions guarantee that the function $\{f(\varepsilon)-f(0)\}/\varepsilon$ is a monotonic nonincreasing function of ε in $0<\varepsilon<1$.

Let $g(\varepsilon) = f(\varepsilon)-f(0)$. If $\psi(\cdot)$ is concave then $g(\varepsilon)$ is concave over $0\le\varepsilon\le1$ and passes through the origin. Let $0<\varepsilon_1<\varepsilon_2<1$. Then $\{(\varepsilon_2-\varepsilon)/\varepsilon_2\ ,\ \varepsilon_1/\varepsilon_2\}$ are weights, $(1-\lambda, \lambda)$ such that $\varepsilon_1=(1-\lambda)0+\lambda\varepsilon_2$. Hence $g(\varepsilon_1) \ge (1-\lambda)g(0)+\lambda g(\varepsilon_2) = \lambda g(\varepsilon_2)$ and so $g(\varepsilon_1)/\varepsilon_1 \ge g(\varepsilon_2)/\varepsilon_2$. Hence $g(\varepsilon)/\varepsilon$ is nonincreasing in ε over $0<\varepsilon\le1$ and so $F(x,y) = \operatorname{Lt}_{\varepsilon\downarrow0} g(\varepsilon)/\varepsilon$ is bounded from below if $\psi(\cdot)$ and hence $g(\cdot)$ is finite. A lower bound is given in (G4).

(G2) $F(x,y) = G(x,y-x)$
where

$$G(x,z) = \operatorname{Lt}_{\varepsilon\downarrow0}\left\{\left[\psi(x+\varepsilon z)-\psi(x)\right]/\varepsilon\right\}$$

This is the standard Gâteaux derivative and Kiefer (1974) used this concept in his design theory though he did not call it a directional derivative. It would not be appropriate to do so. Certainly it does not benefit from concavity of $\psi(x)$. However this representation of $F(x,y)$ in terms of $G(\cdot,\cdot)$ proves useful. Again the notation $G_\psi(x,z)$ may be used.

Note that $G(x,e_j) = \partial^+\psi/\partial x_j$, the right hand partial derivative of $\psi(\cdot)$ with respect to the j^{th} component of x, e_j being the j^{th} unit vector.

(G3) If $x,y \in S$, where S is a convex set, then so does $\{(1-\varepsilon)x+\varepsilon y\}$, which is clearly an advantage if one wishes $F(x,y)$ only for $x,y \in S$. In contrast $G(x,y)$ does not particularly benefit from such convexity.

(G4) $F(x,y) \geq \psi(y)-\psi(x)$ if $\psi(\cdot)$ is concave.
We have from the discussions in (G1) that

$$\begin{aligned} F(x,y) &\geq g(\varepsilon)/\varepsilon \quad \text{when} \quad 0 \leq \varepsilon \leq 1 \\ &= \left[\psi\{(1-\varepsilon)x + \varepsilon y\} - \psi(x)\right]/\varepsilon \\ &\geq \left[(1-\varepsilon)\psi(x) + \varepsilon\psi(y) - \psi(x)\right]/\varepsilon \ , \quad 0 \leq \varepsilon \leq 1 \\ &= \psi(y) - \psi(x). \end{aligned}$$

(G5) $F(x,x) = 0$, a desirable property since no change is effected in $\psi(\cdot)$ if one does not move from x. In contrast $G(x,x) = F(x,2x) \neq 0$.

(G6) Intuitively $F(x,y)$ in some sense measures the rate of change in $\psi(\cdot)$ at x in the direction of y. This is emphasised by the relationship in property (G1) that $F(x,y) = f'_+(0)$. However it does so in units of measurement which depend on the distance between x and y. $F(x,y)$ depends on this distance as well as on the said rate of change.

To move from point x in the direction of point y is to move from point x in the direction of the vector $m = y-x$ and hence in the direction of the vector $cm, c>0$. Passing along the full length of the vector cm from x we would, according to the theory of vectors, arrive at the point $\{x+c(y-x)\}$ as the following vector diagram illustrates.

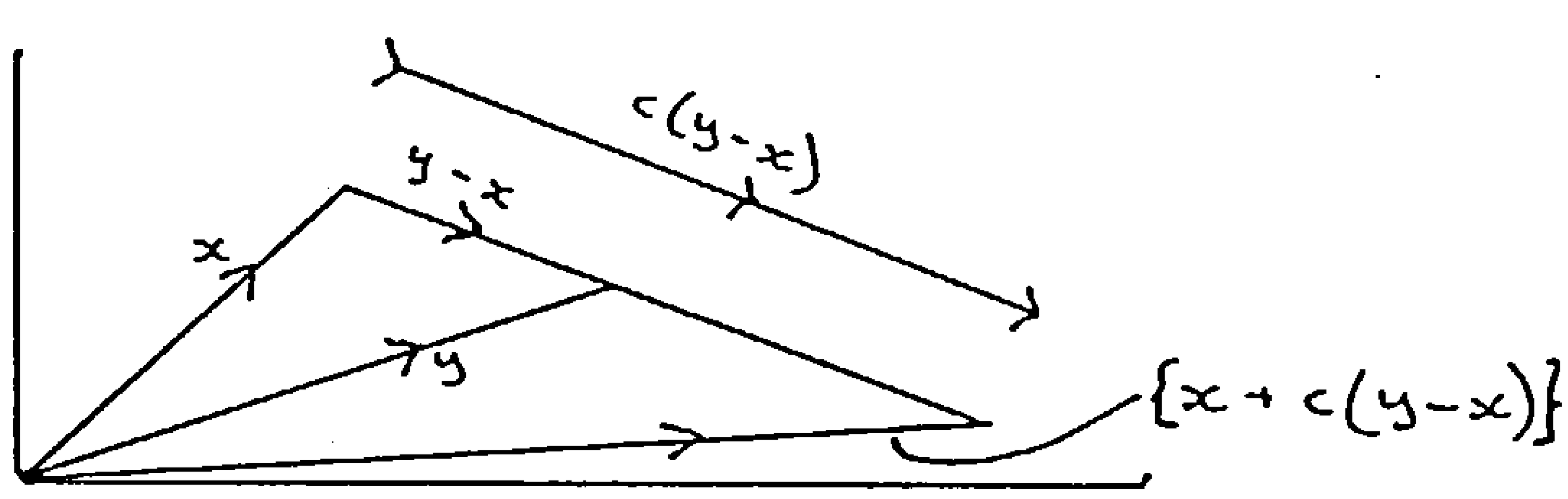


Figure 2.2.1

Hence $F\{x,x+c(y-x)\}$ measures the rate of change in $\psi(\cdot)$ at x in directions which remain the same for all positive c.

Not surprisingly we obtain,

$$\begin{aligned} F\{x, x+c(y-x)\} &= \underset{\varepsilon \downarrow 0}{Lt}\left\{\left[\psi\{(1-\varepsilon)x + \varepsilon[x+c(y-x)]\} - \psi(x)\right]/\varepsilon\right\} \\ &= \underset{\varepsilon \downarrow 0}{Lt}\left\{\left[\psi\{x+c\varepsilon(y-x)\} - \psi(x)\right]/\varepsilon\right\} \\ &= \underset{\delta \downarrow 0}{Lt}\, c\left\{\left[\psi\{x+\delta(y-x)\} - \psi(x)\right]/\delta\right\} \ , \quad \delta = c\varepsilon \\ &= \underset{\delta \downarrow 0}{Lt}\, c\left\{\left[\psi\{(1-\delta)x+\delta y\} - \psi(x)\right]/\delta\right\} \end{aligned}$$

that is

2.2.1 $\quad F\{x, x + c(y - x)\} = cF(x,y)$

This dependence on c indicates that F(x,y) cannot be thought of in any absolute sense as THE derivative of or rate of change in $\psi(\cdot)$ at x in the direction of y. To identify a quantity of such a nature we need to decide on a 'correct' value for c.

What precisely is the 'rate of change' measured by an ordinary derivative of a function of one variable g(x). The right hand derivative

$$g'_+(x) = \underset{\varepsilon \downarrow 0}{\mathrm{Lt}}\left\{\left[g(x+\varepsilon) - g(x)\right]/\varepsilon\right\},$$

is the amount by which a linear approximation to $g(\cdot)$ at x, valid in the region $(x, x + \delta)$, $\delta > 0$, (the tangent plane to $g(\cdot)$ at x in that region) will change for a unit increase in x, i.e. a step of one unit in the direction of x increasing. This argument holds good in both of the following pictures.

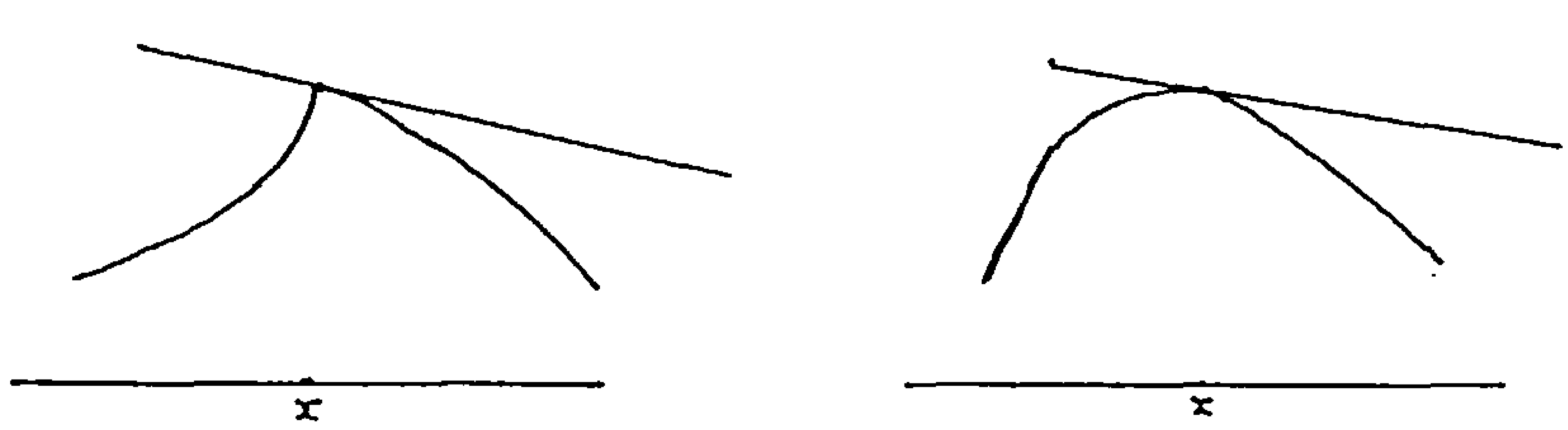


Figure 2.2.2

Recall now that $F(x,y) = f'_+(0)$ where $f(\varepsilon) = \psi\{(1-\varepsilon)x + \varepsilon y\}$. Since $f'_+(0)$ is the amount of change induced in the linear approximation to $f(\cdot)$ at 0 by a unit increase in ε, it follows that F(x,y) defines the amount of change induced in a corresponding linear approximation to $\psi(\cdot)$ at x by a step towards y, the magnitude of which is the distance between y and x, namely $\|y - x\| = \sqrt{(y-x)'(y-x)}$.

This suggests that we should calculate F(x,y) only for a y which is a unit distant from x. Certainly $g'_+(x) = F_g(x, x+1)$ in the case of a single variable function $g(\cdot)$. The problem however is that we will be presented with a y of interest which will not typically be a unit distant from x. Such a y must be scaled up or down appropriately. Clearly the solution is to choose the constant above such that $\{c(y-x)\}$ has unit length, namely c such that $c^{-1} = \|m\|$, $m = y - x$.

This creates the normalised directional derivative

$$F^{I}(x,y) = F(x,y)/\sqrt{m'm} \, .$$

Of course this uses only one particular norm. A more general normalised directional derivative would be

$$F^{A}(x,y) = F(x,y)/\sqrt{m'Am} \, ,$$

where A is symmetric nonnegative definite. This normalised version will be made use of later.

This completes the list of general properties, the remaining comments recording some relatives or generalisations of $F_\psi(x,y)$.

(i) A converse concept would be the directional derivative of $\psi(\cdot)$ at x as x is approached from the direction of y; namely

$$\overline{F}_\psi(x,y) = \operatorname*{Lt}_{\delta\uparrow 0}\left\{\left[\psi\{(1+\delta)x - \delta y\} - \psi(x)\right]/\delta\right\}$$

However

$$\overline{F}_\psi(x,y) = \operatorname*{Lt}_{\delta\uparrow 0}\left\{\left[\psi\{x + \delta(x-y)\} - \psi(x)\right]/\delta\right\} \qquad \varepsilon = -\delta$$

$$= \operatorname*{Lt}_{\varepsilon\downarrow 0} -\left\{\left[\psi\{x + \varepsilon(y-x)\} - \psi(x)\right]/\varepsilon\right\} = -F_\psi(x,y),$$

a result which is to be expected.

$\overline{F}(x,y)$ will clearly enjoy properties analogous to that of $F(x,y)$. In particular $\overline{F}_g(x, x-1) = g'_-(x)$, the left hand derivative at x of a function $g(\cdot)$ of one variable.

(ii) An offspring of $F_\psi(x,y)$ defines higher order directional derivatives of $\psi(\cdot)$ at x in the direction y, namely,

$$F_\psi^{(n)}(x,y) = d^n f(\varepsilon)/d\varepsilon^n\Big|_{\varepsilon=0^+} \, , \quad f(\varepsilon) = \psi\{(1-\varepsilon)x + \varepsilon y\} \, .$$

(iii) The term $G_\psi(x,z)$ clearly enjoys a parallel generalisation in the term $G_\psi^{(n)}(x,z)$ below, but it also extends to the other term $G_\psi^{(n)}(x; z_1, \ldots, z_n)$, and to combinations of them.

$$G_\psi^{(n)}(x,z) = d^n g(\varepsilon)/d\varepsilon^n\Big|_{\varepsilon=0^+} \, , \quad g(\varepsilon) = \psi(x+\varepsilon z) \, ; \qquad G_\psi^{(n)}(x; z_1, \ldots, z_n) = \partial^n g(\varepsilon_1, \ldots, \varepsilon_n)/\partial\varepsilon_1 \ldots \partial\varepsilon_n\Big|_{\varepsilon_i=0} \, , \quad g(\varepsilon) = \psi\{x + \Sigma\varepsilon_i z_i\}$$

The case $G_\psi^{(k)}(x; e_{i_1}, \ldots, e_{i_k})$ identifies higher order partial derivatives of $\psi(\cdot)$ at x, e_j being the j^{th} unit vector.

§2.3 Differentiability Defined; Further Properties Of $F(x,y)$

§2.3.1 So far we have not made any assumptions about differentiability as has been emphasised by the several occasions on which we elected to be specific and state results to hold for right hand derivatives. A function need not be differentiable at a point x in order that it should have well defined directional derivatives in all directions. Whittle (1971,p.61) quotes the following picture in support of this

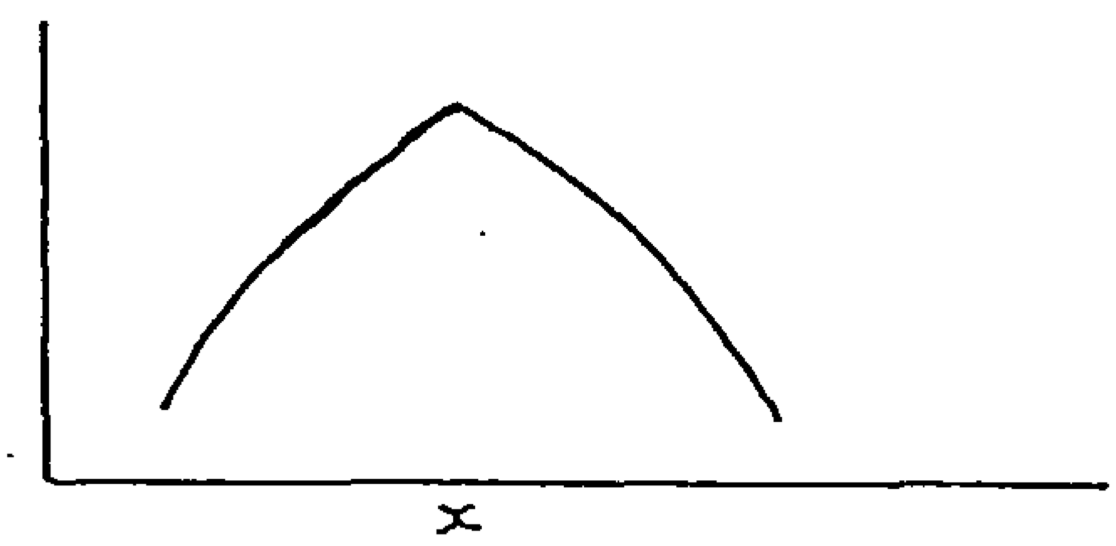


Figure 2.3.1

The picture illustrates the following, as stated by Whittle, "that a function could have a discontinuity in slope at a point x although sloping away from this point in a perfectly smooth fashion in any given direction".

However differentiability plays an important simplifying role in the calculus of optimisation and this is no less the case here. We now proceed to redefine the concept in terms of $F(x,y)$.

Of course the idea is that, at x, $\psi(\cdot)$ should be smoothly changing in all directions. A more precise definition is that, at x, the ψ-surface should just touch or possibly "cross in parallel" a unique linear hyper-plane, the tangent plane to $\psi(\cdot)$ at x, or the supporting hyperplane at x if the two surfaces do not cross. This plane would then provide a linear approximation to $\psi(\cdot)$ at x in any direction, so that the linear approximation to $\psi(\cdot)$ at x which it would suggest in the direction of y and in the opposite direction would be the "same" apart from a difference in sign.

For two surfaces to coincide in such a manner they must have some common characteristics at the point of contact x. In particular, apart from sharing a common value they must be changing at the same rate otherwise they will not be in parallel and will definitely cross. They must have common first derivatives, partial, directional, Gâteaux or whatever and hence whatever properties are enjoyed by the derivatives of one function at x must be enjoyed by those of the other function.

Consider then the form of the directional derivative of a linear function $L(x) = a'x + b$

$$F_L(x,y) = \underset{\varepsilon \downarrow 0}{Lt}\left\{\left[L\{x+\varepsilon(y-x)\} - L(x)\right]/\varepsilon\right\}$$
$$= \underset{\varepsilon \downarrow 0}{Lt}\left\{\left[a'[x+\varepsilon(y-x)] - a'x\right]/\varepsilon\right\}$$
$$= a'(y-x) = a'y - a'x = L(y) - L(x)$$

Similarly $G(x,y) = a'y$, and the vector of partial derivatives is $\partial L/\partial x = a$.

This suggests that for $\psi(\cdot)$ to be differentiable at x it must be that

$$F(x,y) = (y-x)'d \qquad \text{for all } y$$

or

$$G(x,y) = y'd \qquad \text{for all } y$$

where $d = \partial\psi/\partial x$, $d = (d_1,..,d_n)'$, $d_j = \partial\psi/\partial x_j$.

The condition on $G(x,y)$ is a familiar definition of differentiability. If we were to accept this as such a definition an equivalent and as we shall see a more useful form would be definition 2.3.1 applied to any function $\psi(\cdot)$. However, we restrict application of the definition to only concave functions to accord with Rockafeller (1970,p.244).

<u>Definition 2.3.1 Differentiability (of a concave function)</u>

A concave function $\psi(\cdot)$ is differentiable at a point x if

2.3.1(a) $$F(x, \textstyle\sum c_r y_r) = \sum c_r F(x,y_r) + (\sum c_r - 1)F(x,2x)$$

or

2.3.1(b) $$G(x, \textstyle\sum c_r y_r) = \sum c_r G(x,y_r)$$

These two conditions are equivalent as we shall see below. The latter, clearly the simpler, states that $G(x,y)$ must be linear in the second argument y, while in general this need not be the case with $F(x,y)$.

§<u>2.3.2</u> A whole host of properties follow from this definition. All of them assume differentiability at x and that $x \in S$.

(D1) $G(x,y) = y'd$ since $y = (y_1,\dots,y_n)' = \sum_{j=1}^{n} y_j e_j$ and $d_j = \partial\psi/\partial x_j = G(x,e_j)$.

Conversely suppose $G(x,y) = y'd$ for all y

Then

$$G\left(x,\sum c_r y_r\right) = \left(\sum c_r y_r\right)'d = \sum c_r y_r' d = \sum c_r G(x,y_r).$$

Hence condition 2.3.1(b) is equivalent to requiring that $G(x,y) = y'd$.

Of interest is that, according to theorem 25.2 of Rockafeller (1970,p.244), a sufficient condition for 2.3.1(b) to hold in the case of concave functions, is that the n two-sided partial derivatives exist at x and are finite.

(D2) $F(x,y) = G(x,y - x) = G(x,y) - G(x,x) = (y - x)'d$

i.e. $F(x,y) = m'd \quad , \quad m = y - x$

If, as in the design context, we regard the argument of the function $\psi(\cdot)$ as a symmetric $k\times k$ matrix A, then this result can be re-expressed in the form

$$F(A,B) = \text{trace}\left\{(B - A)\nabla\psi(A)\right\}$$

where $\nabla\psi(A)$ is the $k\times k$ matrix whose $(i,j)^{th}$ element is $\left(\partial\psi/\partial a_{ij}\right)(1+\delta_{ij})/2$.

(D3) A proof that 2.3.1(b) implies 2.3.1(a) is,

$$\begin{aligned} F\left(x,\sum c_r y_r\right) &= G\left(x,\sum c_r y_r\right) - G(x,x) \\ &= \sum c_r G(x,y_r) - G(x,x) && \text{(by 2.3.1(b))} \\ &= \sum c_r\left\{G(x,y_r) - G(x,x)\right\} + \left(\sum c_r - 1\right)G(x,x) \\ &= \sum c_r F(x,y_r) + \left(\sum c_r - 1\right)F(x,2x) && \text{(by G5)} \end{aligned}$$

A proof that 2.3.1(a) implies 2.3.1(b) is

$$\begin{aligned} G\left(x,\sum c_r y_r\right) &= F\left(x,x+\sum c_r y_r\right) \\ &= F(x,x) + \sum c_r F(x,y_r) + \left(1+\sum c_r - 1\right)F(x,2x) && \text{(by 2.3.1(a))} \\ &= \sum c_r F(x,y_r) + \left(\sum c_r\right)F(x,2x) && \text{(by G5)} \\ &= \sum c_r\left\{F(x,y_r) + F(x,2x)\right\} \\ &= \sum c_r\left\{F(x,x) + F(x,y_r) + F(x,2x)\right\} \\ &= \sum c_r F(x,x+y_r) && \text{(by 2.3.1(a))} \\ &= \sum c_r G(x,y_r) . \end{aligned}$$

(D4) $F(x,2x - y) = 2F(x,x) - F(x,y) + \left\{(2 - 1) - 1\right\}F(x,2x)$

i.e. $F(x,2x - y) = -F(x,y)$ (by G5)

or $\bar{F}(x,2x - y) = F(x,y)$

This really is the distinguishing feature of differentiability from nondifferentiability; that, as you pass through x in the direction of y, the rate of change in $\psi(\cdot)$ should be the same on the approach to and the departure from x. In the case of a function g(x) of a one-dimensional variable x, a consequence is that there is no need to distinguish between right hand and left hand derivatives, for

$$\underset{\varepsilon\uparrow 0}{\mathrm{Lt}}\left\{\left[g(x+\varepsilon)-g(x)\right]/\varepsilon\right\} = \underset{\varepsilon\downarrow 0}{\mathrm{Lt}}\left\{\left[g(x+\varepsilon)-g(x)\right]/\varepsilon\right\}$$

i.e., $$\overline{F}(x, x-1) = F(x, x+1).$$

(D5) If $\sum c_r = 1$ then clearly $F(x, \sum c_r y_r) = \sum c_r F(x, y_r)$.

This proves to be a very useful result for us, for suppose $S = \mathcal{P}(\mathcal{U})$, then if $y \in \mathcal{P}(\mathcal{U})$ then $y = x(q) = \sum q_j u_j$, $\sum q_j = 1$, $q_j \geq 0$.

$$\text{2.3.2} \quad \therefore \begin{cases} F(x, y) = \sum q_j F(x, u_j) \\ F\{x(p), x(q)\} = \sum q_j F\{x(p), u_j\} \end{cases}.$$

Note that the result would be true at any differentiable x whether or not $x \in \mathcal{P}(\mathcal{U})$. It only requires that $y \in \mathcal{P}(\mathcal{U})$. Also the constraints $q_j \geq 0$ could be relaxed.

We note also that this condition might only hold for certain types of y, say y such that $\mathrm{Sup}(q) \subseteq \mathrm{Sup}(p)$ in the case $x = x(p)$. We will see an example of this later.

This possibility might seem more reasonable when it is considered that we can regard x(p) as an element of the convex hull $\mathcal{P}(\omega)$ of any arbitrary finite discrete set ω such that $\mathrm{Sup}(p) \subseteq \omega$. The above linearity may not hold for $y = x(q)$ where q assigns non-zero weight to $\{\omega - \mathrm{Sup}(p)\}$. Conceivably this could also be the case if q assigns non-zero weight to a particular subset of Sup(p). This suggests that we should use the phrase '$\psi(\cdot)$ is differentiable at x(p) with respect to $\mathcal{U}$ or ω or Sup(p)' as the case may be. Only very occasionally will we use this phraseology, when it seems necessary, although strictly it is lacking in mathematical soundness.

We will now consider particular cases of 2.3.2.

If $S = \mathcal{P}$ and $\phi(\cdot)$ is a function of $p \in \mathcal{P}$ then $F_\phi(p, q) = \sum q_j F_\phi(p, e_j)$.

From this we obtain that

$$F_\phi(p,q) = \sum q_j\{G_\phi(p,e_j) - G_\phi(p,p)\}$$
$$= \sum q_j G_\phi(p,e_j) - G_\phi(p,p) \qquad (\sum q_j = 1)$$
$$= \sum(q_j - p_j)G_\phi(p,e_j) = \sum(q_j-p_j)\partial\phi/\partial p_j \, .$$

This result however also follows directly from (D2). A particular case of which we will make use is

2.3.3 $$F_\phi(p,e_j) = \partial\phi/\partial p_j - \sum p_i \partial\phi/\partial p_i$$

Note that if $\phi(p) = \psi\{x(p)\}$, $x(p) \in \mathcal{D}(\mathcal{U})$, $\partial\phi/\partial p_j = G_\phi(p,e_j) = G_\psi\{x(p),u_j\}$.

Considering $S = \mathcal{M}$ we conclude that

$$F_\psi\{M(p),M(q)\} = \sum q_j F_\psi\{M(p), v_j v_j'\} \, .$$

In a final example we take $S = C$, the cuboid in n-dimensional space with, for M>0, the 2^n vertices $M(\pm1,\pm1,,,,,\pm1)'$ which in the limit as $M \to \infty$ would yield the whole n-dimensional space. Let u_j denote one of these vertices. Then

$$F(x,u_j) = G(x,u_j) - G(x,x)$$
$$= M\{\sum_{j=1}^{n}(\pm)\partial\psi/\partial x_j\} - \sum_{j=1}^{n} x_j\partial\psi/\partial x_j$$

i.e.

2.3.4 $$F(x,u_j) = \sum_{j=1}^{n}(x_j \pm M)\partial\psi/\partial x_j$$

(D6) For $S = \mathcal{D}(\mathcal{U})$, $\sum p_j F\{x(p),u_j\} = 0$

This is so since $\sum p_j F\{x(p),u_j\} = F\{x(p),x(p)\}$.

In particular $\sum p_j F\{M(p), v_j v_j'\} = 0$

We note two particular consequences of this result.

(i) For at least one $u_r \in Sup(p)$, $F\{x(p),u_r\} > 0$ while for at least one $u_s \in Sup(p)$, $F\{x(p),u_s\} < 0$, unless $F\{x(p),u_j\} = 0$ for all $u_j \in Sup(p)$.

(ii) If a vector m is such that $\sum_{j=1}^{J} m_j = 0$, i.e. $\underset{\sim}{1}'m = 0$, where $\underset{\sim}{1}$ is a vector of 1's., then

2.3.5 $$F\{x(p), x(p) + \sum m_j u_j\} = \sum m_j F\{x(p),u_j\} = m'\underset{\sim}{F}$$

<u>Proof</u> $x(p) + \sum m_j u_j = x(q) = \sum q_j u_j$, $q_j = p_j + m_j$, $\sum q_j = 1$

$\therefore$ (by D5) $F\{x(p), x(p) + \sum m_j u_j\} = \sum p_j F\{x(p),u_j\} + \sum m_j F\{x(p),u_j\}$.

This result is of relevance in the formulation of algorithms for solving (P2).

(D7) If $S = \mathcal{P}(\mathcal{U})$ then

$$\max_{y \in S} F(x,y) = \max_{1 \le j \le J} F(x,u_j) = F(x,u^*)$$
$$\min_{y \in S} F(x,y) = \min_{1 \le j \le J} F(x,u_j) = F(x,u_*).$$

The proof is elementary. Since $y = x(q)$ for some $q \in \mathcal{P}$ we have

$$F(x,y) = \sum q_j F(x,u_j)$$

and hence

$$\left(\sum q_j\right) \min_{1 \le t \le J} F(x,u_t) \le F(x,y) \le \left(\sum q_j\right) \max_{1 \le r \le J} F(x,u_r)$$

The results follow since $\sum q_j = 1$ and the bounds can be attained by taking $y = u^*$ and $y = u_*$ respectively.

(D8) For $S = \mathcal{P}(\mathcal{U})$, $\max_{y \in S} F(x,y) \ge 0$, $\min_{y \in S} F(x,y) \le 0$.

This follows from (D6) and (D7).

(D9) An explicit solution can be obtained for the value of m which maximises $F_\phi^A(\theta, \theta + m) = d'm / \sqrt{m'Am}$ subject to $Cm = 0$. The solution can be

2.3.6 $$m = \pm\{A^- d - A^- C'(CA^- C')^- CA^- d\}. \qquad (d = \partial\phi/\partial\theta)$$

In view of 2.3.5, d could be replaced by $\underline{E}$ when $C = \underline{1}'$, $\theta = p$.

It is also the case that 2.3.6 can minimise $\{\phi(\theta) + d'm - (m'Am)/2\}$ subject to $Cm = 0$.

These results are relevant to the formulation of an algorithm for problem (P3) and in the case of $C = \underline{1}'$ for problems (P1) and (P2).

This completes the list of properties.

§2.4 Constrained Stationarity Defined

We have already said that differentiability plays an important simplifying role in optimisation. This is so because differentiability of a point x^* demands in the case of unconstrained optimisation that x^* can be a stationary value if it is to be an optimum. It can also prove to be the case in a constrained optimisation problem that a differentiable point x^* requires to be what we will call a constrained stationary value. In the case of problem (P2) this leads to simple optimality conditions.

In this section we concentrate on defining the simple concept of constrained stationarity and a complete list of formal optimality theorems will be given in the next section. However we mention the first of these after defining some terminology.

We repeat formally an assumption we made in section 2.1.

<u>Assumption 2.4.1</u>

Assume S to be a closed bounded continuous subset of n dimensional space.

An example is $S = \mathcal{P}(\mathcal{U})$.

For any $x \in S$ let S_x, B_x denote the following sets

2.4.1 $S_x = \{y : y \in S \text{ and } \exists\, \bar{\alpha}_y > 0 \text{ s.t. } [(1-\alpha)x + \alpha y] \in S \text{ for } 0 < \alpha \leq \bar{\alpha}_y\}$

2.4.2a $B_x = \{y : y \in S \text{ and } \exists\, \bar{\bar{\alpha}}_y > 0 \text{ s.t. } [(1 \mp \alpha)x \pm \alpha y] \in S \text{ for } 0 < \alpha \leq \bar{\bar{\alpha}}_y\}$

The set S_x is that collection of points $y \in S$ such that the path x to y remains immediately in S. The set B_x is the set of points y such that if one moves from x towards y or in the opposite direction one stays immediately in S.

By boundedness and closure of S we mean respectively that the elements of S are finite and that the resultant points of closure are included in S; they are the boundary points. Continuity of S will be guaranteed if the set S_x is nonempty for each $x \in S$, including boundary points.

If x is an interior point of S then $B_x = S_x = S$.

If x is a boundary point then

(i) S_x may or may not be a strict subset of S; see respectively x_2, x_3 of figure 2.1; if S is convex $S_x = S$ for all $x \in S$.

(ii) B_x may or may not be empty; for x_2, x_3 of figure 2.1, B_{x_2} is nonempty, B_{x_3} is empty; if x were to lie on a linear boundary plane to S, then B_x would contain only other boundary points; if S is convex then B_x reduces to

2.4.2b $$B_x = \{y : y \in S \text{ and } \exists\, \bar{\bar{\alpha}}_y > 0 \text{ s.t.}, [(1+\alpha)x - \alpha y] \in S \text{ for } 0 < \alpha \leq \bar{\bar{\alpha}}_y\}.$$

The set B_x could be thought of as defining 'S-internal approach to x paths'; that is, a path through x, whose direction is such that the immediate approach to x is contained in S. This is the case for paths through x in the direction x to y for $y \in B_x$.

Conversely the set $(S_x - B_x)$ defines 'S-external approach to x paths'; that is, paths through x, whose direction is such that the immediate approach to x is outside S; S is not entered until arrival at x.

<u>Theorem 2.4.1 A constrained local maximum</u>

The point x^* is a constrained local maximum of $\psi(x)$ on S if

$$F_\psi(x^*, y) \leq 0 \quad \forall\, y \in S_{x^*},$$

or equivalently,

$$\max_{y \in S_{x^*}} F_\psi(x^*, y) \leq 0. \qquad \square$$

The truth of this theorem is self evident, provided the practical interpretation of $F_\psi(x^*,y)$ as a rate of change is accepted. It is simply a restatement in mathematical language of the condition that $\psi(\cdot)$ must decrease away from x^* along paths leading immediately into S. No conditions are imposed on $F(x^*,y)$ for $y \in S - S_{x^*}$, i.e. those y's such that the direction x^* to y leads immediately out of S.

A reversal of the inequalities would define a constrained local minimum.

The theorem however is of little practical value. This though would be true of a corresponding theorem for an unconstrained local optimum for we have not made any assumption of differentiability.

Theorem 2.4.2

If x^* is a differentiable point of $\psi(\cdot)$ and is a constrained local maximum of $\psi(\cdot)$ on S then

$F(x^*,y) = 0$ for all $y \in B_{x^*}$.

Proof The set B_{x^*} defines those y's such that the direction x to y and the opposite direction both lead immediately into S. Such an opposite direction is in particular the direction x^* to $(2x^* - y)$. Hence while $(2x^* - y)$ may not belong to S it must be, for all small positive β, that $u=\{(1-\beta)x^* + \beta v\} \in S$, $v = 2x^* - y$.

From 2.2.1 $\qquad F(x^*, u) = \beta F(x^*, v)$

From (D4) $\qquad F(x^*, y) = -F(x^*, 2x^* - y) = -F(x^*, v)$

From theorem 2.4.1 $\qquad F(x^*, y) \leq 0$

and $\qquad F(x^*, u) \leq 0$

$\Longrightarrow \qquad F(x^*, v) \leq 0$

Hence we must have $F(x^*,y) = F(x^*,2x^* - y) = 0$. $\square$

This shows that a local maximum x^* of $\psi(\cdot)$ in S, must, if B_{x^*} is not empty, be an example, of the following.

Definition 2.4.1 A Constrained Stationary Value

A differentiable point $x \in S$ of $\psi(\cdot)$ is a constrained stationary value of $\psi(\cdot)$ with respect to the set S if

$F(x,y) = 0$ for all $y \in B_x$,

assuming the set B_x is not empty.

Theorem 2.4.3

Let $f(\varepsilon) = \psi\{(1-\varepsilon)x + \varepsilon y\}$, $\quad F^{(2)}(x,y) = d^2f(\varepsilon)/d\varepsilon^2\big|_{\varepsilon=0^+}$.

A constrained stationary value x of $\psi(\cdot)$ with respect to S must be one and only one of the following; a constrained local maximum in S, a constrained local minimum in S, a constrained saddle point with respect to S, a boundary point which is none of these. These types have the respective definitions

(a) (i) $F^{(2)}(x,y) \leq 0 \quad \forall\, y \in B_x$

(ii) $F(x,y) \leq 0 \quad \forall\, y \in (S_x - B_x)$

(b) (i) $F^{(2)}(x,y) \geq 0 \quad \forall\, y \in B_x$

(ii) $F(x,y) \geq 0 \quad \forall\, y \in (S_x - B_x)$

(c) (i) For at least one, $y \in B_x$ $F^{(2)}(x,y) \geq 0$, $F^{(2)}(x, 2x-y) \leq 0$
or " ≤ 0, -- ≥ 0

(ii) No condition on $F(x,y)$, $y \in S_x - B_x$.

(d) (i) $F^{(2)}(x,y) \leq 0 \quad \forall y \in B_x$
or " $\geq 0 \quad \forall y \in B_x$

(ii) $\exists\, y_1, y_2 \in S_x - B_x$ s.t. $F(x,y_1) < 0$, $F(x,y_2) > 0$. □

There is no need for a formal proof. In the case of (a), (b) and (c) conditions (i) are a version of the usual second order conditions needed to define the appropriate type of stationary point in a function of one variable. They must be satisfied here in the stationarity associated 's-internal approach to x paths'. The conditions (ii) are necessary extra boundary associated 'first order' conditions which must be satisfied by the appropriate directional derivatives along the 's-external approach to x paths'.

Class (d) acknowledges that while the right second order conditions for a local maximum or a local minimum may be satisfied along paths x to y having an 's-internal approach to x', it may be that the right 'first order' conditions will not be satisfied along all paths x to y having an 's-external approach to x'.

<u>Theorem 2.4.4</u>

Suppose $S = \mathcal{P}(\mathcal{U})$ and that $x(p) \in S$ $(p \in \mathcal{P})$ is a differentiable point of $\psi(\cdot)$. Then $x(p) = \sum_{j=1}^{J} p_j u_j$ is a constrained stationary value of $\psi(\cdot)$ with respect to S if and only if

2.4.3 $$F\{x(p), u_j\} = 0 \quad \text{when } p_j > 0, \text{ i.e. } u_j \in Sup(p).$$

<u>Proof</u> We require to show that 2.4.3 is necessary and sufficient to guarantee $F\{x(p),y\} = 0$ for all $y \in B_{x(p)}$, which, since S is convex, can be taken in the form 2.4.2(b). Let $x = x(p)$ throughout.

<u>Necessity</u> We show that $u_r \in B_x$ iff $p_r > 0$,

i.e, $\exists\ \bar{\bar{\alpha}}_r > 0$ s.t., $z_r = [(1+\alpha)x - \alpha u_r] \in S$ for $0 < \alpha \leq \bar{\bar{\alpha}}_r$ iff $p_r > 0$.

We will have $z_r = \sum \beta_j u_j$, $\beta_r = (1+\alpha)p_r - \alpha = p_r + \alpha(p_r - 1)$,
$\beta_i = (1+\alpha)p_i$, $i \neq r$.

Clearly $\sum \beta_j = 1$ and $\beta_i \geq 0$ for $i \neq r$, while $\beta_r = -\alpha$ if $p_r = 0$.
But if $p_r > 0$, $\beta_r \geq 0$ for $0 < \alpha \leq \bar{\bar{\alpha}}_r$, $\bar{\bar{\alpha}}_r = p_r/(1-p_r) > 0$.

So for $y = u_r$, $\exists\, \bar{\bar{\alpha}}_y (= \bar{\bar{\alpha}}_r) > 0$ s.t., $[(1+\alpha)x - \alpha y] \in S$ for $0 < \alpha \leq \bar{\bar{\alpha}}_y$ iff $p_r > 0$.
i.e. $u_r \in B_x$ iff $p_r > 0$

We note that this illustrates that $x(p)$ is a boundary point of $\mathcal{P}(\mathcal{U})$ if at least one p_r is zero.

Sufficiency Let a typical $y = x(\gamma) = \sum \gamma_j u_j \in S$.

We first show that $y \in B_x$ ie $\exists\, \bar{\bar{\alpha}}_y > 0$ s.t. $z = [(1+\alpha)x - \alpha y] \in S$ for $0 < \alpha \leq \bar{\bar{\alpha}}_y$ iff $\gamma_j = 0$ whenever $p_j = 0$

We will have $z = \sum \beta_j u_j$, $\beta_j = (1+\alpha)p_j - \alpha\gamma_j = p_j + \alpha(p_j - \gamma_j)$

Clearly $\sum \beta_j = 1$. Also it follows that

(a) $\beta_j = 0$ if $p_j = \gamma_j = 0$

(b) $\beta_j = -\alpha\gamma_j$ if $p_j = 0$ and hence $\beta_j < 0$ if $\gamma_j > 0, \alpha > 0$ in which case $z \notin S$, $y \notin B_x$

(c) $\beta_j \geq 0$ $\forall \alpha > 0$ if $p_j \geq \gamma_j \geq 0$ while $\beta_j > 0$ for $0 < \alpha < p_j/(\gamma_j - p_j)$ if $0 < p_j < \gamma_j$

$\therefore \forall r$ s.t. $p_r > 0$, $\beta_r \geq 0$ if $0 < \alpha \leq \bar{\bar{\alpha}}_y$, $\bar{\bar{\alpha}}_y = \min_{\gamma_j > p_j > 0} \{p_j/(\gamma_j - p_j)\}$.

Result (b) illustrates that it is necessary and results (a) and (c) illustrate that it is sufficient to put $\gamma_j = 0$ whenever $p_j = 0$ in order to ensure $\beta_j \geq 0$ for small $\alpha > 0$ under which condition $z \in S$ and hence $y \in B_x$

The sufficiency of 2.4.3 now follows from the fact that, since x is differentiable, then for $y = x(\gamma)$, $F(x, y) = \sum \gamma_j F\{x(p), u_j\}$. □

Under 2.4.3 we will certainly have $F\{x(p), x(\gamma)\} = 0$ if $\gamma_j = 0$ whenever $p_j = 0$.

Corollaries

(i) If $S = \mathcal{M}$, $M(p)$ is a constrained stationary value of $\psi(\cdot)$ with respect to $\mathcal{M}$ iff

2.4.4 $$F\{M(p), v_j v_j'\} = 0 \text{ . when } p_j > 0 .$$ □

(ii) If $S = \mathcal{P}$, p is a constrained stationary value of $\phi(\cdot)$ with respect to $\mathcal{P}$ iff

2.4.5 $$\partial\phi/\partial p_j = \sum_{i=1}^{J} p_i \partial\phi/\partial p_i \qquad (\text{by } 2\cdot3\cdot3)$$ □

(iii) Take $S = C$ as in (D5) and suppose that M has been chosen so large that x is an interior point of C, in which case $x = \sum p_j u_j$ where $p_j > 0$ for all 2^n vertices $u_j = M(\pm 1, \pm 1, \ldots, \pm 1)'$. Hence if x is to be a stationary value we will want $\sum_{j=1}^{n} (x_j \pm M)\partial\psi/\partial x_j = 0$ for all 2^n vertices whatever M is. Clearly under certain regularity conditions on $\psi(\cdot)$ this will be true iff

$$\partial\psi/\partial x_j = 0 \ , \ j = 1, \cdots, n$$ □

This then recovers the condition for an ordinary unconstrained stationary value. Hence in definition 2.4.1 or at least in equation 2.4.3 we have an explicit generalisation of the latter.

§2.5 Optimality Theorems

§2.5.1 We are now in a position to state a series of optimality theorems. Their truth is, in the main, self evident in the light of the results of the previous section. Assumption 2.4.1 is maintained and $F^{(2)}(x,y)$ is as in Theorem 2.4.3.

Theorem 2.5.1

The point x^* is a constrained local maximum of $\psi(x)$ on S if

$$F(x^*, y) \leq 0 \quad \forall\ y \in S_{x^*}$$

that is

$$\max_{y \in S_{x^*}} F(x^*, y) \leq 0. \qquad \square$$

This is a restatement for completeness of theorem 2.4.1.

Theorem 2.5.2

If x^* is a differentiable point of $\psi(\cdot)$, then it is a constrained local maximum on S if

(i) x^* is a constrained stationary value of $\psi(\cdot)$ on S (assuming the set B_{x^*} is nonempty) such that

$$F^{(2)}(x^*, y) \leq 0 \quad \forall\ y \in B_{x^*}$$

(ii) $$F(x^*, y) \leq 0 \quad \forall\ y \in \left(S_{x^*} - B_{x^*}\right). \qquad \square$$

This is just a restatement of the appropriate part of Theorem 2.4.3.

Corollary: If B_{x^*} is nonempty then $\max_{y \in S_{x^*}} F(x^*, y) = 0$, a consequence of the strict stationarity of x^*. $\square$

These are theorems on local maximality, but of course typically, as in (P1) and (P2), we desire to establish global optimality on a feasible region. In order that we could conclude that a global maximum had been identified we would require further knowledge about $\psi(\cdot)$, such as that there could be only one local maximum, a unique x^* satisfying the appropriate theorem. Such of course is the case if $\psi(\cdot)$ is concave on S, a property which usually goes hand in hand with convexity of S.

If S is convex then $S_x = S$ and $\psi(\cdot)$ will be concave if $F^{(2)}(x,y) \leq 0$ for all $x,y \in S$.

Theorem 2.5.3

If S is convex and $\psi(x)$ is concave on S then x^* maximises $\psi(\cdot)$ on S iff

$$F(x^*, y) \leq 0 \quad \forall y \in S$$

i.e. $\max_{y \in S} F(x^*, y) \leq 0$.

An alternative but equivalent necessary and sufficient condition is that

$$F(y, x^*) \geq 0 \quad \forall y \in S$$

i.e. $\min_{y \in S} F(y, x^*) \geq 0$. □

The first condition clearly follows directly from the preceding theorems and it is one which is expressed in a similar vein to the picture presented in section 2.1. It states that we must not immediately start to climb if we move from x^* towards any point $y \in S$. It is for us the more natural viewpoint from which to express necessary and sufficient optimality conditions.

However concavity allows of the second set of conditions via (G4), for if we are on a single peak ψ-surface, then we cannot start to go downhill if we move from any $y \in S$ towards x^*. If the concavity of $\psi(\cdot)$ is strict then we will have $F(y,x^*) > 0$ for $y \neq x^*$. The result has been included mainly for interest and completeness although use of it will be made in section 4.3.3.

Theorem 2.5.4

If S is convex, $\psi(x)$ is concave on S and x^* is a differentiable point of $\psi(\cdot)$, then x^* maximises $\psi(\cdot)$ on S if

(i) x^* is a constrained stationary value of $\psi(\cdot)$ on S, assuming the set B_{x^*} is nonempty and

(ii) $F(x^*,y) \leq 0$ for all $y \in (S - B_{x^*})$ □

Corollary If B_{x^*} is nonempty then $\max_{y \in S} F(x^*, y) = 0$. □

Note these conditions stem from those of the first part of theorem 2.5.3. There is no 'stationarity analogue' of the alternative conditions there. Differentiability of x^* does not make it possible to simplify them. In general they will be a more complex set of conditions.

The truth of theorem 2.5.4 is fairly clear. So also is that of the other theorems as we have said. Apart from the stationarity conditions they are little short of a restatement in mathematical language of conditions necessary for a maximum. They will only be of practical value if they suggest an explicit solution for x^*, and that is unlikely if x^* lies on the boundary of S. The optimality conditions they define are infinite. If we employ numerical techniques to compute the optimum and the solution they suggest is in fact the correct solution, it may not be easy to verify this. These observations may explain why the directional derivative has not played a more prominent part in constrained optimisation.

The above criticisms can also be made of the next theorem but not of the succeeding one.

<u>Theorem 2.5.5</u>

If $S = \mathcal{P}(\mathcal{U})$ and $\psi(x)$ is concave on S, then $x(p^*)$, $p^* \in \mathcal{P}$, maximises $\psi(\cdot)$ on S iff

$$\left.\begin{aligned} &F\{x(p^*), x(q)\} \le 0 \\ \text{i.e.}\quad &\max_{q \in \mathcal{P}} F\{x(p^*), x(q)\} \le 0 \end{aligned}\right\} \begin{aligned} &\forall\, q \in \mathcal{P} \\ &\text{i.e. } \forall\, x(q) \in S \end{aligned}$$

An alternative but equivalent condition is that

$$F\{x(q), x(p^*)\} \ge 0$$

i.e.
$$\min_{q \in \mathcal{P}} F\{x(q), x(p^*)\} \ge 0 . \qquad \square$$

<u>Theorem 2.5.6 Vertex Direction Optimality Theorem</u>

If $S = \mathcal{P}(\mathcal{U})$, $\psi(x)$ is concave on S and $x(p^*)$ is a differentiable point of $\psi(\cdot)$, then $x(p^*)$ maximises $\psi(\cdot)$ on S iff

(i) $$F\{x(p^*), u_j\} = 0 \quad \text{when} \quad p_j^* > 0$$

(ii) $$F\{x(p^*), u_j\} \le 0 \quad \text{"} \quad p_j^* = 0 .$$

<u>Proof</u> By theorem 2.5.4 $x(p^*)$ must be a constrained stationary value and by theorem 2.4.4 the conditions (i) are necessary and sufficient for this.

Let $y = x(q) = \sum q_j u_j$. We require to show that condition (ii) is necessary and sufficient to ensure the condition of Theorem 2.5.4 that $F(x^*, y) \le 0$ for all $y \in (S - B_{x^*})$.

It is clearly necessary, and sufficiency follows by the argument,

$$F\{x(p^*), x(q)\} = \sum_{r=1}^{J} q_r F\{x(p^*), u_r\} = \sum_{\{p_j = 0\}} q_j F\{x(p^*), u_j\} \le 0. \qquad \square$$

Clearly we have here a much more powerful theorem than the others, specifying a finite set of optimality conditions. It should be easy to check whether or not these are satisfied by a postulated solution obtained by numerical techniques. Differentiability though is an essential requirement.

Corollary (i)

If $S = \mathcal{M}$, $\psi(M)$ is concave on $\mathcal{M}$ and $M(p^*)$ is a differentiable point of $\psi(\cdot)$, then $M(p^*)$ maximises $\psi(\cdot)$ on $\mathcal{M}$ iff

(i) $F\{M(p^*), v_j v_j'\} = 0$ when $p_j^* > 0$

(ii) $F\{M(p^*), v_j v_j'\} \leq 0$ when $p_j^* = 0$ $\square$

Corollary (ii)

If $S = \mathcal{P}$, $\phi(p)$ is (weakly) concave on $\mathcal{P}$ and p^* is a differentiable point of $\phi(\cdot)$ on $\mathcal{P}$, then p^* maximises $\phi(\cdot)$ on $\mathcal{P}$ iff

(i) $\partial\phi/\partial p_j^* = \left\{\sum_{i=1}^{J} p_i^* \partial\phi/\partial p_i^*\right\}$ when $p_j^* > 0$

(ii) $\partial\phi/\partial p_j^* \leq \{$ " $\}$ " $p_j^* = 0$ $\square$

Corollary (iii) p^* solves $\min\limits_{p \in \mathcal{P}} \max\limits_{y \in \mathcal{P}(\mathcal{U})} \{F\{x(p), y\}\}$

Proof

From (D7) $\max\limits_{y \in \mathcal{P}(\mathcal{U})} \{F(x, y)\} = \max\limits_{1 \leq j \leq J} \{F(x, u_j)\}$

Clearly $\max\limits_{1 \leq j \leq J} \{F\{x(p^*), u_j\} = 0$

Consequently $\max\limits_{y \in \mathcal{P}(\mathcal{U})} \{F\{x(p^*), y\}\} = 0$

From (D8) $\max\limits_{y \in \mathcal{P}(\mathcal{U})} \{F\{x(p), y\}\} \geq 0 \quad \forall p.$

Thus p^* attains what is a lower bound for other p. $\square$

Theorem 2.5.6 was in fact derived by Whittle (1973) but only with a general optimum design problem in mind. So also did Kiefer (1974) though using the Gateaux derivative. Wu (1976) derived it by appealing to the Kuhn-Tucker theorem in a more general setting than the design problem. The latter is admittedly a standard result in constrained optimisation but it is not one that is conventionally stated in terms of directional derivatives.

In the design context the result was referred to as the General Equivalence Theorem. This derives from the fact that, in the of the D-optimum version of $\psi(\cdot)$, corollary (iii) establishes the equivalence of D-optimality and G-optimality. This follows since, as we shall see, we then have $F\{M(p), v_j v_j'\} = v_j' M^{-1}(p) v_j - k$, $(\text{rank}(M(p)) = k)$. Hence corollary (iii) implies that p^* solves $\min\limits_{p \in \mathcal{P}} \max\limits_{1 \leq j \leq J} \{v_j' M^{-1}(p) v_j\}$, which is the G-optimal criterion mentioned in section 1.4. Kiefer and Wolfowitz (1960) derived this result directly thereby proving theorem 2.5.6 for D-optimality as well, and this was the first appearance of the theorem.

Other authors too have derived the general theorem but using Lagrangian Theory and Duality. We look at an example of this approach for completeness.

§2.5.2 The concept of duality imagines that to a primal maximisation problem there might correspond a dual minimisation problem in that the two problems would share a common optimum or achieve their optima at the same point, and that this could be established by a 'duality theorem' without knowing the common optimising value or point. This clearly might shed further light on the optimum of either problem and may provide in the dual an easier problem to solve.

Sibson (1974a) and Silvey and Titterington (1974) have respectively established dual problems and corresponding duality theorems for D_A-optimality and for a general design criterion. Also Silvey (1972) and Sibson (1972) were the first to consider duality in the design context; for the D-optimal criterion. Pukelsheim (1979 , 1980) too has considered the concept in depth with particular application to design. His work illustrates that there can be several duals to an optimisation problem. Hence there may exist a most informative dual or one that is simplest to solve.

The principle of the lagrangian approach to constrained optimisation is essentially to replace the constraint problem by an unconstrained optimisation problem in higher dimensions whose solution contains, as a component, a solution to the constrained problem.

Whittle (1971) develops the theory of this approach extensively discussing strong and weak lagrangian principles and conditions under which these hold.

Silvey and Titterington (1974) established the following duality theorem, though in the design context only. They adopt a lagrangian approach to the solution of the dual problem, it being a constrained optimisation. This is typical of duality theorems. An appeal to a theorem of Whittle (1971) establishes the duality and thereby obtains conditions for a solution.

We present the theorem for problem (P2), for which their proof carries over.

Theorem 2.5.7

Let $\psi(\cdot)$ be a differentiable concave function on $S = \mathcal{P}(\mathcal{U})$. Let $x^* = x(p^*) = \sum p_j^* u_j$, $p^* \in \mathcal{P}$, solve the primal maximisation problem of maximising $\psi(\cdot)$ over S. Then x^* solves the dual minimisation problem of minimising $\psi(x)$ subject to $F(x^*,u_j) \leq 0$.

Moreover $F\{x(p^*), u_j\} = 0$ if $p_j^* > 0$. □

We will not prove the theorem but make the following comments on the proof.

In the light of the knowledge that $F(x,y)$ is a directional derivative, and given concavity of $\psi(\cdot)$ on S, the result may seem rather self evident. The theorem therefore makes more sense if we partially imagine that we do not realise that $F(x,y)$ is a directional derivative, for that is effectively the position in which authors were placed when establishing earlier duality theorems for specific criteria such as D-optimality; see Sibson (1972), Silvey (1972).

That this attitude should be adopted is reinforced by the fact that the proof only requires to make use of the condition that $F(x^*,u_j) \leq 0$, and subsequently to assume the differentiability condition 2.3.1 on $F(x,y)$, and finally the result that $F(x,x) = 0$. We can stop short, if only marginally, from concluding directly from the first two results that $F(x^*,u_j) = 0$ if $p_j^* > 0$. It is because of this that the last sentence can be included as part of the theorem which is therefore a stronger one than theorem 2.5.6. If the only conditions on the optimum of the primal problem were that $F(x^*,u_j) \leq 0$, then it is conceivable that there may exist an $\tilde{x}$ such that $\tilde{x} \notin \mathcal{P}(\mathcal{U})$, $F(\tilde{x},u_j) \leq 0$ and $\psi(\tilde{x}) < \psi(x^*)$, in which case the primal and dual problem would not share a common optimising x^*. Assuming differentiability a direct appeal to 2.3.1 with the help of Whittle's theorem establishes that this cannot be the case.

The conditions necessary for valid application of the latter theorem further imply that $F\{x^*,u_j\} = 0$ if $p_j^* > 0$.

§2.5.3 For completion we quote the following three additional theorems.

Theorem 2.5.8

(i) Suppose p^* solves (P2) for $\mathcal{U} = \{u_1, \ldots, u_j\}$. We say then that p^* is optimal for $\mathcal{U}$. Then p^* is optimal for any ω such that $\mathrm{Sup}(p^*) \subseteq \omega \subseteq \mathcal{U}$

(ii) Let $\tilde{p}$ be a probability distribution on $\mathcal{U}$ so that $\tilde{p} \in \mathcal{P}$. Then a necessary condition for $\tilde{p}$ to be optimal for $\mathcal{U}$ is that $\tilde{p}$ be optimal for $\mathrm{Sup}(\tilde{p})$. □

Theorem 2.5.9

Suppose $\mathcal{U} = \{u_1, \ldots u_J\}$ is a subset of n-dimensional euclidean space E_n and that p^* is optimal for $\mathcal{U}$. Let $u \in E_n$ be such that $u \notin \mathrm{Sup}(p^*)$. Let $\omega = \{\mathrm{Sup}(p^*) \cup \{u\}\}$. Regarding p^* as a probability distribution on ω let $p^*(\omega)$ denote the weight assigned to $\omega \in \omega$ by p^* so that $p^*(u) = 0$. For any probability distribution p on ω let $x_\omega(p) = \sum_{\omega \in \omega} \omega p(\omega)$ so that $x(p^*) = x_\omega(p^*)$. Let p^{**} be optimal for ω. Then $\psi\{x_\omega(p^{**})\} \geq \psi\{x(p^*)\} = \psi\{x_\omega(p^*)\}$ in general. However if $\psi(\cdot)$ is differentiable at $x_\omega(p^*)$ with respect to ω then $p^{**} = p^*$ iff

$$F\{x_\omega(p^*), u\} = F\{x(p^*), u\} \leq 0$$ □

The truth of these theorems is either self evident or their proofs are elementary. Implicit use of part (ii) of Theorem 2.5.8 will occur frequently, while we now make use of theorem 2.5.9.

Theorem 2.5.10

Let $\mathcal{U}_c$ be a compact set. Let $\mathcal{P}_m$ denote the set of probability measures on $\mathcal{U}_c$ and let $\mathcal{C}(\mathcal{U}_c)$ denote the convex hull of $\mathcal{U}_c$ so that $x \in \mathcal{C}(\mathcal{U}_c)$ iff $x = x(p) = E_p(u)$ for some $p(\cdot) \in \mathcal{P}_m$. Let $\mathcal{P}_{fd}$ be the set of finite descrete probability distributions on $\mathcal{U}_c$ and for $p \in \mathcal{P}_{fd}$ let $\mathrm{Sup}(p) = \{u \in \mathcal{U}_c : p(u) > 0\}$. Call Sup(p) the support of p. Assume $\psi(\cdot)$ is concave on $\mathcal{C}(\mathcal{U}_c)$. Then

(i) For any $x \in \mathcal{C}(\mathcal{U}_c)$, $x = x(p)$, $p \in \mathcal{P}_{fd}$.

(ii) For $p^* \in \mathcal{P}_{fd}$, $x^* = x(p^*)$ maximises $\psi(\cdot)$ on $\mathcal{C}(\mathcal{U}_c)$ iff
$F\{x(p^*), x(p)\} \leq 0 \;\; \forall \; p \in \mathcal{P}_{fd}$ or iff
$F\{x(p), x(p^*)\} \geq 0 \;\; \forall \; p \in \mathcal{P}_{fd}$

(iii) If for $p^* \in \mathcal{P}_{fd}$, $x(p^*)$ is a differentiable point of $\psi(\cdot)$ then $x(p^*)$ maximises $\psi(\cdot)$ on $\mathcal{C}(\mathcal{U}_c)$ iff

$$F\{x(p^*), u\} \leq 0 \quad \forall \; u \in \mathcal{U}_c ,$$

with equality in the case of $u \in \mathrm{Sup}(p^*)$.

Proof Part (i) follows from Carathéodery's theorem and this justifies the restrictions thereafter that $p \in \mathcal{P}_{fd}, p^* \in \mathcal{P}_{fd}$. In fact in the case $\mathcal{U}_c \subset E_n$ we could have restricted the set $\mathcal{P}_{fd}$ to those p's such that Sup(p) contains at most (n + 1) points.

Part (ii) then follows from theorem 2.5.5.

Part (iii) could be given a direct proof analogous to that of theorem 2.5.6 or the following alternative.

If $x(p^*)$ is to maximise $\psi(\cdot)$ over $\mathcal{C}(\mathcal{U}_c)$ then we must have $\psi\{x(p^*)\} \geq \psi\{x_\omega(p^{**})\}$ where $x_\omega(p^{**})$ maximises $\psi(\cdot)$ over any subset ω of $\mathcal{U}_c$. In particular p^* must be optimal for its support which is the case iff $F\{x(p^*), u\} = 0$ for all $u \in$ Sup(p^*), and also the above must be true for $\omega = \{Sup(p^*) \cup \{u\}\}$, where $u \in \mathcal{U}_c$ but $u \notin Sup(p^*)$. However from theorem 2.5.9 we have, for this particular case of ω, that $\psi\{x_\omega(p^{**})\} \geq \psi\{x_\omega(p^*)\} = \psi\{x(p^*)\}$, $p^* = p^{**}$, iff

$$F\{x_\omega(p^{**}), u\} = F\{x(p^*), u\} \leq 0$$

Hence the theorem. □

In part (iii) of this theorem we clearly have a continuous analogue to theorem 2.5.6. However its only practical value can be that it may yield explicit solutions for an optimising p^*. If as is generally the case numerical techniques must be used to find p^* then one is forced to discretise ; thereby generating a particular case of (P2) for which theorem 2.5.6 is the test of optimality, although admittedly this is a particular case of theorem 2.5.10. The following more general consequence than theorem 2.5.6 is sometimes useful.

Corollary 2.5.10.1

Let the set $\mathcal{U}_c$ in theorem 2.5.10 be a bounded convex set with a finite set of extreme points $\mathcal{U} = \{u_1, .., u_J\}$. Let p^* be a solution to the optimum in that theorem. For at least one p^*, Sup(p^*) $\subset \mathcal{U}$; a p^* solving (P2) for this $\mathcal{U}$.

Proof

If p^* solves (P2) then $x^* = x(p^*) = \sum p_j^* u_j$ maximises $\psi(\cdot)$ over $\mathcal{P}(\mathcal{U})$; but here $\mathcal{U}_c = \mathcal{C}(\mathcal{U}_c) = \mathcal{P}(\mathcal{U})$ so that x^* maximises $\psi(\cdot)$ over $\mathcal{C}(\mathcal{U}_c)$. Hence the result. □

CHAPTER 3

CONSTRAINED OPTIMALITY IN DESIGN

§3.1 Directional Derivatives Of Design Criteria

§3.1.1 So far we have not considered the evaluation of directional derivatives for any specific function. This should not be difficult in the case of simple functions such as those of section 1.1.1, since their partial derivatives are simple and differentiability makes available the formula $F(x,y) = (y-x)'d$. However the design criteria being functions of matrices are more complex and they can be nondifferentiable.

Since $F(x,y) = G(x,y-x)$ we consider evaluation of $G(x,y)$. Some general rules would be useful. Davies (1974) has determined the following:

(a) Let $h(x) = af(x) + bg(x)$

Then $G_h(x,y) = aG_f(x,y) + bG_g(x,y)$

(b) Let $*$ denote a 'multiplication' operation which satisfies $(y_1+y_2)*z = y_1*z+y_2*z$ and $y*(z_1+z_2) = y*z_1+y*z_2$ but need not be associative or cummutative.

Let $h(x) = f(x)*g(x)$

Then $G_h(x,y) = f(x)*G_g(x,y) + G_f(x,y)*g(x)$

(c) Let $h(x) = f\{g(x)\}$

Then $G_h(x,z) = G_f(y,w)$

where $y = g(x)$ and $w = G_g(x,z)$

Special cases of (c) are:

(d) $G_h(x,y) = G_g(x,y)\times\{df(y)/dy\}$

where $df(y)/dy = \underset{\lambda\to 0}{\mathrm{Lt}}\left\{[f(y+\lambda) - f(y)]/\lambda\right\} = G_f(y,1).$

(e) Let $f(y)$ be linear, that is $f(ax + by) = af(x) + bf(y)$.

Then $G_h(x,y) = f\{G_g(x,y)\}$

These results hold good in the case when the functions h,f,g and their arguments are vectors or matrices.

We now proceed to sketch out how to derive the directional derivatives of the design criteria $\psi(\cdot)$ of section 1.4, which depend on the matrix AM^-A'. Firstly Table 3.1.1 is a list of particular ancillary derivatives of which we will make use. All the results were derived by Davies (1974). The variables M,N are k×k matrices.

TABLE 3.1.1

	$f(M)$	$G_f(M,N)$	
(i)	M	N	
(ii)	M^t, (t integral)	$\sum_{j=1}^{t} N_j$, $N_j = M^{j-1} N M^{t-j}$	(by (i),(b))
(iii)	$tr(LM^t)$	$t \cdot tr(LM^{t-1}N)$	(by (ii),(a),(e))
(iv)	$[(1/k)\,tr(M^t)]^{1/t}$	$\dfrac{tr(M^{t-1}N)[(1/k)\,tr(M^t)]^{1/t}}{tr(M^t)}$	(by (iii),(d))
(v)	$det(M)$	$det(M)\cdot tr(NM^{-1})$	(by (iv), $t\to 0$)
(vi)	$\ln[det(M)]$	$tr(NM^{-1})$	(by (v),(d))

§3.1.2 We need to derive one further major result.

Lemma 3.1.1

If M is nonegative definite and the nullspace $\mathcal{N}(M)$ of M is contained in that, $\mathcal{N}(A)$, of A, then the function $g(M) = AM^+A'$ has directional derivative

$$F_g(M,N) = AM^+A' - AM^+XCX'M^+A'$$

where A,M,N,X, and further matrices B,Y,Z,T,P_1,P_2,P satisfy the following relationships.

(i) $XX' = N$

(ii) $P = (P_1 : P_2)$ is orthogonal

(iii) $M = (P_1 : P_2)\begin{bmatrix} T & 0 \\ 0 & 0 \end{bmatrix}\begin{pmatrix} P_1' \\ P_2' \end{pmatrix}$, T diagonal, nonsingular.

(iv) $X = (P_1 : P_2)\begin{pmatrix} Y \\ Z \end{pmatrix}$, $Y = P_1'X$, $Z = P_2'X$

(v) $C = I - Z'(ZZ')^+Z$

(vi) $A = BP_1'$, $B = AP_1$.

Some notes

Note 1 Clearly the elements of T are the non-zero eigenvalues of M and the columns of P_1 are the eigenvectors corresponding to these, while the columns of P_2 are eigenvectors corresponding to the zero eigenvalue.

Note 2 If $N = M(q) \in \mathcal{M}$ then $N = VQV'$ where $Q = \mathrm{diag}\{q_1, \ldots, q_J\}$
Hence we can take $X = VQ^{1/2}$ where $Q^{1/2} = \mathrm{diag}\{\sqrt{q_1}, \ldots, \sqrt{q_J}\}$ whether or not $q_j = 0$.

Note 3 Particular consequences of the above relationships are

$$BT^{-1}B' = AM^{+}A' \quad , \quad BT^{-1}Y = AM^{+}X$$

These follow since

$$M^{+} = (P_1 : P_2)\begin{bmatrix} T^{-1} & 0 \\ 0 & 0 \end{bmatrix}\begin{pmatrix} P_1' \\ \cdots \\ P_2' \end{pmatrix}$$

– Note 4 We have particularly specified M^{+} as opposed to any other generalised inverse. This is for simplicity. However the null-space conditions of the lemma imply, of course, that $AM^{+}A' = AM^{-}A'$. Thus $F_g(M,N)$, as stated in the lemma, must also be the value of the directional derivative of $AM^{-}A'$. In general though, $F_g(M,N)$ will not be the same function of M^{-} as it is of M^{+}.

Proof of lemma Davies proceeds as follows

Let $M_\varepsilon = (1-\varepsilon)M + \varepsilon XX'$, $T_\varepsilon = (1-\varepsilon)T + \varepsilon YY'$.

We have $MP_2 = 0$ from which it follows that $AP_2 = 0$ since $\mathcal{N}(M) \subseteq \mathcal{N}(A)$.
Hence

$$M_\varepsilon = (P_1 : P_2)\begin{bmatrix} T_\varepsilon & \varepsilon YZ' \\ \varepsilon ZY' & \varepsilon ZZ' \end{bmatrix}\begin{pmatrix} P_1' \\ \cdots \\ P_2' \end{pmatrix}$$

and for $0 < \varepsilon < 1$

$$AM_\varepsilon^{+}A = B(I : 0)\begin{bmatrix} T_\varepsilon & \varepsilon YZ' \\ \varepsilon ZY' & \varepsilon ZZ' \end{bmatrix}^{+}\begin{pmatrix} I \\ \cdots \\ 0 \end{pmatrix}B'$$

Now use Rhode (1965) and the fact that T_ε must be non-singular in an argument analogous to that which derived the simplification 1.4.7 of $AM^{+}A'$ in the case $A = [I_s \vdots 0]$ to obtain that

$$\begin{aligned} AM_\varepsilon^{+}A' &= B\left[T_\varepsilon - \varepsilon YZ'(ZZ')^{+}ZY'\right]^{-1}B' \\ &= B\left[T - \varepsilon T + \varepsilon YCY'\right]^{-1}B' \\ &= B\left[T^{-1} - \varepsilon T^{-1}(YCY' - T)T^{-1} + o(\varepsilon)\right]B' \\ &= BT^{-1}B' + \varepsilon BT^{-1}B' - \varepsilon BT^{-1}YCY'T^{-1}B' + o(\varepsilon) \\ &= AM^{+}A' + \varepsilon AM^{+}A' - \varepsilon AM^{+}XCX'M^{+}A' + o(\varepsilon) \end{aligned}$$

Hence

$$\left[AM_\varepsilon^{+}A' - AM^{+}A'\right]/\varepsilon = \left[AM^{+}A' - AM^{+}XCX'M^{+}A'\right] + O(\varepsilon)$$

$$\therefore \lim_{\varepsilon \downarrow 0}\left\{\left[\quad\right]/\varepsilon\right\} = \left[\quad '' \quad\right]$$

$$\text{i.e.} \quad F_g(M,N) = \left[\quad '' \quad\right] \qquad \square$$

Addendum By a similar argument

$$G_g(M,N) = -AM^{+}XCX'M^{+}A'. \qquad \square$$

We now list some corollaries

Corollary (i) If M is nonsingular $G_g(M,N) = -AM^{-1}NM^{-1}A'$.
Proof It follows that T is of order $k \times k$ and hence P_2, Z strictly do not exist or could be regarded as zero. Then $C = I$ and of course $M^+ = M^{-1}$. □

Corollary (ii) If $A = I$ and M is nonsingular so that $g(M) = M^{-1}$ then $G_g(M,N) = -M^{-1}NM^{-1}$ □

Corollary (iii) Let $M,N \in \mathcal{M}$, $M = M(q)$, $N = M(r)$; $r,q \in \mathcal{P}$. Suppose M is singular. If $\mathrm{Sup}(r) \leq L\{\mathrm{Sup}(q)\}$ where $L\{\mathrm{Sup}(q)\}$ is the linear subspace spanned by $\mathrm{Sup}(q)$, the support of q, then

$$G_g(M,N) = -AM^+NM^+A'$$
$$F_g(M,N) = AM^+A' - AM^+NM^+A'$$

Proof In the meaning of formula 1.3.9

$$M = V_q P_q V_q'$$

Now orthogonal $P = (P_1 \vdots P_2)$ is to be such that for T nonsingular diagonal

$$M = P\begin{bmatrix} T & 0 \\ 0 & 0 \end{bmatrix}P'$$

and hence

$$P'V_q P_q V_q' P' = \begin{bmatrix} T & 0 \\ 0 & 0 \end{bmatrix}.$$

Since P_q is positive definite this will be so iff $P_2'V_q = 0$. That this is so under the conditions of the lemma can be seen from the considerations of consequence (vi) of formula 1.4.6.

Since $\mathcal{R}(M) = \mathcal{R}(V_q)$, the columns of V_q, i.e. the support, $\mathrm{Sup}(q)$, of q, belong to $\mathcal{R}(M)$. The latter is the orthogonal complement of $\mathcal{N}(M)$ which is spanned by the eigenvectors corresponding to the zero eigenvalue of M, i.e. the columns of P_2. Hence $P_2'V_q = 0$. In general $P_2'x = 0$ iff $x \in \mathcal{R}(M)$ and hence here, $P_2'x = 0$ iff $x \in \mathcal{R}(V_q) = L\{\mathrm{Sup}(q)\}$.

Consider now $N = M(r)$, $\mathrm{Sup}(r) \leq L\{\mathrm{Sup}(q)\}$. Again in the meaning of formula 1.3.9

$$N = V_r P_r V_r'$$

Hence $N = XX'$ where $X = V_r P_r^{1/2}$ and so $Z = P_2'X = P_2'V_r P_r^{1/2}$

Now $\mathrm{Sup}(r) \subseteq L\{\mathrm{Sup}(q)\}$ implies that the columns of V_r belong to $L\{\mathrm{Sup}(q)\}$ and hence by the preceding argument $P_2'V_r = 0$.

It follows that $Z = 0$ implying $C = I$ and hence $XCX' = N$. Hence the result. $\square$

Note that no value was specified for $n\{\mathrm{Sup}(q)\}$.

The result could also be shown to be true by a similar argument for any nonnegative definite M, N such that $\mathcal{R}(N) \subset \mathcal{R}(M)$.

<u>Corollary (iv)</u> $F_g(M,N) = G_g(M,N) - G_g(M,M)$ for M nonnegative definite.

<u>Proof</u> Clearly M satisfies the condition imposed on the matrix N at the end of corollary (iii), namely that $\mathcal{R}(N) \subset \mathcal{R}(M)$. Hence

$$G_g(M,M) = -AM^+MM^+A' = -AM^+A'.$$

We have already claimed, for general N, that

$$G_g(M,N) = -AM^+XCX'M^+A'$$

Hence the result. Note however that, while this implies

$$G_g(M,N-M) = G_g(M,N) - G_g(M,M),$$

it does not necessarily follow that $G_g(M,N)$ is always linear in N. $\square$

<u>Corollary (v)</u>

(a) $F_g(M, v_jv_j') = AM^{-1}A' - AM^{-1}v_jv_j'M^{-1}A'$ if M is nonsingular.

(b) $F_g\{M(q), v_jv_j'\} = AM^+(q)A' - AM^+(q)v_jv_j'M^+(q)A'$ $\begin{cases} \text{if } q_j > 0 \text{ or} \\ \text{if } q_j = 0,\ v_j \in L\{\mathrm{Sup}(q)\}. \end{cases}$

(c) $F_g\{M(q), v_jv_j'\} = AM^+(q)A'$ if $q_j = 0$, $v_j \notin L\{\mathrm{Sup}(q)\}$.

<u>Proof</u> Case (a) follows directly from corollary (i) taking $N = v_jv_j'$. If $q_j > 0$ then $v_j \in \mathrm{Sup}(q)$ and hence $v_j \in L\{\mathrm{Sup}(q)\}$. Case (b) then follows directly from corollary (iii), for if $v_j \in L\{\mathrm{Sup}(q)\}$, then v_jv_j' is a particular case of the matrix $N = M(q)$ there. Case (c) is obtained as follows.

Clearly we can take $X = v_j$ and hence $Z = P_2'v_j$. Thus Z is a vector and so the matrix C is a scalar, namely $C = 1 - Z'(ZZ')^+Z$

Since $v_j \notin L\{\mathrm{Sup}(q)\}$, $P_2'v_j \neq 0$, and if $Z \neq 0$ then it is the case that

$$(ZZ')^+ = ZZ'/(Z'Z)^2 .$$

Thus $C = 1-1 = 0$. Hence the result. $\square$

It was reported in Note 4 which followed the stating of lemma 3.1.1, that $F_g(M,N)$ must define the directional derivative of

AM^-A' at M towards N. In particular then (b) and (c) must define the directional derivatives, towards the vertices v_jv_j', of $AM^-(q)A'$. In fact, in an exception to the case of general N, we can replace $M^+(q)$ by $M^-(q)$ in (b) and (c). Under the null-space conditions of the lemma, equation 1.4.6 is satisfied. Consequence (v) of that equation tells us that $AM^+(q)A' = AM^-(q)A'$, while, when $v_j \in L\{Sup(q)\}$, it follows that $v_j \in \mathcal{R}(M(q))$ and so, from consequences (vi) of 1.4.6, $AM^+(q)v_j = AM^-(q)v_j$.

§3.1.3 We now make use of the above results to derive the directional derivatives for standard design criteria. Particular use is made of rule (c); that is, if $\psi(M) = f\{g(M)\}$, then

$$G_\psi(M,N) = G_f\{g(M), G_g(M,N)\}.$$

We take $g(M) = AM^+A'$ amd always assume that M (and N) is nonnegative definite so that we make use of the formula

$$F_g(M,N) = G_g(M,N) - G_g(M,M).$$

1. The general trace criteria

$$\psi(M) = -\mathrm{tr}\{g^t(M)\},\ t>0\ ;\quad G_\psi(M,N) = -t\cdot \mathrm{tr}\{g^{(t-1)}(M)\cdot G_g(M,N)\} \quad \begin{pmatrix}\text{by (iii)}\\ \text{Table 3.1.1}\end{pmatrix}$$
$$= t\cdot \mathrm{tr}\{(AM^+A')^{t-1}AM^+XCX'M^+A'\}$$
$$\therefore\ G_\psi(M,M) = t\cdot \mathrm{tr}\{(AM^+A')^t\}.$$

Hence

$$F_\psi(M,N) = t\left[\mathrm{tr}\{(AM^+A')^{t-1}AM^+XCX'M^+A'\} - \mathrm{tr}(AM^+A')^t\right]$$

In particular

$$(1/t)F_\psi(M,N) = \begin{cases} \mathrm{tr}\{(AM^{-1}A')^{t-1}AM^{-1}NM^{-1}A'\} - \mathrm{tr}(AM^{-1}A')^t & ,\ M \text{ nonsingular} \\ \mathrm{tr}\{(AM^+A')^{t-1}AM^+NM^+A'\} - \mathrm{tr}(AM^+A')^t & ,\ \begin{cases} M = M(q),\ N = M(r) \\ Sup(r) \subseteq L\{Sup(q)\}\end{cases}\end{cases}$$

If M is nonsingular

$$(1/t)F_\psi(M, v_jv_j') = \begin{cases} v_j'M^{-1}A'(AM^{-1}A')^{t-1}AM^{-1}v_j - \mathrm{tr}(AM^{-1}A')^t \\ v_j'M^{-(t+1)}v_j - \mathrm{tr}(M^{-t}) & ,\ A = I_k\end{cases}$$

More generally

$$(1/t)F_\psi\{M(p), v_jv_j'\} = \begin{cases} v_j'M^+(p)A'(AM^+(p)A')^{t-1}AM^+(p)v_j - \mathrm{tr}(AM^+(p)A')^t, & \begin{cases}\text{if } p_j > 0 \text{ or} \\ p_j = 0,\ v_j \in L\{Sup(p)\}\end{cases} \\ -\mathrm{tr}(AM^+(p)A')^t & \text{if } p_j = 0,\ v_j \notin L\{Sup(p)\}\end{cases}$$

2. D_A-optimality

$$\psi(M) = -\log_e \det\{g(M)\} \; ; \quad G_\psi(M,N) = -\mathrm{tr}\{g^{-1}(M)\cdot G_g(M,N)\} \qquad \text{by (vi) of Table 3.1.1}$$
$$= \mathrm{tr}\{(AM^+A')^{-1}AM^+XCX'M^+A'\}$$
$$\therefore \quad G_\psi(M,M) = \mathrm{tr}\{(AM^+A')^{-1}(AM^+A')\} = s$$

Hence

$$F_\psi(M,N) = \mathrm{tr}\{(AM^+A')^{-1}AM^+XCX'M^+A'\} - s$$

In particular

$$F_\psi(M,N) = \begin{cases} \mathrm{tr}\{(AM^{-1}A')^{-1}AM^{-1}NM^{-1}A'\} - s, & M \text{ nonsingular} \\ \mathrm{tr}\{(AM^+A')^{-1}AM^+NM^+A'\} - s, & \begin{cases} M=M(q),\ N=M(r) \\ \mathrm{Sup}(r) \subseteq L\{\mathrm{Sup}(q)\} \end{cases} \end{cases}$$

If M is nonsingular

$$F_\psi(M, v_j v_j') = \begin{cases} v_j'M^{-1}A'(AM^{-1}A')^{-1}AM^{-1}v_j - s \\ v_j'M^{-1}v_j - k, & A = I_k \end{cases}$$

More generally

$$F_\psi\{M(p), v_j v_j'\} = \begin{cases} v_j'M^+(p)A'(AM^+(p)A')^{-1}AM^+(p)v_j - s, & \begin{cases} \text{if } p_j > 0 \text{ or} \\ p_j = 0,\ v_j \in L\{\mathrm{Sup}(p)\} \end{cases} \\ -s & \text{if } p_j = 0,\ v_j \notin L\{\mathrm{Sup}(p)\} \end{cases}$$

3. L_A-optimality

$$\psi(M) = -\mathrm{tr}\{L(AM^+A')\} \; ; \quad G_\psi(M,N) = -\mathrm{tr}\{L\cdot G_g(M,N)\} \qquad \text{by (iii), Table 3.1.1, } t=1$$
$$= \mathrm{tr}\{LAM^+XCX'M^+A'\}$$
$$\therefore \quad G_\psi(M,M) = \mathrm{tr}\{L(AM^+A')\}$$

Hence

$$F_\psi(M,N) = \mathrm{tr}\{LAM^+XCX'M^+A'\} - \mathrm{tr}\{L(AM^+A')\}$$

In particular

$$F_\psi(M,N) = \begin{cases} \mathrm{tr}\{LAM^{-1}NM^{-1}A'\} - \mathrm{tr}\{L(AM^{-1}A')\}, & M \text{ nonsingular} \\ \mathrm{tr}\{LAM^+NM^+A'\} - \mathrm{tr}\{L(AM^+A')\}, & \begin{cases} M=M(q),\ N=M(r) \\ \mathrm{Sup}(r) \subseteq L\{\mathrm{Sup}(q)\} \end{cases} \end{cases}$$

If M is nonsingular

$$F_\psi(M, v_j v_j') = v_j'M^{-1}A'LAM^{-1}v_j - \mathrm{tr}\{L(AM^{-1}A')\}$$

More generally

$$F_\psi\{M(p), v_j v_j'\} = \begin{cases} v_j'M^+(p)A'LAM^+(p)v_j - \mathrm{tr}\{L(AM^+(p)A')\}, & \begin{cases} \text{if } p_j > 0 \text{ or} \\ p_j = 0,\ v_j \in L\{\mathrm{Sup}(p)\} \end{cases} \\ -\mathrm{tr}\{L(AM^+(p)A'\} & \text{if } p_j > 0,\ v_j \notin L\{\mathrm{Sup}(p)\}. \end{cases}$$

4. c-optimality $\quad \psi(M) = -c'M^{+}c$, a scalar

Taking $A = c'$ and $t = 1$ in the results for the general trace criteria we obtain

$$F_{\psi}(M,N) = c'M^{+}XCX'M^{+}c - c'M^{+}c$$

In particular

$$F_{\psi}(M,N) = \begin{cases} c'M^{-1}NM^{-1}c - c'M^{-1}c & , \quad M \text{ nonsingular} \\ c'M^{+}NM^{+}c - c'M^{+}c & , \quad \begin{cases} M = M(q), N = M(r) \\ Sup(r) \subseteq L\{Sup(q)\} \end{cases} \end{cases}$$

If M is nonsingular

$$F_{\psi}(M, v_j v_j') = (v_j'M^{-1}c)^2 - c'M^{-1}c$$

More generally

$$F_{\psi}\{M(p), v_j v_j'\} = \begin{cases} \left[v_j'M^{+}(p)c\right]^2 - c'M^{+}(p)c & \begin{cases} \text{if } p_j > 0 \text{ or} \\ p_j = 0, \ v_j \in L\{Sup(p)\} \end{cases} \\ -c'M^{+}(p)c & \text{if } p_j = 0, \ v_j \notin L\{Sup(p)\} \end{cases}$$

5. E_A-optimality

Consider now the case $\psi(M) = -\lambda_{max}\{g(M)\}$

This is less straightforward to deal with. Consider first the function

3.1.1 $$f(M) = \lambda_{max}(M) = \max_{b'b=1} b'Mb$$

Suppose that $\lambda_{max}(M)$ has a multiplicity of $q = 1$. Then there is a unique b solving the above maximisation, namely the normalised eigenvector corresponding to $\lambda_{max}(M)$. Denote this by $b^*(M)$. Assume also that $\lambda_{max}(M+\epsilon N)$ has multiplicity 1 and let $b^*(M + \epsilon N)$ be the counterpart of $b^*(M)$.

Lemma 3.1.2

Assume for any matrix A that

$$[b^*(M+\epsilon N)]'A[b^*(M+\epsilon N)] = [b^*(M)]'A\,b^*(M) + o(\epsilon)$$

Then

$$G_f(M,N) = [b^*(M)]'N\,b^*(M) .$$

Proof

It follows from the assumption made in the lemma that

$$[f(M+\epsilon N) - f(M)]/\epsilon = [b^*(M)]'Nb^*(M) + o(\epsilon) .$$

Hence the result. $\square$

Consider now the case when $\lambda_{max}(M)$ has multiplicitly $q>1$. There is then more than one optimising $b^*(M)$ in 3.1.1, namely any normalised eigenvector corresponding to $\lambda_{max}(M)$; in particular any linear combination $b^*(M) = \sum_{i=1}^{q} \lambda_i b_i$ of q orthonormal eigenvectors $b_1, \ldots, b_q$ corresponding to $\lambda_{max}(M)$ with $\sum_{i=1}^{q} \lambda_i^2 = 1$ to ensure $b^*(M)$ is normalised.

Note that if b_i and b_j are two (orthogonal) eigenvectors of a matrix M then $b_i' M b_j = 0$; b_i, b_j are M-orthogonal. Hence

$$[b^*(M)]' M b^*(M) = \sum \lambda_i^2 b_i' M b_i = \lambda_{max}(M) \cdot \left(\sum \lambda_i^2\right) = \lambda_{max}(M).$$

Let $\mathcal{B}^*$ denote the set of all optimising $b^*(M)$.

Suppose that $\lambda_{max}(M + \epsilon N)$ has multiplicity 1. Then it would seem that it would be the case, in the light of lemma 3.1.2, that

3.1.2 $$G_f(M,N) = \max_{b \in \mathcal{B}^*} b' N b.$$

A similar result should hold if $\lambda_{max}(M + \epsilon N)$ has multiplicity larger than 1.

Let us suppose that 3.1.2 is true. Consider now $\psi(M) = -\lambda_{max}(AM^+A') = -f\{g(M)\}$. Let q now denote the multiplicity of $\lambda_{max}(AM^+A')$ and let $\mathfrak{h}^*$ denote the set of all normalised eigenvectors corresponding to $\lambda_{max}(AM^+A')$. Using rule (c) in the form

$$F_\psi(M,N) = G_\psi(M, N-M) = -G_f\{g(M), G_g(M, N-M)\}$$

we obtain

$$F_\psi(M,N) = -\max_{z \in \mathfrak{h}^*}\{z' G_g(M, N-M) z\} \ ;$$

that is,

3.1.3 $$F_\psi(M,N) = -\max_{z \in \mathfrak{h}^*}\{z'[AM^+A' - AM^+XCX'M^+A']z\}$$

We note then that in the case M nonsingular

3.1.4 $$F_\psi(M,N) = -\max_{z \in \mathfrak{h}^*}\{z' A M^{-1}(M-N)M^{-1}A' z\} \ .$$

This agrees with the derivation obtained by Kiefer (1974) for nonsingular M although his formula was

3.1.5 $$F_\psi(M,N) = -\lambda_{max}\{QAM^{-1}(M-N)M^{-1}A'Q'\}$$

where $Q' = [b_1 \cdots b_q]$, $b_1, \ldots, b_q$ being q orthonormal eigenvectors corresponding to $\lambda_{max}(AM^{-1}A')$.

The agreement between the two formulae can be seen as follows:

$$\lambda_{max}(AM^{-1}A') = \max_{\lambda'\lambda=1}\{\lambda' Q A M^{-1}(M-N)M^{-1}A'Q'\lambda\}$$

that is,

3.1.6 $$\lambda_{max}(AM^{-1}A') = \max_{\lambda'\lambda=1}\{(\Sigma\lambda_j b_j)' AM^{-1}(M-N)M^{-1}A'(\Sigma\lambda_j b_j)\}.$$

We have already observed that $\Sigma\lambda_j b_j$ will, if $\Sigma\lambda_j^2 = 1$, define another normalised eigenvector corresponding to $\lambda_{max}(AM^{-1}A')$, so that an optimising λ for 3.1.6 appears to provide, in $z = \Sigma\lambda_j b_j$, an optimising z for 3.1.4.

Further simplification of 3.1.3 derives from the fact that for $z \in \mathfrak{z}^*$, $z'AM^+A'z = \lambda_{max}(AM^+A')$.

$$\therefore \quad F_\psi(M,N) = -\max_{z\in\mathfrak{z}^*}\{\lambda_{max}(AM^+A') - z'AM^+XCX'M^+A'z\}$$

that is,

$$F_\psi(M,N) = \min_{z\in\mathfrak{z}^*}\{z'AM^+XCX'M^+A'z\} - \lambda_{max}(AM^+A').$$

(A similar argument would yield $G_\psi(M,N) = \min_{z\in\mathfrak{z}^*}\{z'AM^+XCX'M^+A'z\}$)

In particular

$$F_\psi(M,N) = \begin{cases} \min_{z\in\mathfrak{z}^*}\{z'AM^{-1}NM^{-1}A'z\} - \lambda_{max}(AM^{-1}A'), & M \text{ nonsingular} \\ \min_{z\in\mathfrak{z}^*}\{z'AM^+NM^+A'z\} - \lambda_{max}(AM^+A'), & \begin{cases} M = M(q), N = M(r) \\ Sup(r) \subseteq L\{Sup(q)\} \end{cases} \end{cases}$$

If M is nonsingular

$$F_\psi(M, v_j v_j') = \begin{cases} \min_{z\in\mathfrak{z}^*}\{v_j'M^{-1}A'z\}^2 - \lambda_{max}(AM^{-1}A') \\ \min_{z\in\mathfrak{z}^*}\{v_j'M^{-1}z\}^2 - \lambda_{max}(M^{-1}), & A = I_k \end{cases}$$

More generally

$$F_\psi\{M(p), v_j v_j'\} = \begin{cases} \min_{z\in\mathfrak{z}^*}\{v_j'M^+(p)A'z\}^2 - \lambda_{max}(AM^+(p)A'), & \begin{cases} \text{if } p_j > 0 \text{ or} \\ p_j = 0,\ v_j \in L\{Sup(p)\} \end{cases} \\ -\lambda_{max}(AM^+(p)A') & \text{if } p_j = 0,\ v_j \notin L\{Sup(p)\}. \end{cases}$$

§3.2 Support Differentiability; Regression Design Optimality Theorems

§3.2.1 We turn now to consider some optimality theorems for the design criteria of the previous section; the need to do so being, that non-differentiability can occur. The matrix AM^+A' (and hence functions of it), is in general not overall differentiable, as is evident from the fact that $G(M,N)$ depends on N through the matrix XCX'.

The necessary full linearity of $G(M,N)$ in N, in general, only obtains in the case of D_A-optimality, c-optimality and $\psi(M) = -tr(AM^{-1}A')^t$ if M is nonsingular, while, even if the latter is the case, the generalised E-optimal criterion will typically be non-differentiable if the multiplicity q of $\lambda_{max}(AM^+A')$ is larger than 1. This is clear enough from Kiefer's formula for $F_\psi(M,N)$ given in 3.1.4. Typically, only if Q is $1\times s$, which requires $q = 1$, will the necessary linearity obtain, for then $\lambda_{max}\{QAM^{-1}(M-N)M^{-1}A'Q'\}$ is simply the scalar $QAM^{-1}(M-N)M^{-1}A'Q'$.

However we have seen in the case of D_A-optimality, L_A-optimality, c-optimality and $\psi(M) = -tr(AM^+A')^t$, a restricted linearity of $G(M,N)$ in N; namely, for $M = M(p) \in \mathcal{M}$, $G(M,N)$ is linear on the set $\{N : N = M(r), Sup(r) \subseteq L\{Sup(p)\}\}$.

Definition 2.7.1 Support Differentiability

We say that a function $\psi(M)$ has support differentiability at $M = M(p)$ if $G(M,N)$ is linear on the set $\{N : N = M(r), Sup(r) \subseteq L\{Sup(p)\}\}$. That is, $\psi(\cdot)$ is "differentiable at $M(p)$ with respect to $L\{Sup(p)\}$".

In retrospect the above criteria must enjoy something like this property since for them full differentiability fails if and only if θ is inestimable; that is, if $M(p)$ is singular. Consider the following notion.

The linear model 1.2.1, $E(y) = v'\theta$, is equivalent to

$$E(y) = \omega'\gamma \qquad (3.2.1)$$

where ω, γ are of length $k' > k$ with $\gamma' = (\theta', \beta')$ for arbitrary β and $\omega' = (v', 0')$; that is, the last $(k'-k)$ components of ω are zero and hence are redundant for y does not depend on β; observations on y will never yield information about β.

However in principle an appropriate expansion of $\mathcal{V}$ will define a 'design space' ω for 3.2.1, such that any design matrix $M_\omega(p)$, corresponding to a 'design p on ω', will be of the form

$$M_\omega(p) = \begin{bmatrix} M(p) & 0 \\ 0 & 0 \end{bmatrix}$$

Always $M_\omega(p)$ will be singular with the resultant loss of 'full differentiability'. Clearly though 'full differentiability' under 1.2.1 at nonsingular $M(p)$ is a particular 'support differentiability' under 3.2.1 at $M_\omega(p)$.

§3.2.2 We now consider what can be achieved in respect of optimality conditions under the possibility of nondifferentiability in the design context.

Of course the problem is that we will not have available the simple finite set of optimality conditions which, under the design version of theorem 2.5.6, we would have at our command at differentiable $M(p)$. At best this theorem can only identify non-optimality. In principle we must test the optimality of a postulated solution by checking for the conditions of the design version of theorem 2.5.5.

However one would have thought that support differentiability might point the way to a simple test of optimality and indeed this does lead to some simplification. For instance we can derive a theorem specifying a finite set of sufficient conditions for optimality. They are however not necessary conditions. At best the theorem can in some instances identify an optimum while in other instances it will identify non-optimality. As yet there still remains the task of identifying a simple enough finite set of necessary and sufficient conditions although Silvey (1978) and Pukelsheim (1979, 1980) have come close to a solution. It could be said that the latter has established a finite set of necessary and sufficient but invisible conditions; he has established the existence of such a set.

Not surprisingly other authors have considered the problem. These include Karlin and Studden (1966), Atwood (1969), Sibson (1974a) Silvey and Titterington (1973), Fellman (1974). However in general their results are too complicated to be of practical value and we will not report their work.

We will confine ourselves to an appraisal of a collection of theorems which includes design versions of theorems 2.5.5 and 2.5.6, two simple theorems defining sufficient optimality conditions based on support differentiability and finally the theorems of Silvey and Pukelsheim which are the most recent in the literature and have been closest to a simple solution to the underlying problem.

§3.2.3

Theorem 3.2.1 Suppose $\mathcal{V} = \{v_1, \cdots\cdots, v_J\}$, $\mathcal{M} = \{M(p) = \sum p_j v_j v_j', \ p \in \mathcal{P}\}$. If $\psi(M)$ is concave on $\mathcal{M}$ and $M(p^*)$ is a differentiable point of $\psi(\cdot)$, then $M(p^*)$ maximises $\psi(\cdot)$ on $\mathcal{M}$ iff

(i) $F\{M(p^*), v_j v_j'\} = 0$ when $p_j^* > 0$

(ii) $F\{M(p^*), v_j v_j'\} \leq 0$ when $p_j^* = 0$

This follows directly from theorem 2.5.6 and is more properly the general equivalence theorem. □

We now consider an example which illustrates that the above conditions do not guarantee optimality at a nondifferentiable $M(p)$. The example was considered by Silvey (1974).

Ex. 3.1.1 The design space is $\mathcal{V} = \{v_1, v_2, v_3\} = \{(1,0)', (4,1)', (4,2)'\}$. The criterion is c-optimality with $c = (1,0)'$, so that

$$F\{M(p), v_j v_j'\} = \begin{cases} (v_j' M^+(p) c)^2 - c' M^+(p) c & v_j \in L\{Sup(p)\} \\ -c' M^+(p) c & v_j \notin L\{Sup(p)\} \end{cases}$$

Consider the design $p = (1,0,0)'$,

$$M(p) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = M^+(p)$$

$$Sup(p) = \{v_1\}, \ L\{Sup(p)\} = \{x = (x_1, x_2)' : x_2 = 0\}, \ c' M^+(p) x = x_1^2, \ c' M^+(p) c = 1 .$$

Hence

$$F\{M(p), xx'\} = \begin{cases} x_1^2 - 1 & , \ x_2 = 0 \\ -1 & , \ x_2 \neq 0 \end{cases}$$

Therefore $F\{M(p), v_j v_j'\} = 0, -1, -1$ respectively for j = 1,2,3 and so the conditions of theorem 3.2.1 are satisfied, except, of course, that $M(p)$ is not a point of differentiability.

Consider now the design $q = (0, \alpha, 1-\alpha)$ In the terminology of lemma 3.1.1

$$N = M(q) = \begin{bmatrix} 4 & 4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ 0 & (1-\alpha) \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} 16 & 8-4\alpha \\ 8-4\alpha & 4-3\alpha \end{bmatrix},$$

$$X = \begin{bmatrix} 4 & 4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \sqrt{\alpha} & 0 \\ 0 & \sqrt{1-\alpha} \end{bmatrix}, \quad P_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$Z = (\sqrt{\alpha}, 2\sqrt{1-\alpha}), \quad ZZ' = (4-3\alpha), \quad (ZZ')^+ = (4-3\alpha)^{-1},$$

$$y = XZ' = \begin{pmatrix} 8-4\alpha \\ 4-3\alpha \end{pmatrix}, \quad D = XCX' = XX' - yy'/(4-3\alpha)$$
$$= N - yy'/(4-3\alpha).$$

$$\begin{aligned} F_\psi\{M(p), M(q)\} &= c'M^+(p)DM^+(p)c - c'M^+(p)c \\ &= D_{11} - 1 \\ &= \{16 - (8-4\alpha)^2/(4-3\alpha)\} - 1 \\ &= \{16\alpha(1-\alpha)/(4-3\alpha)\} - 1 \begin{cases} > 0, \alpha \in [.2735, .9139] \\ \leq 0, \alpha \notin [\ .\ ,\ .\ ] \end{cases} \end{aligned}$$

This latter derivative was also obtained by Ford (1976) by an alternative method. That the derivative can be positive is evidence of the non-optimality of $M(p)$.

<u>Theorem 3.2.2</u> Assume $\mathcal{U}$ and $\mathcal{M}$ as in theorem 3.2.1.

If $\psi(M)$ is concave on $\mathcal{M}$, then $M(p^*)$, for $p^* \in \mathcal{P}$, maximises $\psi(\cdot)$ on $\mathcal{M}$ iff

$$F\{M(p^*), M(q)\} \leq 0 \quad \forall\ q \in \mathcal{P},$$

that is,

$$\max_{q \in \mathcal{P}} F\{M(p^*), M(q)\} \leq 0.$$

This follows directly from theorem 2.5.5. □

We now state a design version of theorem 2.5.10.

<u>Theorem 3.2.3</u>

Let $\mathcal{U}_c$ be a compact design space. Let $\mathcal{P}_m$ denote the set of probability measures on $\mathcal{U}_c$ and let $\mathcal{M}_c$ denote the convex hull of the set $\{vv' : v \in \mathcal{U}_c\}$ so that $M \in \mathcal{M}_c$ iff $M = M(p) = E_p\{vv'\}$, for some $p(\cdot) \in \mathcal{P}_m$. Let $\mathcal{P}_{fd}$ denote the set of finite discrete probability distributions on $\mathcal{U}_c$ and for $p \in \mathcal{P}_{fd}$ let $\mathrm{Sup}(p) = \{v \in \mathcal{U}_c : p(v) > 0\}$. Call $\mathrm{Sup}(p)$ the support of p. Assume $\psi(\cdot)$ is concave on $\mathcal{M}_c$. Then

(i) For any $M \in \mathcal{M}_c$, $M = M(p)$, $p \in \mathcal{P}_{fd}$.

(ii) For $p^* \in \mathcal{P}_{fd}$ $M^* = M(p^*)$ maximises $\psi(\cdot)$ on $\mathcal{M}_c$ iff

$$F\{M(p^*), M(p)\} \leq 0 \quad \forall\ p \in \mathcal{P}_{fd}$$

or iff

(iii) If for $p^* \in \mathcal{P}_{fd}$, $M(p^*)$ is a differentiable point of $\psi(\cdot)$ then $M(p^*)$ maximises $\psi(\cdot)$ on $\mathcal{M}_c$ iff

$$F\{M(p^*), vv'\} \leq 0 \quad \forall\ v \in \mathcal{V}_c,$$

with equality in the case of $v \in \mathrm{Sup}(p^*)$. □

A proof would be analogous to that of theorem 2.5.10 and would use a design version of theorem 2.5.9. We have already used a corresponding analogue of theorem 2.5.8 and will further do so in the next section.

Again the practical value of theorem 3.2.5 is that if part (iii) applies this can yield explicit solutions for p^* or verify the optimality of an indirectly derived 'formula' for p^*. We shall see this in the next section.

If however p^* is obtained numerically then in the case of differentiable $M(p^*)$ theorem 3.2.1 is the one to which we shall appeal, the design space having been discretised.

Again the following analogue to corollary 2.5.10.1 sometimes proves useful.

<u>Corollary 3.2.3.1</u>

Let the set $\mathcal{V}_c$ in theorem 3.2.3 be a bounded convex set with a finite set of extreme points $\mathcal{V} = \{v_1, \ldots, v_J\}$. Let p^* be a solution to the optimum in that theorem. If $\psi(M)$ is increasing on the set of nonnegative definite matrices then for at least one p^*, $\mathrm{Sup}(p^*) \subset \mathcal{V}$; a p^* solving the design problem (P4) for this $\mathcal{V}$.

<u>Proof</u> Since $p^* \in \mathcal{P}_{fd}$ we can restrict attention to discrete probability distributions on $\mathcal{V}_c$.

Let $\tilde{M} = \sum_{i=1}^{m} q_i \tilde{v}_i \tilde{v}_i'$, $q_i \geq 0$, $\sum_{i=1}^{m} q_i = 1$,

where $\tilde{v}_1, \ldots, \tilde{v}_m$ are distinct elements of $\mathcal{V}_c$. Since $\mathcal{V}$ defines the extreme points of $\mathcal{V}_c$ we will have

$$\tilde{v}_i = \sum_{j=1}^{J} \lambda_{ij} v_j, \quad \lambda_{ij} \geq 0, \ \sum_{j=1}^{J} \lambda_{ij} = 1 .$$

Hence

$$\tilde{M} = \sum_{i=1}^{m} q_i \tilde{M}_i$$

where

$$\tilde{M}_i = \left(\sum_{j=1}^{J} \lambda_{ij} v_j\right)\left(\sum_{j=1}^{J} \lambda_{ij} v_j'\right)$$

Consider $$M = \sum_{i=1}^{m} q_i M_i$$

where $$M_i = \sum_{j=1}^{J} \lambda_{ij} v_j v_j'$$

Now $$x' M_i x = \sum_{j=1}^{J} \lambda_{ij} (x' v_j)^2$$

$$x' \widehat{M}_i x = \left(\sum_{j=1}^{J} \lambda_{ij} x' v_j\right)^2 \leq x' M_i x .$$

Hence $$x' M x \geq x' \widetilde{M} x ,$$

and the matrices $M_i - \widehat{M}_i$, $M - \widetilde{M}$ are nonnegative definite. In consequence $\psi(M) \geq \psi(\widetilde{M})$

Now
$$M = \sum p_j v_j v_j' , \quad p_j = \sum_{i=1}^{m} q_i \lambda_{ij} \geq 0 , \quad \sum_{j=1}^{J} p_j = 1 .$$

Hence $M = M(p)$ with $\mathrm{Sup}(p) \subset \mathcal{V}$ and hence we have shown that to each discrete design q on $\mathcal{V}_c$ there corresponds a design p on $\mathcal{V}$ such that $\psi(\cdot)$ takes on a larger value at the design matrix $M(p)$ of p than at that of q. Hence the maximum of $\psi(\cdot)$ over $\mathcal{M}_c$ must be given by the maximum of $\psi(\cdot)$ over $\mathcal{M} = \{M : M = E_p\{vv'\}, p \in \mathcal{P}\}$. □
In practice then we can regard $\mathcal{V}$ as the design space.

We now obtain some slight simplifications of theorem 3.2.2 in the case of the standard design criteria.

Theorem 3.2.4

Assume $\mathcal{V}$ and $\mathcal{M}$ as in theorem 3.2.1. Let $\psi(M) = -\lambda_{max}(AM^+A')$. Then $M(p^*)$, $p^* \in \mathcal{P}$, maximises $\psi(M)$ on $\mathcal{M}$ iff

$$\max_{q \in \mathcal{P}} G\{M(p^*), M(q)\} = G\{M(p^*), M(p^*)\}$$

i.e. $$G\{M(p^*), M(q)\} \leq G\{M(p^*), M(p^*)\} \quad \forall q \in \mathcal{P}$$

OR iff

$$G\{M(q), M(p^*)\} \geq G\{M(q), M(q)\} \quad \forall q \in \mathcal{P}$$

Proof The result follows from theorem 3.2.2 when we note that
$$F\{M(p), M(q)\} = G\{M(p), M(q)\} - G\{M(p), M(p)\} ,$$
a consequence of corollary (iv) of lemma 3.1.1. As we have said this does not imply differentiability. □

It is to be noted that equality in the first two conditions of the theorem may only be attained when q = p*. In contrast a stronger result can be proved for the other design criteria.

Theorem 3.2.5

Assume $\mathcal{V}$ and $\mathcal{M}$ as in theorem 3.2.1.

Let $\psi(M)$ be the D_A-optimal or c-optimal criterion or $\psi(M) = -tr(AM^+A')^t$. Then $M(p^*)$, $p^* \in \mathcal{P}$, maximises $\psi(M)$ on $\mathcal{M}$ iff

$$\max_{q \in \mathcal{P}} G\{M(p^*), M(q)\} = G\{M(p^*), v_j v_j'\} \quad \forall\ v_j \in Sup(p^*).$$

Proof These functions again enjoy the property that

$$F\{M(p), M(q)\} = G\{M(p), M(q)\} - G\{M(p), M(p)\}.$$

Hence conditions as in theorem 3.2.3 must hold. In particular we must have

$$\max_{q \in \mathcal{P}} G\{M(p^*), M(q)\} \le G\{M(p^*), M(p^*)\}.$$

Equality however must be obtained here for $v_j \in Sup(p^*)$, since the functions listed enjoy support differentiability. If p^* is to be the optimum for the whole design space, then it must be the optimum for its own support. Regarding the latter temporarily as the design space, so that we now have full differentiability at $M(p^*)$ on the appropriately redefined $\mathcal{M}$, and appealing to theorem 3.2.1 we conclude that

$$F\{M(p^*), v_j v_j'\} = 0 \quad , \quad v_j \in Sup(p^*)$$

ie

$$G\{M(p^*), v_j v_j'\} = G\{M(p^*), M(p^*)\} \quad , \quad v_j \in Sup(p^*).$$

Hence the theorem. □

Kiefer (1974) obtains results similar to these two theorems and also the result that p^* is that value p' which minimises $\sup_{p \in \mathcal{P}} G\{M(p'), M(p)\}$ under a general condition. This is that on $\mathcal{M}' = \{M : M \in \mathcal{M}, \psi(M) < \infty\}$ it be the case that $\psi(M) = P\{H(M)\}$, where $H(\cdot)$ is positive, homogenous of positive degree h,

and continuously differentiable in a neighbourhood of $\mathcal{M}$, and where $P(\cdot)$ is strictly decreasing and continuously differentiable on $H(\mathcal{M}')$ while $\log\{P^{-1}(\phi)\}$ is convex in ϕ. Also still $\psi(M)$ or some function $G\{\psi(M)\}$ must be concave on $\mathcal{M}$.

§3.2.4 The above two necessity theorems though slight simplifications of theorem 3.2.2 are still not of much practical value. In contrast the next two theorems are much simpler but they are however just sufficiency theorems.

The first theorem covers D_A-optimality and c-optimality.

Theorem 3.2.6

If the null space of $M_* = M(p^*)$ is contained in that of A, then a sufficient condition for M_* to maximise $\psi(M) = -\mathrm{tr}(AM^+A')^t$ over $\mathcal{M}$ is that

3.2.2 $$\mathrm{tr}(AM_*^+A')^t = \max_j v_j'M_*^+A'(AM_*^+A')^{t-1}AM_*^+v_j .$$

This includes c-optimality if we take t = 1 when A = c', and D_A-optimality is covered by the case t = 0 when $\mathrm{tr}(AM_*^+A')^t = \mathrm{tr}(I_s) = s$.

Proof We require to show that this condition guarantees the condition of theorem 3.2.5, namely that

3.2.3 $$\max_{q\in\mathcal{P}} G\{M_*, M(q)\} = G(M_*, M_*) = \mathrm{tr}(AM_*^+A')^t$$

For any M such that $\mathcal{N}(M) \subseteq \mathcal{N}(A)$ we have, in the notation of lemma 3.2.1

$$G(M,N) = \mathrm{tr}[(AM^+A')^{t-1}AM^+XCX'M^+A']$$

where, in particular, $N = XX'$, $C = I - D$, $D = Z'(ZZ')^+Z$

Hence

$$G(M,N) = \mathrm{tr}[(AM^+A')^{t-1}AM^+NM^+A'] - \mathrm{tr}[(AM^+A')^{t-1}AM^+XDX'M^+A']$$

Now the matrix XDX' is nonnegative definite and so can be expressed in the form RR'. Taking $N = M(q) = \sum q_jv_jv_j'$. we obtain

$$G(M,M(q)) = \sum_{j=1}^{J} q_j v_j'M^+A'(AM^+A')^{t-1}AM^+v_j - \mathrm{tr}\{R'M^+A'(AM^+A')^{t-1}AM^+R\}$$

The last term on the right is positive and hence

$$G(M,M(q)) \le \max_j v_j'M^+A'(AM^+A')^{t-1}AM^+v_j \quad \forall q \in \mathcal{P}$$

that is

3.2.4 $$\sup_{q\in\mathcal{P}} G\{M, M(q)\} \le \max_j v_j'M^+A'(AM^+A')^{t-1}AM^+v_j$$

Now particular cases of M(q) are of course the vertices v_jv_j' of $\mathcal{M}$ and we have

$$G\{M(p), v_jv_j'\} = \begin{cases} v_j'M^+(p)A'(AM^+(p)A')^{t-1}AM^+(p)v_j, & v_j \in L\{\mathrm{Sup}(p)\} \\ 0 & v_j \notin \text{''} \end{cases}$$

Hence the inequality in 3.2.4 may well be strict, whatever p, be it p^* or not. This though will not be the case if $v_* \in L\{\mathrm{Sup}(p)\}$, where v_* is any solution to $\max_j v_j'M_*^+A'(AM_*^+A')^{t-1}AM_*^+v_j$,

a condition which will be guaranteed under 3.2.2. This follows from the support differentiability of $\psi(M)$ for it is because of this that 3.2.2 can occur at p^*. It cannot occur at any other p and it need not occur at p^*.

The global p^* must of course be optimum for its support and so appealing to theorem 3.2.1, as in the proof of theorem 3.2.5, we have

$$\mathrm{tr}\left(AM_*^+A'\right) = v_j'M_*^+A'\left(AM_*^+A'\right)^{t-1}AM_*^+v_j \qquad \forall\ p_j^* > 0$$

If for any $M = M(p)$ the corresponding equalities do not hold then p cannot be optimal for its support and hence cannot be the optimal p^*.

Equation 3.2.2 will not hold if v_* is not given positive weight by p^* but if it does hold then $v_* = v_j$ for any v_j such that $p_j^* > 0$.

Equality in 3.2.2. and the resultant equality in 3.2.4 will guarantee 3.2.3. $\square$

The next theorem covers the generalised E-optimal criterion.

<u>Theorem 3.2.7</u>

If the null space of $M_* = M(p^*)$ is contained in that of A, then a sufficient condition for M to maximise $\psi(M) := -\lambda_{max}\left(AM^+A'\right)$ is that for some normalised eigenvector z corresponding to $\lambda_{max}(AM_*^+A')$

3.2.5 $$\lambda_{max}\left(AM_*^+A'\right) \geq \max_j v_j'M_*^+A'zz'AM_*^+v_j ,$$

If equality holds in 3.2.5 then

3.2.6 $$\lambda_{max}\left(AM_*^+A'\right) = v_j'M_*^+A'zz'AM_*^+v_j \qquad \forall\ p_j^* > 0.$$

<u>Proof</u> In principle we have to prove that 3.2.5 guarantees the conditions of theorem 3.2.4, namely that

$$G\left(M_*, M(q)\right) \leq G\left(M_*, M_*\right) \qquad \forall\ q \in \mathcal{P} ;$$

that is, for $M(q) = N = XX'$ we have to prove that

3.2.7 $$\min_{z\in\mathfrak{h}^*} z'AM_*^+XCX'M_*^+A'z \leq \lambda_{max}\left(AM_*^+A'\right)$$

where $\mathfrak{h}^*$ is the set of all normalised eigenvectors corresponding to $\lambda_{max}\left(AM_*^+A'\right)$.

Clearly 3.2.7 will be true if for some $z \in \mathfrak{h}^*$

3.2.8 $$z'AM_*^+XCX'M_*^+A'z \leq \lambda_{max}\left(AM_*^+A'\right)$$

Now since, in the notation of lemma 3.2.1, $C = I - D$, $D = Z'(ZZ')^+Z$, the left hand term decomposes to

$$\begin{aligned} & z'AM_*^+NM_*^+A'z - z'AM_*^+XDX'M_*^+A'z \\ & \leq z'AM_*^+NM_*^+A'z \\ & = \sum q_j z'AM_*^+v_jv_j'M_*^+A'z \qquad , \quad N = \sum q_jv_jv_j'. \end{aligned}$$

Hence 3.2.8 will follow if

$$\lambda_{max}\left(AM_*^+A'\right) \geq \sum q_j z'AM_*^+v_jv_j'M_*^+A'z$$

This in turn will be true if 3.2.5 holds.

Equality in 3.2.5 implies 3.2.6 since for any vector z

$$\sum p_j^* v_j' M_*^+ A' z z' A M_*^+ v_j = z' A M_*^+ M_* M_*^+ A' z = z' A M_*^+ A' z$$

Hence $z' A M_*^+ A' z \leq \max_j v_j' M_*^+ A' z z' A M_*^+ v_j$, but $z \in \mathfrak{z}^* \Longrightarrow z' A M_*^+ A' z = \lambda_{max}(A M_*^+ A')$ □

Note that 3.2.6 is not a consequence of support differentiability, for the directional derivatives $F\{M(p), v_j v_j'\}$ depend on the terms $v_j' M^+ A' z z' A M^+ v_j$, $M = M(p)$. Only if a common z, say $\tilde{z}$, solved $\min_{z \in \mathfrak{z}^*} z' A M^+ N M^+ A' z$ for all $N = M(q) \in \mathcal{M}$ would support differentiability obtain, and this $\tilde{z}$ would then satisfy the conditions of theorem 3.2.7.

§3.2.5

Theorems 3.2.6 and 3.2.7 prove sufficiency of conditions 3.2.2, 3.2.5, 3.2.6 for optimality. Unfortunately they are not necessary conditions although they can certainly be attained. Theorem 3.2.1 illustrates for instance that condition 3.2.2 holds when a whole design space of $J < k$ points forms the support of the optimum. Empirical investigations have also discovered 3.2.2 to hold in many other examples.

A counter example to necessity of the latter is very simple though. Suppose a design space is $\mathcal{V} = \{v_1, v_2\}$ where $v_1 = (1, 0)'$ and $v_2 = (x_1, x_2)'$, $x_2 \neq 0$. Take $A = c' = (1, 0)'$, as in Ex.3.2.1 so that the criterion is c-optimality and we take $t = 1$. We show that $p^* = (1, 0)'$.

Let $q = [(1 - \epsilon), \epsilon]'$. This for $0 < \epsilon \leq 1$ defines all other possible designs.

$$\begin{aligned} F\{M(p), M(q)\} &= F\{v_1 v_1', (1-\epsilon) v_1 v_1' + \epsilon v_2 v_2'\} \\ &= F\{v_1 v_1', v_1 v_1' + \epsilon(v_2 v_2' - v_1 v_1')\} \\ &= \epsilon F\{v_1 v_1', v_2 v_2'\} \qquad \text{from } 2.2.1 . \end{aligned}$$

$v_1 v_1' = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and from the results of Ex.3.2.1, $F\{v_1 v_1', v_2 v_2'\} = -1$ and hence $F\{M(p), M(q)\} = -\epsilon < 0$ for all q. Hence p is optimal.

The countering of the necessity of condition 3.2.2 is complete when we note that $(v_2' M^+(p) c)^2 = x_1^2$ and there was no restriction imposed on x_1.

However intuitively it would seem, if only because of support differentiability, that surely conditions similar to 3.2.2, 3.2.5, 3.2.6 would prove both necessary and sufficient for optimality. Indeed this is the case for as we have said Pukelsheim establishes the existence of such a set. Also Silvey obtains corresponding conditions for a general class of functions which in some instances are necessary and which he believes are necessary more generally.

We examine Silvey's results first. Consider that one further rather neat way of establishing optimality of a nondifferentiable $x^* = x(p^*) \in \mathcal{D}(\mathcal{U})$ for concave $\psi(\cdot)$ would be to find a differentiable z such that $\psi\{x(p^*)\} = \psi(z)$ and which satisfied $F(z,u_j) \leq 0$. Then of course

3.2.9 $$F\{z, x(q)\} = \sum q_j F(z,u_j) \leq 0$$

and so less interestingly we have established the optimality of x^* and of z if $z \in \mathcal{D}(\mathcal{U})$.

However, as we noted in (D5) of section 2.3, 3.2.9 does not require differentiable z to belong to $\mathcal{D}(\mathcal{U})$ and provided the concavity of $\psi(\cdot)$ extends appropriately beyond $\mathcal{D}(\mathcal{U})$ the above conditions will still establish the optimality of x^* from such a z. Silvey (1978) discovered the form that such a z should take in the case of functions $\psi(M) = f(AM^-A')$. His result is the following.

<u>Theorem 3.2.8</u>

Assume f is concave on $\mathcal{M}_A = \{M : M \in \mathcal{ND}, \mathcal{N}(M) \subseteq \mathcal{N}(A)\}$. Recall that $M \in \mathcal{M}_A$ is such that $Y'M = A$ for some Y. Let $M_* \in \mathcal{M} \wedge \mathcal{M}_A$ be singular with rank r and let $\mathcal{H}(M_*)$ be the set of matrices H of order $k \times (k-r)$ which are such that $M_{**} = M_* + HH'$ is nonsingular. Then a sufficient condition that $\psi(M)$ be maximal over $\mathcal{M}$ at $M_* = M(p^*)$ is that there exists an $H \in \mathcal{H}(M_*)$

$$F_\psi(M_{**}, v_j v_j') \leq 0 \quad , \quad j = 1, \cdots, J$$

<u>Proof</u>

Clearly M_{**} satisfies conditions implying that $\psi(M_{**}) = f(AM_{**}^{-1}A') \geq f(AM^-A')$ for all $M \in \mathcal{M}$. Optimality of M_* follows since $(M_* + HH')^{-1}$ is a generalised inverse of M_* when $H \in \mathcal{H}(M_*)$. See Searle (1971,p.23). Hence $AM_*^-A' = AM_{**}^{-1}A'$. □

Note that for $\psi(M) = -tr(AM^-A')^t$,

$$F(M_{**}, v_j v_j') = v_j' M_{**}^{-1} A'(AM_{**}^{-1}A')^{t-1} AM_{**}^{-1} v_j - tr(AM_{**}^{-1}A')^t .$$

Hence the result says that a sufficient condition for optimality of M is that for some generalised inverse M_*^- of M_* of the form $M_*^- = (M_* + HH')^{-1}$, $H \in \mathcal{H}(M_*)$ it should be the case that

3.2.10 $$tr(AM_*^-A')^t \geq \max_j v_j' M_*^- A'(AM_*^-A')^{t-1} AM_*^- v_j .$$

Clearly we have in 3.2.10 a sufficient inequality condition similar to 3.2.2, although one which has the disadvantage that it is invisible; one has to find an M_*^- from the many generalised inverses of M_*. Silvey however reports that it is also a necessary condition in the case of D_A-optimality and $\psi(M) = -tr(AM^-A')$. He believes it to be necessary more generally. In fact this is true for Pukelsheim obtains necessity in the case of $\psi(M) = -tr(AM^-A')^t$. He also obtains a necessary and sufficient condition similar to 3.2.5 for the generalised E_A-optimal criterion.

Pukelsheim's result uses the notion of a t-contracting generalised inverse of a matrix. He proves the following theorem in Pukelsheim (1980).

<u>Theorem 3.2.9</u>

Let $M_* = M(p^*) \in \mathcal{M} \cap \mathcal{M}_A$.

(a) In order that M_* maximise $\psi(M) = -tr\{(AM^+A')^t\} = -tr\{(AM^-A')^t\}$, in the case $-1<t<\infty$ it is necessary and sufficient that there exist a g-inverse (generalised inverse) G of M which satisfies

3.2.11 $$v_j' G' A'(AM_*^-A')^{t-1} AG v_j \leq tr(AM_*^+A')^t \quad , \quad j = 1, \cdots, J.$$

If optimality holds then every t-contracting g-inverse G of M satisfies these inequalities, and if these inequalities are satisfied then equality holds for all points of support of every optimal design p^*.

(b) Let $\lambda_{max} = \lambda_{max}(AM_*^+A') = \lambda_{max}(AM_*^-A')$.

Let $\mathfrak{z}^*$ denote the set of all normalised eigenvectors corresponding to λ_{max}. Let S denote the set of matrices $\{zz' : z \in \mathfrak{z}^*\}$.

In order that M_* maximise $[-\lambda_{max}(AM^+A')]$ it is necessary and sufficient that there exist a g-inverse G of M and a matrix E belonging to the convex hull of S which satisfy

3.2.12 $$v_j' G' A' E A G v_j \leq \lambda_{max}(AM_*^+A') \quad , \quad j = 1, \cdots, J .$$

If optimality holds then for every ∞-contracting g-inverse G of M_* there exists a matrix E in the convex hull of S satisfying these inequalities, and if these inequalities are satisfied then equality holds for all points of support of every optimal design p^*. $\square$

What is a t-contracting g-inverse G of M? The definition is linked to a dual problem. We note only that such a generalised inverse G of M satisfies the condition that

$$MG(C) \subset C \quad , \quad C = \{ c \in R^k : c'Nc \leq 1 \} ,$$

and N solves a dual problem. Thus Pukelsheim's approach is a duality one and so we will not pursue his proof. Clearly he has established a finite though invisible set of optimality conditions similar to those that would hold in the case of differentiability. Note that if $E = \sum_{j=1}^{r} q_j z_j z_j'$ for $z_j \in \mathfrak{Z}^*$ with $\Sigma q_j = 1$ then $E = zz'$ where $z = \sum_{j=1}^{r} \lambda_j z_j$, $\lambda_j^2 = q_j$, $\Sigma \lambda_j^2 = 1$ and so in equation 3.2.12 we have a condition similar to conditions 3.2.5, 3.2.6.

Accepting Pukelsheim's result it is now possible to establish the necessity conjectured by Silvey above because Pukelsheim shows that there always exists a t-contracting G-inverse of M of the form $G = (M + HH')^{-1}$ with H of the appropriate type, when $M \in \mathcal{M} \cap \mathcal{M}_A$.

This completes our list of results on optimality. We conclude with the following comments.

(i) In the light of theorem 3.2.6 for $\psi(M) = -tr(AM^-A')^t$, a check for Pukelsheim's test of optimality need only be considered at a design p if p is optimal for its support, i.e. if

$$v_j' M^+(p) A' (AM^+(p)A')^{t-1} AM^+(p) v_j = tr(AM^+(p)A')^t \quad \forall \; p_j > 0 .$$

The check will then only be necessary if the inequality

$$v_i' M^+(p) A' (AM^+(p)A')^{t-1} AM^+(p) v_i > tr(AM^+(p)A')^t$$

holds for some $v_i \notin L\{Sup(p)\}$ and does not hold for any zero weighted $v_i \in L\{Sup(p)\}$. This is a consequence of the support differentiability of $\psi(M)$.

(ii) We will typically determine postulated optimal designs using iterative procedures. If such a procedure is sensible we should feel confident about the optimality of the point of convergence even if a check for Pukelsheim's condition does prove necessary. Such confidence could be based on the properties of $\psi\{M(p)\}$ at the approaching iterations to the proposed solution, $\tilde{p}$

In particular suppose we have support differentiability and that $\tilde{p}$ is optimum for $U \cap L\{Sup(\tilde{p})\}$. We have to decide whether or not weight should be put at any of the remaining nonsupport points. Suppose $v_t \notin L\{Sup(\tilde{p})\}$ but that at a late iteration $p_t^{(n)} > 0$ but small. Then, if $F\{M(p^{(n)}), v_t v_t'\} << 0$, this strongly suggests in the light of theorem 3.2.1 that we should have $p_t^* = 0$. This idea will be considered in more detail in section 4.3.

This type of information could also be obtained from $F(N_j, v_j v_j')$ or $F(N, v_j v_j')$ where, for small $\varepsilon > 0$,

$$N_j = (1-\varepsilon)M(\tilde{p}) + \varepsilon v_j v_j' \quad , \quad N = (1-\varepsilon)M(\tilde{p}) + \varepsilon M(q)$$

where q shares out the weight ε equally between those $v_j \notin L\{Sup(\tilde{p})\}$

In particular if $\psi\{M(p)\} = -tr(AM^+(p)A')^t$ then the condition

$$v_j' N_j^+ A' (AN_j^+A')^{t-1} AN_j^+ v_j << tr(AN_j^+A')^t$$

would strongly suggest that it should be that $p_j^* = 0$.
Note that while $tr(AN_j^+A')^t \longrightarrow tr(AM^+(\tilde{p})A')^t$ as $\varepsilon \to 0$ it can be that

3.2.13 $$\underset{\varepsilon \to 0}{Lt}\, v_j' N_j^+ A' (AN_j^+A')^{t-1} AN_j^+ v_j \neq v_j' M^+(\tilde{p}) A' (AM^+(\tilde{p})A')^{t-1} AM^+(\tilde{p}) v_j$$

Hence, while the left hand term might be smaller than $tr(AN_j^+A')^t$, the right hand term could be larger than $tr(AM^+(\tilde{p})A')^t$.

The justification for these ideas is that if $\psi(\cdot)$ is differentiable at $\{(1-\varepsilon)M(\tilde{p}) + \varepsilon v_j v_j'\}$ for all small $\varepsilon > 0$ then it must be that

3.2.14 $$\underset{\varepsilon \to 0}{Lt}\, F\{(1-\varepsilon)M(\tilde{p}) + \varepsilon v_j v_j', v_j v_j'\} = F\{M(\tilde{p}), v_j v_j'\} ,$$

whether or not $M(\tilde{p})$ is a fully differentiable point of $\psi(\cdot)$.

Result 3.2.13 and 3.2.14 are not inconsistent since, if $v_j \notin L\{Sup(\tilde{p})\}$, then $F\{M(\tilde{p}), v_j v_j'\}$ does not involve the right hand term of 3.2.13.

(iii) Suppose we opt to check Pukelsheim's condition. A disadvantage is that there is the problem of finding a suitable G. How can this be done? First we note that Silvey (1978) hints at an iterative scheme for finding, if one exists, an H satisfying his condition.

An additional comment, of interest, is that in the case of the E_A-optimal criterion Pukelsheim's result seems to suggest a further example of problem (P1). The matrix E in condition 3.2.12 requires to be of the form $\sum p_j z_j z_j'$ for some $z_i \in \mathfrak{h}^*$. For a given generalised inverse G of M_* this condition will certainly be satisfied if

$$\min_p \max_j v_j' G' A' \left(\sum p_i z_i z_i'\right) A G v_j \leq \lambda_{max}(AM_*^+A') .$$

The minimising p^* solves (P1) for $\phi(p) = \max_j v_j' G' A' \left(\Sigma p_i z_i z_i'\right) A G v_j$.
It might be possible to formulate an iterative scheme for finding a suitable G and zz' by combining Silvey's proposal for finding a matrix G of the form $(M^* + HH')^{-1}$ with iterative schemes for solving (P1). This is admittedly a rather speculative suggestion.

§3.3 Examples Of Optimal Designs

§3.3.1 In these sections we derive the solution to some optimal design problems for which explicit solutions can be obtained, quote the solution in other cases to which we will refer and mention some related results.

First two examples of functions $\phi(p)$ are quoted which offer explicit solutions to (P1) and which have a form which particular design criteria $\psi\{M(p)\}$ take on in specific circumstances.

Ex. 3.3.1(i)

$$\phi_1(p) = -c(p_1 p_2 \quad p_J)^{-1} \quad , \quad c > 0$$

The solution to (P1) is given by $p_j^* = 1/J$. This function is the form which the D_A-optimal criterion would take on under a design p which assigns weight to every point in a design space $\mathcal{V}$ consisting of $J = s \leq k$ linearly independent vectors $v_1, .., v_J$.

From 1.3.6

$$M(p) = VPV'$$

where $V = [v_1\ v_2 \quad v_s]$ has rank(V) $= s$, while this is the full rank value of $P = \mathrm{diag}\{p_1 \cdots \cdot p_s\}$. Hence by theorem 6.2.18 of Graybill (1969)

$$M^+(p) = (V')^+ P^+ V^+ = (V^+)' P^{-1} V^+$$

where, by theorem 6.2.16 of Graybill, $V^+ = (V'V)^{-1} V'$.

Hence

$$A M^+(p) A' = W P^{-1} W'$$

where $W = A(V'V)^{-1}V'$ is $s \times s$. Therefore

$$-\det\{A M^+(p) A'\} = -\det(P^{-1})[\det(W)]^2 = -c(p_1 p_2 \quad p_s)^{-1}, \ c = [\det(W)]^2.$$

This illustrates that the D_A-optimal design on s linearly independent points (A being of order $s \times k$ and of rank s) assigns weight $1/s$ to each point. In particular this is the case with D_s-optimality, while by taking $A = I_k$ we can conclude that the D-optimal design on k linearly independent points assigns equal weighting $1/k$ to each point. This could also have been established by observing that since V is then $k \times k$ we have

$$\det\{M(p)\} = \det(P) \times [\det(V)]^2$$

which is therefore a particular case of

$$\phi_2(p) = c(p_1 p_2 \cdots p_J) \quad , \quad c > 0$$

These results are useful in practice for often it is the case that the support of a D-optimum p^* consists of k points and similarly with D_s-optimality, D_A-optimality.

More generally it has been established that a D-optimum p^* satisfies $p_j \le 1/k$ when Sup(p^*) contains more than k points. Similarly a D_A-optimum p^* satisfies $p_j^* \le 1/s$. We will prove this result later in section 9.1.

Ex.3.3.1(ii)

$$\phi_3(p) = -\sum_{j=1}^{J} a_j p_j^{-t} \quad , \quad a_j > 0 .$$

The solution for (P1) is $p_j^* = a_j^{1/(t+1)} \Big/ \left(\sum_{i=1}^{J} a_i^{1/(t+1)} \right)$,

as a check of the conditions of theorem 2.5.6 will verify. Since $\phi_3(p)$ is a homogeneous function of degree $(-t)$, these conditions are equivalent to $\partial\phi_3/\partial p_j^* = -t\phi_3(p^*)$.

Several examples of design problems are particular cases of $\phi_3(p)$.

Suppose that the design space consists of k linearly independent points $v_1, v_2, \ldots, v_k$, so that

$$M(p) = VPV' ,$$

where $V = [v_1 v_2 \ldots\ldots v_k]$ is $k \times k$ nonsingular, $P = \text{diag}\{p_1, \ldots, p_k\}$.

Then

$$\begin{aligned} tr\{M^{-1}(p)\} &= tr\{(V')^{-1} P^{-1} V^{-1}\} \\ &= tr\{P^{-1} V^{-1} (V^{-1})'\} = \sum a_j p_j^{-1} \end{aligned}$$

where $a_i = w_i' w_i$, w_i being the i^{th} column of $(V^{-1})'$.

If the k design points were orthogonal, that is $v_i' v_j = 0, i \ne j$ so that $V'V = D = \text{diag}\{d_1, \ldots, d_k\}$, $d_j = v_j' v_j$, and if t were integral then

$$M^t(p) = VD^{t-1} P^t V'$$

since the matrices D,P commute. For example,

$$M^3(p) = (VPV')^3 = VPV'VPV'VPV' = VPDPDPV' = VD^2P^3V'.$$

Hence

$$\begin{aligned} tr\{M^{-t}(p)\} &= tr\{(V')^{-1} P^{-t} D^{-(t-1)} V^{-1}\} \\ &= tr\{P^{-t} D^{-(t-1)} V^{-1} (V^{-1})'\} \\ &= \sum a_j p_j^{-t} . \end{aligned}$$

where $a_j = d_j^{-(t-1)} w_j' w_j$.

Similar reductions of $tr\left(AM^t(p)A'\right)^t$ can be made under corresponding conditions. In particular if the design space consists of s linearly independent points then

$$tr\{AM^{+}(p)A'\} = tr\{P^{-1}W'W\}$$

where W is as in example 3.3.1.

§ 3.3.2 We note D-optimal designs for two classes of regression models.

(a) polynomial regression; the regression is

$$E(y_x) = v_x'\theta \quad , \quad v_x \in \mathcal{V} = \{ v = (1, x, x^2, \dots, x^{k-1})' : -1 \le x \le 1 \} .$$

We have a standard continuous design space. Fedorov (1972,p.88,89) reports that the discrete D-optimum design is unique, having as its support the k roots of the polynomials $(1-x^2)P'_{k-1}(x)$ where $P_k(x)$ is the k^{th} Legendre polynomial

$$P_k(x) = \sum_{n=0}^{N} \left\{ \frac{(-1)^n (2k-2n)!\, x^{k-2n}}{2^k n!\,(k-n)!\,(k-2n)!} \right\} \quad , \quad N = \begin{cases} k/2 & , \ k \text{ even} \\ (k-1)/2 & , \ k \text{ odd} \end{cases}$$

Since Sup(p^*) contains k points the D-optimum design on it assigns weight (1/k) to each of these. In the case k = 5 we obtain the simplification

$$(1-x^2)\,P_4'(x) = (5/2)\,x\,(7x^2-3)(1-x^2) \ ,$$

so the support of p^* is given by $x = 0, \pm 1, \pm\sqrt{3/7} = \pm\cdot 655$.

Fedorov also reports unique solutions to the D-optimum design for polynomial regression when the constant variance assumption is replaced by $\mathrm{Var}(y_x) = \sigma^2\lambda(x)$ where $\lambda(x)$ is of known form. Again Sup(p^*) consists of k points for four different examples of $\lambda(x)$.

(b) trigonometric regression; the regression is

$$E(y_x) = v_x'\theta \quad , \quad v_x \in \mathcal{V},$$

$$\mathcal{V} = \{ v = (w_1, \dots\dots, w_k)' : w_1 = 1 , w_{2r} = \cos(rx) , w_{2r+1} = \sin(rx) ; r = 1, \dots, m ; \\ k = 2m+1 ; 0 \le x \le 2\pi \}$$

Fedorov (1972,p.96) reports that any design assigning equal weighting to a support containing at least k equally spaced points in $(0, 2\pi)$ is D-optimum e.g.

$$\mathrm{Sup}(p) = \{ v_x : x = 2\pi(i-1)/n + Q \ ; \ i = 1, \dots, n \ ; \ n \ge k = 2m+1 ; x_n + Q \le 2\pi \}$$

As a consequence any design, D-optimum for a given value of m is D-optimum for smaller values of m .

§3.3.3 We now examine some consequences of a geometry of D-optimality which is given more prominence in the duality approach to optimal design.

The conditions of theorems 3.2.1, 3.2.3 imply that p* is D-optimum for $\mathcal{V}=\{v_1,..,v_J\}$ or for $\mathcal{V}$ continuous if

$$v'M^{-1}(p^*)v = k \quad , \quad v \in \mathrm{Sup}(p^*)$$
$$v'M^{-1}(p^*)v \le k \quad , \quad v \notin \mathrm{Sup}(p^*) .$$

Since $M(p^*) \in \mathcal{PD}$, this says that the points in Sup(p*) must lie on the boundary of the ellipsoid $E(p^*)=\{x : x'M^{-1}(p^*)x \le k\}$, which is centred on the origin, while the non-optimal support points should typically lie inside E(p*). Of course for the general problem (P2) we can similarly say that Sup(p*) must lie on the boundary of $R(p^*)=\{y : F\{x(p^*), y\} \le 0\}$ and that this set should otherwise contain $\mathcal{U}$. We will refer to this later in respect of an algorithm. However the set E(p*) is a particularly simple transformation of the corresponding set R(p*) and, in combination with the known D-optimum solution on a support of k points, can help to derive or shed light on D-optimum designs. We consider two examples.

Ex. 3.3.3(i) Suppose $\mathcal{V}$ is an ellipsoid, say $\mathcal{V}=\{v : v'Nv \le c\}$. Clearly a D-optimum design p* must be such that E(p*) coincides with $\mathcal{V}$. In view of the above conditions this will be the case if

$$M^{-1}(p^*) = (k/c)N \quad , \quad M(p^*) = (c/k)N^{-1} .$$

This can be achieved by a p* with a particular type of k-point support.

Suppose $\mathrm{Sup}(p) = \{v_1,..,v_k\}$ and that $p_j = 1/k$, which of course must be the case if p is to be at all optimal. Then in the meaning of formula 1.3.9, $V_p = [v_1.....v_k]$ is of order k×k and $P_p = (1/k)I_k$, and hence

$$M(p) = (1/k)V_pV_p'$$

Hence if p is to be p* we must have

$$V_pV_p' = cN^{-1}$$

Since $N \in \mathcal{PD}$ we will have $N^{-1} = QD^{-1}Q'$ where the columns of orthogonal Q are the normalised eigenvectors of N, and D is diagonal, with entries the corresponding positive eigenvalues. Hence the solution is to choose Sup(p) such that

$$V_p = \sqrt{c}\,QD^{-1/2}$$

This implies that
$$V_p'V_p = cD^{-1/2}Q'QD^{-1/2} = cD^{-1}$$
$$V_p'NV_p = cD^{-1/2}Q'QDQ'QD^{-1/2} = cI$$

and hence that the support $v_1,\ldots,v_k$ must be such that
$$v_i'v_j = v_i'Nv_j = 0 \ , \quad i \neq j \quad ; \quad v_j'Nv_j = c$$

Any design p then which assigns weight $1/k$ to a set of k orthogonal points and in fact N-orthogonal points lying on the boundary of $\mathcal{U}$ is D-optimal. So also is any convex combination of such designs.

In particular suppose $N = I$, $c = 1$ so that $\mathcal{U}$ is the unit sphere and suppose $k = 4$. Let

$$\begin{aligned}
v_1 &= (1,0,0,0)' \\
v_2 &= (0,1,0,0)' \\
v_3 &= (0,0,1,0)' \\
v_4 &= (0,0,0,1)' \\
v_5 &= (.5, .5, .5, .5)' \\
v_6 &= (.5, .5, -.5, -.5)' \\
v_7 &= (.5, -.5, .5, -.5)' \\
v_8 &= (.5, -.5, -.5, .5)'
\end{aligned}$$

Then for any $\lambda \in [0,1]$ the design p allocating weights in the manner
$$\left\{ \begin{array}{cccccccc} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 & v_8 \\ \lambda/4 & \lambda/4 & \lambda/4 & \lambda/4 & (1-\lambda)/4 & (1-\lambda)/4 & (1-\lambda)/4 & (1-\lambda)/4 \end{array} \right\}$$

is D-optimum for $\mathcal{U}$ and of course for $\mathrm{Sup}(p) = \{v_1,\ldots,v_8\}$

Ex. 3.3.3(ii) The following is a more realistic problem. Suppose that a design space $\mathcal{U}$ consists of $(k + 1)$ points. There are two possibilities. Either $\mathrm{Sup}(p^*) = \mathcal{U}$ or one of the design points is a nonsupport point in which case p^* assigns weight $(1/k)$ to the remainder. In total there are $(k+2)$ different possible optimum supports. However it is possible to identify the correct support set.

In view of the invariance of the D-optimal criterion under a linear transformation of $\mathcal{U}$ we can without loss of generality assume that $\mathcal{U} = \{e_1, \ldots, e_k, \underline{y}\}$ where $e_1, \ldots, e_k$ are the k unit vectors and $y = (y_1, \ldots, y_k)'$.

We first prove a subsidiary lemma.

Lemma 3.3.1

Let S denote the unit sphere $\{\underline{x} : \underline{x}'\underline{x} \le 1\}$.

Let R $= \{\underline{x} = (x_1, \ldots, x_k)' : \underline{x}'\underline{x} > 1 \; ; \; (2x_j^2 - \underline{x}'\underline{x}) \le 1, \; j = 1, \ldots, k\}$

Let R_j $= \{\underline{x} = (x_1, \ldots, x_k)' : (2x_j^2 - \underline{x}'\underline{x}) > 1\}$, $j = 1, \ldots, k$.

These (k+2) sets form a partition of k-dimensional euclidean space E_k.

Proof First the following diagram for k = 2 well illustrates the result

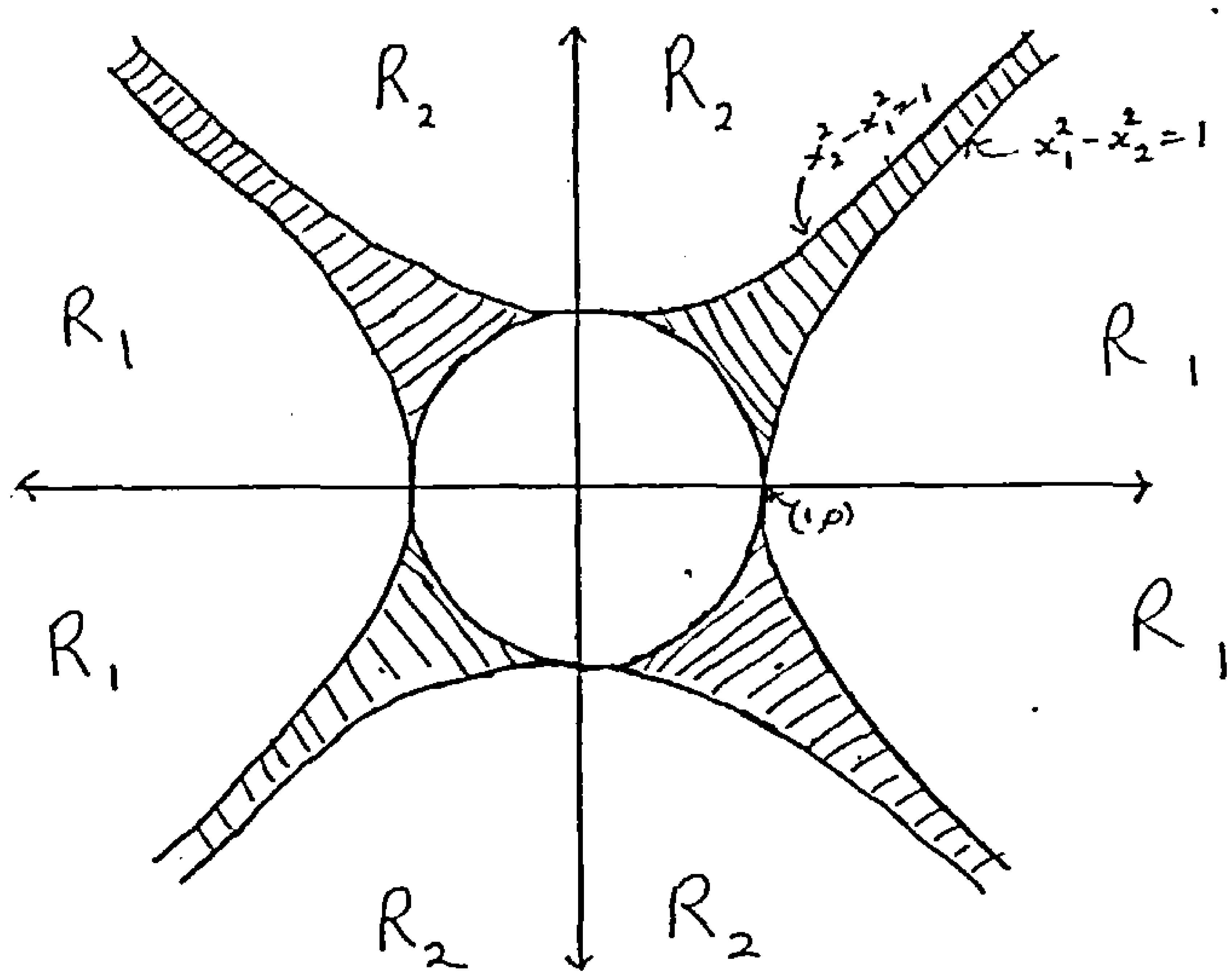


Figure 3.3.1

The set R is the shaded region bounded by the unit circle and the hyperbolas $x_1^2 - x_2^2 = 1$, $x_2^2 - x_1^2 = 1$ and hence lying 'between' S and R_1 and R_2.

It is clear that $R \cap S = \phi$ and that $R \cap R_j = \phi$. We need to show that $R_j \cap S = \phi$, $R_i \cap R_j = \phi$, $(i \neq j)$.

Suppose $\underline{x}'\underline{x} \le 1$. Then $\underline{x}'\underline{x} + 1 \ge 2\underline{x}'\underline{x} \ge 2x_j^2$ with equality only if $x_j = 1$. Hence $R_j \cap S = \phi$.

Now suppose that $(2x_j^2 - \underline{x}'\underline{x}) > 1$ or $2x_j^2 > (1 + \underline{x}'\underline{x})$. Then $x_j^2 > \left(1 + \sum_{t \neq j} x_t^2\right)$. Hence $(1 + \underline{x}'\underline{x}) > \left(2 + 2\sum_{t \neq j} x_t^2\right) > 2 > 2x_i^2$, $i \neq j$
Thus $R_i \cap R_j = \phi$, $i \neq j$.

Finally the union of the (k+2) sets must be E_k since either $\underline{x}'\underline{x} \le 1$ or the reverse or since the inequality $2x_j^2 - \underline{x}'\underline{x} > 1$ either holds for j = 1, or for j = 2,..., or for j = k, or does not hold for any j. □

In conclusion a point $\underline{y} = (y_1, \ldots, y_k)'$ must belong to one of these sets and only to one of them. We can now prove the result of interest.

<u>Lemma 3.3.2</u>

Let p^* denote the D-optimum design on $\mathcal{U} = \{e_1, \ldots, e_k, \underline{y}\}$

Then

(a) $\underline{y} \notin \mathrm{Sup}(p^*)$ iff $\underline{y} \in S$

(b) $\underline{e}_j \notin \mathrm{Sup}(p^*)$ iff $\underline{y} \in R_j$, $j = 1, \ldots, k$

(c) $\mathcal{U} = \mathrm{Sup}(p^*)$ iff $\underline{y} \in R$

<u>Proof</u> Part (a) is simple. The uniform weighting D-optimum design $\breve{p}$ on $\{e_1, \ldots, e_k\}$ has $M(\breve{p}) = (1/k)I$. Hence if $\underline{y}'\underline{y} \le 1$ then $\underline{y}'M^{-1}(\breve{p})\underline{y} \le k$. Hence the conditions of theorem 3.2.1 are satisfied. Conversely if $\underline{y} \notin S$ then p is not optimum and further it must be that $\underline{y} \in \mathrm{Sup}(p^*)$, for no other feasible optimum design can exclude it from its support.

The proof of (a) and (c) is as follows:

Let A_1 be the k×k matrix

$$A_1 = \begin{bmatrix} 1/y_1 & 0 & & 0 \\ -y_2/y_1 & 1 & & \\ \vdots & & 1 & \\ \vdots & & & 0 \\ -y_k/y_1 & 0 & 0 & 1 \end{bmatrix}$$

Then $A_1\underline{y} = e_1$, $A_1 e_j = e_j$, $j = 2, \ldots, k$; i.e. A_1 maps $\{\underline{y}, e_2, \ldots, e_k\}$ onto $\{e_1, \ldots, e_k\}$ while e_1 is mapped onto $z_1 = A_1 e_1 = (1/y_1, -y_2/y_1, \ldots\ldots, -y_k/y_1)'$.

Then $z_1'z_1 = (1 + \underline{y}'\underline{y} - y_1^2)/y_1^2$ and hence $z_1 \in S$ iff $\underline{y} \in R_1$. Now if $z_1 \in S$ then by appealing again to the invariance property of D-optimality we can claim that $\mathrm{Sup}(p^*)$ excludes e_1 and that p^*

is the uniform design on $\{e_2,\ldots,e_k,y\}$. Conversely if $z_1 \notin S$ then this cannot be the optimal design and moreover, again it must be that $e_1 \in \text{Sup}(p^*)$, for no other feasible optimum design can exclude e_1 from its support.

Similarly $z_j = (A_j e_j) \in S$ iff $\underline{y} \in R_j$ where A_j maps y onto e_j and e_i $(i \neq j)$ onto itself. Hence we can conclude that $e_j \notin \text{Sup}(p^*)$ iff $y \in R_j$.

Part (c) follows since if $\underline{y} \notin S$ and $\underline{y} \notin R_j$ then $\underline{y} \in R$, and, since $\mathcal{U}$ is the support for the only other possible optimum design. Hence the lemma. □

The lemma could be applied to an arbitrary set of (k+1) vectors by say taking $y = V^{-1}v_{k+1}$ where $V = [v_1 \ldots v_k]$, although the matrix inversion may be undesirable.

There still may be the necessity to evaluate p* for, while the optimum has been found if Sup(p*) is one of the k-point subsets, this is not so if $\text{Sup}(p^*) = \mathcal{U}$. It would seem that such a p* must be determined numerically except in special circumstances.

If $\{v_1,v_2,\ldots,v_{k+1}\} = \{e_1,\ldots,e_k,\underline{y}\}$ then the equations $v_j' M^{-1}(p) v_j = k$, of theorem 2.7.1, simplify to

$$p_j^{-1} - p_{k+1}(y_j^2/p_j^2)/(1+q\,p_{k+1}) = k \quad , \quad j=1,\ldots,k \, ,$$

and

$$q/(1+q\,p_{k+1}) = k \, ,$$

where

$$q = \sum_{j=1}^{k} y_j^2/p_j \, .$$

Simple though these may seem, an explicit solution appears to elude them in general.

One exception is the case $\underline{y} = (y_1,\ldots,y_k)$ where $y_j = \pm x$, $j=1,\ldots,k$ then $p_{k+1} = (kx^2-1)/(k^2x^2-1)$, $p_j = (k-1)x^2/(k^2x^2-1)$, $j=1,\ldots,k$.

Another instance of an explicit solution is seen below.

§3.3.4 Wynn's example

We end by citing what has now become a rather celebrated example and which originated in Wynn (1970,1972).

This takes k = 3 and $\mathcal{V} = \{v_1, v_2, v_3, v_4\}$ where

$$v_1 = (1,-1,-1)', \quad v_2 = (1,-1,1)', \quad v_3 = (1,1,-1)', \quad v_4 = (1,2,2)'$$

and so we have a particular case of the problem considered in the previous section. In fact though, Wynn had in mind the more realistic regression model

$$E(y) = v'\theta \quad , \quad v \in \mathcal{W}$$

where $\mathcal{W} = \{ \underset{\sim}{x} = (x_1, x_2, x_3)' : x_1 = 1, \ (x_2, x_3) \in Q \}$,
where Q is the quadrilateral with vertices $(-1,-1), (-1,1), (1,-1), (2,2)$.
However $\mathcal{W}$ is clearly a bounded convex set with extreme points v_1, v_2, v_3, v_4. An appeal to corollary 3.2.3.1 establishes that $Sup(p^*) \subseteq \mathcal{V}$.

We note that for $V = [v_1 \ v_2 \ v_3]$

$$V^{-1} = \begin{bmatrix} 0 & -1/2 & -1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$

and hence that $\underset{\sim}{y} = V^{-1}v_4 = (-2, 3/2, 3/2)' \in R$.

Hence by lemma 3.3.2 $Sup(p^*) = \mathcal{V}$. Calculation of p^* is however simplified by employing symmetry arguments to justify $p_2^* = p_3^*$. Making the appropriate substitutions in the equations $v_j' M^{-1}(p) v_j = 3$, an explicit solution can be obtained; namely

$$p_1^* = 4/32 \ , \ p_2^* = p_3^* = 9/32 \ , \ p_4^* = 10/32 .$$

Many authors have since made use of this example and we also refer to it in several contexts. Indeed its use is so widespread that it is deserving of an evaluation of $M^{-1}(p)$ for general p.
Let $M(p) = M = \{m_{ij}\}$, $M^{-1}(p) = N = \{n_{ij}\}$.
Then

$$\begin{aligned}
m_{11} &= (p_1+p_2+p_3+p_4) = 1 \\
m_{12} = m_{21} &= (-p_1-p_2+p_3+2p_4) \\
m_{13} = m_{31} &= (-p_1+p_2-p_3+2p_4). \\
m_{23} = m_{32} &= (p_1-p_2-p_3+4p_4) \\
m_{22} = m_{33} &= (p_1+p_2+p_3+4p_4) = 1 + 3p_4 \\
\det(M) &= 4^2 p_1p_2p_3 + 6^2 p_1p_2p_4 + 6^2 p_1p_3p_4 + 8^2 p_2p_3p_4
\end{aligned}$$

$$n_{11} = (4p_1p_2+4p_1p_3+16p_2p_4+16p_3p_4)/\det(M)$$
$$n_{12} = n_{21} = (4p_1p_2+4p_2p_4-12p_3p_4)/\det(M)$$
$$n_{13} = n_{31} = (4p_1p_3+4p_3p_4-12p_2p_4)/\det(M)$$
$$n_{23} = n_{32} = (4p_2p_3-9p_1p_4-3p_2p_4-3p_3p_4)/\det(M)$$
$$n_{22} = (4p_1p_2+9p_1p_4+4p_2p_3+p_2p_4+9p_3p_4)/\det(M)$$
$$n_{33} = (4p_1p_3+9p_1p_4+4p_2p_3+9p_2p_4+p_3p_4)/\det(M)$$

We will make use of these formulae later. That for det(M) is an illustration of the fact that $\det\{M(p)\}$ is a polynomial of degree k in the components of p with the format that the coefficient of the term $p_{i_1} \times p_{i_2} \times \ldots \times p_{i_k}$, where $i_r \neq i_s$, has the typically positive value of $\{\det(V)\}^2$, where $V = \begin{bmatrix} v_{i_1} & v_{i_2} & \ldots & v_{i_k} \end{bmatrix}$, while all other possible terms have zero coefficients. See theorems 1.1.1, 1.1.2 of Fedorov (1972, p.15). This renders $\det\{M(p)\}$ a homogeneous polynomial of degree k with positive coefficients.

CHAPTER 4

ALGORITHMS, SOME PRELIMINARIES

§4.1 Introduction

As has been already stated it is typically not possible to evaluate an explicit solution p^* to problem (P2) or in particular to derive an optimal regression design explicitly. Iterative techniques must be employed and so special algorithms have been devised particularly for the design problem, though also for other examples of (P1). With one exception we will consider these in subsequent chapters.

The main concern in this chapter is to consider a few points about algorithms, some general, some particular to our problem.

Recall first that in chapter 1 we have indicated why there is a need for special algorithms. It is assumed that we wish to identify an optimising p^*. Typically p^* will lie on the boundary of $\mathcal{P}(\mathcal{U})$. Certainly this will be the case if $\mathcal{U}$ or $\mathcal{V}$ is a discretisation of a continuous space.

The problem is that we do not know Sup(p^*), otherwise we could take $\mathcal{U}$ to be the latter, in which case the only active constraint would be $\sum p_j = 1$. It should be simple enough to devise modifications to standard iterations to ensure this constraint.

However coping with boundary constraints such as $p_j \geqslant 0$ cannot be so readily achieved, although attempts to use standard algorithms have been made. The iterative technique considered at the end of this chapter is a particular example. Conceived for the design context it, however, first finds an optimising M_* which is indirectly guaranteed to belong to $\mathcal{M}$. There is therefore an additional problem of finding p^* such that $M(p^*) = M_*$.

Really what this serves to demonstrate is that algorithms for constrained optimisation are to some extent thin on the ground.

Wu, a numerical analyst, has acknowledged these points to the extent of making his own contribution to the design problem. See Wu (1976, 1978a, 1978b). This will be considered later.

§4.2 Basics And Motivation Of An Algorithm

§4.2.1 An algorithm for an optimisation problem is of course a sequence of successive approximations to a solution x^*. One makes an initial guess x_0 to x^* and tries by some means or other to derive from x_0, a hopefully improved approximation x_1. Then by the same means a further improvement x_2 is derived from x_1 and so on.

A sequence $x_0, x_1, \ldots\ldots\ldots\ldots$ is thus generated in the belief or hope that the sequence will converge to the optimum x^*, or at the very least will be such that for some finite r, $x_r \doteq x^*$. One may require to add the qualification that x_0 be 'near' x^*.

We consider now some general ideas for formulating an algorithm for the problem of maximising a concave function $\psi(x)$ over a set S.

In the above sequence we will have $x_{r+1} = g(x_r)$ for some function $g(\cdot)$ implicit or explicit in form. A specific function $g(\cdot)$ will define an algorithm. What sort of function $g(\cdot)$ will make for a good algorithm?

Given that we are considering the maximisation of a concave function $\psi(\cdot)$, it would seem that it should suffice to have $\psi(x_{r+1}) \geq \psi(x_r)$, or that this inequality should hold almost always. At the very least there should be an upward trend in the values $\psi(x_0)$, $\psi(x_1)$, $\psi(x_2)\ldots\ldots$. In conventional phraseology we would wish the algorithm to be a hill climbing technique. Concavity guarantees that there are no wrong hills to climb.

We must also, have iterates remaining in the feasible region, i.e. $x_r \in S$.

There are various ways in which we could derive an x_{r+1} from an x_r, for which we could believe (much) more often than not that $\psi(x_{r+1}) \geq \psi(x_r)$. We consider some now.

(a) We could take x_{r+1} to be that value which maximises a simple approximation to $\psi(\cdot)$ at x_r, subject to $x_{r+1} \in S$. A Taylor series expansion readily springs to mind.

(b) Similarly x_{r+1} could be determined by a restricted but simple maximisation of $\psi(\cdot)$ about x_r; for example, find z_r to maximise $\psi(x_r+z)$ over the set $S' = \{z : (x_r+z)\in S\}$; set $x_{r+1} = x_r + z_r$. Clearly we will have $\psi(x_{r+1}) \geq \psi(x_r)$.

(c) Suppose $x_r = x_r^*$, the value maximising $\psi(\cdot)$ over some subset S_r of S or some approximation S_r to S. In the light of x_r^*, expand S_r to some larger subset S_{r+1} of S so that $S_r \subset S_{r+1} \subset S$, or derive an improved approximation S_{r+1} to S. Then take x_{r+1} to be the maximum of $\psi(\cdot)$ over S_{r+1}. If $S_r \subset S_{r+1}$ we will again have $\psi(x_{r+1}) \geq \psi(x_r)$.

In the case of problem (P2) we would be adopting this approach if we were to solve (P2) for a well chosen subset $\mathcal{U}_0$ of $\mathcal{U}$. Then for $\mathcal{U}_1$, $\mathcal{U}_2, \ldots$, where $\mathcal{U}_{r+1}$ augments $\mathcal{U}_r$ by the addition of one or more of the vertices excluded by $\mathcal{U}_r$.

In any instance above when we have $\psi(x_{r+1}) > \psi(x_r)$, then, if $\psi(\cdot)$ is concave, it must be that $F(x_r,x_{r+1}) > 0$, and this must be true even in those instances when the former inequality is not satisfied, if the technique employed is a sensible one, for this is the only reasonable definition of the latter, that $F(x_r,x_{r+1}) > 0$.

Hence we could devise iterative techniques which chooses x_{r+1} using the latter as a criterion. This can be done by looking around for a direction m_r in which to step from x_r, m_r being chosen then to satisfy $F(x_r,x_r+m_r) > 0$. Then it is necessary to decide on the magnitude $\alpha_r > 0$ of the step to be taken in the m_r- direction. Thereby

$$x_{r+1} = x_r + \alpha_r m_r \quad , \quad r = 0, 1, 2, \ldots\ldots .$$

(Note that in this notation m_r, α_r are defined for $r = 0$; $x_1 = x_0 + \alpha_0 m_0$).

The chosen m_r must be such that $(x_r + \alpha m_r) \in S$, for all small positive α.

We will have $F(x_r,x_{r+1}) > 0$, since $F(x_r,x_{r+1}) = \alpha_r F(x_r,x_r+m_r)$.

In the case of problem (P2) a convenient way of determining m_r is to discover a point $y_r \in \mathcal{P}(\mathcal{U})$ towards which it is good to move, namely a point y_r such that $F(x_r,y_r) > 0$. This defines m_r as $m_r = y_r - x_r$ so that

$$x_{r+1} = (1-\alpha_r)x_r + \alpha_r y_r .$$

Optimal choices of y_r could be a value of y which solves or nearly solves either $\max_{y\in S} F(x_r,y)$ or $\max_{y\in S} F^A(x_r,y)$. If it is difficult to determine such optima but easy to identify a subset of y's such that $F(x_r,y) > 0$, then we might opt to take y_r to be a value from this set. A particular example of approach (a) will be seen to fall into this general class.

A converse approach could be to identify a point $y_r \in S$, away from which it is good to move from x_r. A formula for m_r is then $m_r = x_r - y_r$, so that, still for $\alpha_r > 0$,

$$x_{r+1} = (1+\alpha_r)x_r - \alpha_r y_r .$$

If we are to have $F(x_r, x_r+m_r) > 0$, it must be that $F(x_r, 2x_r-y_r) > 0$. At differentiable x_r this will be the case if $F(x_r,y_r) < 0$.

§4.2.2 The choice of α_r is clearly important. At the very least we must have $x_{r+1} \in S$. If S is convex then this will mean that α_r must satisfy $\alpha_r < \bar{\alpha}_r$ where $\bar{\alpha}_r > 0$ is the smallest value such that $(x_r + \alpha m_r) \notin S$ for all $\alpha > \bar{\alpha}_r$. This is assuming that $F(x_r, x_r+m_r) > 0$ and that $(x_r+\alpha m_r) \in S$ for small positive α.

An optimal choice would clearly be that value α_r^* which solves

4.2.1 "maximise $\psi(x_r + \alpha m_r)$ over $0 < \alpha \leq \bar{\alpha}_r$".

We note that we will also use the notation $\bar{\alpha}_r(m_r)$, $\alpha_r^*(m_r)$ to denote the above terms, or $\bar{\alpha}_r(y_r)$, $\alpha_r^*(y_r)$ when $m_r = \pm(y_r - x_r)$. For the latter case consider $z_r(\gamma, y_r) = (1-\gamma)x_r + \gamma y_r$. There will be a largest value $\gamma_\ell(y_r)$, $\gamma_\ell(y_r) < 0$, and a smallest value $\gamma_u(y_r)$, (>0), such that $z_r(\gamma, y_r) \notin S$ for $\gamma < \gamma_\ell(y_r)$ and $\gamma > \gamma_u(y_r)$. In terms of these

4.2.2 $$\bar{\alpha}_r(y_r) = \begin{cases} \gamma_u(y_r) & \text{if } m_r = (y_r - x_r),\ F(x_r,y_r) > 0 \\ -\gamma_\ell(y_r) & \text{if } m_r = (x_r - y_r),\ F(x_r,y_r) < 0 . \end{cases}$$

Clearly 4.2.1 is a one variable function maximisation, namely for $f(\alpha) = \psi(x_r + \alpha m_r)$. If $\psi(\cdot)$ is strictly concave then so is $f(\alpha)$ and α_r^* must be unique. Either $\alpha_r^* = \bar{\alpha}_r$ or $\alpha_r^* < \bar{\alpha}_r$. If $\psi(\cdot)$ is differentiable at $(x_r+\alpha m_r)$ for $0<\alpha\leq\bar{\alpha}_r$, then the latter will be the case if and only if $f(\alpha)$ has a (unique) stationary value in $(0,\bar{\alpha}_r)$. This will be a possibility, since $f'(0)>0$ when $F(x_r, x_r+m_r) > 0$.

It follows then that

4.2.3 $$\alpha_r^* = \min(\breve{\alpha}_r^+, \bar{\alpha}_r) ,$$

where $\breve{\alpha}_r^+$ denotes the smallest positive solution, if solutions exist, to the equation

4.2.4 $$f'(\alpha) = 0 .$$

If no solution exists then $\alpha_r^* = \bar{\alpha}_r$ and 4.2.3 would still be defined if we let $\breve{\alpha}_r^+ = \infty$ in this instance.

Of interest is that equation 4.2.4 is equivalent to

4.2.5 $$F_\psi\left[x_r+\alpha m_r, x_r+\beta m_r\right] = 0 \quad \text{for all } \beta .$$

That this equation should be satisfied by $\alpha = \alpha_r^*$ in the case $\alpha_r^* < \bar{\alpha}_r$ would in fact follow from theorem 2.5.2, taking S there to be the set $\{x : x = x_r+\alpha m_r, 0<\alpha\leq\bar{\alpha}_r\}$. As with other differentiable 'interior' maxima, α_r^* must be a constrained stationary value of $f(\cdot)$.

The equivalence however has the following more direct derivation.

By definition, for any γ,
$$f'(\gamma) = \operatorname*{Lt}_{\varepsilon\to 0}\left[f(\gamma+\varepsilon) - f(\gamma)\right]/\varepsilon .$$

Hence
$$f'(\gamma) = \operatorname*{Lt}_{\varepsilon\to 0}\left[\psi\{x_r+(\gamma+\varepsilon)m_r\} - \psi\{x_r+\gamma m_r\}\right]/\varepsilon = G_\psi\left[x_r+\gamma m_r, m_r\right].$$
while
$$F_\psi\left[x_r+\gamma m_r, x_r+\beta m_r\right] = G_\psi\left[x_r+\gamma m_r, (\beta-\gamma)m_r\right] = (\beta-\gamma)\, G_\psi\left[x_r+\gamma m_r, m_r\right].$$
So we obtain that
$$f'(\gamma) = (\beta-\gamma)^{-1} F_\psi\left[x_r+\gamma m_r, x_r+\beta m_r\right] .$$
In particular

4.2.6 $$f'(\gamma) = (1-\gamma)^{-1} F_\psi\left[(1-\gamma)x_r+\gamma y_r, y_r\right] , \quad m_r = y_r - x_r$$

while $\breve{\alpha}_r$ will satisfy
$$F_\psi\left[(1-\breve{\alpha}_r)x_r + \breve{\alpha}_r y_r, y_r\right] = 0 \qquad \text{if } m_r = y_r - x_r$$
$$F_\psi\left[(1+\breve{\alpha}_r)x_r - \breve{\alpha}_r y_r, y_r\right] = 0 \qquad \text{if } m_r = x_r - y_r$$

Consider again $Z_r(\gamma, y_r) = (1-\gamma)x_r + \gamma y_r$, and consider the equation

4.2.7 $$F_\psi[z_r(\gamma, y_r), y_r] = 0 .$$

Let $\gamma_r^+(y_r)$ denote the smallest positive solution to 4.2.7 or $+\infty$ if such does not exist, and let $\gamma_r^-(y_r)$ denote the numerically smallest negative solution to 4.2.7 or $-\infty$ if such does not exist. Then it follows from above that

4.2.8 $$\alpha_r^*(y_r) = \begin{cases} \min\{\gamma_r^+(y_r), \bar{\alpha}_r(y_r)\} & \text{if } F_\psi(x_r, y_r) > 0 \\ \min\{-\gamma_r^-(y_r), \bar{\alpha}_r(y_r)\} & \text{if } F_\psi(x_r, y_r) < 0 . \end{cases}$$

The only advantage really in drawing the above observations is that we will already have evaluated the formula $F_\psi(x,y)$, while, in some instances, explicit formulae may be more readily derived for the solution to equation 4.2.7. Silvey (1974,p.9) also noted that $\alpha_r^*(y_r)$ had to satisfy this equation in a particular case of the algorithm which will be considered in chapter 5.

However we will not always be able to make use of 4.2.8. Advantages would be gained from an examination of the solutions of 4.2.7 only in the following instances, assuming $\psi(\cdot)$ concave and differentiable.

(i) If explicit formulae do exist for its solutions then we will be able to identify either that $\alpha_r^*(y_r)$ is one of those solutions or that $\alpha_r^*(y_r) = \bar{\alpha}_r(y_r)$.

(ii) If we can identify that no solutions to 4.2.7 exist then it must be again that $\alpha_r^*(y_r) = \bar{\alpha}_r(y_r)$.

(iii) If no explicit formulae exist for any solutions to 4.2.7 but we can establish indirectly that $\psi(\cdot)$ must have a constrained stationary value of the above type, then we could solve the equation numerically by means such as Newton-Raphson techniques. We would have an instance of this if we had $F(x_r,y_r) > 0$ and could establish that $\psi(y_r) < \psi(x_r)$. An example of this occurs in the design context.

Failing such information, the only sensible action would seem to be direct numerical solution of 4.2.1 for which efficient algorithms such as Golden Section or Armijo techniques exist. The early stages of such calculations may identify the information assumed in (iii).

However any numerical computation that might be required in identifying an optimal α_r^* may be considered not to be worth the effort.

Two other methods of choosing α_r have been employed in the design context.

Suppose it were the case that the point y_r chosen to determine the direction m_r was such that $y_r \in S$, $\psi(y_r) > \psi(x_r)$. Then one may opt to take $x_{r+1} = y_r$. If $\psi(\cdot)$ is concave on convex S, then it must be that $F(x_r,y_r) > 0$, so that $m_r = y_r - x_r$ and we are selecting $\alpha_r = 1$.

A variation on this is that in the case of L-optimality Fedorov (1972) obtains for a particular type of $y_r \in S$, satisfying $F(x_r,y_r) > 0$, an upper bound $\hat{\alpha}_r(y_r)$ on the range of values α such that $\psi[x_r + \alpha(y_r - x_r)] > \psi(x_r)$. This bound $\hat{\alpha}_r(y_r) < \bar{\alpha}_r(y_r)$ and in view of concavity is then such that $\psi\{x_r + \hat{\alpha}_r(y_r)[y_r - x_r]\} = \psi(x_r)$. Fedorov recommends taking $\alpha_r = \beta\hat{\alpha}_r(y_r)$ for some β satisfying $0 < \beta < 1$.

Failing these possibilities or lacking knowledge of them we might simply determine the sequence α_r in the following arbitrary manner. Choose, a-priori, a sequence $\beta_0, \beta_1, \beta_2, \ldots,$ such that $0 < \beta_r < 1$ or $0 < \beta_r \leq 1$ and take $\alpha_r = \beta_r \bar{\alpha}_r$. This in fact is a convention adopted in some particular cases of the class of algorithm to be considered in chapter 5. These always have $\bar{\alpha}_r = 1$ so that $\alpha_r = \beta_r$.

The restrictions on β_r are supplied as insurance that $0 < \alpha_r < \bar{\alpha}_r$. However clearly β_r cannot be just any random number in (0,1). On its choice will depend the attainment of convergence to x^*. Certainly it is possible that $\psi(x_{r+1}) < \psi(x_r)$. In the case of the algorithms mentioned above, for which $\bar{\alpha}_r = 1$, it has been shown that conditions necessary for convergence are that $\beta_r \to 0$ and $\sum \beta_r = \infty$.

§4.2.3 This completes our general comments. The algorithms which we will consider for problem (P2) fall into one or other of the above categories although these clearly are not mutually exclusive or exhaustive. The algorithms were originally formulated for the design problem. It is worthy of mention that when some of them were initially formulated, a specific criteria, mainly D-optimality, was under consideration and the directional derivative tool was not yet visible.

We will maintain the notation used above for the next few chapters.

In particular x_r will denote a current iteration, x_{r+1} the subsequent iteration. The values $y_r, m_r, \alpha_r (\alpha_r > 0)$ and s will be such that

4.2.9 $$s = F(x_r, y_r)/|F(x_r, y_r)| \qquad (\text{so that } s = \pm 1)$$

4.2.10 $$m_r = s(y_r - x_r)$$

4.2.11 $$x_{r+1} = x_r + \alpha_r m_r$$

4.2.12 $$x_{r+1} = (1 - s\alpha_r)x_r + (s\alpha_r)y_r$$

that is

4.2.13 $$x_{r+1} = (1-\alpha_r)x_r + \alpha_r y_r \qquad \text{if } F(x_r, y_r) > 0$$

4.2.14 $$x_{r+1} = (1+\alpha_r)x_r - \alpha_r y_r \qquad \text{if } F(x_r, y_r) < 0 .$$

If the function $\psi(\cdot)$ has a matrix argument, M_r, M_{r+1} will denote current and subsequent iterates while s, N_r will be such that

4.2.15 $$s = F(M_r, N_r)/|F(M_r, N_r)|$$

4.2.16 $$M_{r+1} = (1 - s\alpha_r)M_r + (s\alpha_r)N_r$$

that is

4.2.17 $$M_{r+1} = (1-\alpha_r)M_r + \alpha_r N_r \qquad \text{if } F(M_r, N_r) > 0$$

4.2.18 $$M_{r+1} = (1+\alpha_r)M_r - \alpha_r N_r \qquad \text{if } F(M_r, N_r) < 0 .$$

If problem (P1) is under specific consideration, $p^{(r)}$, $p^{(r+1)}$ will denote current and subsequent iterates while $q^{(r)}$, $m^{(r)}$, s will be such that

4.2.19 $$s = F_\phi(p^{(r)}, q^{(r)}) / |F_\phi(p^{(r)}, q^{(r)})|$$

4.2.20 $$m^{(r)} = s(q^{(r)} - p^{(r)})$$

4.2.21 $$p^{(r+1)} = p^{(r)} + \alpha_r m^{(r)}$$

4.2.22 $$p^{(r+1)} = (1 - s\alpha_r) p^{(r)} + (s\alpha_r) q^{(r)}$$

that is

4.2.23 $$p^{(r+1)} = (1-\alpha_r) p^{(r)} + \alpha_r q^{(r)} \qquad \text{if} \quad F_\phi(p^{(r)}, q^{(r)}) > 0$$

4.2.24 $$p^{(r+1)} = (1+\alpha_r) p^{(r)} - \alpha_r q^{(r)} \qquad \text{if} \quad F_\phi(p^{(r)}, q^{(r)}) < 0 .$$

We will of course require that $p^{(r+1)} \in \mathcal{P}$, i.e. that (i) $p_j^{(r+1)} \geq 0$, (ii) $\sum p_j^{(r+1)} = 1$.

A necessary and sufficient condition for the latter in view of 4.2.20 is that $\sum m_j^{(r)} = 0$, which in turn will be the case, in view of 4.2.19, if $\sum q_j^{(r)} = 1$. Typically we will have $q^{(r)} \in \mathcal{P}$ if direct selection of $q^{(r)}$ has been the method adopted by which to determine $m^{(r)}$.

Since $p_j^{(r+1)} = p_j^{(r)} + \alpha_r m_j^{(r)}$ then condition (i) will be guaranteed if

(a) $\alpha_r > p_j^{(r)} / (-m_j^{(r)})$ when $m_j^{(r)} > 0$ and if

(b) $\alpha_r < p_j^{(r)} / (-m_j^{(r)})$ when $m_j^{(r)} < 0$

Condition (a) will always be satisfied if we already have the restriction $\alpha_r > 0$, while condition (b) is not always guaranteed.

To summarise, we will have $p^{(r+1)} \in \mathcal{P}$ iff $\sum m_j^{(r)} = 0$ and if we keep the magnitude α_r of the steplength below the bound,

4.2.25 $$\bar{\alpha}_r(m^{(r)}) = \min_{m_j^{(r)} < 0} \{ p_j^{(r)} / (-m_j^{(r)}) \} .$$

Equivalently we will have $p^{(r+1)} \in \mathcal{P}$ iff $\sum q_j^{(r)} = 1$ and if we keep the magnitude of the steplength α_r below the bound

4.2.26 $$\bar{\alpha}_r(q^{(r)}) = \min_{p_j^{(r)} > q_j^{(r)}} \{ p_j^{(r)} / (p_j^{(r)} - q_j^{(r)}) \} \qquad \text{if } F_\phi(p^{(r)}, q^{(r)}) > 0$$

or

4.2.27 $$\bar{\alpha}_r(q^{(r)}) = \min_{q_j^{(r)} > p_j^{(r)}} \{ p_j^{(r)} / (q_j^{(r)} - p_j^{(r)}) \} \qquad \text{if } F_\phi(p^{(r)}, q^{(r)}) < 0 .$$

It is easy to see geometrically or pictorially the necessity for these upper bounds $\bar{\alpha}_r$. Condition 4.2.25 guards against overstepping in the following way. If the weight p_j, under a design p, at vertex u_j, is larger than the weight q_j, under q, at u_j, then a step from p in the direction of q is a step that leads to a reduction in the weight at u_j. Since $(1-\alpha)p_j + \alpha q_j = 0$ implies that $\alpha = p_j/(q_j - p_j)$, the value $\bar{\alpha}_r$ in 4.2.25 is the shortest such step that will just put to zero the weight at, at least one such vertex, so that any larger a step would induce negative components in $p^{(r+1)}$.

In the case of (P2) or (P3) we will have the relationships $x_r = x(p^{(r)})$, $M_r = M(p^{(r)})$, and if $q^{(r)} \in \mathcal{D}$, $y_r = x(q^{(r)})$, $N_r = M(q^{(r)})$. We will maintain the latter notation even if $q^{(r)}$ satisfies only $\sum_j q_j^{(r)} = 1$.

Hence for the terms $\gamma_r^+(y_r)$, $\gamma_r^-(y_r)$ defined in relationship to equation 4.2.7 we will have $\gamma_r^+(q^{(r)}) = \gamma_r^+(y_r)$ or $\gamma_r^+(N_r)$, $\gamma_r^-(q^{(r)}) = \gamma_r^-(y_r)$ or $\gamma_r^-(N_r)$.

Finally it follows that

4.2.28 $$\alpha_r^*(q^{(r)}) = \begin{cases} \min\left[\gamma_r^+(q^{(r)}), \min\limits_{p_j^{(r)} > q_j^{(r)}} \{ p_j^{(r)} / (p_j^{(r)} - q_j^{(r)}) \} \right] & \text{if } F_\phi(p^{(r)}, q^{(r)}) > 0 \\ \min\left[-\gamma_r^-(q^{(r)}), \min\limits_{q_j^{(r)} > p_j^{(r)}} \{ p_j^{(r)} / (q_j^{(r)} - p_j^{(r)}) \} \right] & \text{if } F_\phi(p^{(r)}, q^{(r)}) < 0 \end{cases}$$

As we have said, with one exception, the algorithms that we study for (P2) will always seek an optimising p^* directly. This means that they will pass through a sequence $p^{(0)}, p^{(1)}, p^{(2)}, \ldots\ldots$ They will thereby generate a sequence x_r, M_r only indirectly through the above relationships.

§4.3 Considerations Specific To Problem (P2)

§4.3.1 Initial Approximations

The starting value x_0 or $p^{(0)}$ is important to an algorithm for any optimisation problem. Convergence to the optimum will stand or fall depending on the chosen value. In the case of problem (P2) it is convenient to view the making of this choice as requiring two decisions.

The first is to decide on $\mathrm{Sup}(p^{(0)})$. The answer to this will depend on the type of algorithm which we employ, an option which in turn will largely depend on $\mathcal{U}$ or $\mathcal{V}$. It might seem that we should take $\mathrm{Sup}(p^{(0)})$ to be $\mathcal{U}$ or $\mathcal{V}$. Certainly this will be a sensible choice when J is relatively small, while it will indeed be an essential choice in the case of some algorithms.

However this will not be a sensible choice when $\mathcal{V}$ is large, in particular a discretisation of a continuous space, when $\mathrm{Sup}(p^*)$ is likely to be a small subset of $\mathcal{V}$. Particularly in the case of the algorithm of the next chapter we will require to make a subjective decision about $\mathrm{Sup}(p^{(0)})$, choosing it to be a small subset of a large $\mathcal{V}$. In the contexts in which we are prepared to use other algorithms, the choice of $\mathrm{Sup}(p^{(0)})$ will be fairly clear.

The choice of $p^{(0)}$ is more obvious, In the absence of any information about p^* it seems natural that $p^{(0)}$ should allocate uniform weights to its support points so that we have the neutral $p^{(0)} = (1/J,\dots,1/J)$ if $\mathrm{Sup}(p^{(0)}) = \mathcal{U}$.

A final point is that almost always it will be essential that $\psi(\cdot)$ be differentiable at $x(p^{(0)})$ or $M(p^{(0)})$. As we shall see it will only be wise to invite nondifferentiability if $\mathrm{Sup}(p^*)$ is known and this indicates that $x(p^*)$ is nondifferentiable, and this only in the case of a function $\psi(\cdot)$ enjoying support differentiability.

§4.3.2 Stopping Rules

It would seem from the appearance of theorem 2.5.6 that a reasonable stopping rule at a differentiable $p^{(r)}$ would be to stop if

(R1) $$\max_{1 \le j \le J} F\left[x(p^{(r)}), u_j\right] \le \delta$$

for some appropriately small $\delta > 0$.

In fact such a rule is given much stronger weight by the following theorem proved by Silvey (1974) for the design context.

Theorem 4.3.1

Let $\mathcal{U} = \{u_1, \ldots, u_J\}$ and let $\psi(\cdot)$ be concave on $\mathcal{P}(\mathcal{U})$. If $\psi(\cdot)$ is differentiable at x(p) and if, for $\delta > 0$, $F[x(p), u_j] \le \delta$ at all $u_j \in \mathcal{U}$, then $\psi[x(p)] \ge \psi[x(p^*)] - \delta$.

Proof Since $\psi(\cdot)$ is differentiable at x(p) we have

$$F[x(p), x(p^*)] = \sum p_j^* F[x(p), u_j] \le \delta .$$

The result trivially follows by appealing to (G4) of section 2.2.2 to claim, in view of the concavity of $\psi(\cdot)$, that

$$F[x(p), x(p^*)] \ge \psi[x(p^*)] - \psi[x(p)] . \qquad \square$$

Hence if we stop according to (R1) then we can claim "to be within δ of an optimising p*".

If a design criterion enjoys support differentiability then we could similarly argue that a realisation of the condition

(R2) $$\max_{p_j^{(r)} > 0} F\left[M(p^{(r)}), v_j v_j'\right] \le \delta$$

indicates that $p^{(r)}$ is a good approximation to the optimum on Sup($p^{(r)}$), whether or not $p^{(r)}$ is overall differentiable. While this does not guarantee that $p^{(r)}$ solves (P4) for the full set $\mathcal{V} = \{v_1, v_2, \ldots, v_J\}$, we may well believe that p* has been identified if the zero weights which have given rise to the lack of overall differentiability have been arrived at by the following process.

§4.3.3 Setting weights to zero

As we have said before we have in problem (P2) a constrained optimisation problem whose solution p* may lie on the boundary of the constraint region. This will certainly be the case if $\mathcal{U}$ or $\mathcal{V}$ is a discretisation of a continuous space. Then J is large and many optimum weights will be zero. Convergence to such an optimum would be slow by any of the algorithms which we will consider but could be speeded up by formally setting weights to zero. Again theorem 2.5.6 suggests that the following would be a judicious rule to adopt at a differentiable $p^{(r)}$.

(R3) " Set $p_t^{(r+1)} = 0$ if (i) $F[x(p^{(r)}), u_t] \leq -\epsilon_1$
and (ii) $p_t^{(r)} \leq \epsilon_2$ "

for some small $\epsilon_1, \epsilon_2 > 0$, say $\epsilon_1 = \epsilon_2 = \cdot 01$.

If the two conditions in (R3) hold it would seem likely that $p_t^* = 0$.

Again the rule would still be reasonable if $\psi(\cdot)$ enjoyed only support differentiability at $p^{(r)}$ since attention is focused only on those vertices u_t such that $p_t^{(r)} > 0$.

Note that we could regard an application of (R3) as a change from one example of (P2) to another. We have eliminated one of the vertices and as it were have passed from one set of vertices to another with a smaller value of J.

Of course one would always have to keep in mind the possibility that $p_t^* > 0$, that we had wrongly set zero weight on u_t. Where differentiability makes it appropriate the value of $F[x(p^{(r)}), u_t]$ should still be examined for (some) subsequent iterates or at least for a postulated optimum. We must clearly be careful about setting weights to zero. How we have arrived at a $p_t^{(r)}$ satisfying rule (R3) will be important and this will be considered in due course.

§4.3.4 Collapsing clusters

When a set $\mathcal{U}$ of vertices or a design space $\mathcal{V}$ is a discretisation of a continuous space then what are called clusters can form.

Recall from section 3.3.2 that the D-optimal design for the polynomial regression model

$$E(y_{(x)}) = v'_{(x)}\theta \quad , \quad v_{(x)} \in \mathcal{V} = \{v_{(x)} = (1, x, x^2, x^3, x^4)' : -1 \le x \le 1\}$$

is

x	-1	$-\sqrt{3/7}$	0	$\sqrt{3/7}$	1
p^*_x	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

$(\sqrt{3/7} = 0{\cdot}655)$

However the D-optimal design p^{**}, obtained numerically for the model

$$E(y_{(x)}) = v'_{(x)}\theta \quad , \quad v_{(x)} \in \mathcal{V}_D = \{v_{(x)} = (1, x, x^2, x^3, x^4)' : x = -1(\cdot 01)1\}$$

is

x	-1	$-{\cdot}66$	$-{\cdot}65$	0	${\cdot}65$	${\cdot}66$	1
p^{**}_x	${\cdot}200$	${\cdot}085$	${\cdot}115$	${\cdot}200$	${\cdot}115$	${\cdot}085$	${\cdot}200$

The set $\mathcal{V}_D$ containing J = 201 points would be the usual discretisation of $\mathcal{V}$.

However the two optima are clearly different. The design p^{**} has a larger support than p^* although in this instance the two supports happen to share points in common, namely those corresponding to $x = 0, \pm 1$. That this is possible here is because $\mathcal{V}_D$ contains these support points of p^*. In contrast $\mathcal{V}_D$ does not contain the two points of Sup(p^*) corresponding to $x = \pm\sqrt{3/7}$. This must dictate a difference between the two optima.

At the same time though one would expect the two designs to be similar, and this indeed readily appears to be the case in view of the fact that Sup(p^{**}) seems, as it were, to have opted to replace each of the values $x = \pm\sqrt{3/7}$ by the cluster of two points belonging to $\mathcal{V}_D$ which lie on either side of them. In fact the weightings of the two designs leave one in no doubt that they are counterparts.

Where their supports share points in common they allocate the same weights, namely 1/5 = .200. Where a point of Sup(p^*) corresponds to a cluster in Sup(p^{**}), the weight of the former seems as it were to have been shared out among the points in its companion cluster, for (.115 + .085) = .200 = 1/5; and the sharing has not been arbitrary for

$$({\cdot}200)^{-1}\left[({\cdot}115)({\cdot}65) + ({\cdot}085)({\cdot}66)\right] = {\cdot}654 \doteqdot \sqrt{3/7}$$

Consider the convex combination of the points in a cluster for which the convex weight of a point is proportional to its optimum weight under p^{**}. Then the latter result says that we have this convex combination approximately equal to the cluster's counterpart point in $Sup(p^*)$.

In general optima on continuous spaces $\mathcal{U}$ or $\mathcal{V}$ and corresponding optima on discretised versions of such spaces will be different but related in the above fashion. They can only be identical if $\mathcal{U}_D$ or $\mathcal{V}_D$ contains the whole support of the optimum p^* for $\mathcal{U}$ or $\mathcal{V}$.

Typically then, when calculating an optimum on a discretisation of a continuous space, clusters will eventually form and a natural suggestion is that, at an appropriate iterate $p^{(r)}$, a cluster should be replaced by a convex combination of the above type based on $p^{(r)}$ and that all the weight under $p^{(r)}$ at the vertices in the cluster should be allocated to this new single vertex. See Fedorov (1972, p100) or Atwood (1976a, 1976b, 1980).

Clearly this is a device aimed at obtaining a solution as simple as the optimum on the continuous space and as a result must speed up convergence. It has the same effect as does setting weights to zero, namely we pass to an alternative example of (P2) with a different set of vertices and a reduced value of J.

We note that if we used a finer discretisation than above, clusters may be larger than two vertices and this would also seem to be a likely event if a regressor variable x was a vector. The above rule though for collapsing clusters would still be applicable.

§4.3.5 Differentiability

The main point to make here is that all of the algorithms which we will consider assume differentiability at x_r or $p^{(r)}$. This is not as serious a restriction as it might seem, at least in the design context when we can have support differentiability.

However nondifferentiability should not be courted too early. Certainly an initial $p^{(0)}$ should when possible be differentiable and since nondifferentiability in the design context arises when the support of a design contains fewer than k vertices we should be less inclined to set weights to zero or collapse clusters if this would mean passing to such a support. Only if we are convinced either on theoretical grounds or from early calculations that the optimum lies in such a direction should we commit ourselves to nondifferentiability. This is clearly sensible for a number of reasons but a particular consideration regarding putting weights to zero is that there is, as we shall see, no simple way of going back; that is, while we can proceed further in this direction by setting more weights to zero, thereby further reducing the support, we shall see that none of the ideas forming the basis of the algorithms to come will provide us with a simple rule for identifying, at a non-differentiable $p^{(r)}$, a vertex with which to augment the support of $p^{(r)}$ if it were thought necessary to do so.

We may then persist in retaining through successive iterations some weight at a vertex when we might otherwise have chosen to eliminate it. Fedorov and Tukey (1976) make some formal recommendations of this nature for the design context. See also Atwood (1976b).

When can we be convinced about passing to a nondifferentiable point? We have discussed relevant issues in section 3.2. We might further say that if our calculations proceed in a sensible way then we could have confidence about taking such a step. The following might be considered a sensible approach.

§4.3.6 Algorithms for all types

We have said that we will consider various algorithms. Inevitably these vary in attribute. Some are simple computationally; some in appropriate circumstances are highly efficient but heavy in computation, requiring the inversion of $J \times J$ matrices; some in contrast can cope with large values of J; some do not like non-support points.

Each algorithm has advantages and disadvantages depending on the particular case of (P2) under consideration. We therefore have no intention of drawing any rigorous comparison between iterative schemes. In contrast we shall recommend a composite scheme which would use two or three of the algorithms which we will consider. The idea is that as we gradually eliminate vertices as in sections 4.3.3 and 4.3.4, and thereby reduce the value of J, we could switch from an opening algorithm which copes easily with large J or with large numbers of nonsupport points to one of the more efficient algorithms. The ideas then of sections 4.3.3 and 4.3.4 are not just merely cosmetic. It will be important to bring down the value of J as much as possible before we contemplate inverting $J \times J$ matrices. The idea will be considered in more detail once we have studied the said algorithms of the ensuing chapters. First we now consider a type of algorithm to which most of the above does not apply since it does not generate a sequence of designs $p^{(r)}$ only a sequence of values x_r or design matrices M_r.

§4.4 Cutting Plane Algorithms, An Atypical Class

§4.4.1 Cutting plane algorithms are appropriate for the maximisation of a nonlinear concave function $\psi(\cdot)$ over a convex set or subject to linear inequalities. They generate a sequence of approximating linear programs. That this is possible can be seen on realising that maximising a concave function $\psi(x)$ over a set S is equivalent to maximising the variable y subject to $y \leq \psi(x)$ for all $x \in S$, which in turn is equivalent to maximising y subject to the constraints that the ordinate y should lie below the supporting hyperplanes to $\psi(\cdot)$ at all $x \in S$; a linear programming problem with infinitely many constraints, or an infinite linear program with possibly some additional basic linear inequality constraints.

Kelley (1960) perceived that the solution to such a problem could be realised by a sequence of finite linear programs. The progression from one linear program to the next is best described diagrammatically. Consider the following picture.

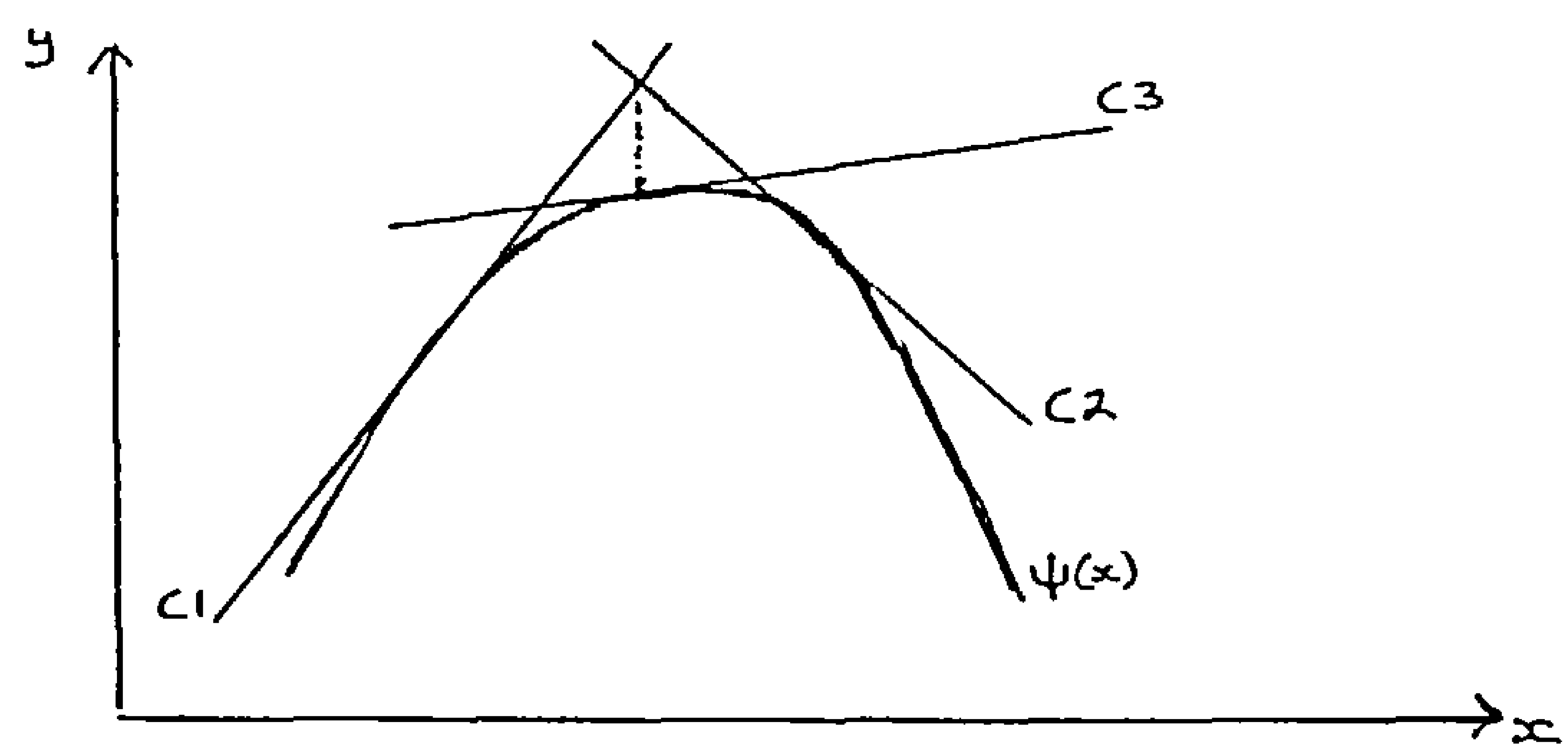
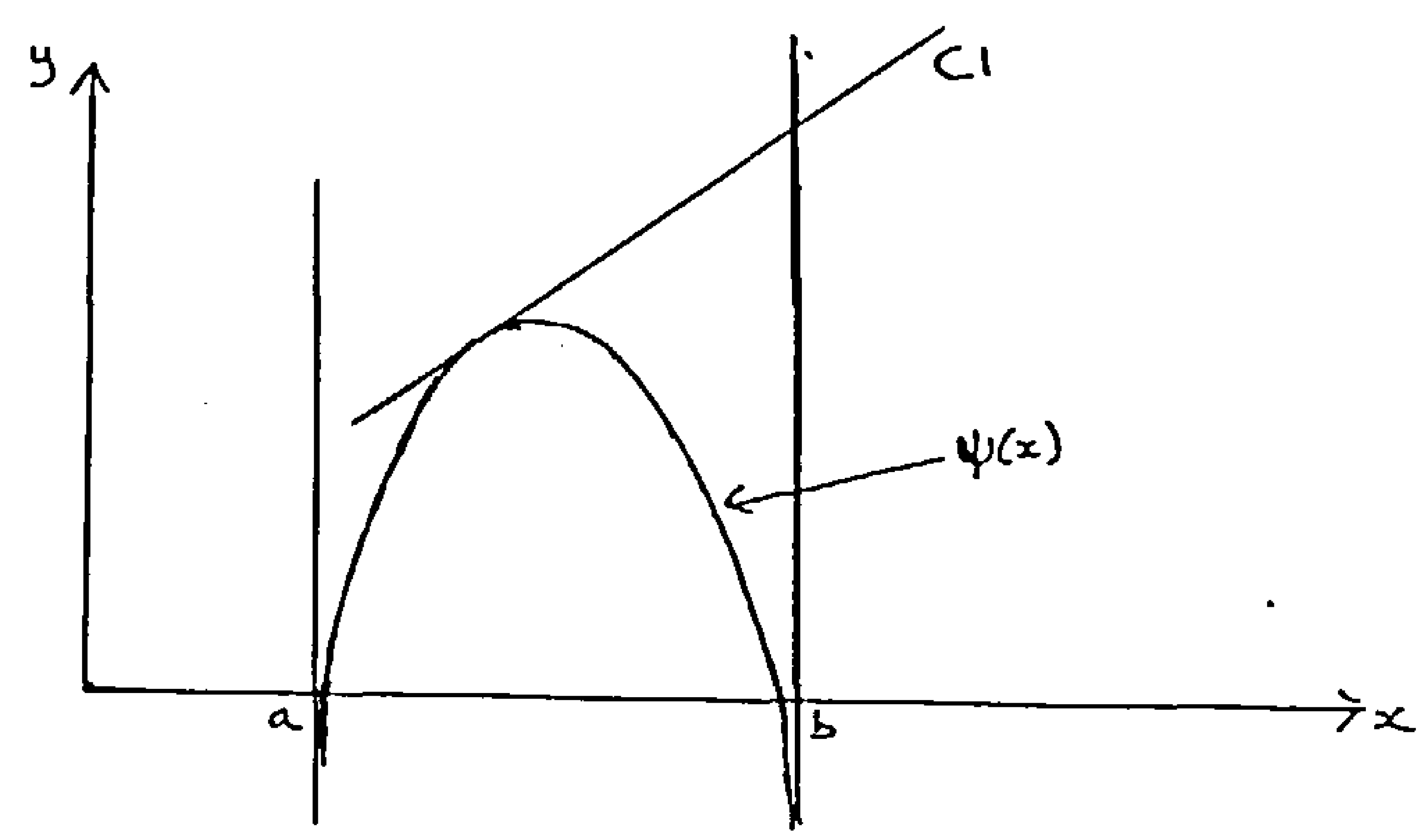


Figure 4.4.1

From the linear program "maximise y subject to the constraint that y lie below lines C1 and C2" proceed to the more accurate linear program "maximise y subject to the constraint that y lie below lines C1, C2 and C3" ;the additional line C3 being the tangent to $\psi(\cdot)$ at the current optimising value of the independent variable x. This additional constraint cuts off some previously possible optimum values for y including the current one. A full linear program does not require to be solved each time since the dual simplex algorithm makes possible efficient updating of the current solution to accommodate additional constraints.

Wolfe (1961) proved convergence of this procedure in a general context but further concluded that convergence would be improved by adopting as the additional improving constraint, not that suggested above, but a weighted linear combination of current linear constraints, the weights being their lagrange multipliers. If $\psi(\cdot)$ is quadratic this process leads to the exact solution. Further improvement would be obtained by deleting constraints that become redundant or slack as does (C1) in figure 4.4.1.

Note that the method would fail in the following case



Here we have a function $\psi(x)$ with $\psi(a) = \psi(b) = -\infty$. Consider maximisation of $\psi(x)$ over $[a,b]$. Suppose the initial approximating linear program is 'maximise y subject to the constraint that y lie below line C1'. The solution to this linear program in $[a,b]$ occurs at x = b. This suggests the asymptote 'x = b' as the additional constraint and thereby we do not obtain a new approximating linear program.

§<u>4.4.2</u> Most of the above is in fact outlined by Sibson (1974b) where he proposes using a cutting plane algorithm for the D-optimal criterion. He considers three approaches and we shall see that this criterion is particularly amenable to such algorithms.

Computing D-optimum solutions in this way was first proposed by Sibson and Kenny (1975) and this would seem to be the first instance of using a cutting plane algorithm for solving (P2), which should in principle be amenable to the approach if $\psi(\cdot)$ is concave. Similarly with (P1) if $\phi(\cdot)$ is concave.

The first of Sibson's proposals considers D-optimality in its explicit form and is aimed at finding a p^*. He takes

$$\phi(p) = \psi[M(p)] \quad , \quad \psi(M) = \log_e \det(M)$$

There is no unusual difficulties in applying the method except that it could fail as outlined above since at singular $M(p)$ $\psi[M(p)] = -\infty$. However Sibson realised that this could almost always be avoided by including the additional (linear) constraints $p_j \leq 1/k$, $j = 1, \cdots, J$, it being justifiable and sensible any way to do so since $p_j^* \leq 1/k$, as we observed in section 3.3.1.

That this device can guarantee avoidance of singular $M(p)$ follows because any design matrix $M(p)$, with p such that $p_j \leq 1/k$, $j = 1, \cdots\cdots, J$, is a convex combination of what Sibson and Kenny call extremal designs. The latter are designs assigning uniform weighting to the points in their support, their supports being all those subsets of $\mathcal{U}$ containing k points. If such a support consists of k linearly independent points then its extremal design is D-optimum for that support and of course the associated design matrix is nonsingular. Further all extremal designs and any convex combination of them will define nonsingular design matrices iff all k-point subsets of $\mathcal{U}$ form a basis for $\mathcal{U}$.

Sibson further states that if the optimum is given by one of these extremal designs then this modified version of the algorithm will identify it in a finite number of steps. If a non-extremal design is optimal then $p_j^* < 1/k$ and hence eventually these additional constraints will become slack.

Sibson's second proposal is based on eliminating the constraint $\sum p_i = 1$. This he does by considering the lagrangian appropriate to such a constraint, namely

$$L(p) = \phi(p) + \lambda\left(1 - \sum p_j\right) .$$

For D-optimality $\lambda = k$, a known constant, and hence an equivalent problem to the D-optimal version of (P1) is 'maximise $L(p) = \phi(p) + k(1 - \sum p_j)$ subject to $p_j \geq 0$'

Concavity of $\phi(p)$ and linearity of Σp_j guarantee concavity of $L(p)$ and the constraints are linear inequalities. A cutting plane algorithm can be formulated, also incorporating the constraints $p_j \leq 1/k$, which will clearly find a p^*.

Sibson points out that the linear approximations in the above two approaches will be different and it is not clear what the relative efficiency of the two methods will be.

The answer to this latter query would seem though to be clear for any criterion $\phi(p)$, other than that corresponding to D-optimality. The second proposal would seem the more complex since typically the lagrange multiplier will not be a known constant but will depend on p^*. If $\phi(\cdot)$ is differentiable at p^* then

$$\lambda = \Sigma p_j \, \partial\phi/\partial p_j \Big|_{p=p^*}$$

However it still may not be an easy matter to employ a cutting plane algorithm on other examples of (P1) particularly for other design versions of the problem. In general we will not have information on p^* that could be utilised to avoid failure of the method. We do know that $p_j^* \leq 1/s$ in the case of D_A-optimality and, while the set of matrices $M(p)$ such that $p_j \leq 1/t$ is again the convex hull of extremal designs, it still could be that $\log_e \det[AM^+(p)A'] = -\infty$, even although $p_j \leq 1/s$, for the latter conditions will not always guarantee that $\mathcal{N}[M(p)] \subseteq \mathcal{N}(A)$.

§4.4.3 Sibson's last proposal for D-optimality is a cutting plane algorithm which differs from the two above in that it does not find a p^* but $M^{-1}(p^*)$.

In Sibson (1972) he proved as a duality theorem that the solution to the problem 'maximise $\psi(N) = \log_e \det(N)$ over $N \in \mathcal{PD}$ subject to $v_j'Nv_j \leq k$' is $N^* = M^{-1}(p^*)$. This proved a conjecture of Silvey (1972).

This dual problem is a maximisation of a concave function on a convex set subject to linear inequality constraints since $v_j'Nv_j$ is linear in N. Sibson proposed solving it by a cutting plane algorithm.

Again it is an approach which does not readily extend to other criteria. In view of theorem 2.5.7 a corresponding dual for the general problem (P2) would seem to be 'maximise $\psi(x)$ subject to $F(x,u_j) \leq 0$', assuming $\psi(\cdot)$ is differentiable at x^*. However only if $F(x,u_j)$ depended on x only through being linear in a one to one function $h(\cdot)$ of x, i.e. $F(x,u_j) = L_j\{h(x)\}$, would we have the necessary linear inequality constraint ingredient for subjection to a cutting plane algorithm. At differentiable x, $F(x,u_j) = G(x,u_j) - G(x,x)$, and so depends on x through two components.

Gribek and Kortanek (1977) however get round this difficulty by using the alternative necessary and sufficient condition of theorem 2.5.5 for optimality of $x(p^*)$ in (P2); namely that

$$F[x(p), x(p^*)] \geq 0 \quad \text{for all } x(p) \in \mathcal{P}(\mathcal{U})$$

In view of theorem 3.2.4 this further reduces to

$$G[M(p), M(p^*)] \geq G[M(p), M(p)]$$

in the case of all the design criteria of chapter three whether or not $\psi(M)$ is differentiable at $M(p)$.

If however $\psi(\cdot)$ is differentiable at $M(p)$ then $G[\cdot, M(p^*)]$ is linear in $M(p^*)$ and the above states that $M(p^*)$ should minimise $G[M(p), N]$ with respect to N subject to $G[M(p), N] \geq G[M(p), M(p)]$ for all p, a constrained minimisation of a multilinear objective function, the constraints being linear inequalities.

Gribek and Kortanek formulate two very elaborate algorithms for the design context, one for $\psi(\cdot)$ closed on $\mathcal{M}$ and one in which $M(p)$ is normalised with respect to a matrix norm for general $\psi(\cdot)$. For the former they state that a cutting plane can still be found at a point where $\psi(\cdot)$ is finite. In general they prove convergence under a wide set of assumptions.

These algorithms are designed to discover $M(p^*)$ rather than p^* although Gribek and Kortanek say that they could be altered to find p^*. This would however obscure the basic ideas in the algorithms. Indeed an advantage in seeking $M^* = M(p^*)$, as Sibson

(1974b) points out is that if $\psi(\cdot)$ is strictly concave on $\mathcal{M}$ then M^* is unique whereas there may be infinitely many optimal p^*'s. The practical design problem though desires to know a p^* and thus we would require to go on to solve the equation,

$$M(p^*) = M^* \text{ subject to } \Sigma p_j = 1,\ p_j \geq 0 .$$

. While this can be solved by linear programming type methods, lack of knowledge of $\mathrm{Sup}(p^*)$ would be a hindrance.

To conclude we observe that Wu (1976) in a comparison of algorithms says that if a cutting plane algorithm is viewed as a sequence of maximisations over a sequence of polygons containing the polygon $\mathcal{P}(\mathcal{U})$ or $\mathcal{M}$, then some of the other algorithms that we are about to study can be viewed as a sequence of maximisations over a sequence of polygons contained in $\mathcal{P}(\mathcal{U})$ or $\mathcal{M}$ or as interior polygonal approximation methods.

CHAPTER 5

FORWARD AND REVERSE VERTEX DIRECTION ALGORITHMS

§5.1 Definition, General Comments

This first class of algorithm for finding a p^* is tailor made for problem (P2). The vertices $\{u_1,\ldots,u_J\}$ are the basic ingredients of that problem without which it clearly lacks definition. This is also the case with these algorithms. Their distinguishing feature is that an iteration consists of moving towards or, in some instances, away from a vertex. They were originally conceived in the design context, although Wu (1976, 1978a, 1978b), who originated the phrase 'vertex direction', suggests that they are similar to methods devised by Frank and Wolfe (1956) for quadratic programming problems. The basic principle then, in terms of the notation of section 4.2, is that we select a vertex u_t to be the point y_r towards which or away from which to move from from $x_r = x(p^{(r)})$.

Referring to the equations 4.2.7,...,4.2.25 we therefore have for

5.1.1 $$s = F(x_r, u_t)/|F(x_r, u_t)| \quad ,$$

5.1.2 $$x_{r+1} = (1-s\alpha_r)x_r + (s\alpha_r)u_t$$

or

5.1.3 $$x_{r+1} = \begin{cases} (1-\alpha_r)x_r + \alpha_r u_t & , \; F(x_r, u_t) > 0 \\ (1+\alpha_r)x_r - \alpha_r u_t & , \; F(x_r, u_t) < 0 \end{cases} .$$

where $\alpha_r \geq 0$ is the magnitude of the steplength taken.

Similarly in the design context when $u_t = v_t v_t'$ we will have for

5.1.3 $$s = F(M_r, v_t v_t')/|F(M_r, v_t v_t')| \quad ,$$

5.1.4 $$M_{r+1} = (1-s\alpha_r)M_r + (s\alpha_r)v_t v_t'$$

or

5.1.5 $$M_{r+1} = \begin{cases} (1-\alpha_r)M_r + \alpha_r v_t v_t' & , \; F(M_r, v_t v_t') > 0 \\ (1+\alpha_r)M_r - \alpha_r v_t v_t' & , \; F(M_r, v_t v_t') < 0 \end{cases} .$$

This will mean that if $x_r = x(p^{(r)})$, $M_r = M(p^{(r)})$

5.1.6 $$p^{(r+1)} = (1 - s\alpha_r)p^{(r)} + (s\alpha_r)e_t$$

or

5.1.7 $$p^{(r+1)} = \begin{cases} (1-\alpha_r)p^{(r)} + \alpha_r e_t, & F\{p^{(r)}, e_t\} > 0 \\ (1+\alpha_r)p^{(r)} - \alpha_r e_t, & F\{p^{(r)}, e_t\} < 0 \end{cases}$$

Hence in the notation of section 4.2 $q^{(r)} = e_t$, this being the t^{th} unit vertex in J dimensions.

From 4.2.26, 4.2.27 we then have that the upper bound on the steplength is

5.1.8 $$\bar{\alpha}_r(u_t) = \begin{cases} 1, & F\{x_r, u_t\} > 0 \\ p_t^{(r)}/(1-p_t^{(r)}), & F\{x_r, u_t\} < 0 \end{cases}$$

Finally from 4.2.9 or 4.2.28

5.1.9 $$\alpha_r^*(u_t) = \begin{cases} \min\{\gamma_r^+(u_t), 1\}, & F\{x_r, u_t\} > 0 \\ \min\{-\gamma_r^-(u_t), [p_t^{(r)}/(1-p_t^{(r)})]\}, & F\{x_r, u_t\} < 0. \end{cases}$$

The above will still be denoted by $\bar{\alpha}_r(v_t)$, $\alpha_r^*(v_t)$ in the design context when $u_t = v_t v_t'$.

If $F(x_r, u_t) > 0$ we will refer to u_t as a forward vertex direction and u_t will be called a reverse vertex direction if $F(x_r, u_t) < 0$.

This defines the basic terminology. There remains the central issue of deciding on a rule for selecting u_t and the choice of α_r. These will be discussed in section 5.3 where a number of suggestions for u_t will be made. Some general issues are examined in the meantime.

(i) It would not always be possible to opt for the above type of iteration. An essential implicit assumption is differentiability at x_r, although support differentiability would suffice at $p^{(r)}$ with the exception in the design context of a $p^{(r)}$ which is optimal for $L\{Sup(p^{(r)})\} \cap \mathcal{U}$.

If $\psi(\cdot)$ is differentiable at x_r then it is easy to see that either of the above options would be available to us except if x_r is optimal. At a nonoptimal differentiable x_r there must exist at least one vertex u_k such that $F(x_r,u_k) > 0$ and at least one vertex u_i such that $F(x_r,u_i) < 0$. In fact, as observed in (D7) of section 2.3, there must exist at least one pair $u_k, u_l \in \mathrm{Sup}(p^{(r)})$ in this category. Further in view of the differentiability at x_r it would follow that $F(x_r, 2x_r-u_\ell) = -F(x_r,u_\ell) > 0$. Under these circumstances the iteration x_{r+1} above can be defined and is sensible since, as we have already observed, it will be sensible to move towards y_r if $F(x_r,y_r) > 0$ and sensible to move from differentiable x_r in the direction opposite to y_r if $F(x_r,y_r) < 0$. This is assuming $(x_r + \alpha m_r) \in \mathcal{P}(\mathcal{U})$ for small positive α, $m_r = \pm(y_r - x_r)$, whichever the case may be.

(ii) Note that if we know $p^{(0)}$, then we will know $p^{(r)}$ throughout and, assuming convergence, the algorithm will thereby identify a postulated p^*. Discussion on the choice of $p^{(0)}$ will be deferred until section 5.6.

(iii) Further viewing (P2) as a particular case of (P1) it emerges from 5.1.6, 5.1.7 that the distinguishing feature, with respect to (P1), of such an iteration is that it makes a change in only one of the weights, namely that of u_t, apart from a proportional change in the others. More precisely if $F(x_r,u_t) > 0$, when, from 5.1.8, we must have $\alpha_r < 1$, it removes a proportion α_r of the weight at those (other) vertices belonging to $\mathrm{Sup}(p^{(r)})$ and assigns the total removed weight to u_t. If $F(x_r,u_t) < 0$, when, from 5.1.8, we must have $\alpha_r < \left\{p_t^{(r)}/(1-p_t^{(r)})\right\}$, it removes weight $\alpha_r(1-p_t^{(r)}), \left(< p_t^{(r)}\right)$, from u_t and distributes it proportionately to those other vertices of $\mathrm{Sup}(p^{(r)})$ according to their weights $p_j^{(r)}$. This highlights two important points.

First, to offer a suitable forward vertex direction, u_t need not belong to $\mathrm{Sup}(p^{(r)})$ but if weight is to be removed from u_t then we must have $u_t \in \mathrm{Sup}(p^{(r)})$. There must be weight at u_t if any is to be removed, for, if $p_t^{(r)} = 0$, then $p_t^{(r+1)} = -\alpha_r$ and then $p^{(r+1)} \notin \mathcal{P}$. This is also clear from 5.1.8 since we would have $\bar{\alpha}_r(u_t) = 0$ if

$p_t^{(r)} = 0$, $F(x_r,u_t) < 0$ and so it would be that $(x_r + \alpha m_r) \notin \mathcal{P}(\mathcal{U})$ for any small positive α if $m_r = (x_r - u_t)$.

Secondly in view of the fact that any weight removed from u_t is distributed only among the support points of $p^{(r)}$, we could not consider taking a reverse vertex direction step at every iteration x_r unless we knew that $\mathrm{Sup}(p^*) \subseteq \mathrm{Sup}(p^{(0)})$. Typically this would require that $\mathrm{Sup}(p^{(0)}) = \mathcal{U}$.

(iv) That an iteration of the form 5.1.1 can still be defined at an $x_r = x\left(p^{(r)}\right)$ offering only support differentiability and where $p^{(r)}$ is not optimal for its support, follows by the same argument as in (i), namely that there still must be at least one pair $u_k, u_l \in \mathrm{Sup}(p^{(r)})$ such that $F(x_r,u_k) > 0$, $F(x_r,u_l) < 0$.

Further since, in the case of those design criteria of section 3.1, which enjoy support differentiability, we have $F\{M(p), v_j v_j'\} = -G\{M(p), M(p)\}$ if $v_j \notin L\{\mathrm{Sup}(p)\}$, and since $G(M,M) > 0$ in such instances, it must be that $F\{M(p^{(r)}), v_k v_k'\} > 0$ iff $v_k \in L\left\{\mathrm{Sup}(p^{(r)})\right\}$.

Suppose now, in the case of such a design criterion, that $p^{(r)}$ is optimal for $L\left\{\mathrm{Sup}(p^{(r)})\right\}$. Then $F\left\{M(p^{(r)}),\ v_j v_j'\right\} \le 0$ with equality if $p_j^{(r)} > 0$ and with strict inequality if $v_j \notin L\left\{\mathrm{Sup}(p^{(r)})\right\}$. In such circumstances, even if $p^{(r)}$ is not the solution p^* to (P2), an iteration of the form 5.1.2 cannot be defined for there is no v_j such that $F\left\{M\left(p^{(r)}\right),\ v_j v_j'\right\} > 0$ and no $v_i \in \mathrm{Sup}\left\{p^{(r)}\right\}$ such that $F\left\{M(p^{(r)}),\ v_i v_i'\right\} < 0$. Put verbally there is no vertex towards which to move and no suitable one from which to move away from x_r, and even if it were sensible to move away from a $v_j \notin L\left\{\mathrm{Sup}(p^{(r)})\right\}$, lack of differentiability means that it is not guaranteed that $F\left\{M\left(p^{(r)}\right), 2M(p^{(r)}) - v_j v_j'\right\} > 0$. Any iteration from $M\left(p^{(r)}\right)$ must move in a direction which is neither towards nor away from a particular vertex.

(v) Why this particular type of simplified algorithm? Almost certainly other algorithms not restricted to vertex directions would be more powerful. We will see however a role for this type of technique in that it can roughly identify Sup(p^*), an advantage if J is large as when $\mathcal{U}$ or $\mathcal{V}$ is a discretisation of a continuous space.

Also in the design context iterations of the form 5.1.4, 5.1.5 are much simpler to execute than other iterations. We will see that in the process of selecting v_t, either we will wish to evaluate $F\{M_r, v_j v_j'\}$ for all j, or possibly to evaluate $\psi(\cdot)$ at the different possible values that M_{r+1} would take if v_t is allowed to range over $\mathcal{V}$. Almost certainly we will wish to evaluate $\psi(\cdot)$ at M_{r+1} for the selected value of v_t. Consider that for the criteria of section 3.1 the quantities $\psi(M)$, $F(M,N)$ depend on M through M^- or M^{-1}. Hence if it is nonsingular then in general a full inversion of $M_{r+1} = (1-s\alpha_r)M_r + (s\alpha_r)N_r$ is required. However if $N_r = v_t v_t'$ then a matrix result makes possible a simple update of M_{r+1}^{-1} from M_r^{-1}. As a result simple updating of $\psi(M_{r+1})$, $F\{M_{r+1}, v_t v_t'\}$ is also possible in the case of D_A-optimality and L_A-optimality, since for these criteria the two quantities $\psi(M)$, $F(M, v_j v_j')$ depend simply enough on M^{-1} if M is nonsingular. Also explicit formulae for optimal steplengths α_r^* can also be derived by use of this matrix result in the case of these two design criteria. We derive these results in the next section.

§5.2 A Matrix Result And Some Useful Corollaries

§5.2.1 The following matrix result simplifies fairly substantially execution of a vertex direction algorithm in the case of D_A-optimality and L_A-optimality.

Lemma 5.2.1

Let G, M be k×k matrices and let a, b be k×1 vectors which satisfy the relationship

5.2.1 $$G = M + ab'$$

(a) If M is nonsingular then

5.2.2 $$\det(G) = [1 + a'M^{-1}b] \cdot \det(M)$$

(b) If also G is nonsingular then

5.2.3 $$G^{-1} = M^{-1} - (M^{-1}ab'M^{-1})/(1 + a'M^{-1}b).$$

Proof This is theorem 8.9.3 of Graybill (1969). He does not give a formal proof although establishing (b) is trivially done by checking that $GG^{-1} = I$, while in theorem 8.4.3 he proves (a) for $M = D$, where D is a diagonal matrix. His proof though is cumbersome since it appeals to the basic expansion of det $(\cdot)$. Graybill also devotes theorems 8.3.3, 8.5.2 and 8.5.3 to this particular case of G.

From theorem 1.6.8 of Graybill there exist nonsingular matrices P,Q such that

$$PMQ = I$$

Hence

$$M = P^{-1}Q^{-1}$$

$$\det(M) = \{\det(P)\det(Q)\}^{-1}$$

while

$$PGQ = I + cd'$$

where $c = Pa$, $d = Qb$ and hence $c'd = a'M^{-1}b$.

Since we then have

$$\det(G) = [\det(P), \det(Q)]^{-1}\det[I + cd']$$

we need to show that $\det[I + cd'] = 1 + c'd$.

That this is the case follows from the fact that $(I + cd')$ has two distinct eigenvalues, namely 1 with multiplicity (k-1) and $(1 + c'd)$. This in turn is a consequence of the fact that these eigenvalues exceed those of the matrix (cd') by a value of 1. That the latter matrix is of rank 1 means that it has only one non-zero eigenvalue which trivially is c'd. □

This simple lemma will be seen to have wide ranging implications for us. We will be able to use it several times over mainly because the form 5.2.1 of G is one which is 'preserved' under matrix multiplication as is evident in the above proof and also under matrix inversion, for 5.2.3 states that $G^{-1} = N + cd'$.

In view of the fact that formula 5.1.4 states that $M_{r+1} = (1-s\alpha_r)M_r + (s\alpha_r)v_t v_t'$, a consequence of opting to 'change' only one weight, the lemma is clearly a potentially useful result to a vertex direction algorithm in the design context. We will also see in section 6.1 that it can in fact play a simplifying role in other algorithms which change all the weights simultaneously.

<u>Corollary 5.2.2</u>

Lemma 5.2.1 is true with signs reversed as is seen by putting b = -c. □

Let $\quad Z(\gamma, v) = (1-\gamma)M + \gamma v v'$
$\quad\quad Z_r(\gamma, v) = (1-\gamma)M_r + \gamma v v'$.

The following equations are easily verified derivations from lemma 5.2.1. They assume nonsingularity of M and, where necessary of $Z(\gamma, v)$. Note that if $M_r = M(p^{(r)})$ is nonsingular then so is $Z_r(\gamma, v_t)$ when $\gamma = s\alpha_r$, $0 \leq \alpha_r < \bar{\alpha}_r(v_t)$ for then $p = (1-\gamma)p^{(r)} + \gamma e_t$ is a design for which $\mathrm{Sup}(p) \supseteq \mathrm{Sup}(p^{(r)})$, which must contain k linearly independent design points.

5.2.4 $$\det\{Z(\gamma, v)\} = (1-\gamma)^{k-1}[1-\gamma+\gamma v'M^{-1}v]\det(M).$$

5.2.5 $$Z^{-1}(\gamma, v) = \frac{M^{-1}}{(1-\gamma)} - \frac{\gamma M^{-1}v v' M^{-1}}{(1-\gamma)[1-\gamma+\gamma v'M^{-1}v]}$$

5.2.6 $$\mathrm{tr}\{Z^{-1}(\gamma, v)\} = \frac{\mathrm{tr}\, M^{-1}}{(1-\gamma)} - \frac{\gamma v'M^{-2}v}{(1-\gamma)[1-\gamma+\gamma v'M^{-1}v]} .$$

That these results are useful to D-optimality and to A-optimality cannot be in doubt, but this does not only derive from their simplicity. They depend on the vector $w = M^{-1}v$ which has to be calculated anyway in the evaluation of $G(M, vv')$. In fact the expression for each criterion includes this derivative as a term. Otherwise the dependence on ω is through at least one of the terms $\omega\omega', \omega'\omega, v'\omega$, as is the case with the following expressions for the respective $G\{Z^{-1}(\gamma, v), vv'\}$ terms which will be useful in the evaluation of optimal steplengths.

5.2.7 $$v'Z^{-1}(\gamma, v)v = v'M^{-1}v/[1-\gamma+\gamma v'M^{-1}v]$$

5.2.8 $$v'Z^{-2}(\gamma, v)v = v'M^{-2}v/[\quad " \quad]^2$$

More generally

5.2.9 $$v_j'Z^{-1}(\gamma, v_t)v_j = \frac{v_j'M^{-1}v_j}{(1-\gamma)} - \frac{\gamma(v_j'M^{-1}v_t)^2}{(1-\gamma)[1-\gamma+\gamma v_t'M^{-1}v_t]}$$

5.2.10 $$v_i'Z^{-1}(\gamma, v_t)v_j = \frac{v_i'M^{-1}v_j}{(1-\gamma)} - \frac{\gamma v_i'M^{-1}v_t v_t'M^{-1}v_j}{(1-\gamma)[1-\gamma+\gamma v_t'M^{-1}v_t]}$$

In theory these latter two equations with $Z(\gamma, v_t)$ replaced by $Z_r(\pm\alpha_r, v_t)$ could be used to evaluate $G(M_{r+1}, v_j v_j')$ for all v_j in the case of D-optimality, and they could also, along with corresponding expressions for $v_j'Z_r^{-2}(\gamma, v_t)v_j$, $v_i'Z_r^{-2}(\gamma, v_t)v_j$, $\gamma = \pm\alpha_r$, be similarly used in the case of A-optimality. However having updated M_{r+1}^{-1} from M_r^{-1} by means of 5.2.5 it would be more efficient in general, to calculate the vectors $w_j = M_{r+1}^{-1} v_j$ and the appropriate scalar products $v_j'w_j$ or $w_j'w_j$.

§5.2.2 The above results are fairly well known. They are particular cases however of corresponding results for D_A-optimality and L_A-optimality which are just as simple but which are less common. They are direct consequences of the following corollary to lemma 5.2.1

Corollary 5.2.3

Let G, M be k×k nonsingular matrices; A, B be s×k matrices, H an s×s matrix and a, b be k×1 vectors which satisfy the relationships

5.2.11 $G = M + ab'$

5.2.12 $H = AG^{-1}B'$

(a) If $(AM^{-1}B')$ is nonsingular then

5.2.13 $\det(H) = \{1 - [a'M^{-1}A'(AM^{-1}B')^{-1}BM^{-1}b]/[1 + a'M^{-1}b]\}\det(AM^{-1}B')$.

(b) If also H is nonsingular then

5.2.14 $$H^{-1} = AM^{-1}B' + \frac{(AM^{-1}B')^{-1}AM^{-1}ab'M^{-1}B'(AM^{-1}B')^{-1}}{[1 + a'M^{-1}b - a'M^{-1}A'(AM^{-1}B')^{-1}BM^{-1}b]}$$

Proof These results are corollaries to lemma 5.2.1 because the matrix H has the form of the matrix G as we now see. From 5.2.3

$$H = A\left\{M^{-1} - \frac{M^{-1}ab'M^{-1}}{[1 + a'M^{-1}b]}\right\}B'$$

$$= AM^{-1}B' - \frac{AM^{-1}ab'M^{-1}B'}{[1 + a'M^{-1}b]}$$

$$= N - \beta cd'$$

where $N = AM^{-1}B'$, $c = AM^{-1}a$, $d = BM^{-1}b$, $\beta = [1 + a'M^{-1}b]^{-1}$.

Lemma 5.2.1 therefore yields expressions for det (H) and H^{-1} in terms of N, c, d, β from which equations 5.2.13, 5.2.14 follow. $\square$

The relevance of this result to us is that the D_A-optimal and L_A-optimal design criteria are functions of the nonsingular matrix $(AM^{-1}A')$ at a nonsingular M, while the Gâteaux derivative of the former depends on the inverse of $(AM^{-1}A')$. Recall that for these criteria $G(M, vv')$ takes the respective values $v'M^{-1}A(AM^{-1}A')^{-1}AM^{-1}v$ and $v'M^{-1}ALA'M^{-1}v$. We now quote counterpart equations to those of statements 5.2.4,...,5.2.8. They follow from corollary 5.2.3 with the aid of not a little algebra in some cases.

5.2.15 $$\begin{cases} \text{Let} \\ w = M^{-1}v \\ g = Aw = AM^{-1}v \\ h = (AM^{-1}A')^{-1}g = (AM^{-1}A')^{-1}AM^{-1}v \\ x = g'h = v'M^{-1}A'(AM^{-1}A')^{-1}AM^{-1}v \\ y = v'w = v'M^{-1}v. \end{cases}$$

5.2.16 $$\det\{A[Z^{-1}(\gamma,\upsilon)]A'\} = \frac{[1-\gamma+\gamma y-\gamma x]\det(AM^{-1}A')}{(1-\gamma)^s[1-\gamma+\gamma y]}$$

5.2.17 $$[A\{Z^{-1}(\gamma,\upsilon)\}A']^{-1} = (1-\gamma)[AM^{-1}A']^{-1} + \gamma(1-\gamma)hh'/[1-\gamma+\gamma y-\gamma x].$$

5.2.18 $$\mathrm{tr}\{LA\{Z^{-1}(\gamma,\upsilon)\}A'\} = \frac{\mathrm{tr}\{LAM^{-1}A'\}}{(1-\gamma)} - \frac{\gamma g'Lg}{(1-\gamma)[1-\gamma+\gamma y]}$$

5.2.19 $$\upsilon'\{Z^{-1}(\gamma,\upsilon)\}A'[A\{Z^{-1}(\gamma,\upsilon)\}A']^{-1}A\{Z^{-1}(\gamma,\upsilon)\}\upsilon$$
$$= (1-\gamma)x/[(1-\gamma+\gamma y)(1-\gamma+\gamma y-\gamma x)]$$

5.2.20 $$\upsilon'\{Z^{-1}(\gamma,\upsilon)\}A'LA\{Z^{-1}(\gamma,\upsilon)\}\upsilon = g'Lg/[1-\gamma+\gamma y]^2$$

Again the usefulness of all five expressions derives not only from their simplicity, but also from the fact that they depend on terms, namely w, g, h, which will have to be calculated anyway in the evalution of directional derivatives. The particular use to which we can put the first three equations is in updating the values of corresponding terms from one iteration of a vertex direction algorithm to the next, while the last two, as we shall now see, are of a form which makes it possible to evaluate explicitly the optimal steplength for such an iteration.

Recall that an ascertainment of the existence or otherwise of explicit forms for optimal steplengths requires examination of equation 4.2.7. Here this equation becomes

5.2.21 $$F\{Z(\gamma,\upsilon),\upsilon\upsilon'\} = 0$$

In the case of D_A-optimality this becomes

5.2.22 $$\upsilon'\{Z^{-1}(\gamma,\upsilon)\}A'[A\{Z^{-1}(\gamma,\upsilon)\}A']^{-1}A\{Z^{-1}(\gamma,\upsilon)\}\upsilon - s = 0$$

that is

5.2.23 $$[(1-\gamma)x]/[(1-\gamma+\gamma y)(1-\gamma+\gamma y-\gamma x)] - s = 0$$

In the case of L_A-optimality 5.2.21 becomes

5.2.24 $$\upsilon'\{Z^{-1}(\gamma,\upsilon)\}A'LA\{Z^{-1}(\gamma,\upsilon)\}\upsilon - \mathrm{tr}\{LA\{Z^{-1}(\gamma,\upsilon)\}A'\} = 0$$

that is

5.2.25 $$\frac{g'Lg}{[1-\gamma+\gamma y]^2} + \frac{\gamma g'Lg}{(1-\gamma)(1-\gamma+\gamma y)} - \frac{\mathrm{tr}\{LAM^{-1}A'\}}{(1-\gamma)} = 0$$

Both equations 5.2.23 and 5.2.25 reduce to the form

5.2.26 $$\left(a\gamma^2 + b\gamma + c\right)/d(\gamma) = 0 \ .$$

In the case of D_A-optimality we have

$$\begin{aligned} a &= s(y-1)(1+x-y) \\ b &= (s-1)x + 2s(1-y) \\ c &= (x-s) = F(M, vv') \\ d(\gamma) &= (1-\gamma+\gamma y)(1-\gamma+\gamma y-\gamma x) \end{aligned}$$

In the case of L_A-optimality we have

$$\begin{aligned} a &= (y-1)\left[g'Lg - (y-1)\mathrm{tr}\{LAM^{-1}A'\}\right] \\ b &= -2(y-1)\mathrm{tr}\{LAM^{-1}A'\} \\ c &= g'Lg - \mathrm{tr}\{LAM^{-1}A'\} = F(M, vv') \\ d(\gamma) &= (1-\gamma)\left[1-\gamma+\gamma y\right]^2 \end{aligned}$$

If real solutions exist to 5.2.26 they must clearly be solutions of

5.2.27 $$Q(\gamma) = a\gamma^2 + b\gamma + c = 0$$

This of course has the roots real or otherwise

5.2.28 $$\gamma_1 = \{-b + \sqrt{b^2-4ac}\}/(2a) \ , \quad \gamma_2 = \{-b - \sqrt{b^2-4ac}\}/(2a) \ .$$

It is possible though, when these are real, that only one of them or possibly neither or them would be a solution to 5.2.26.

Suppose now that γ_1, γ_2 are both non-zero solutions to 5.2.26 for $Z(\gamma, v) = Z_r(\gamma, v_t)$. We must identify $\gamma_r^+(v_t), \gamma_r^-(v_t)$ the smallest positive solution and the largest negative solution or which of the latter does not exist. Clearly it will be a simple matter to do this. However in some instances we can identify in advance which of γ_1, γ_2 is $\gamma_r^+(v_t)$ and so on. This can be done by making use of the result that $\gamma_1\gamma_2 = c/a = \left(F(M_r, v_t v_t')\right)/a$ and also of the fact that if $F(M_r, v_t v_t') > 0$, in which case $\bar{\alpha}_r(v_t) = 1$, then one of the solutions γ_1, γ_2 must in general lie in $(0,1)$

The latter result is due to the fact that in general both for $\psi(M) = -\det(AM^+A')$ and $\psi(M) = -\mathrm{tr}\{L(AM^+A')\}$, $\psi\{Z_r(\gamma, v_t)\} \longrightarrow -\infty$ as $\gamma \longrightarrow 1$. Since $\psi(\cdot)$ is concave in both cases it must be that $\alpha_r^*(v_t) < \bar{\alpha}_r(v_t)$ if $F(M_r, v_t v_t') > 0$.

More precisely the following are possibilities.

It could conceivably be that a = b = c = 0 in which case any γ is a solution of 5.2.26, 5.2.27.

More realistically equation 5.2.26 will have no (real) solutions in the following instances; if a = b = 0, c $\neq$ 0; if 5.2.27 has no real roots; if $Q(\gamma)$ does have real roots but fully cancels with the quadratic $d(\gamma)$ in the case of D_A-optimality or with a quadratic factor of the cubic $d(\gamma)$ in the case of L_A-optimality; or if a = 0 and the term $(b\gamma + c)$ cancels with one of the factors of $d(\gamma)$.

The equation will have only one solution in the following instances: if a = 0 and the term $(b\gamma + c)$ does not cancel as above; if 5.2.27 has a double root without $Q(\gamma)$ fully cancelling as above; or if 5.2.27 has two distinct real roots but one of the resultant factors of $Q(\gamma)$ cancels with a factor of $d(\gamma)$.

Finally the equation will have two real solutions if 5.2.27 has two distinct real roots and neither of the resultant factors of $Q(\gamma)$ factors as above. If c = 0 then one of the solutions is zero but then this should be the case since $c = F(M, vv') = F\{Z(0, v), vv'\}$.

We will encounter examples of most of the above.

Whichever obtains however, we can clearly identify it by checking the values of a, b, c and the roots of 5.2.27. Then taking $Z(\gamma, v) = Z_r(\gamma, v_t)$ we can identify the optimal steplength from $M(p^{(r)})$ to or away from $v_t v_t'$. We require to appeal to 5.1.9 and to the ideas enumerated at the end of section 4.2.

In particular if there are no solutions to 5.2.26 then we should take $\alpha_r^*(v_t) = \bar{\alpha}_r(v_t)$. If zero is a solution or if $Q(\gamma)/d(\gamma) = 0$ for all γ, then we should take $\alpha_r^*(v_t) = 0$, i.e. we should not take a step from $M(p^{(r)})$ towards or away from $v_t v_t'$. This is clear enough in the former case since by dint of concavity zero must be the maximising value of α in 4.2.1. If any γ is a solution of 4.2.26 then $\alpha_r^*(v_t)$ could in fact be any value $\alpha > 0$ subject to $p^{(r+1)}) = \{(1-\alpha)p^{(r)} + \alpha e_t\}$ or $\{(1+\alpha)p^{(r)} - \alpha e_t\}$ belonging to $\mathcal{P}$.

This is because the criterion $\psi(\cdot)$ is constant on the 'line' running through $M(p^{(r)})$ and $v_t v_t'$. No change in the value of $\psi(\cdot)$ will be achieved by such a step. However in the absence of any other information it would seem sensible not to take such action, but to take a vertex direction iteration corresponding to an alternative vertex to v_t assuming $p^{(r)} \neq p^*$.

If $F(M_r, v_t v_t') > 0$ then $\psi(\cdot)$ at first increases away from M_r towards $v_t v_t'$ before turning down to achieve the limiting value of $-\infty$. Since nonsingularity of M_r guarantees that of $Z_r(\gamma, v_t)$ and hence differentiability of $\psi(\cdot)$ at $Z_r(\gamma, v_t)$, the function $f(\gamma) = \psi\{Z_r(\gamma, v_t)\}$ must have a stationary value in $(0,1)$. By a similar argument if $F\{M_r, v_t v_t'\} < 0$ then neither γ_1 nor γ_2 can lie in $(0,1)$. This would be true whether or not the above limit held.

That $\psi\{Z_r(\gamma, v_t)\} \longrightarrow -\infty$ as $\gamma \to 1$ is not surprising, and this is in general true for all design criteria. $Z_r(1, v_t)$ is the design matrix corresponding to the one point design p with $\mathrm{Sup}(p) = \{v_t\}$, which in general will not guarantee estimability of $\alpha = A\theta$ save in the exceptional case of $A = c'$, $v_t = c$. The criterion is then c-optimality for which $L = 1$, $s = 1$, $x = y = \mathrm{tr}\{L(AM^{-1}A')\}$, and for $v_t = c$, $g'Lg = y^2$, and, whether viewed as a particular case of D_A-optimality or of L_A-optimality,

$$\psi\{Z_r(\gamma, v_t)\} = -c'Z_r^{-1}(\gamma, c)c = -y/(1-\gamma+\gamma y) \longrightarrow -1 \text{ as } \gamma \to 1.$$

Returning to determination of $\gamma_r^+(v_t), \gamma_r^-(v_t)$, this can be done as in the following instance. If $F\{M_r, v_t v_t'\} > 0$, then we have that one of $\gamma_1, \gamma_2 \in (0,1)$ and hence at least one of them is positive. If further $a > 0$ so that $\gamma_1 \gamma_2 > 0$, then both roots are positive with $\gamma_1 > \gamma_2$ and hence $0 < \gamma_2 < 1$ and $\gamma_r^+(v_t) = \gamma_2$ while $\gamma_r^-(v_t)$ is undefined. Proceeding in this way, and allowing for the possibility that $c = F(M_r, v_t v_t') = 0$, we obtain, with

$$\gamma_1 = \left(-b + \sqrt{b^2 - 4ac}\right)/(2a), \quad \gamma_2 = \left(-b - \sqrt{b^2 - 4ac}\right)/(2a),$$

5.2.29
$$\alpha_r^*(v_t)=\begin{cases} 0 & \text{if } c = F(M_r, v_t v_t') = 0 \\ \gamma_2 & \text{if } a > 0,\ F(M_r, v_t v_t') > 0 \\ \gamma_2 & \text{if } a < 0,\ F(M_r, v_t v_t') > 0 \\ \min\{-\gamma_2,\ p_t^{(r)}/(1-p_t^{(r)})\} & \text{if } a > 0,\ F(M_r, v_t v_t') < 0 \\ \min\{-\gamma_2,\ p_t^{(r)}/(1-p_t^{(r)})\} & \text{if } a < 0,\ F(M_r, v_t v_t') < 0,\ \gamma_1, \gamma_2 < 0. \\ p_t^{(r)}/(1-p_t^{(r)}) & \text{if } a < 0,\ F(M_r, v_t v_t') < 0,\ \gamma_1, \gamma_2 > 0 \\ p_t^{(r)}/(1-p_t^{(r)}) & \text{if } (b^2 - 4ac) < 0 \end{cases}$$

Suppose now either as a result of one factor of $Q(\gamma)$ cancelling with one of $d(\gamma)$ or because $a = 0$ that equation 5.2.26 reduces to a linear equation

$$a_1 \gamma - a_2 = 0$$

with therefore the unique solution

$$\tilde{\gamma} = a_2 / a_1$$

In view of concavity this can be zero iff $F\{M_r, v_t v_t'\} = 0$. By arguments analogous to the above we will have

5.2.30
$$\alpha_r^*(v_t) = \begin{cases} a_2/a_1 & \text{if } F(M_r, v_t v_t') \geq 0 \\ \min\{-a_2/a_1,\ p_t^{(r)}/(1-p_t^{(r)})\} & \text{if } F(M_r, v_t v_t') < 0,\ a_2/a_1 < 0 \\ p_t^{(r)}/(1-p_t^{(r)}) & \text{if } F(M_r, v_t v_t') < 0,\ a_2/a_1 > 0. \end{cases}$$

Such cancelling happens in the case of D-optimality and in the case of c-optimality for $v_t = c$, when this criterion is considered either as a particular case of D_A-optimality or of L_A-optimality.

In this latter case $x = y = c'M^{-1}c$ so that in the D_A-optimal case

$$Q(\gamma) = (y-1)(\gamma-1)^2$$
$$d(\gamma) = (1-\gamma+\gamma y)(1-\gamma)\ .$$

We must note that if $y = 1$ then $Q(\gamma)/d(\gamma) = 0$ for all γ, that is any γ is a solution of 5.2.26, and so $F\{Z(\gamma, c), cc'\} = 0$, and the criterion takes the constant value of 1 on the 'line' through $M(p)$ and cc'.

We have already considered this possibility at nonsingular M and concluded that in the absence of further information we should take $\alpha_r^*(v_t) = 0$

However if $y \neq 1$ the factor $(\gamma-1)$ is clearly common to both $Q(\gamma)$, $d(\gamma)$. Remove it and we have $a_1 = a_2 = (y-1)$ and hence $\gamma = 1$ whatever the non-zero value of y. There is thus never any negative solution here.

If we view c-optimality as a particular case of L_A-optimality $(L = I)$ we additionally have $\mathrm{tr}\{L(AM^{-1}A')\} = y$, $g'Lg = y^2$ and we get

$$Q(\gamma) = y(y-1)(\gamma-1)^2 ,$$

which will have exactly the same implications.

We conclude that for $v_t = c$

$$\alpha_r^*(v_t) = \begin{cases} 1 & \text{if } F(M_r, v_t v_t') > 0 \\ -p_t^{(r)}/(1-p_t^{(r)}) & \text{if } F(M_r, v_t v_t') < 0 \\ 0 & \text{if } F(M_r, v_t v_t') = 0 \end{cases}$$

The D-optimal criterion is that particular case of D_A-optimality for which $A = I_k$, $s = k$. In consequence $x = y$ and we obtain

$$Q(\gamma) = (\gamma-1)[k(y-1)\gamma - (y-k)]$$
$$d(\gamma) = (1-\gamma)(1-\gamma+\gamma y)$$

Again the factor $(\gamma-1)$ is common to both. However if $y = 1$, then, for all γ, $Q(\gamma)/d(\gamma) = (1-k) < 0$, as should $F(M, vv')$. There is therefore no solution to 5.2.26, in which case, for $Z(\gamma, v) = Z_r(\gamma, v_t)$ we should have $\alpha_r^*(v_t) = \bar{\alpha}_r(v_t) = p_t^{(r)}/(1-p_t^{(r)})$.

Note that it is possible that $y = v'M^{-1}(p)v \leq 1$.

Take $\mathcal{V} = \{(1,0)', (0,1)', (x,x)'\}$,

$p = (1/3, 1/3, 1/3)$, so that

$$M(p) = (1/3)\begin{bmatrix} 1+x^2 & x^2 \\ x^2 & 1+x^2 \end{bmatrix}$$

and for $v = (x,x)'$, $v'M^{-1}(p)v = 6x^2/(9+2x^2) < 1$ if $|x| < 3/2$.

Excepting the case $y = 1$ then 5.2.26 will have in the case $Z(\gamma,v) = Z_r(\gamma,v_t)$ the unique solution

$$5.2.31 \qquad \tilde{\gamma}_r(v_t) = \left(v_t'M_r^{-1}v_t - k\right)\Big/\left\{k\left(v_t'M_r^{-1}v_t - 1\right)\right\}$$
$$= F\left(M_r, v_tv_t'\right)\Big/\left\{ \quad '' \quad \right\}$$

If $F\left(M_r, v_tv_t'\right) > 0$ then we do have $0 < \tilde{\gamma}_r(v_t) < 1$ since $v_t'M_r^{-1}v_t > k$ implies both that $\left(v_t'M_r^{-1}v_t - k\right) < \left(kv_t'M_r^{-1}v_t - k\right)$ and also that $kv_t'M_r^{-1}v_t > k^2 > k$.

Similarly $\tilde{\gamma}_r(v_t) > 1$ if $v_t'M_r^{-1}v_t < 1$, since then $\left(v_t'M_r^{-1}v_t - k\right) < \left(kv_t'M_r^{-1}v_t - k\right) < 0$.

Appealing to 5.2.30 we obtain

$$5.2.32 \qquad \alpha_r^*(v_t) = \begin{cases} \left(v_t'M_r^{-1}v_t - k\right)\Big/\left[k\left(v_t'M_r^{-1}v_t - 1\right)\right] & \text{if } F\left(M_r, v_tv_t'\right) > 0 \\ \min\left\{\dfrac{k - v_t'M_r^{-1}v_t}{k\left(v_t'M_r^{-1}v_t - 1\right)}, \dfrac{p_t^{(r)}}{1 - p_t^{(r)}}\right\} & \text{if } 1 < v_t'M_r^{-1}v_t < k \\ p_t^{(r)}\Big/\left(1 - p_t^{(r)}\right) & \text{if } v_t'M_r^{-1}v_t < 1 \end{cases}$$

This could have been more readily derived by equating to k, the expansion given for $v_t'\left\{Z_r^{-1}(\gamma, v_t)\right\}v_t$ in equation 5.2.7.

Fedorov (1972), among others derived 5.2.32, but by a direct maximisation of the expansion given for $\det\left\{Z_r(\gamma, v_t)\right\}$ in 5.2.4. He noted (p.114) that the optimising α_r^* satisfied $v_t'\left\{Z_r^{-1}(\alpha_r^*, v_t)\right\}v_t = k$.

Using exactly the same approach for D_s-optimality Atwood (1973) obtains a quadratic equation in terms of $\beta = \gamma/(1-\gamma)$ 'for the optimal steplength' in the particular case $F\left(M_r, v_tv_t'\right) > 0$. He also established that in this instance

$$v_t'\left\{Z_r^{-1}(\alpha_r^*, v_t)\right\}A'\left[A\left\{Z_r^{-1}(\alpha_r^*, v_t)\right\}A'\right]^{-1}A\left\{Z_r(\alpha_r^*, v_t)\right\}v_t = s. \quad , \alpha_r^* = \alpha_r^*(v_t).$$

Finally again using this approach for L_A-optimality in the case $A = I$, Atwood (1976a) obtains an equation for γ exactly of the form $Q(\gamma)/d(\gamma) = 0$ where $Q(\gamma)$ is the same quadratic as in 5.2.26 but now $d(\gamma) = (1-\gamma)^2(1-\gamma+\gamma u)^2$ which in view of 4.2.6 is what should be the case.

While, in view of the equivalence theorem, the two authors must have found the equalities which they observed intuitively appealing, they did not fully realise why these equalities should hold. They did not however have the directional derivative tool at their disposal then.

This completes the results concerning vertex direction algorithms that we wish to derive from lemma 5.2.1. They are restricted to the above criteria and a crucial assumption in applying the lemma to these algorithms is that $M_r = M(p^{(r)})$ should be nonsingular. While we must have this for D-optimality and A-optimality it need not be the case that $M(p^*)$ be nonsingular for the general D_A-optimality or L_A-optimality criteria.

We will still however in general select $p^{(0)}$ such that $M(p^{(0)})$ is nonsingular and, while we could ensure at succeeding iterates that M_r be nonsingular, we will see later that we may be happy to pass to a singular M_r, the timing of such a step being such that we would subsequently wish to use an algorithm not of the vertex direction type. Moreover we can still have a simpler than usual task in in evaluating M_{r+1}^{-1} or M_{r+1}^{+}, whatever algorithm we use, as shall be seen in section 6.1.

§5.3 Examples Of Vertex Direction Algorithms

We now consider some examples of vertex direction algorithms. We have to decide both on a method for selecting a vertex towards which or away from which to move from iterate $x_r = x(p^{(r)})$, and on a formula for the steplength α_r, $\alpha_r > 0$

As has been said the optimal choice $\alpha_r^*(u_t)$ of α_r is given in 5.1.8, and we have just seen instances when we can evaluate this explicitly, but typically numerical techniques would be required for its evaluation. In the case of a forward vertex direction algorithm we know that $\alpha_r^*(u_t) < 1 = \bar{\alpha}_r(u_t)$ and hence must solve equation 4.2.7 with $Z(\delta, u) = Z_r(\delta, u_t) = (1-\delta)x_r + \delta u_t$, so that we could determine $\alpha_r^*(u_t)$ by solving the latter in $(0,1)$ using Newton Raphson techniques. In the case of a reverse vertex direction the best course of action would be to solve 4.2.1 using techniques such as Golden Section.

However it was for a forward vertex direction algorithm that the concept of arbitrary steplengths was conceived; i.e. $\alpha_r = \beta_r \bar{\alpha}_r(u_t)$ for an arbitrary but reasonable preassigned value β_r in $0 < \beta_r < 1$. This puts $\alpha_r = \beta_r$ in the case of a forward vertex direction. The choice of β_r will be examined in section 5.6.

However we choose α_r, it will typically be a value $\alpha_r(u_t) > 0$ depending on u_t. In this context either $\alpha_r(u_t) = \alpha_r^*(u_t)$ or $\alpha_r(u_t) = \beta_r \bar{\alpha}_r(u_t)$

We will denote a vertex direction algorithm by $V\{u_t, \alpha_r(u_t)\}$ where u_t indicates the type of vertex to be selected. Hence we will typically be referring to algorithm $V\{u_t, \alpha_r^*(u_t)\}$ or algorithm $V\{u_t, \beta_r \bar{\alpha}_r(u_t)\}$ or algorithm $V\{u_t, \beta_r\}$ if u_t is a forward vertex direction.

Recall that the basic iteration is

$$x_{r+1} = Z\{\alpha_r(u_t)\}$$

where

$$Z\{\alpha_r(u_t)\} = \begin{cases} \{1 - \alpha_r(u_t)\}x_r + \{\alpha_r(u_t)\}u_t & \text{if } F(x_r, u_t) > 0 \\ \{1 + \alpha_r(u_t)\}x_r - \{\alpha_r(u_t)\}u_t & \text{if } F(x_r, u_t) < 0 \end{cases}$$

We now present a not exhaustive list of suggestions on how to choose u_t from $\mathcal{U} = \{u_1, \ldots, u_J\}$ (or v_t from $\mathcal{V} = \{v_1, \ldots, v_J\}$).

Let $x_r = x(p^{(r)})$, $M_r = M(p^{(r)})$

Let

5.3.1
$$\begin{cases} \mathcal{U}_r^+ = \{u_j : u_j \in \mathcal{U},\ F(x_r, u_j) > 0\} \\ \mathcal{U}_r^- = \{u_j : u_j \in \mathrm{Sup}(p^{(r)}),\ F(x_r, u_j) < 0\} \\ \mathcal{V}_r^+ = \{v_j : v_j \in \mathcal{V},\ F(M_r, v_j v_j') > 0\} \\ \mathcal{V}_r^- = \{v_j : v_j \in \mathrm{Sup}(p^{(r)}),\ F(M_r, v_j v_j') < 0\} \end{cases}$$

A first selection is

5.3.2
$$\begin{cases} u^{(1)} = \text{the vertex which solves } \max_{u_j \in \mathcal{U}_r^+} \{F(x_r, u_j)\} \\ u^{(2)} = \quad\quad \min_{u_j \in \mathcal{U}_r^-} \{F(x_r, u_j)\} \\ u^{(3)} = \quad\quad \max_{u_j \in \{\mathcal{U}_r^+ \cup \mathcal{U}_r^-\}} \{|F(x_r, u_j)|\} \end{cases}$$

Clearly
$$u^{(3)} = \begin{cases} u^{(1)} & \text{if } F(x_r, u^{(1)}) > |F(x_r, u^{(2)})| \\ u^{(2)} & \text{if } \quad '' \quad < | \quad '' \quad | \end{cases}$$

Secondly

5.3.3
$$\begin{cases} u^{(4)} = \text{the vertex which solves } \max_{u_j \in \mathcal{U}_r^+} \{F^A(x_r, u_j)\} \\ u^{(5)} = \quad\quad \min_{u_j \in \mathcal{U}_r^-} \{F^A(x_r, u_j)\} \\ u^{(6)} = \quad\quad \max_{u_j \in \{\mathcal{U}_r^+ \cup \mathcal{U}_r^-\}} \{|F^A(x_r, u_j)|\} \end{cases}$$

Clearly
$$u^{(6)} = \begin{cases} u^{(4)} & \text{if } F^A(x_r, u^{(4)}) > |F^A(x_r, u^{(5)})| \\ u^{(5)} & \text{if } \quad '' \quad < | \quad '' \quad | \end{cases}$$

The following are choices for which an arbitrary steplength is appropriate

5.3.4
$$\begin{cases} u^{(7)} = \begin{cases} u^{(1)} & \text{if } \psi(Z\{\beta_r \bar{\alpha}_r(u^{(1)})\}) > \psi(Z\{\beta_r \bar{\alpha}_r(u^{(2)})\}) \\ u^{(2)} & \text{if } \quad '' \quad < \quad '' \end{cases} \\ \\ u^{(8)} = \begin{cases} u^{(3)} & \text{if } \psi(Z\{\beta_r \bar{\alpha}_r(u^{(3)})\}) > \psi(Z\{\beta_r \bar{\alpha}_r(u^{(4)})\}) \\ u^{(4)} & \text{if } \quad '' \quad < \quad '' \end{cases} \end{cases}$$

5.3.4 (cont'd)

$$\begin{cases} u^{(9)} = \text{the vertex which solves } \max_{u_j \in \mathcal{U}_r^+} \{ \psi(Z\{\beta_r \bar{\alpha}_r(u_j)\}) \} \\ u^{(10)} = \qquad \max_{u_j \in \mathcal{U}_r^-} \{ \quad " \quad \} \\ u^{(11)} = \qquad \max_{u_j \in \{\mathcal{U}_r^+ \cup \mathcal{U}_r^-\}} \{ \quad " \quad \} \end{cases}$$

Clearly

$$u^{(11)} = \begin{cases} u^{(9)} & \text{if } \psi(Z\{\beta_r \bar{\alpha}_r(u^{(9)})\}) > \psi(Z\{\beta_r \bar{\alpha}_r(u^{(10)})\}) \\ u^{(10)} & \text{if } \quad " \quad < \quad " \end{cases}$$

The following are choices for which an optimal steplength is appropriate

5.3.5

$$\begin{cases} u^{(12)} = \begin{cases} u^{(1)} & \text{if } \psi(Z\{\alpha_r^*(u^{(1)})\}) > \psi(Z\{\alpha_r^*(u^{(2)})\}) \\ u^{(2)} & \text{if } \quad " \quad < \quad " \end{cases} \\ u^{(13)} = \begin{cases} u^{(3)} & \text{if } \psi(Z\{\alpha_r^*(u^{(3)})\}) > \psi(Z\{\alpha_r^*(u^{(4)})\}) \\ u^{(4)} & \text{if } \quad " \quad \quad " \end{cases} \\ u^{(14)} = \text{the vertex which solves } \max_{u_j \in \mathcal{U}_r^+} \{ \psi(Z\{\alpha_r^*(u_j)\}) \} \\ u^{(15)} = \qquad \max_{u_j \in \mathcal{U}_r^-} \{ \quad " \quad \} \\ u^{(16)} = \qquad \max_{u_j \in \{\mathcal{U}_r^+ \cup \mathcal{U}_r^-\}} \{ \quad " \quad \} \end{cases}$$

Clearly

$$u^{(16)} = \begin{cases} u^{(14)} & \text{if } \psi(Z\{\alpha_r^*(u^{(14)})\}) > \psi(Z\{\alpha_r^*(u^{(15)})\}) \\ u^{(15)} & \text{if } \quad " \quad < \quad " \end{cases}$$

In the design context the notation will change to $v^{(1)},\ldots,v^{(16)}$.

Obviously the above cannot be an exhaustive list. Each is clearly optimal in some sense, some being more optimal than others, with $V\{u^{(16)}, \alpha_r^*(u^{(16)})\}$ being such that $\psi(x_{r+1})$ is larger than that for any of the other $u^{(j)}$. Note that there need not be a unique $u^{(j)}$ for any j.

Of course the more optimal ones may demand more computation and this will certainly be the case if $\alpha_r^*(u_t)$ has to be evaluated numerically. Hence the variety of choices. In general it will be simpler to evaluate $F(x_r, u_j)$ than $\psi(Z\{\alpha_r(u_j)\})$ and the former has to be evaluated anyway to test for optimality.

All choices require a maximisation or minimisation over a finite set. If this were large one may settle for a u_t satisfying weaker conditions than those of the $u^{(j)}$ that would have been our choice. Also with a view to cutting down on computations we might opt to determine $u^{(4)}$ or $u^{(5)}$ as opposed to $u^{(6)}$ and similarly with $u^{(9)}$, $u^{(10)}$, $u^{(11)}$ and with $u^{(14)}$, $u^{(15)}$, $u^{(16)}$.

We now consider the origins of some of the above $u^{(j)}$ and comment on them.

The idea of a (forward) vertex direction algorithm seems to have been conceived and only used in the design context and the first such recommendation was for D-optimality. Independently both Fedorov (1972) (see earlier references in that text) and Wynn (1970, 1972) recommended $v^{(1)}$ and they did so despite the fact that they did not have the directional derivative tool at their disposal. Wynn used $V\{v^{(1)}, 1/(k+r+1)\}$ while Fedorov, realising that the optimal steplength could be evaluated explicitly, favoured $V\{v^{(1)}, \alpha_r^*(v^{(1)})\}$. Wynn also employed $V\{v^{(1)}, 1/(k+r+1)\}$ in the case of D_s-optimality while Atwood (1973) recommended $V\{v^{(1)}, \alpha_r^*(v^{(1)})\}$ in view of his quadratic equation for finding $\beta^* = \alpha^*/1-\alpha^*$, $\alpha^* = \alpha_r^*(v^{(1)})$.

This choice of $v^{(1)}$ has since been a predominant one in the design context and is about the simplest choice, its identification not requiring much extra computation in view of the fact that, as we have said, $F(M_r, v_j v_j')$ must be calculated anyway to test for optimality. Several authors have considered $V\{v^{(1)}, \beta_r\}$ with a general design criterion in mind as shall be seen in section 5.6.

The choice $u^{(1)}$ is one which at differentiable x_r has the superior credential that it solves $\max_{y \in \mathcal{R}(U)} F(x_r, y)$ as seen in (D7) of section 2.3.

One might have thought then that as a result $u^{(1)}$ and $u^{(3)}$ would be the same, that is that $F(x_r, u^{(1)}) > |F(x_r, u^{(2)})|$ always, for at differentiable $F(x_r, 2x_r - u^{(2)}) = |F(x_r, u^{(2)})| > 0$.

If $u^{(1)}$ maximises $F(x_r, y)$ with respect to y then surely $F(x_r, u^{(1)}) > F(x_r, 2x_r - u^{(2)})$.

However this is only guaranteed to be the case if $(2x_r - u^{(2)}) \in \mathcal{D}(\mathcal{U})$ for the maximisation is restricted to $y \in \mathcal{D}(\mathcal{U})$. It need not be the case that $(2x_r - u^{(2)}) \in \mathcal{D}(\mathcal{U})$. What we will have is $F(x_r, u^{(1)}) > F(x_r, x_r + \alpha m_r)$, $m_r = x_r - u^{(2)}$ for $0 < \alpha < \bar{\alpha}(m_r)$. A design example illustrates the point.

Take $\mathcal{U}$ to be Wynn's design space quoted in section 3.3.4. That is $\mathcal{U} = \{v_1, v_2, v_3, v_4\} = \{(1,-1,-1)', (1,-1,1)', (1,1,-1)', (1,2,2)'\}$ and let $p = (1/4\ 1/4\ 1/4\ 1/4)'$. For j = 1,2,3,4 respectively $F\{M(p),\ v_j v_j'\} = -104/152,\ 8/152,\ 8/152,\ 88/152$. Therefore $v^{(1)} = v_4$, $v^{(3)} = v^{(2)} = v_1$, while $\bar{\alpha}(v^{(2)}) = 1/4(1-1/4)^{-1} = 1/3$. Finally $F\{M(p), (1+\alpha)M(p) - \alpha v_1 v_1'\} = -\alpha F\{M(p), v_1 v_1'\} = (104/152)\alpha$ and $(104/152)\alpha \leq (34 + 2/3)/152$ if $0 < \alpha \leq 1/3$.

This seems an undesirable feature, one due to the dependence of $F(x,y)$ on the distance between x and y.

Possibly because it satisfies the above maximisation, $u^{(1)}$ has been erroneously referred to as a steepest ascent direction. See Atwood (1976b). **In view of (D9) of section 4.3, a 'steepest' ascent direction is not a** vertex direction. Had it been otherwise there would not have been the need, justification or scope for a range of alternatives as extensive as the above list. Choice $u^{(4)}$ might be viewed as a compromise; the steepest forward vertex direction (with respect to the norm or its generating matrix A). Wu (1976) appears to consider $u^{(4)}$ in certain circumstances. We will see below though, that $u^{(1)}$ will, in the case of D-optimality, still seem to be a superior choice than $u^{(4)}$. We would similarly refer to $u^{(5)}$ as the steepest reverse vertex direction. Clearly $u^{(6)}$ selects the steeper of the two directions corresponding to $u^{(4)}$ and $u^{(5)}$. This may be superior to $u^{(1)}$.

The above misconception may alternatively be due to the fact that in the case of D-optimality $v^{(1)}, v^{(9)}, v^{(14)}$ are identical so that $V\{v^{(1)}, \beta_r\}$ and $V\{v^{(9)}, \beta_r\}$, generate identical iterates and so do $V\{v^{(1)}, \alpha_r^*(v^{(1)})\}$ and $V\{v^{(14)}, \alpha_r^*(v^{(14)})\}$.

This is easily seen from the fact that in 5.2.4 $\det\{Z_r(\gamma, v_j)\} = (1-\gamma)^{k-1}\cdot(1-\gamma+\gamma v_j' M_r^{-1} v_j)\cdot\det(M_r)$. This depends on v_j only through the term $v_j' M_r^{-1} v_j = G(M_r, v_j v_j')$ and in fact is proportional to the linear term $\{1-\gamma+\gamma G(M_r, v_j v_j')\}$.

Hence for a given $\gamma > 0$, this will be maximised over $\mathcal{V}_r^+$ by $v^{(1)}$. It is also because of this that, for a given $v_j \in \mathcal{V}_r^+$, the value of γ which maximises $\det\{Z_r(\gamma, v_j)\}$ over $0 \le \gamma < 1$ also depends on v_j only through $G(M_r, v_j v_j')$, it being $(v_j' M_r^{-1} v_j - k)/k(v_j' M_r^{-1} v_j - 1)$. Hence $\det\{Z_r(\gamma, v_j)\}$ is simultaneously maximised over $\mathcal{V}_r^+$ and $0 \le \gamma < 1$ by the pair $\{v^{(1)}, \alpha_r^*(v^{(1)})\}$ for any $v^{(1)}$ which solves $\max_{v_j \in \mathcal{V}_r^+}\{F(M_r, v_j v_j')\}$.

If $v^{(j)}$ is a forward vertex direction then in general $V\{v^{(9)}, \beta_r\}$ is superior to $V\{v^{(j)}, \beta_r\}$ while $V\{v^{(14)}, \alpha_r^*(v^{(14)})\}$ is superior to $V\{v^{(j)}, \alpha_r^*(v^{(j)})\}$.

It is this which suggests that $v^{(1)}$ is a better choice than $v^{(4)}$ in this D-optimal case.

It is also the case for this criterion that $v^{(2)} = v^{(10)}$ always and consequently $v^{(3)} = v^{(1)}$.

However such equivalence will not hold in general as is readily seen in the case of A-optimality where $\mathrm{tr}\{Z_r(\gamma, v_j)\}$ not only depends on v_j through $G(M_r, v_j v_j')$, but also through the terms $v_j' M_r^{-1} v_j$. Wynn (1970) also observes that $V\{v^{(1)}, 1/(k+r+1)\}$ and $V\{v^{(19)}, 1/(k+r+1)\}$ would not be the same in the case of D_s-optimality. He points out that Silvey (1969) effectively selects $v^{(8)}$ for what is an example of c-optimality.

The above discussion so far has been concerned with forward vertex directions. The idea of a reverse vertex direction was first mooted by Atwood (1973). He did not however consider the idea that one might a priori opt to move in a reverse vertex direction. Anyway as we have noted one could not consistently choose such a direction at every iteration unless $\mathrm{Sup}(p^{(0)}) \supseteq \mathrm{Sup}(p^{*})$. Atwood in fact suggested $V\{v^{(12)}, \prec_r^{*}(v^{(12)})\}$ for D-optimality. In general the algorithms which are based on $u^{(3)}$, $u^{(6)}$, $u^{(7)}$, $u^{(8)}$, $u^{(12)}$, $u^{(13)}$, $u^{(16)}$ adopt the policy of sparingly choosing a reverse vertex direction, when to do so seems more optimal than to opt for a converse forward vertex direction.

Apart from an idea of Ford (1976) which will be considered in the next section no other author has developed the concept of a reverse vertex direction.

§5.4 Bi-vertex Direction Algorithm

This iterative procedure does not take a vertex direction step at each iteration. Instead it adopts the slightly more complicated policy of simultaneously moving towards a vertex which would define a forward vertex direction and away from a vertex which would define a reverse vertex direction. By this we mean that

$$x_{r+1} = [1-(\beta_1-\beta_2)]x_r + \beta_1 u_s - \beta_2 u_t ,$$

for some positive β_1, β_2 which ensure that $x_{r+1} \in \mathcal{P}(\mathcal{U})$, where $F(x_r, u_s) > 0$, $u_t \in Sup(p^{(r)})$ and $F(x_r, u_t) < 0$. Such a step then makes a significant change in only two weights, apart from a proportional change in the others.

Ford (1976) first conceived the idea and had in mind $(u_s, u_t) = (u^{(1)}, u^{(2)})$. We consider first his approach to the choice of β_1, β_2. Let

$$Z_r(\gamma_1, \gamma_2, u_s, u_t) = [1-(\gamma_1-\gamma_2)]x_r + \gamma_1 u_s - \gamma_2 u_t .$$

Ford aims to find γ_1^*, γ_2^* to maximise $\psi\{Z_r(\gamma_1, \gamma_2, u_s, u_t)\}$ subject to $(\gamma_1-\gamma_2) < 1$.

One would expect $\gamma_1^*, \gamma_2^* > 0$ and also, if the concavity of $\psi(\cdot)$ extends appropriately beyond $\mathcal{P}(\mathcal{U})$, we should have $F(x_r, y_r) > 0$, where $y_r = Z_r(\gamma_1^*, \gamma_2^*, u_s, u_t)$, at least if u_t, u_s are well chosen. Note that it need not be that $y_r \in \mathcal{P}(\mathcal{U})$ but we will have $y_r = x(q^{(r)})$ where $\sum q_j^{(r)} = 1$. It would be natural, given the choice of γ_1^*, γ_2^*, to take an optimal step in the direction of y_r and if $F(x_r, y_r) > 0$, then we will have

5.4.1
$$\begin{aligned} x_{r+1} &= [1-\alpha_r^*(y_r)]x_r + [\alpha_r^*(y_r)]y_r \\ p^{(r+1)} &= [1-\alpha_r^*(q^{(r)})]p^{(r)} + [\alpha_r^*(q^{(r)})]q^{(r)} . \end{aligned}$$

It must be that

5.4.2
$$\alpha_r^*(y_r) = \begin{cases} 1 & \text{if } \bar{\alpha}_r(y_r) > 1 \\ \bar{\alpha}_r(y_r) & \text{if } \bar{\alpha}_r(y_r) < 1 \end{cases} \quad \text{if } F(x_r, y_r) > 0,$$

since if $y_r \in \mathcal{P}(\mathcal{U})$ then the solution to 4.2.1 for $m_r = y_r - x_r$ must be given by $\alpha_r^*(m_r) = 1$

Ford derives explicit formulae for γ_1^*, γ_2^* in the case of D-optimality. That this is possible is due to the fact that by twice appealing to lemma 5.2.1 we obtain for

$Z_r(\gamma_1, \gamma_2, v_s, v_t) = \{[1-(\gamma_1-\gamma_2)]M_r + \gamma_1 v_s v_s' - \gamma_2 v_t v_t'\}$ that

5.4.3 $$\det\{Z_r(\gamma_1,\gamma_2,v_s,v_t)\} = \{\det(M_r)\}\gamma_0^k\{1+\gamma_1 d_1/\gamma_0 - \gamma_2 d_2/\gamma_0 - \gamma_1\gamma_2 d_1 d_2/\gamma_0^2 + \gamma_1\gamma_2 d_{12}^2/\gamma_0^2\}$$

where $\gamma_0 = 1-(\gamma_1-\gamma_2)$, $d_1 = v_s' M_r^{-1} v_s$, $d_2 = v_t' M_r^{-1} v_t$, $d_{12} = v_1' M_r^{-1} v_2$.

Similar updates could be obtained for relevant terms in the case of D_A-optimality and L_A-optimality. Indeed that this is so is the reason for contemplating this type of algorithm.

Ford first fixes $(\gamma_1-\gamma_2) = x$ and solves the derivative equations of the lagrangian to obtain explicit expressions for γ_1, γ_2 in terms of x. Substituting in 5.4.3 he equates the derivative of the function f(x) thus obtained to zero. This requires the solution of a quadratic; the solution which maximises f(x) can be identified. As a result γ_1^*, γ_2^* are defined by a hierarchy of expressions and it is not clear that we will always have $\gamma_1^*, \gamma_2^* > 0$ or even that $\gamma_1^* - \gamma_2^* < 1$ for there is no constraint that x satisfy $x < 1$. The solution γ_1^*, γ_2^* in this respect then is an unconstrained one and from the above must be a stationary value of

$$f(\gamma_1,\gamma_2) = \det\{Z_r(\gamma_1,\gamma_2,v_s,v_t)\}.$$

By an argument analogous to that establishing that 4.2.4 and 4.2.5 must have a common solution, it must be here that

$$F_\psi\{Z_r(\gamma_1^*,\gamma_2^*,v_s,v_t),\ Z_r(\gamma_1,\gamma_2,v_s,v_t)\} = 0, \quad \forall\ \gamma_1,\gamma_2.$$

This will be true iff γ_1^*, γ_2^* solve

5.4.4 $$\begin{cases} F_\psi\{Z_r(\gamma_1,\gamma_2,v_s,v_t),\ v_s v_s'\} = 0 \\ F_\psi\{Z_r(\gamma_1,\gamma_2,v_s,v_t),\ v_t v_t'\} = 0 \end{cases}$$

These of course are equivalent to equating the ordinary partial derivatives with respect to γ_1, γ_2 to zero. In the case of D-optimality they become, with $M_r = M$,

5.4.5 $$v_j'[\{1-(\gamma_1-\gamma_2)\}M + \gamma_1 v_s v_s' - \gamma_2 v_t v_t']^{-1} v_j = k, \quad j = s,t$$

Clearly these could be simplified by twice appealing to lemma 5.2.1. In general both will simplify to quadratics in γ_1, γ_2. This should also occur in the case of D_A and L_A-optimality.

In view of Ford's results, equations 5.4.5 enjoy a solution in explicit form. This may also be true of D_A- and L_A-optimality. However it does not seem any simpler to derive γ_1^*, γ_2^* directly from 5.4.5 than to obtain them by Ford's method.

Our main reason for pointing out that Ford's γ_1^*, γ_2^* must satisfy 5.4.4 is that this serves to shed light on results he obtains in two examples in which he takes $(v_s, v_t) = (v^{(1)}, v^{(2)})$.

Ex. 5.4.1

Again we consider Wynn's design space i.e.

$\mathcal{V} = \{v_1, v_2, v_3, v_4\} = \{(1,-1,-1)'(1,-1,1)', (1,1,-1)', (1,2,2)'\}$.

For $p^{(0)} = (1/3, 1/3, 1/3, 0)'$, $v^{(1)} = v_4$, $v^{(2)} = v_1$,

$\gamma_1^* = 10/32$, $\gamma_2^* = 5/32$, $Z_0(\gamma_1^*, \gamma_2^*, v^{(1)}, v^{(2)}) = y(q^{(0)})$,

where $q^{(0)} = (4/43, 9/32, 9/32, 10/32)$ which not only belongs to $\mathcal{P}$ but is in fact p^*.

Ex. 5.4.2

Here $\mathcal{V} = \{v_1, v_2, v_3, v_4\} = \{(1,0,0)', (0,1,0)', (0,0,1)', (1/2, 1/2, 1/2)'\}$

For $p^{(0)} = (1/4, 1/4, 1/4, 1/4)'$, $v^{(1)} = v_1$ or v_2 or v_3, $v^{(2)} = v_4$,

$\gamma_1^* = 0$, $\gamma_2^* = 0.6$, $Z_0(\gamma_1^*, \gamma_2^*, v^{(1)}, v^{(2)}) = y(q^{(0)})$, where

$q^{(0)} = (2/5, 2/5, 2/5, -1/5)' \notin \mathcal{P}$, $F(p^{(0)}, q^{(0)}) = 27/35 > 0$

$\alpha_0^*(q^{(0)}) = \bar{\alpha}_0(q^{(0)}) = 5/9$ and hence, from 5.4.1,

$p^{(1)} = (1/3, 1/3, 1/3, 0) = p^*$.

Note that the iteration reduces to taking the optimal step in the reverse vertex direction $v^{(2)}$. Also $v^{(3)} = v^{(2)}$.

So in these two examples we have identified the optimum in one step. Clearly this can only occur if $x^* = Z_0(\tilde{\gamma}_1, \tilde{\gamma}_2, u_s, u_t)$ for some $\tilde{\gamma}_1, \tilde{\gamma}_2$ and this will only be the case if the weights under $p^{(0)}$ of all vertices except u_s, u_t are, relative to each other, in the same proportion as the weights of those vertices under p^*.

Suppose x^* does satisfy the above. Then we will not necessarily identify it in one step. Only if u_t and u_s are included in the support of p^* are we guaranteed to do this, for then $F\{x(p^*), u_t\} = F\{x(p^*), u_s\} = 0$ and, that $\tilde{\gamma}_1, \tilde{\gamma}_2$ is the solution to 5.4.4, is clear.

However if $p^*_t = 0$, as might be the case, then typically $F\{x(p^*), u_t\} < 0$ and so $\tilde{\gamma}_1, \tilde{\gamma}_2$ cannot solve 5.4.4. Then $x^* = Z_0(\tilde{\gamma}_1, \tilde{\gamma}_2, u_s, u_t)$ solves only $\max_{\gamma_1, \gamma_2} \{\psi\{Z_0(\gamma_1, \gamma_2, u_s, u_t)\}\}$ subject to $Z_0(\tilde{\gamma}_1, \tilde{\gamma}_2, u_s, u_t) \in \mathcal{D}(\mathcal{U})$ and is not a corresponding unrestricted maximising value.

In this case we will only identify $x^* = x(p^*)$ in one step if x^* lies on the 'line' running through $x_0 = x(p^{(0)})$ and $y_0 = Z_0(\gamma_1^*, \gamma_2^*, u_s, u_t)$

This was the case in example 5.4.2. Possibly this will always be the case with D-optimality, a manifestation of a property similar to the equivalence of $u^{(1)}$ and $u^{(9)}$, maybe a generalisation of that equivalence.

However it would seem likely that in general x^* will not lie on the above line, in which case we will not pass to x^* from x_0. The iteration 5.4.1 would still be in general a sensible one to take, though it is conceivable that a greater increase would be obtained by taking the optimal step towards u_s or the optimal step away from u_t.

Ford acknowledges that choices other than $(u^{(1)}, u^{(2)})$ for (u_s, u_t) could be employed. Another reasonable choice might be $(u^{(4)}, u^{(5)})$. Those $u^{(j)}$ though, which require evaluation of $\psi(\cdot)$ in their identification, would be less natural choices.

Still assuming $F(x_r, u_s) > 0$, $F(x_r, u_t) < 0$ other simpler examples of bi-vertex iterations include

5.4.5 $\quad m_r = u_s - u_t \quad$, ie, $\; y_r = x_r + u_s - u_t$,

and more generally

5.4.6 $\quad y_r = [1 - (\gamma_1 - \gamma_2)]\, x_r + \gamma_1 u_s - \gamma_2 u_t \quad$,

where γ_1, γ_2 are chosen in some way other than 'optimally' and satisfy $\gamma_1, \gamma_2 > 0$, $\gamma_1 - \gamma_2 < 1$.

For example we might take

$$\gamma_1 = F(x_r, u_s)/D \;,\; \gamma_2 = |F(x_r, u_t)|/D \,,$$

where $D = F(x_r, u_s) + |F(x_r, u_t)|$.

In both cases $F(x_r, y_r) > 0$. We would then take

$x_{r+1} = (1-\alpha_r)x_r + \alpha_r y_r = x_r + \alpha_r m_r$, $m_r = y_r - x_r$, with $\alpha_r > 0$ chosen by one of the methods already considered either $\alpha_r = \alpha_r^*(y_r)$ or $\alpha_r = \beta_r \bar{\alpha}_r(y_r)$.

Note that selection 5.4.5 simply switches weight α_r from u_t to u_s. This is a technique used to compute exact designs.

The above suggestions as well as further improvements on vertex direction algorithms suggested by Atwood (1973), St. John and Draper (1975) and in fact the basic vertex direction algorithm, will be seen to be special cases of iterations to be considered in chapter 7.

§5.5 On Convergence Of Vertex Direction Algorithms

We report now on the literature on convergence of vertex direction algorithms. Many authors have contributed but all in the design context.

Wynn (1970) and Fedorov (1972) prove convergence respectively of $V\{v^{(1)}, 1/(k+r+1)\}$ and of $V\{v^{(1)}, \alpha_r^*(v^{(1)})\}$ for the case of D-optimality. Wynn (1972) proves convergence of the former technique in the case of D_s-optimality. Again for D-optimality Atwood (1973) proves convergence of $V\{v^{(7)}, \alpha_r^*(v^{(7)})\}$ while Fedorov (1972) proves the same of $V\{v^{(1)}, \beta\tilde{\alpha}_r(v^{(1)})\}$ in the case of L-optimality, where $0 < \beta < 1$ and $\tilde{\alpha}_r(v_j)$ is an upper limit which he derives for the range of values of α such that $\psi\{M_r + \alpha(v_j v_j' - M_r)\} \geq \psi(M_r)$. Tsay (1977) suggests that convergence also holds for a modification to $\tilde{\alpha}_r(v^{(1)})$.

These results are rather particular having a specific type of criterion in mind and also, with the exception of Wynn's technique, the iterations concerned are monotonic in which case proof of convergence will be easier.

There is though an extensive literature on convergence of arbitrary steplength procedures in the case of a general design criterion. This however is only for the case of $v^{(1)}$. It was for $v^{(1)}$ that the notion of taking $\alpha_r = \beta\bar{\alpha}_r(v^{(1)})$, where $0 < \beta_r < 1$ and $\beta_0, \beta_1, \beta_2 \cdots$ is a predetermined sequence, was originally conceived. Then $\bar{\alpha}_r(v^{(1)}) = 1$, $\alpha_r = \beta_r$.

Necessary conditions for convergence in this case appear to be that (i) $\beta_r \rightarrow 0$ as $r \rightarrow \infty$ and (ii) $\sum_{r=0}^{\infty} \beta_r = \infty$, the latter condition preventing convergence before reaching the optimum. They are conditions quoted in proofs of convergence by Fedorov (1972), Fedorov and Malyutov (1972), Tsay (1976b), Wu and Wynn (1978). Other contributors include Pazman (1974a) and Silvey (1974).

The latter's result is in fact not a demonstration of convergence. Instead he shows that if $V\{v^{(1)}, \beta_r\bar{\alpha}_r(v^{(1)})\}$ is used with β_r satisfying (i) and (ii) above, then it is the case that

$p^{(r)} \doteq p^*$ for some finite r. The following is a generalisation of his theorem applicable not just to vertex direction algorithms and not just to problem (P2).

Theorem 5.5.1

Suppose that a function $\psi(\cdot)$ is bounded and concave on a bounded convex set S and that, with a view to maximising $\psi(\cdot)$ on S, a sequence of iterates is defined by the relationship

$$x_{r+1} = [1-\alpha_r(y_r)]x_r + [\alpha_r(y_r)]y_r .$$

where $y_r \in S$, $F(x_r, y_r) > 0$, $\alpha_r(y_r) = \beta_r \bar{\alpha}_r(y_r)$

for β_r a given number satisfying $0 < \beta_r < 1$.

Let $\mathcal{F}_r^+ = \{y \in S : F(x_r, y) > 0\}$.

If (i) $\beta_r \longrightarrow 0$ as $r \to \infty$ and (ii) $\Sigma \beta_r = \infty$ then for some n, $\mathcal{F}_n^+ = \phi$.

Proof Appealing to a Taylor series expansion

$$\psi(x_{r+1}) - \psi(x_r) = \alpha_r(y_r)F(x_r, y_r) + z_r$$

where $z_r = o(\alpha_r(y_r))$.

Now $\alpha_r(y_r)F(x_r, y_r) = \beta_r F(x_r, x_r + \tilde{m}_r)$ where $\tilde{m}_r = \bar{\alpha}_r(y_r)[y_r - x_r]$, and so is such that $\bar{\alpha}_r(\tilde{m}_r) = 1$ while if $F(x_r, y_r) > 0$ then $F(x_r, x_r + \tilde{m}_r) > 0$.

$$\text{Thus} \quad \psi(x_{r+1}) - \psi(x_r) = \beta_r F(x_r, x_r + \tilde{m}_r) + z_r ,$$

and now since, by boundedness of S, $\bar{\alpha}_r(y_r)$ must be finite we can claim that $z_r = o(\beta_r)$ and conclude that $z_r/\beta_r \longrightarrow 0$ as $r \longrightarrow \infty$.

In consequence

$$\psi(x_{r+1}) - \psi(x_r) = \sum_{s=0}^{r} \beta_s F(x_s, x_s + \tilde{m}_s) + \sum_{s=0}^{r} z_s$$

Let $\tilde{\mathcal{F}}_r^+ = \{\tilde{m} : x_r + \tilde{m} \in S, \bar{\alpha}_r(\tilde{m}) = 1, F(x_r, x_r + \tilde{m}) > 0\}$.

and let $\tilde{m}_r^*$ solve $\max_{\tilde{m} \in \tilde{\mathcal{F}}_r^+} \{F(x_r, x_r + \tilde{m})\}$

Suppose that there exists δ such that for every r, $F(x_r, x_r + \tilde{m}_r^*) > \delta$.

Then there must exist q, $0 < q < 1$, such that for each selected y_r satisfying $F(x_r, y_r) > 0$, $F(x_r, x_r + \tilde{m}_r) > q F(x_r, x_r + \tilde{m}_r^*) > q\delta$.

Hence

$$\psi(x_{r+1}) - \psi(x_0) > q\delta \sum_{s=0}^{r} \beta_r + \sum_{s=0}^{r} z_s$$
$$\longrightarrow \infty \quad \text{as} \quad r \longrightarrow \infty$$

since $\sum_{s=0}^{r} z_s / \sum_{s=0}^{r} \beta_s \longrightarrow 0$ as $r \to \infty$, because $\sum \beta_s$ diverges and $z_s/\beta_s \to 0$.

Therefore if $\psi(x_0)$ is finite which we can ensure then $\psi(x_r) \to \infty$ as $r \to \infty$, which is impossible since $\psi(x_r)$ is bounded above by the finite maximum $\psi(x^*)$.

Hence there cannot exist a δ satisfying the above and the result follows since any y_r^* solving $\max_{y \in \mathcal{F}_r^+} \{F(x_r, y)\}$ must be such that $y_r = x_r + \tilde{m}_r^*$ for some $\tilde{m}_r^*$. □

The theorem says that the necessary conditions for optimality will be attained by an iteration of the above type. Clearly it is a fairly general result. However it does not guarantee that any specific algorithmic rule will enjoy the result for it assumes that at each x_r we will be prepared to move towards one of the current y's satisfying $F(x_r, y) > 0$. The rule used to select y_r may preclude all the values in $\mathcal{F}_r^+$, if that were nonempty. We will see this to be so in the case of algorithms for problem (P2) at a point $x(p^{(r)})$, where $p^{(r)}$ is optimal for its support. We have noted in the design context that we could not take a forward vertex direction iteration in the case of a criterion enjoying support differentiability at a $p^{(r)}$ which is optimal for $L\{\mathrm{Sup}(p^{(r)})\} \cap \mathcal{U}$. No vertex would then belong to $\mathcal{F}_r^+$

Of course the theorem indicates the remedy in such instances, namely to adopt some other rule (temporarily) that will take $y_r = y$ for some $y \in \mathcal{F}_r^+$. For instance when it is possible we might take a forward vertex direction iteration. At a differentiable $p^{(r)}$ some vertices u_j will always belong to $\mathcal{F}_r^+$ with $S = \mathcal{P}(\mathcal{U})$.

§5.6 The Initial Support

We consider now the choice of $p^{(0)}$.

In section 4.3.1 we discussed starting values and indicated that, in the case of the algorithms of this chapter, we would wish to take $Sup(p^{(0)})$ to be a small subset of a large (discretising) $\mathcal{V}$, with the proviso that $x(p^{(0)})$ or $M(p^{(0)})$ be points of differentiability. Also $p^{(0)}$ should allocate uniform weight to its support points.

We have in mind forward vertex direction algorithms in particular. Certainly there is no doubt about the feasibility of such algorithms proceeding from such an initial design. By their very nature these algorithms can augment a current $Sup(p^{(r)})$ by one vertex. Of course we would require to take $Sup(p^{(0)}) = \mathcal{V}$ if reverse vertex direction iterations only were to be considered.

However why the desirability of such an initial approximation? Primarily because $Sup(p^*)$ will almost certainly be a small subset of a large $\mathcal{V}$. Also it is the case that convergence of forward vertex direction algorithms is slow, particularly when arbitrary steplengths are taken; and moreover it is likely to be slower the more non-optimal-support points there are in $Sup(p^{(0)})$. What might have seemed a natural choice, namely to take $Sup(p^{(0)})$ to be $\mathcal{V}$, is an initial support which contains all non-optimal-support points. It is also one for which the calculation of the design matrix $M(p^{(0)})$ and subsequent $M(p^{(r)})$ could be time consuming and unnecessarily so if $Sup(p^*)$ is a small subset of $\mathcal{V}$.

An ideal would then seem to be that $Sup(p^{(0)})$ should contain as few points as possible. In the design context the constraint of differentiability on $M(p^{(0)})$ requires that $Sup(p^{(0)})$ contain k linearly independent vertices. Thus we take $Sup(p^{(0)})$ to be k such points. The empirical results at the end of this section suggest this to be an open minded choice.

We note that a natural choice of a k-point $Sup(p^{(0)})$ in the case of a regression model, when the design space $\mathcal{V}$ is a discretisation of the image $f(\mathcal{X})$ of a continuous space $\mathcal{X}$, would be derived from a uniform grid which identifies k points of $\mathcal{X}$. The grid should include boundary points of $\mathcal{X}$. If the latter is a finite interval on the real line the grid consists of k equally spaced points including the end points of $\mathcal{X}$.

If $\mathrm{Sup}(p^{(0)})$ does contain k points then a $p^{(0)}$ which distributes weight uniformly to its support points assigns weight $1/k$ to each of them. This clarifies Wynn's choice of $\beta_r = 1/(k+r+1)$, in the case of $v^{(1)}$. Then $p^{(r+1)} = (1-\beta_r)p^{(r)} + \beta_r v^{(1)}$ for $r = 0,1,2,\ldots$ In particular $p^{(1)} = (1-\beta_0)p^{(0)} + \beta_0 v^{(1)}$ with $\beta_0 = 1/(k+1)$. The weights under $p^{(r)}$ are then the relative frequencies with which a vertex is selected as the forward vertex direction, with inclusion in $\mathrm{Sup}(p^{(0)})$ being regarded as such a selection.

Of course the choice of such a $p^{(0)}$ does not mean that convergence will be good. $\mathrm{Sup}(p^{(0)})$ will still inevitably contain non-optimal-support points, while for early iterations $\mathrm{Sup}(p^{(r)})$ will inevitably contain others. Convergence will be retarded because the effect of weight allocated to a non-optimal-support point will linger on in view of the fact that, at each iteration of a forward vertex direction algorithm, there is only a normalising proportionate decrease in the weights of all or, of all but one of, the current support vertices. Reverse vertex direction steps, in making a decisive reduction in weight at a vertex, can therefore speed up convergence dramatically.

However there are some favourable results to report on the initial behaviour of Wynn's original algorithm $V\{v^{(1)}, 1/(k+r+1)\}$, when calculating D-optimal designs in a number of examples.

Consider first the discretised trigonometric regression design space

$$\mathcal{V} = \{ v_x = (x, x^2, \sin(2\pi x), \cos(2\pi x))' : x \in \mathcal{X}_d \}$$
$$\mathcal{X}_d = \{ 0, .01, .02, \qquad , 1\}.$$

In the case of D-optimality

$$\mathrm{Sup}(p^*) = \{ v_{(x)} : x = .08, .09, .38, .73, .74, 1 \}.$$

these points having respectively the optimal weights .22, .03, .25, .17, .08, .25.

Wynn's algorithm was started from a series of 112 initial designs. In each case $\mathrm{Sup}(p^{(0)})$ comprised four linearly independent vertices with $p^{(0)}$ assigning weight 1/4 to these. These initial supports were those corresponding to the following subsets of $\mathcal{X}_{\mathcal{L}}$; $\{0,.33,.67,1\}$, $\{.97,.98,.99,.1\}$; each of the 15 subsets of size 4 of the set $\{0,.2,.4,.6,.8,1\}$; each of the 5 subsets of size 4 of the set $\{x,x+.2,x+.4,x+.6,x+.8\}$ for each of the 19 values $x = .01,.02,\ldots,.18,.19$.

At each iteration the algorithm selects a vertex towards which to move. In each case the algorithm was run until four points were selected at least ten times, presence in $\mathrm{Sup}(p^{(0)})$ being regarded as one selection, and in each case the index Z of the iteration, from which the algorithm selects only members of $\mathrm{Sup}(p^*)$ or their immediate neighbours, was observed. Here Z is the index of the iteration from which the algorithm selected only vertices in the set

$\{v_{(x)} : x = .08,.09,.38,.73,.74,1,.07,.10,.37,.39,.72,.75,.99,1\}$.

The frequencies f(Z) of the various values which Z achieved over the 112 examples is as follows:

Z	:	1	3	4	5	6	7	8	9	10	11	12	14
f(z)	:	1	5	1	20	11	26	9	12	15	10	1	1

The unique instance of Z = 1 occurs in the case of what might have seemed the unlikely set $\{0,.2,.4,1\}$. The single occurrences of Z = 12, Z = 14 occur respectively with the sets $\{.37,.57,.77,.97\}$, $\{.08,.48,.68,.88\}$ both of which seem fairly moderate. In fact both these sets would have yielded Z = 9 had the algorithm not selected $v_{(x)}$, for x = .06 at the 11th iteration in the first case and for x = .36 at the 13th iteration in the other case , and note that both these vectors are immediate neighbours but one to members of $\mathrm{Sup}(p^*)$.

The set $\{0,.33,.67,1\}$ yielded Z = 5 as did the set $\{0,.2,.8,1\}$ while the set $\{.97,.98,.99,1\}$ yielded Z = 8 which is particularly respectable considering that this initial support is somewhat extreme.

Typical features throughout the different runs were the following:

(i) the vertices $v_{(x)}$ for $x = .08, .38, .73, 1$ were always the first four to be selected at least 10 times the number of iterations needed to achieve this ranging from about 50 to 80.

(ii) often the optimal support point corresponding to $x = .09$ was not selected while that for $x = .74$ was selected fairly infrequently.

(iii) of the immediate neighbours $v_{(x)}$, only $x = .37$ persists in being selected in late iterations with $x = .07$ being commonly selected in very early iterations; the other immediate neighbours were rarely selected at all.

Clearly the evidence is that, whatever, $\text{Sup}(p^{(0)})$ the algorithm takes only a few iterations to identify the members of $\text{Sup}(p^*)$ and their immediate neighbours or more loosely to identify clusters containing the members of $\text{Sup}(p^*)$. On average $Z = 7.4$ which is about $2k$ iterations.

Similar results occur in other examples; for instance, in the case of the finer discretised trigonometric design space

$$\mathcal{V} = \{v_{(x)} = (x, x^2, \sin 2\pi x, \cos 2\pi x)' : x \in \mathcal{X}_d\}$$

$$\mathcal{X}_d = \{0, .001, .002, \ldots, .998, .999, 1\}$$

Now $\text{Sup}(p^*) = \{v_{(x)} : x = .082, .083, .381, .734, .735, 1\}$.

One would have thought that the algorithm would take longer to identify clusters as above in this example, but with

$$\text{Sup}(p^{(0)}) = \{v_{(x)} : x = 0, .334, .667, 1\}, \quad Z = 11.$$

Finally consider the discretised polynomial regression design space

$$\mathcal{V} = \{v_{(x)} = (1, x, x^2, \ldots, x^{k-1})' : x \in \mathcal{X}_d\}$$

$$\mathcal{X}_d = \{-1, -.99, \ldots, .99, 1\}$$

With $\text{Sup}(p^{(0)})$ taken to be the k vertices $v_{(x)}$ which correspond to the set of k equally spaced values of x in $\mathcal{X}_d$ which includes $\{-1,1\}$, the index Z achieves the values 7,8,6,9,6,2 respectively in the cases of $k = 4,5,6,7,8,9$.

These results suggest three ideas.

(i) One might simply use a forward vertex direction algorithm initially in order to identify $Sup(p^*)$ at least to within nearest neighbours. This is an idea to which we will return at the end of chapter 6.

(ii) If one is to persist in the use of a forward vertex direction algorithm in order to identify p^*, then convergence might be marginally improved by the following. Run the algorithm for so long and then put to zero the weights of those vertices not recently selected, while scaling up the weights of the other vertices.

(iii) This third notion returns to the choice of $Sup(p^{(0)})$. The above empirical results seem to imply that if weight is assigned, in an initial support, to a small number of non-optimal-support vertices, the effect is much the same whatever vertices are chosen. We have the design context in mind, partly with a view to avoiding the inversion of an initial design matrix $M(p^{(0)})$.

Imagine that the design space $\mathcal{V} = \{v_1,\ldots,v_J\}$ is augmented to $\mathcal{W} = \{v_1,\ldots,v_J, w_1,\ldots,w_k\}$ where $w_i = ce_i$, e_i being the i^{th} unit vector. Let p^* be the optimal design on $\mathcal{V}$. Then this will be the optimal design on $\mathcal{W}$ if $F\{M(p^*), w_j w_j'\} \leq 0$. If p^* is differentiable this implies that

$$c^2 G\{M(p^*), e_j e_j'\} \leq G\{M(p^*), M(p^*)\}.$$

Clearly this will be true if c^2 is small.

Suppose that interest is in all the parameters so that $Sup(p^*)$ must contain at least k of the v_j's. Our suggestion is to choose c such that c^2 is small and to take $Sup(p^{(0)})$ to be $\{w_1,\ldots,w_k\}$. Thus we have

$$M^{-1}(p^{(0)}) = (k/c^2)I.$$

Having started a vertex direction algorithm from this initial approximation one would restart as above. This would be at a point by which time at least k of the v_j's had been selected. The weights at the vertices w_j would be put to zero.

Better still one could employ the following k bi-vertex direction iterations initially. At iteration $r = 1,\ldots,k$, switch the weight $1/k$ at some $w_t \in \mathrm{Sup}(p^{(r)})$ to some $v_s \notin \mathrm{Sup}(p^{(r)})$, w_t, v_s being chosen to satisfy restricted versions of one of the criteria defining $u^{(1)},\ldots,u^{(16)}$. For example take v_s to be the vertex solving $\max_{v_j \notin \mathrm{Sup}(p^{(r)})} \left\{F\{M(p^{(r)}),\ v_j v_j'\}\right\}$ and take w_t to be the vertex

solving $\min_{w_i \in \mathrm{Sup}(p^{(r)})} \left\{F\{M(p^{(r)}),\ w_i w_i'\}\right\}$.

Note that if $M(p^{(r)})$ is nonsingular then, in the latter case, $w_t = ce_t$ where e_t solves $\min_{ce_i \in \mathrm{Sup}(p^{(r)})} \left\{G\{M(p^{(r)}),\ e_i e_i'\}\right\}$.

Vuchkov (1977) has proposed the following suggestion for D-optimality. He takes $c = 1$ and at the first k iterations he chooses v_s, e_t to solve

$$\max_{\substack{v_j \notin \mathrm{Sup}(p^{(r)}) \\ e_i \in \mathrm{Sup}(p^{(r)})}} \left\{ \det\left[M(p^{(r)}) + (1/k)v_j v_j' - (1/k)e_i e_i'\right]\right\},$$

and switches the weight $1/k$ at e_t to v_s. Since lemma 5.2.1 yields a simple expansion of this function it is easy to find v_s, e_t and, since he is maximising the determinant at each stage he must, after the k steps, have transferred the k weights $(1/k)$ to k linearly independent v_j's, a design for which the design matrix is nonsingular. His idea however would not be so easy to apply in cases where evaluation of the criterion $\psi(\cdot)$ was not simple. One based on directional derivatives would be easier to apply, but may require appropriate restrictions to guarantee producing a nonsingular design matrix after k switchings of weights.

This completes our discussion of vertex direction algorithms. We close by restating that on empirical grounds they seem to have the useful property of being able to quickly identify clusters containing the members of $\mathrm{Sup}(p^*)$.

CHAPTER 6

CONSTRAINED STEEPEST ASCENT AND NEWTON RAPHSON TYPE ALGORITHMS

§6.1 Further Matrix Results

Before we turn to consider the algorithms to which this chapter is devoted, we consider first some matrix results which have relevance to all of the algorithms that we have yet to study. This is so because all of them, in contrast to vertex direction algorithms, make significant changes at each iteration to all of the weights of a current support. In the design context this would seem to dictate a full inversion of the design matrix $M(p)$, if that is nonsingular, at each iteration, the benefits enjoyed by vertex or bi-vertex direction algorithms, as a result of lemma 5.2.1, being unavailable. However the following results, further consequences of that lemma, suggest that comparable benefits can be available to what we might refer to as multi-vertex direction algorithms.

Lemma 6.1.1

Suppose that the design matrix $M(p) = \sum_{j=1}^{J} p_j v_j v_j'$, where $J \geq k$, and also the matrix $V_k = [v_1 v_2 \ldots\ldots v_k]$ are nonsingular, and that $p_j > 0$, $j = 1,\ldots,J$. Then

6.1.1 $$M^{-1}(p) = \sum_{j=1}^{J} t_j w_j w_j' ,$$

where the t_j are scalars and the w_j are $k\times1$ vectors such that

6.1.2 $$\begin{cases} (V_k')^{-1} = [w_1\, w_2 \cdots\cdots w_k] & \\ t_j = p_j^{-1} \quad \text{for } j = 1,2,\cdots,k & \\ \left.\begin{aligned} t_j &= -\Big[p_j^{-1} + \sum_{i=1}^{j-1} t_i (w_i' v_j)^2\Big]^{-1} \\ w_j &= \sum_{i=1}^{j-1} t_i (w_i' v_j) w_i \end{aligned}\right\} & \text{for } j = (k+1), \cdots\cdots, J. \end{cases}$$

Proof Let $M_n = \sum_{j=1}^{n} p_j v_j v_j'$ for $n = 1,2,\ldots,J$

Hence

$$M_n = M_{n-1} + p_n v_n v_n' \quad \text{for } n = 2,\ldots,J$$

$$M(p) = M_J = M_{J-1} + p_J v_J v_J'$$

Appealing to lemma 5.2.1 in the case n > k we have

$$M_n^{-1} = M_{(n-1)}^{-1} - p_n M_{(n-1)}^{-1} v_n v_n' M_{(n-1)}^{-1} / \{1 + p_n v_n' M_{(n-1)}^{-1} v_n\}$$

Assuming that $M_{(n-1)}^{-1} = \sum_{j=1}^{n-1} t_j w_j w_j'$ for some t_j, w_j we then have

$$M_n^{-1} = \sum_{j=1}^{n} t_j w_j w_j'$$

where

$$w_n = M_{(n-1)}^{-1} v_n = \sum_{i=1}^{n-1} t_i (w_i' v_n) w_i$$

$$t_n = -p_n / \{1 + p_n v_n' M_{(n-1)}^{-1} v_n\} = -\left[p_n^{-1} + \sum_{i=1}^{n-1} t_i (w_i' v_n)^2\right]^{-1}$$

The lemma is established when we note, from 1.3.9, that

$$M_k = V_k P_k V_k' \quad , \quad P_k = \mathrm{diag}\{p_1, \dots, p_k\}$$

Since P_k, V_k are of order k×k nonsingular, we have

$$M_k^{-1} = (V_k')^{-1} P_k^{-1} V_k^{-1} = \sum_{j=1}^{k} t_j w_j w_j'$$

$$t_j = p_j^{-1} \quad , \quad V_k^{-1} = [w_1 \; w_2 \cdots w_k] \qquad \square$$

We make some comments on the lemma.

(i) With hindsight the result may seem a fairly natural sequel to the results derived from lemma 5.2.1 for vertex direction iterations. However it does not appear in the literature although there would seem to be a continuous analogue well hidden in Pazman (1974b). He does not however dwell on consequent computational advantages of his result although these would, of course, be less clear in the continuous context.

In contrast, from the above discrete results, we can derive updating formulae both for $\psi\{M(p)\}$ and for $G\{M(p), v_j v_j'\}$ in the case of D-optimality and A-optimality.

(ii) In view of the sequential nature of formulae 6.1.2, such updating rules will only be advantageous for small J. How small? For J = k+1 they will be most advantageous. Indeed for such a case one could argue that the above result bestows on an algorithm, which, at each iteration, makes disproportionate changes to all of k+1 weights, the same benefits as do the updating formulae of section 5.2 on vertex direction algorithms.

As a rough rule there may be little advantage in utilising lemma 6.1.1 if $J > 2k$. This is suggested by the fact that any $k \times k$ matrix $M = [\underline{m}_1 \underline{m}_2 \ldots \underline{m}_k]$, where $\underline{m}_j$ is a $k \times 1$ vector can be expressed as

$$M = I + \sum_{j=1}^{k} (\underline{m}_j - \underline{e}_j)\underline{e}_j'$$

where $\underline{e}_j$ is the j^{th} unit vector.

Thus nonsingular M can be inverted by k successive appeals to lemma 5.2.1, a method which is recommended by numerical analysts.

Often, as has been said, the size J of the optimum support is small, $J = k$ or $(k+1)$ or in the case of a discretised design space $J = 2k-2$.

(iii) The lemma of course demands that $M(p)$ be nonsingular.

However the following analogous result holds for $M^+(p)$ which assumes that nonsingularity of $M(p)$ is due to the fact that $J < k$.

<u>Lemma 6.1.2</u>

If $M(p) = \sum_{j=1}^{J} p_j v_j v_j'$, where $p_j > 0$ for all j and where $\text{rank}(V_J) = J$, $V_J = [v_1 v_2 \ldots v_J]$, then

6.1.3 $$M^+(p) = \sum_{j=1}^{J} t_j w_j w_j'$$

where

6.1.4 $$t_j = p_j^{-1}$$
$$[w_1 w_2 \quad w_J] = V_J(V_J'V_J)^{-1}$$

<u>Proof</u> From equation 1.3.9

$$M(p) = V_J P_J V_J'$$

where $P_J = \text{diag}\{p_1, p_2, \ldots, p_J\}$, and so rank $(P_J) = J$.

From theorem 6.2.18 of Graybill (1969) , $M^+(p) = (V_J')^+ P_J^+ V_J^+$
$$= (V_J')^+ P_J^{-1} V_J^+$$

and from theorem 6.2.16 of Graybill (1969)

$$W_J = (V_J')^+ = (V_J^+)' = V_J(V_J'V_J)^{-1}.$$

Hence the result. □

(iv) The above two lemmas together make the following result possible.

<u>Lemma 6.1.3</u>

For $J > s$ let $M(p) = \sum_{j=1}^{J} p_j v_j v_j'$ where $p_j > 0$ for all j, and suppose that $\mathcal{N}\{M(p)\} \subseteq \mathcal{N}(A)$, where A is of order $s \times k$. Let

$$B = \begin{cases} A M^{-1}(p) A' , & J \geq k , \ \operatorname{rank}(V_k) = k \\ A M^{+}(p) A' , & J < k , \ \operatorname{rank}(V_J) = J \end{cases}$$

where $V_k = [v_1, \ldots, v_k]$, $V_J = [v_1, \ldots, v_J]$.

Hence B is nonsingular and $B = \sum_{j=1}^{J} t_j g_j g_j'$

where $g_j = Aw_j$, and t_j, w_j are defined by 6.1.2 or 6.1.4 as appropriate.

Then assuming that $G_s = [g_1, \ldots, g_s]$ is nonsingular we have

6.1.5 $$B^{-1} = \sum_{j=1}^{J} f_j h_j h_j'$$

where

6.1.6 $$\begin{cases} [h_1 \cdots h_s] = (G_s')^{-1} & \\ f_j = t_j^{-1} , & j = 1, \ldots, s \\ \left.\begin{array}{l} f_j = -\left[t_j^{-1} + \sum_{i=1}^{j-1} f_i (h_i' g_j)^2\right]^{-1} \\ h_i = \sum_{i=1}^{j-1} f_i (h_i' g_j) h_i \end{array}\right\} & j = (s+1), \ldots, J \end{cases}$$

The proof is a direct appeal to lemma 6.1.1. □

Clearly the relevance of this result is that it can yield updating formulae for D_A-optimality and for L_A-optimality, comparable to those of section 5.2 including updates of $G\{M(p), v_j v_j'\}$ for each j. However again such formulae will not always be computationally the most efficient.

(v) A point, which we have not yet emphasised, is that the main reason why these lemmas can yield more efficient evaluation of relevant terms, is due to the fact that the vectors $\{w_1, \ldots, w_k\}$ or $\{w_1, \ldots, w_J\}$ or $\{h_1, \ldots, h_s\}$ do not depend on p. They can therefore be the same from iteration to iteration, provided corresponding weights are not put to zero; for example in the case of lemma 6.1.1 we would want $p_j^{(r)} > 0$ for $1 \leq j \leq k$. If however such weights are put to zero then the matrix V_k, V_J or G_s would have to be redefined and appropriate inverses or generalised inverses recalculated. This may be unavoidable in the case of lemma 6.1.2, but when it is known that the optimum support must consist of at least k or s linearly independent vertices then, judiciously chosen, the matrix V_k or G_s should remain the same throughout. One such choice of V_k is likely to be the matrix whose columns are those k vertices most frequently selected by a long enough series of forward vertex direction iterations. The matrix G_s could be similarly chosen. This anticipates a proposal to be made at the end of this chapter.

§6.2 Constrained Steepest Ascent Iterations

§6.2.1 The algorithms which we consider in this chapter for problem (P2) are in the main suited to sets $\mathcal{U} = \{u_1,\ldots,u_J\}$ which contain only a few vertices which are not in the support of an optimum. In one instance we really must have $\mathcal{U} = \mathrm{Sup}(p^*)$.

The first class of iteration, which we look at, adopts a fairly natural choice of direction. Consider the maximisation of a function $\psi(\cdot)$ over a (convex) set S, when $\psi(\cdot)$ is concave on that set. Intuitively an optimal direction in which to move from an iterate x_r is in that direction m_r^* in which $\psi(\cdot)$ is increasing most rapidly from x_r, subject to $(x_r + \alpha m_r^*) \in S$ for all small positive α. Thus

$$x_{r+1} = x_r + \alpha_r m_r^* ,$$

where $\alpha_r > 0$ can be chosen according to any one of the rules which we have already considered. It is arguably more natural to choose α_r optimally in this instance.

We call the direction m_r^* a direction of constrained steepest ascent. Clearly the directional derivative tool must be able to identify such a direction. However we have already observed that, since $F(x,y)$ depends not only on the rate at which $\psi(\cdot)$ changes at x in the direction of y, but also on the 'distance' between x and y, then the direction m^* of constrained steepest ascent is not, in the case of convex S, given by $m = y^*-x$, where y^* solves $\max\limits_{y\in S} F(x,y)$. In particular, in the case of problem (P2), $u^{(1)}$ cannot be claimed to provide such a direction. A constrained steepest ascent m^* must maximise a normalised directional derivative, namely $F^A(x,x + m)$, subject to $(x + \alpha m) \in S$ for all small positive α.

Consider the problem of maximising a function $\Phi(\underline{\theta})$ subject to just the linear constraints $C\underline{\theta} = b$. Suppose that $\underline{\theta}_r$ satisfies $C\underline{\theta}_r = b$. Then it is necessary that $C\underline{m}_r = 0$ if it is to be guaranteed that $C\underline{\theta}_{r+1} = b$, $\underline{\theta}_{r+1} = \underline{\theta}_r + \alpha \underline{m}_r$. If further all the components of $\underline{\theta}_r$ are strictly positive, then so also will be the components of $\underline{\theta}_{r+1}$ for small positive α. If though one of the components of $\underline{\theta}_r$ were zero, and if the corresponding component of $\underline{m}_r$ were negative, then the corresponding component of $\underline{\theta}_{r+1}$ would also be negative if $\alpha > 0$.

Hence if

$$\underset{\sim}{\theta}_r \in S = \{\underset{\sim}{\theta} = (\theta_1, \cdots, \theta_J) : \theta_j > 0 , C\underset{\sim}{\theta} = b\} ,$$

then the constraint $C\underset{\sim}{m}_r = 0$ is necessary and sufficient to guarantee that $\underset{\sim}{\theta}_{r+1} \in S$ for small positive α. This however is not the case, if the strict inequality $\theta_j > 0$ is replaced by $\theta_j \geq 0$. The following lemma then restating formula 2.3.6 quoted in (D9) of section 2.3 is guaranteed to state a constrained steepest ascent direction for problem (P3) only in the former case. It views (P3) as a generalisation of (P1) and so we relabel as $\phi(\cdot)$, the objective function $\Phi(\cdot)$, of (P3).

Lemma 6.2.1

Suppose $\phi(\cdot)$ is differentiable at the value θ. Then the direction m* which maximises $F_\phi^A(\theta, \theta + m) = d'm/\sqrt{m'Am}$ subject to $Cm = 0$ can be given by

6.2.1 $$m^* = \pm\{A^-d - A^-C'(CA^-C')^-CA^-d\} ,$$

whichever makes $F_\phi^A(\theta, \theta + m) > 0$

The vector d has components $d_j = \partial\phi/\partial\theta_j$.

Assume the matrix C to be of order t×J with rank(C) = t.

If our only interest were problem (P1) or (P2) we would only need to consider the case C = $\underline{1}'$. However we will wish to consider general C in chapter 10. We establish the lemma by following a proof used by Wu (1976) for the case C = $\underline{1}'$.

Proof Let $f(m) = d'm/\sqrt{m'Am}$.

We require to maximise f(m) subject to Cm = 0.

However the function f(m) is homogeneous of degree zero. To restrict ourselves to a particular solution we need to impose a linear constraint on m. If we take this to be d'm = 1 our problem becomes equivalent to

6.2.2 'minimise m'Am subject to d'm = 1, Cm = 0'

i.e.

6.2.3 'minimise m'Am subject to Bm = g'

where $B = \begin{bmatrix} C \\ d' \end{bmatrix}$ $g = \begin{bmatrix} \underline{0} \\ 1 \end{bmatrix}$, $\underline{0}$ is the t×1 zero vector.

This is a quadratic programming problem with linear constraints.

Recall that A must be such that $x'Ay$ defines an inner product. Typically A will be nonsingular and hence positive definite, but Wu proceeds with the possibility in mind that A might be singular. We shall see this to be the case for some reasonable choices of A.

Using a Lagrangian approach we consider

$$L(m) = m'Am - 2\lambda'(Bm - g).$$

Equating the derivatives with respect to m and λ to zero, we obtain the respective equations

6.2.4 $\quad Am = B'\lambda$

6.2.5 $\quad Bm = g$

Again using theorem 7.3.1 of Graybill (1969) solutions, if they exist, to 6.2.4 are given by

6.2.6 $\quad m = A^{-}(B'\lambda) = (A^{-}A - I)z$ for any z.

However if this is to be such that Bm is uniquely equal to g then we must have $BA^{-}A = B$ or equivalently that

6.2.7 $\quad \mathcal{R}(B') \subseteq \mathcal{R}(A)$,

where $\mathcal{R}(A)$ denotes the range space or column space of A.

It therefore follows that

6.2.8 $\quad (BA^{-}B')\lambda = g$

Once more using Graybill's result

$\lambda = (BA^{-}B')^{-}g + [(BA^{-}B')^{-}(BA^{-}B') - I]y$, for any y, solves 6.2.8.

Substituting in 6.2.6 we obtain

6.2.9 $\quad m = A^{-}B'(BA^{-}B')^{-}g + A^{-}B'[(BA^{-}B')^{-}(BA^{-}B') - I]y + (A^{-}A - I)z.$

Putting y = 0, z = 0 we obtain the particular solution

6.2.10 $\quad m = A^{-}B'(BA^{-}B')^{-}g$

Property 6.2.7 will certainly hold if A is positive definite in which case any solution must take the form 6.2.10 with $A^{-} = A^{-1}$. See theorem 8 of Searle (1971, p. 26).

If 6.2.7 does not hold, Rao and Mitra (1971) proceed as follows:

From 6.2.4

$Am + B'Bm = B'\lambda + B'Bm$

that is,

6.2.11 $\quad (A + B'B)m = B'(\lambda + g) \quad , \quad Bm = g$

Now $\mathcal{R}(B') \subset \mathcal{R}(A + B'B)$ and hence

$$m = (A + B'B)^{-}B'(\lambda + g)$$

$$\therefore \quad g = Bm = B(A + B'B)^{-}B'(\lambda + g)$$

$$\Longrightarrow (\lambda + g) = [B(A + B'B)^{-}B']^{-}g$$

$$\Longrightarrow m = \{(A + B'B)^{-}B'[B(A + B'B)^{-}B']^{-}\}g$$

$$= \{G\}g$$

In fact Rao and Mitra (1971) prove in their theorem 3.1.4 that a necessary and sufficient condition for $m = Gg$ to minimise $m'Am$, where A is positive semidefinite, subject to $Bm = g$ is that

$BGB' = B$ and $(GB)'A = AGB$

Wu (1976) references these results.

Assume that 6.2.10 is a solution. Substituting for B and g we obtain

$$m = A^{-}(C' \; d)\begin{bmatrix} CA^{-}C' & CA^{-}d \\ d'A^{-}C' & d'A^{-}d \end{bmatrix}^{-}\begin{pmatrix} \underset{\sim}{0} \\ 1 \end{pmatrix}$$

$$= A^{-}(C' \; d)\begin{bmatrix} D & \underset{\sim}{u} \\ \underset{\sim}{u}' & b \end{bmatrix}\begin{pmatrix} \underset{\sim}{0} \\ 1 \end{pmatrix}$$

$$= A^{-}(C' \; d)\begin{pmatrix} \underset{\sim}{u} \\ b \end{pmatrix}$$

$$= A^{-}(C'\underset{\sim}{u} + b\underset{\sim}{d})$$

$$= A^{-}b\underset{\sim}{d} + A^{-}C'\underset{\sim}{u}$$

where, using the results of Rhode (1965) that were outlined in the derivation of equation 1.4.7,

6.2.12
$$\begin{cases} b = a^{-} \\ a = d'A^{-}d - d'A^{-}C'(CA^{-}C')^{-}CA^{-}d \\ \underset{\sim}{u} = -b(CA^{-}C')^{-}CA^{-}d \end{cases}$$

Hence

6.2.13 $\quad m = b[A^{-}d - A^{-}C'(CA^{-}C')^{-}CA^{-}d]$

Note that in the above argument we have made use of the fact that b is a scalar.

If $a \neq 0$ then $b = a^{-1}$ and so clearly 6.2.13 defines an m satisfying the constraint $d'm = 1 > 0$.

However suppose that b might be negative and recall that $f(m) = d'm/\sqrt{m'Am}$ is homogeneous of degree zero so that if m maximises $f(m)$ subject to $Cm = 0$, then so also does tm for t a positive scalar. We can therefore drop $|b|$ and treat the solution as

6.2.14 $$m^* = \pm m(A,C,d) \quad , \quad m(A,C,d) = A^-d - A^-C'(CA^-C')^-CA^-d,$$

choosing the sign according as $a > 0$, $a < 0$ to ensure $d'm^* > 0$. □

In the case $C = \underline{1}'$, $$m(A,\underline{1}',d) = A^-d - A^-\underline{1}(\underline{1}'A^-\underline{1})^-\underline{1}'A^-d,$$
$$a = d'A^-d - d'A^-\underline{1}(\underline{1}'A^-\underline{1})^-\underline{1}'A^-d.$$

The term $\underline{1}'A^-\underline{1}$ is now a scalar and unless it is zero $(\underline{1}'A^-\underline{1})^- = (\underline{1}'A^-\underline{1})^{-1}$ so that

6.2.15 $$m(A,\underline{1}',d) = A^-d - A^-\underline{1}\underline{1}'A^-d/(\underline{1}'A^-\underline{1}).$$

It also follows that an alternative solution to 6.2.14 is

6.2.16 $$m^* = \pm \tilde{m}(A,d) \quad , \quad \tilde{m}(A,d) = (\underline{1}'A^-\underline{1})A^-d - (\underline{1}'A^-d)A\underline{1}.$$

Recall that in view of 2.3.5 we can replace d by F.

With problem (P1) in mind we will denote by $S(A,\alpha_r)$ the 'steepest ascent' algorithm which takes a step α_r in the direction $m^{(r)} = \pm m(A,\underline{1}',d^{(r)})$.

§6.2.2

We now make some notes on the result.

Note (i) If all matrices of which generalised inverses have been taken are nonsingular then we have
$$a = d'A^{-1}d - d'A^{-1}C'(CA^{-1}C')^{-1}CA^{-1}d,$$

which is positive as shall be seen in Note (ii).

Hence
$$m^* = m(A,C,d) = A^{-1}d - A^{-1}C'(CA^{-1}C')^{-1}CA^{-1}d.$$

Note (ii) If A is singular there is not a unique A^-. Since A is symmetric it would seem desirable to demand the same of A^- and of $(CA^-C')^-$. Wu elects to use the Moore-Penrose A^+ which has the advantage of being uniquely defined and nonnegative definite. If one also uses $(CA^+C')^+$ then $a > 0$. This follows because for $z = (A^+)^{1/2} d$, $H = C(A^+)^{1/2}$

$$a = z'z - z'H'(HH')^+Hz = z'Dz \;, \quad D = I - H'(HH')^+H.$$

The matrix $H'(HH')^+H$ is idempotent since $(HH')^+HH'(HH')^+ = (HH')^+$. Hence so also is D which is therefore nonnegative definite.

In general, for any nonnegative definite matrix Λ,

$$\tilde{m} = \Lambda d - \Lambda C'(C\Lambda C')^+ C\Lambda d$$

satisfies $d'\tilde{m} > 0$.

Note (iii) The following result of interest for its own sake would seem to illustrate that in general $a > 0$ if 6.2.7 holds.

The result is that if $\mathcal{R}(C') \subset \mathcal{R}(A)$ then $m(A,C,d)$ solves the problem 'maximise $Q(m)$ subject to $Cm = 0$' where $Q(m)$ is the quadratic

$$Q(m) = \phi(\theta) + d'm - (1/2)m'Am \,.$$

If A is nonnegative definite then $Q(m)$ is concave.

The term $\phi(\theta)$ could be any number and d any vector, but the idea of course is that $Q(m)$ should be an approximation to $\phi(\theta + m)$, in which case d would be the gradient vector of ϕ at θ and $A = -H$, H being the hessian matrix at θ. In fact Atwood (1976a, 1980) derives the following result for $A = -H$.

An appropriate lagrangian is

$$L(m,\mu) = \phi(\theta) + d'm - (1/2)m'Am - \mu'Cm$$

Equating derivatives with respect to m and μ to zero we obtain the equations,

$$Am = d - C'\mu \qquad , \qquad Cm = 0 \,.$$

If $\mathcal{R}(C') \subset \mathcal{R}(A)$ then a particular solution is given by the following now familiar argument:

$$
\begin{aligned}
& m = A^{-}(d - C'\mu) \\
\Rightarrow \quad & CA^{-}(d - C'\mu) = 0 \\
\Rightarrow \quad & \mu = (CA^{-}C')^{-}CA^{-}d \\
\Rightarrow \quad & m = A^{-}d - A^{-}C'(CA^{-}C')^{-}CA^{-}d = m(A,C,d).
\end{aligned}
$$

This then generalises Atwood's result and, as he observes, since $Cm = 0$ for $m = 0$ and $Q(m)$ is concave, it must be that $Q(0) = \phi(\theta) \leq Q\{m(A,C,d)\}$. If in turn $\phi(\theta)$ is concave it must be that $F_{\phi}\{\theta, \theta + m(A,C,d)\} = d'm(A,C,d) \geq 0 \Rightarrow a \geq 0$. Note that we require only $\mathcal{R}(C') \subset \mathcal{R}(A)$, which is not quite as demanding as 6.2.7. Atwood also derives a solution when the latter condition does not hold but we will not pursue this.

Note (iv) It can be the case that $a = 0$. In this case b can be any real number. However more importantly it follows that $d'm(A,C,d) = 0$ and hence neither $m(A,C,d)$ nor any multiple of it can satisfy the constraint $d'm = 1$. Similarly with $\tilde{m}(A,d)$.

In the case of $C = \underline{1}'$ this will happen if all components d_j of d are equal, indicating optimality of θ. As a consequence, the constraints $\underline{1}'m = 0$ and $d'm = 1$ are inconsistent. However it is then the case that $\tilde{m}(A,d)$ or $m(A,\underline{1}',d) = 0$, which is what the value of m^* should be at an optimal θ^*. We note that $a = \det\{BA^{-}B'\}/(\underline{1}'A^{-}\underline{1})$ in this case, so that $a = 0$ implies singularity of $BA^{-}B'$.

For general C it will again be the case that, at an optimal θ^* maximising $\phi(\theta)$ subject to $C\theta = b$, m^* should be zero in which case $d'm^* \neq 1$.

However $a = 0$ implying $d'm = 0$ can still occur at non-optimal θ. In particular this can occur in the case $C = \underline{1}'$ if A is singular and $A^{-} = A^{+}$. This simply means that no multiple of $m(A,C,d)$ or $\tilde{m}(A,d)$ will satisfy the constraint $d'm = 1$. This may be as above because $Cm = 0$ and $d'm = 1$ are a priori inconsistent, or it may be that 6.2.7 is not satisfied, or possibly θ is not a point of full differentiability, although $f(m)$ would then not be $F^{A}(\theta, \theta + m)$.

Whatever the reason 6.2.14 cannot define a direction of constrained steepest ascent at non optimal Θ. If Θ is differentiable we have only identified a constrained stationary value with respect to the line running through Θ in the direction $\pm m(A,C,d)$. If $\psi(\cdot)$ is concave this stationary value can only be such that a step in either direction cannot lead to an increase in $\psi(\cdot)$. However this should rarely occur and, rather than attempt to evaluate the correct formula for m^*, Wu (1976) contemplates, in the case $C = \underline{1}'$, $A^- = A^+$, a step in the direction of $+m(A,C,d)$, a device presumably as simple as any other for extricating oneself from such a position.

However at such a Θ_r we are in a position anticipated in the discussion to lemma 5.5.1. A better alternative would be to adopt some other simple rule for choosing m_r, a vertex direction iteration if possible or one of the rules of chapter 7.

§6.2.3 We have not considered the choice of the matrix A. Wu (1976, 1978a) recommends the following choices in the case of problem (P2), $C = \underline{1}'$.

6.2.17 $$\begin{cases} A_1 = U'U\ , & U = [u_1\ u_2 \cdots u_J] \\ A_2 = I \\ A_3 = \text{diag}\{u_1'u_1, \cdots\cdots, u_J'u_J\} \end{cases}$$

The choice A_1 is such that $m'A_1m$ is the length of the vector $\sum_{j=1}^{J} m_j u_j$. Clearly it is a choice which takes full account of the nature of the feasible region $\mathcal{X}(\mathcal{U})$ of (P2). Wu chooses to regard this as yielding the steepest ascent direction.

The case $A_2 = I$ describes as a gradient projection method for the resultant m^* is

$$m^* = \underline{d} - \bar{d}\underline{1}$$

where $\bar{d} = (\Sigma d_j)/J$. Clearly m^* is the projection of the gradient vector $\underline{d}$ on the space $\underline{1}'m = 0$. It is therefore a constrained steepest ascent direction obtained by a projection of the unconstrained steepest ascent direction vector $\underline{d}$. In general $m(A,\underline{1}',d)$ could be regarded as a similar type of projection.

The choice A_2 means that we normalise $F_\phi(p,q) = F_\psi\{x(p), x(q)\}$ by the length of $m = (p-q)$. It takes note only of the fact that we have a function $\phi(p)$ for which we wish to solve (P1). The resultant m^* in no way depends on the vertices $\{u_1, \ldots, u_J\}$. It ignores the fact that $\phi(p) = \psi\{x(p)\}$ or as Wu puts it, it ignores the geometry of $\mathcal{P}(\mathcal{U})$.

The choice A_3 is a compromise between A_1 and A_2. Wu describes the resultant m^* as a diagonalised constrained steepest ascent. He also refers to it as producing a normalised gradient projection direction for now m^* has components

$$m_j^* = \left(d_j / u_j' u_j\right) - \left(1/u_j' u_j\right)\left[\left\{\sum d_i / (u_i' u_i)\right\} / \left\{\sum (u_i' u_i)^{-1}\right\}\right]$$

which is a projection onto the space $\underline{1}'m = 0$ of the vector $\{d_1/(u_1'u_1), \ldots, d_J/(u_J'u_J)\}'$, which in turn can be viewed as a normalisation of the gradient vector d.

It is not entirely clear what the above choices of A suggest for the design context when $\mathcal{P}(\mathcal{U}) = \mathcal{M}$ and $u_j = v_j v_j'$, except that $A = A_2 = I$ would again be a possible choice. We consider the following matrices.

$$6.2.18 \quad \left\{ \begin{array}{lll} A_4 = [a_{ij}] & , & a_{ij} = \left\{(v_i' v_j)^2/2 + (v_i^2)'(v_j^2)/2\right\} \\ A_5 = [\tilde{a}_{ij}] & , & \tilde{a}_{ij} = (v_i' v_j)^2 \\ A_6 = \mathrm{diag}\{(v_1'v_1)^2, \ldots\ldots, (v_J'v_J)^2\} & & \end{array} \right.$$

where v_j^2 is a vector whose components are the squares of those of v_j.

The problem is that we can view the matrices $v_j v_j'$ in several guises. In view of their symmetry we could regard them as vectors or points in $k(k+1)/2$ dimensions, or ignoring the symmetry, as points in k^2 dimensions. Finally we could simply regard them as functions of the vectors v_j, these being of length k.

If we adopt the first of these viewpoints then A_4 would seem to be the counterpart of A_1, for $m'A_4m$ is the length of the vector whose components form the upper (or lower) triangular part of the matrix $\sum m_j v_j v_j'$. Similarly $m'A_5m$ is the length of the $k^2 \times 1$ vector whose components are those of the latter matrix. Wu (1976) would appear to be in favour of A_5, for, as a normalised gradient projection choice, he favours A_6 which is a diagonalised version of A_5, as is A_3 of A_1.

The above are choices of a constant matrix A. It might seem natural that one would select the same A_i at each iteration. However this must be qualified as outlined in the next subsection 6.2.4, and, anyway, it is conceivable that it would be possible to choose a different A from one iteration to the next in an advantageous if not an optimal fashion. One may select a different A, for instance, as a means of moving from the sort of position considered in Note (iv) of subsection 6.2.2, while it would be achieved most naturally by letting A depend on p; that is, $A = A(p)$. A number of algorithms conceived by alternative approaches do in fact select directions m_r such that $m_r = m\{A(p^{(r)}),\ \underline{1}',\ d\}$. The matrices $A(p)$ include

$$6.2.19 \quad \begin{cases} A_7(p) = \left[\operatorname{diag}(p_1,\ldots,p_J)\right]^{-1} \\ A_8(p) = -H(p) \\ A_9(p) = -\operatorname{diag}\{h_{11}(p),\ldots,h_{JJ}(p)\}, \end{cases}$$

where $H(p) = \{h_{ij}(p)\}$ is the hessian matrix of $\phi(p)$.

The case $A_7(p)$ corresponds to an algorithm which will appear in chapter 7, while the choice $A_8(p)$ has already been heralded in Note (iii) of subsection 6.2.2. The direction $m\{-H(p),\ \underline{1}',\ d\}$ is the direction which maximises the quadratic approximation to $\phi(\cdot)$ at p in the direction of m, a result which Atwood (1976a, 1980) derived. The choice of $A_8(p)$ would be a conventional choice of varying A and Wu calls the resultant technique a quasi-Newton method.

As a simplification of $A_8(p)$, Atwood (1980) suggests $A_9(p)$. Of interest is that Atwood (1980) also considers the case of a design criterion at a singular matrix $M(p)$ when full differentiability might fail. He contemplates the possibility that the hessian matrix might also be singular then, thereby implying a singular A.

Also of interest is that in the case of H nonsingular $m(-H^{-1},\underline{1}',d)$ and more generally $m(-H^{-1},C,d)$ are directions of iteration suggested by full Newton Raphson iterations for maximising

$$L(p,\lambda) = \phi(p) + \lambda'(Cp-b)$$

which would be the lagrangian for maximising $\phi(p)$ subject to $Cp = b$. We have.

$$\partial L/\partial p = d + C'\lambda\ , \quad \partial L/\partial\lambda = Cp - b\ ,$$

and the second derivative matrix of L is given by

$$D_2(p) = \begin{bmatrix} H(p) & C' \\ C & 0 \end{bmatrix}$$

If H is nonsingular and $CH^{-1}C'$ is nonsingular then

$$D_2^{-1}(p) = \begin{bmatrix} Z(p) & H^{-1}C'(CH^{-1}C')^{-1} \\ (CH^{-1}C')^{-1}CH^{-1} & -(CH^{-1}C')^{-1} \end{bmatrix}$$

where

$$Z(p) = H^{-1} - H^{-1}C'(CH^{-1}C')^{-1}CH^{-1}\ , \quad H = H(p)\,.$$

Full Newton Raphson iterations for p and λ are

$$\begin{bmatrix} p^{(r+1)} \\ \lambda^{(r+1)} \end{bmatrix} = \begin{bmatrix} p^{(r)} \\ \lambda^{(r)} \end{bmatrix} - D_2^{-1}(p^{(r)}) \begin{bmatrix} d^{(r)} + C'\lambda^{(r)} \\ Cp^{(r)} - b \end{bmatrix}$$

If $Cp^{(r)} = b$ these equations express $p^{(r+1)}$ as a function of $p^{(r)}$ only; namely they imply

$$p^{(r+1)} = p^{(r)} - Z(p^{(r)})[d^{(r)} + C'\lambda^{(r)}]$$

Taking now $H = H(p^{(r)})$ we have

$$\begin{aligned} p^{(r+1)} &= p^{(r)} - Z(p^{(r)})d^{(r)} + H^{-1}C'\lambda^{(r)} - H^{-1}C'(CH^{-1}C')^{-1}(CH^{-1}C')\lambda^{(r)} \\ &= p^{(r)} - Z(p^{(r)})d^{(r)} \\ &= p^{(r)} + m\{-H(p^{(r)}), C, d^{(r)}\} \end{aligned}$$

What we have in fact is the iterative rule for iterates under $S\{-H(p^{(r)}), 1\}$.

Titterington in a private communication derived this result for D-optimality, that is, $\phi(p) = -\log_e \det\{(M(p))\}$, $C = \underset{\sim}{1}'$. He also quotes in Titterington (1977) the examples of A listed above. He also suggests $A_9(p)$ as a simplification of $A_8(p)$. As alternative simplification in the case of D-optimality, he also considers

$$A_{10}(p) = k^2 I_J$$

$$A_{11}(p) = k_* \operatorname{diag}\{d_1(p), \ldots, d_J(p)\},$$

where $d_1(p), \ldots, d_J(p)$ are the components of d, $d_j(p) = v_j' M^{-1}(p) v_j$. The motivation for these two particular choices is that $h_{jj}(p) = -d_j^2(p)$, while at D-optimum p^*, $d_j(p^*) = k$ if $p_j^* > 0$. We know then the diagonal of the Hessian at the optimum. It is $A_{10}(p)$, and should be an improvement on $A_9(p)$. Proceeding from the fact that for D-optimality $G_\psi\{M(p), M(p)\} = k$, possible generalisations of $A_{10}(p)$, $A_{11}(p)$ to other criteria might be

$$A_{12}(p) = G^2 I_J$$

$$A_{13}(p) = G \times \text{diag}\{d_1(p), \ldots, d_J(p)\}$$

where $G = G_\psi\{M(p), M(p)\}$. If $\phi(p) = \psi\{M(p)\}$ is homogeneous of degree $(-t)$, then $G = -t\phi(p)$.

It is to be noted however that, if $A = cD$ for scalar c, then iterates under $S\{A,1\}$ are iterates under $S\{D,1/c\}$. Hence for $A = A_{10}(p)$, $A_{12}(p)$ iterates under $S\{A,1\}$ are respectively iterates under $S\{I,\alpha\}$, $\alpha = (1/k)^2$, $(1/G)^2$, while for $A = A_{11}(p)$, $A_{13}(p)$ iterates under $S\{A,1\}$ are respectively iterates under $S\{D,\alpha\}$, $\alpha = 1/k$, $1/G$, where $D = \text{diag}\{d_1(p), \ldots, d_J(p)\}$. Clearly if optimal steps are to be taken such a constant is superfluous.

Of course allowing $A(p)$ to depend on p means the calculation of $A^{-1}(p)$ or $A^+(p)$ at every iteration, and $A(p)$ is a matrix of order $J \times J$. This was the reason for considering the diagonalised versions of $H(p)$ above. An alternative moderating choice in the case of $A(p)$ nonsimple, would be to take A to be $A(p^{(0)})$ or to change A only at every n^{th} iteration. Such conventions are common in Newton-Raphson techniques; using $H(p^{(0)})$ throughout a sequence of iterations is particularly common.

An ideal choice of constant A is probably $A = -H(p^*)$ if this were known. Alternatively if some features of $H(p^*)$ were known, this might suggest modifying $H(p^{(0)})$ to a better choice. This is what motivated the choice of $A_{10}(p)$ above.

§6.2.4 The above results provide a rule for selecting $m^{(r)}$ at $p^{(r)}$ when (P1) is under consideration, namely

$$m^{(r)} = m_*^{(r)} = \pm m(A, \underline{1}', d^{(r)}),$$

where $d^{(r)}$ has components $d_j^{(r)} = \partial\phi/\partial p_j^{(r)}$.

We must still decide on the value of the steplength α_r. This can still be chosen by any of the rules we have previously considered. However it is possibly more natural here to take $\alpha_r = \alpha_r^*(m_*^{(r)})$, which is what Wu opts to do. Almost always though this will require in the design context direct numerical solution of 4.2.1.

Thus this employs the steepest ascent algorithm $S\{A, \alpha_r^*(m_*^{(r)})\}$.

We conclude this section with the following discussion.

It is important to note that the only restriction imposed on m^*, in lemma 6.2.1, is that $Cm^* = 0$, or $\underline{1}'m^* = 0$ if (P2) is under consideration.

If $p^{(r)}$ is on the boundary of $\mathcal{P}$, it is therefore possible that $(p^{(r)} + \alpha m_*^{(r)}) \notin \mathcal{P}$ for any $\alpha > 0$. Indeed this could conceivably be true of both $\{p^{(r)} + \alpha \cdot m(A, \underline{1}', d^{(r)})\}$ and of $\{p^{(r)} - \alpha \cdot m(A, \underline{1}', d^{(r)})\}$.

Because of this the results of this section are in practice applied in the following modified fashion. Namely let $m_j^{(r)} = 0$ if $p_j^{(r)} = 0$ and otherwise let $m_j^{(r)}$ be defined by the formula $m(A, 1', d^{(r)})$ by temporarily imagining that, in (P2), $\mathcal{U}$ is $\mathrm{Sup}(p^{(r)})$ and hence that J is the size of that support. The matrices $A_j, A_j(p)$ are correspondingly restricted. This approach is implicit in Wu's discussions and is also the way in which we will apply other algorithms.

Of course the choice of $\mathrm{Sup}(p^{(r)})$ must be made carefully. Ideally $\mathrm{Sup}(p^{(r)})$ should also be $\mathrm{Sup}(p^*)$. In practice we could let $\mathrm{Sup}(p^{(r)})$ or $\mathrm{Sup}(p^{(0)})$ be determined by other simpler algorithms such as vertex direction techniques. We then gain the advantage of keeping J relatively small.

Note that if we put the weight at a current support point to zero then we must redefine the value of J and also the matrix A. Hence even if that is 'constant' A^{-1} or A^{+} must be recalculated.

This will also be the case if we conclude that $\text{Sup}(p^{(r)})$ excludes support points of the optimum p^*. A vertex direction or some other iteration must then be taken to augment the current support, before continuing with the steepest ascent iterations corresponding to the revalued J and A.

Note also that in the case of problem (P3) only that submatrix of C corresponding to the current support would play an active role in the formula.

Wu applied some of the above schemes to a number of examples. His results will be reported at the end of chapter 7.

(c) these directions satisfy $m_r' Q m_s = 0,\ r \neq s$; here $(-Q)$ is the hessian matrix and the points $m_0, m_1, \ldots, m_{(n-1)}$ are said to be Q-conjugate, from which it follows that they are linearly independent (see Luenberger (1973, p.169)).

The above sequence then has been obtained by taking, what is in fact, an optimal step of unit length in the successive directions $m_0, m_1, \ldots, m_{n-1}$.

A similar result holds for any arbitrary set of Q-conjugate directions, but the optimal steps need not now be of unit length, for if $m_0, \ldots, m_{n-1}$ are Q-conjugate then so are $c_0 m_0, \ldots, c_{n-1} m_{n-1}$ for any constants $c_0, \ldots, c_{n-1}$. The result is that the sequence $\{x_r\}$, where $x_{r+1} = x_r + \alpha_r^*(m_r) m_r$, $\alpha_r^*(m_r) = (g_r' m_r)/(m_r' Q m_r)$, $g_r = \partial f/\partial x_r$, attains the optimum x^* in the n steps. See the Conjugate Direction Theorem of Luenberger (1973, p.170).

It is natural to consider if the directions $m_0, m_1, \ldots, m_{n-1}$ could be chosen optimally. Luenberger (1973) proposes the formula

$$\text{6.3.1} \qquad \begin{cases} m_0 = g_0 \\ m_{r+1} = g_{r+1} + \beta_r m_r \\ \beta_r = -(g_{r+1}' Q m_r)/(m_r' Q m_r). \end{cases}$$

This is called the conjugate gradient algorithm. The directions m_r are generated sequentially as the method progresses, and are conjugate versions of the successive gradients obtained, in contrast to unconstrained steepest ascent which would move only in the direction of the current gradient vector.

Luenberger shows that these choices of m_r are Q-conjugate, and that

$$\text{6.3.2} \qquad \alpha_r^*(m_r) = (g_r' g_r)/(m_r' Q m_r), \quad \beta_r = -(g_{r+1}' g_{r+1})/(g_r' g_r).$$

See the Conjugate Gradient Theorem, Luenberger (1973, p.174).

That these m_r depend on the gradient vectors g_r suggest that they must be a good choice. In fact the optimum could be reached in less than n steps. Unless the optimum has been obtained g_r is orthogonal to the subspace generated by $m_o,..,m_{n-1}$ and hence is linearly independent of these.

§6.3.2 These latter optimality considerations are not too important for the quadratic problem but they are important for the unconstrained maximisation of a nonquadratic concave criterion $\psi(\cdot)$, for it is natural to suppose that a corresponding approach would also work well in such a case, especially when quadratic approximations are usually good in the region of an optimum. Of course we cannot hope to reach the optimum in n steps.

Various conjugate gradient algorithms have been proposed for nonquadratic maximisation problems, all based on the idea of cycles of n conjugate gradient directions.

One is based on the quadratic approximation idea and simply mimics the above formulae but with Q replaced by $(-H_r)$, $H_r = H(x_r)$ being the hessian matrix, so that $(-H_r)$ is nonnegative definite if $\psi(\cdot)$ is concave.

The iteration is

$$x_{r+1} = x_r + \alpha_r m_r$$

where

$$6.3.3 \quad \left\{ \begin{array}{l} \left. \begin{array}{l} m_{nk} = g_{nk} \\ m_{(nk+t+1)} = g_{(nk+t+1)} + \beta_{(nk+t)} m_{(nk+t)} \end{array} \right\} \quad \begin{array}{l} k = 0,1,2, \cdots \\ t = 0,1,\cdots,(n-2). \end{array} \\ \beta_r = -\left(g'_{r+1} H_r m_r\right)/\left(m'_r H_r m_r\right) \\ \alpha_r = -\left(g'_r m_r\right)/\left(g'_r H_r g_r\right) \end{array} \right.$$

Hence at the first iteration and at every n^{th} iteration thereafter a steepest ascent direction is taken while at the other iterations an 'H_r-conjugate gradient direction' is taken. The steplength chosen is suggested by the optimal steplength for the quadratic case, though it will typically not now be optimal. Nor will the simpler formulae 6.3.2 for α_r, β_r obtain.

A disadvantage of the iteration 6.3.3 is the need to calculate H_r at each iteration. With a view to simplification the Fletcher-Reeves method takes

6.3.4 $$\beta_r = -(g'_{r+1}g_{r+1})/(g'_r g_r) \quad ; \quad \alpha_r = \alpha^*_r(m_r)$$

where $\alpha^*_r(m_r)$ is the optimal unconstrained steplength in the direction m_r.

Other variations scale down this choice of β_r to

6.3.5 $$\beta_r = -(g'_{r+1}g_{r+1})/\{c(g'_r g_r)\}$$

where $c > 1$, or to

6.3.6 $$\beta_r = -(g'_{r+1}g_{r+1})/\{c_r(g'_r g_r)\}$$

where c_r is a variable or adaptive choice of scaling factor, its value being chosen according to the current value of some criterion such as $F^A(x_r, x_r + g_r)$.

§6.3.3 These ideas have been concerned with unconstrained optimisation, but Luenberger (1973) says that analogues of conjugate gradient methods or Partan could be generated by handling constraints through reduction or projection. We consider these ideas now and in the next section.

The former is based on the idea of temporarily substituting for a suitable subset of the variables y in terms of the remaining ones z according to the constraints. The above methods are then used on the gradient vector with respect to z to decide on a direction of iteration. The elements of y and z interchange as need be.

Wu opts for the projection approach. The idea is that we project the above directions onto the constraint space to which direction vectors must belong, assuming that the current iterate is feasible. As we have seen this space would be $Cm = 0$, $\underline{1}'m = 0$ if the main constraints are $C\theta = b$, $\underline{1}'p = 1$, as in the maximisations (P3), (P1) of a function $\phi(\cdot)$.

We redefine some notation. Let $D^{(r)}$ replace m_r, $d^{(r)}$ replace g_r in the formulae above, so that with now $n = J$

$$\begin{cases} D^{(Jk)} = d^{(Jk)} \\ D^{(Jk+t+1)} = d^{(Jk+t+1)} + \beta_{(Jk+t)} D^{(Jk+t)} \\ \beta_r = -\{d^{(r+1)'} d^{(r+1)}\} / \{c_r d^{(r)'} d^{(r)}\} \end{cases}$$

If problem (P1) is under consideration we wish to project $D^{(r)}$ onto $\underset{\sim}{1}'m = 0$. The relevant orthogonal projection is of course

6.3.8 $$m^{(r)} = D^{(r)} - \bar{D}^{(r)} \underset{\sim}{1}$$

where $\bar{D}^{(r)}$ is the average of the components of $D^{(r)}$.

Hence

6.3.9 $$m^{(r)} = m(I, \underset{\sim}{1}', D^{(r)}).$$

If our intent is to take $A = I$ then this formula selects the constrained steepest ascent direction for that choice of A at $r = Jk$ for then $D^{(r)} = d^{(r)}$ and so is the orthogonal projection of the unconstrained steepest ascent direction $d^{(r)}$. Wu uses 6.3.9 in a number of numerical examples with the Fletcher Reeves conjugate gradient approach in mind, employing various choices of the scaling constant c and also an adaptice c_r sequence. We will consider his results at the end of chapter 7.

For general A we would clearly opt for the direction $\pm m\{A, \underset{\sim}{1}', d^{(r)}\}$ at $r = Jk$. It seems more natural then to adopt, for all r the alternative projection.

6.3.10 $$m^{(r)} = \pm m(A, \underset{\sim}{1}', D^{(r)})$$

and for general C the projection

6.3.11 $$m^{(r)} = \pm m(A, C, D^{(r)}).$$

One might also entertain the formula

6.3.12 $$\beta_r = -\left(d^{(r+1)'} A^+ d^{(r+1)}\right) / \left(d^{(r)'} A^+ d^{(r)}\right)$$

We note that for the case $A = I$, Wu (1976, p.41) shows that if $\beta_r < 0$ then $F\{p^{(r)}, p^{(r)} + m(I, \underset{\sim}{1}', D^{(r)})\}$ is not smaller than $F\{p^{(r)}, p^{(r)} + m(I, \underset{\sim}{1}', d^{(r)})\}$ and hence must be positive.

We will denote a conjugate gradient algorithm selecting direction 6.3.9 or 6.3.10 by $C(A, \beta_r, \alpha_r)$.

Again a basic assumption must be that $p_j^{(r)} > 0$.

§6.4 Adapting Unconstrained Iterations; Solving Equations

§6.4.1 Suppose we know Sup(p^*). Without loss of generality suppose that Sup(p^*) = $\mathcal{U}$. Suppose also that $\psi(\cdot)$ is differentiable or enjoys support differentiability at p^*. Then the following is true.

Firstly we know from theorem 2.5.6 that the solution p^* to problem (P2) solves the equations

6.4.1 $$F\{x(p), u_j\} = 0 \quad , \quad j = 1, \cdots, J \; .$$

Furthermore not only have the constraints $p_j^* \geq 0$ been rendered inactive, but also the following lemma is true.

Lemma 6.4.1

Suppose that $\psi(\cdot)$ is differentiable at $x(p) = \Sigma p_j u_j$ and suppose that p solves 6.4.1, then, assuming $G\{x(p), x(p)\} \neq 0$, $\underline{1}'p = 1$

Proof At differentiable x(p) we have

$$F\{x(p), u_j\} = G\{x(p), u_j\} - G\{x(p), x(p)\} \, .$$

Thus 6.4.1 implies that

$$G\{x(p), u_j\} = G\{x(p), x(p)\} \, , \; j = 1, \cdots\cdots, J$$

Hence $$\Sigma p_j G\{x(p), u_j\} = G\{x(p), x(p)\} \cdot (\Sigma p_j)$$

However in general

$$\Sigma p_j G\{x(p), u_j\} = G\{x(p), \Sigma p_j u_j\} = G\{x(p), x(p)\} \, .$$

Hence we have

$$G\{x(p), x(p)\} = G\{x(p), x(p)\} \cdot (\Sigma p_j) \, ,$$

and the result is proved. □

These observations clearly alter problem (P2) to a much simpler optimisation problem, an unconstrained one. It is natural to consider other algorithms for solving (P2), namely to employ standard techniques such as Newton Raphson or Fletcher Powell to solve 6.4.1. With a view to speeding up convergence though we might consider modifying the iterates of such techniques to ensure that $\Sigma p_j^{(r)} = 1$

If, in the absence of further information, we take $p_j^{(0)} = 1/J$ then, since $p_j^* > 0$, one would expect iterates $p^{(r)}$ such as these not to stray far from the feasible region $\mathcal{P}$ for problem (P1), and indeed to remain in $\mathcal{P}$ if all p_j^* are not too small, although, as we shall see, typically there will also be negative solutions to 6.4.1; that is, a solution p such that some p_j are negative.

Even when one is not sure that Sup(p^*) and $\mathcal{U}$ are the same, when say it is possible that $\mathcal{U}$ might contain at most one or two vertices not in Sup(p^*), one might contemplate setting out to solve 6.4.1 using iterates as above with $p_j^{(0)} = 1/J$. For if it is the case that Sup(p^*) $\subset \mathcal{U}$, this will manifest itself in the emergence of negative iterates, assuming that these converge, in view of the fact that there then can be only negative solutions to 6.4.1. If iterates turn negative and persist in remaining so, or assume large negative weights, it can reasonably be concluded that $\mathcal{U}$ does contain non-optimum-support points. In general one would require to switch to another algorithm although, as we shall see from empirical results in section 6.4.2, one might be able to identify the non-optimun-support points from the iterates themselves. It should be emphasised that this approach should only, if at all, be considered if one is sure that $\mathcal{U}$ contains at most one or two non-members of Sup(p^*). If we are entirely vague about the members of this set it would be better to employ some other algorithm at least initially. This theme will be reconsidered in section 6.5.

The contents reported in section 6.4.2 are the results of using adapted and unadapted Newton Raphson iterates to solve equation 6.4.1. There are various possibilities that could be employed to adapt or constrain an iterate of a technique for an unconstrained problem, to iterates $p^{(r)}$ which satisfy $\sum p_j^{(r)} = 1$.

Suppose that $\sum p_j^{(r)} = 1$ and let m_r denote the unconstrained direction in which a standard technique would step from $p^{(r)}$. Let x_{r+1} be the subsequent unconstrained iterate that would emerge from the step that the technique would take in the direction m_r. The value of that steplength is not important.

Two simple possibilities spring to mind. These are, adapting by means of projection and, adapting by means of dilation. The former we have already considered in the previous section where the idea of reduction was also mentioned. Three suggestions emerge.

(i) By projection we mean projecting m_r onto $1'm = 0$. Let $m^{(r)}$ be the resultant mapping of m_r. Then take as before

$$p^{(r+1)} = p^{(r)} + \alpha_r m^{(r)} .$$

The most natural procedure to use is an orthogonal transformation so that $m^{(r)} = m_r - \bar{m}_r \underline{1}$ where $\bar{m}_r$ is the average of the components of m_r. We have however seen other alternatives in the previous section, namely $m^{(r)} = \pm m(A, \underline{1}', m_r)$.

(ii) By dilation we simply mean to take

$$p^{(r+1)} = x_{(r+1)} / (\underline{1}' x_{(r+1)})$$

This has the advantage that if the components of x_{r+1} are each positive, then $p^{(r+1)} \in \mathcal{P}$.

One might consider other transformations of x_{r+1} onto $\underline{1}'p = 1$.

(iii) As a variation on the latter one could let

$$q^{(r)} = \{p^{(r)} + m_r\} / \{\underline{1}'(p^{(r)} + m_r)\}$$

Then $\underline{1}' q^{(r)} = 1$ and a possible iterate is

$$p^{(r+1)} = \begin{cases} (1-\alpha_r) p^{(r)} + \alpha_r q^{(r)} & , \quad F_\Phi(p^{(r)}, q^{(r)}) > 0 \\ (1+\alpha_r) p^{(r)} - \alpha_r q^{(r)} & , \quad F_\Phi(p^{(r)}, q^{(r)}) < 0 . \end{cases}$$

The numerical results of section 6.4.2 include reports on the performance of suggestion (ii) when used to adapt Newton Raphson iterates to solve equation 6.4.1.

Let $\underline{f}(p)$ be the vector whose j^{th} component is $F\{x(p), u_j\}$ and let $D(p)$ be the matrix whose $(i, j)^{th}$ element is $\partial F\{x(p), u_i\} / \partial p_j$. Then unmodified Newton Raphson iterates satisfy the relation

6.4.2 $$p^{(r+1)} = p^{(r)} - D^{-1}(p^{(r)}) \underline{f}(p^{(r)}) .$$

Thus under suggestion (ii) dilated Newton Raphson iterates are defined by the relation

6.4.3 $$p^{(r+1)} = \left\{ p^{(r)} - D^{-1}(p^{(r)}) \underline{f}(p^{(r)}) \right\} \Big/ \left\{ \underline{1}' \left[p^{(r)} - D^{-1}(p^{(r)}) \underline{f}(p^{(r)}) \right] \right\}$$

Of course these formulae require again the inversion of a J×J matrix at each iteration, namely $D(p^{(r)})$. A standard simplifying convention is then to replace $D^{-1}(p^{(r)})$ by $D^{-1}(p^{(0)})$, while secant methods would replace the former by some other constant matrix.

We note that an alternative to Newton Raphson techniques for solving equations are Gauss Seidel techniques which would not incur the above problem. One example of such a technique does the following, as described in Ortega and Rheinboldt (1970). Cycling j through the values 1, 2, ..., J in turn, solve numerically for p_j the single equation.

$$F\{x(p),\ u_j\} = 0\ ,$$

while holding the other components p_i fixed at current values. Clearly we could adapt such a technique as above. This procedure would change significantly only one weight at each iteration and so would enjoy the same benefits that vertex direction iterations derive from lemma 5.2.1 in the design context.

§6.4.2 We now report the results of using iterates 6.4.2 and 6.4.3 to solve equation 6.4.1 in design contexts, that is equation 6.4.4 below.

The results are mainly for D-optimality with some additional results for A-optimality. The equation

6.4.4 $$F\{M(p),\ v_j v_j'\} = 0$$

becomes respectively equations 6.4.5, 6.4.6 in these two cases.

6.4.5 $$v_j' M^{-1}(p) v_j - k = 0$$

6.4.6 $$v_j' M^{-2}(p) v_j - tr\{M^{-1}(p)\} = 0$$

Not surprisingly these equations have several solutions. Examples are the following ones.

Ex. 6.4.2(i) Take $\mathcal{U} = \{(1,-1,-1)',(1,-1,1)',(1,1,-1)',(1,2,2)'\}$. Equation 6.4.5 has the three solutions

$p = (8/9, (1\pm\sqrt{11})/6, (1\mp\sqrt{11})/6, -2/9)$ and $p = (4/32, 9/32, 9/32, 10/32)$ and possibly others. The latter solution is of course the D-optimal design.

Ex. 6.4.2(ii) Take $\mathcal{U} = \{(1,0,0)', (0,1,0)', (0,0,1)', (4,5,6)', (3,4,5)'\}$. Two solutions to 6.4.5 are p = (.359,.213,.229,.767,-.569), and p = (.651,-.318,.629,-.139,.178). There is no positive solution in this case, as one vertex, namely $(3,4,5)'$, does not belong to the support of the D-optimum design.

Ex. 6.4.2(iii) Now take $\mathcal{U} = \{\underline{e}_1, \ldots, \underline{e}_k, x\underline{1}\}$, $x \neq 0$, where $\underline{e}_j$ is the $k\times 1$ unit vector, while $\underline{1}$ is a $k\times 1$ vector of 1's. We can quote a wide class of solutions to 6.4.5 here. For any $t \in \{0,1,2,\ldots,k\}$ there can be one or two solutions $p = p(D)$ of the form

$$p_i = \begin{cases} (1+D)/2k & , i = 1,\ldots,t \\ (1-D)/2k & , i = (t+1),\ldots,k \end{cases} , \quad p_{k+1} = \{k + (k-2t)D\}/2k$$

Solutions can only be of this form allowing for the variation that any t of the first k p_i's could be assigned the value $(1+D)/2k$.

The quantity D is given by

$$D = \begin{cases} \dfrac{-x^2(k-2t)(k-1) \pm \left[x^4(k-2t)^2 + \{k(k-2) - (k-2t)^2\}x^2 + 1\right]^{1/2}}{\{(k-2t)^2x^2 - 1\}} , & \text{if } (k-2t)^2x^2 \neq 1 \\ -\{1 + 2t(t-1)/(k-1)(k-2t)\} & \text{if } (k-2t)^2x^2 = 1 \end{cases}$$

These values of D definitely render p(D) a solution of 6.4.5 provided $p_i \neq 0$ for $i = 1,\ldots,(k+1)$.

Particular results are as follows.

(a) If $x = 1$ there are two solutions for each $t = 2,\ldots,k-2$ and $k \geq 4$, namely

$$p_i = -(t-1)/\{k(k-2t+1)\} \quad , i = 1,\ldots,t$$
$$p_i = (k-t)/\{k(k-2t+1)\} \quad , i = (t+1),\ldots,k+1$$

and

$$p_i = p_{k+1} = -t/\{k(k-2t-1)\} \quad , i = 1,\ldots,t$$
$$p_i = (k-t-1)/\{k(k-2t-1)\} \quad , i = (t+1),\ldots,k$$

(b) In the case $t = k-1$ or $t = 1$ and $k \geq 3$ there is the one solution

$$p_i = (k-1)(k-2)x^2/\{k[(k-2)^2x^2 - 1]\} \quad , i = 1,\ldots,(k-1)$$

$$p_k = -[(k-2)x^2 + 2]/\{k[(k-2)^2x^2 - 1]\} , \quad p_{k+1} = -(k-1)/\{k[(k-2)^2x^2 - 1]\}.$$

(c) In the case t = k or t = 0 there is the one solution

$$p_i = (k-1)x^2/(k^2x^2-1)\ ,\ i = 1,\cdots,k\ ;\ p_{k+1} = (kx^2-1)/(k^2x^2-1)\ ,\ |x| \neq 1/k\ .$$

When $|x| > 1/\sqrt{k}$ this identifies a D-optimal design while it is still a solution to 6.4.5 in the case $|x| = 1/\sqrt{k}$ although $p_{k+1} = 0$. Another instance of this possibility is seen in the last example.

Ex. 6.4.2(iv) Finally take k = 4 and

$$\mathcal{U} = \{\, \underline{e}_1, \underline{e}_2, \underline{e}_3, \underline{e}_4, (.5,.5,.5,.5)', (-.5,-.5,.5,.5)', (-.5,.5,-.5,.5)', (-.5,.5,.5,-.5)', (-.5,.5,.5,.5) \}.$$

Here there is an infinity of solutions to equation 6.4.5, namely $p = (\lambda,\lambda,\lambda,\lambda,\mu,\mu,\mu,\mu,0)$ with $\lambda+\mu = 1/4$.

Consider now employing iterations 6.4.2 or 6.4.3 taking $p_j^{(0)} = 1/J$, $j = 1,\ldots,J$. The following issues are of interest.

(1) If solutions to equation 6.4.1 or 6.4.4 are also solutions to the appropriate example of problem (P1), will such iterations converge to one of these solutions?

(2) Will such iterations $p^{(r)}$ tend to satisfy $p_j^{(r)} > 0$?

(3) How will such iterates behave when no solutions to 6.4.1 or 6.4.4 solves problem (P1)?

We have only empirical results to offer. Clearly if $\mathrm{Sup}(p^*) = \mathcal{U}$, then we would expect almost always, that iterations 6.4.2 or 6.4.3 would converge to p^* and further, that $p_j^{(r)} > 0$ for each j , r, if no optimal weights are too small. Not a few examples could be quoted to bear this out. In the case of D-optimality examples are iterations 6.4.3 in each of the following cases.

(i) $\mathcal{U}$ as in example 6.4.2(i) for which $p_j^* \geq p_1^* = .125$.

(ii) $\mathcal{U} = \{(x, x^2, \sin 2\pi x, \cos 2\pi x)' : x = .082, .083, .381, .734, .735, 1\}$ for which $p_j^* \geq .07688$.

(iii) $\mathcal{U} = \{(1,0,0,0)',(0,1,0,0)',(0,0,1,0)',(0,0,0,1)',(.75,.75,0,0)'$
$(.75,0,.75,0)',(.75,0,0,.75)',(0,.75,.75,0)'(0,.75,0,.75)'$
$(0,0,.75,.75)'\}$
for which $p_j^* \geqslant 3/20 = .067$.

(iv) $\mathcal{U} = \{(1,x_1,x_2,x_1^2,x_1x_2,x_2^2)' : x_i = -1,0,1 \ ; \ i = 1,2\}$ for which $J = 9$, $k = 6$, $p_j^* \geqslant .08015$. This $\mathcal{U}$ is the support of the D-optimum design for quadratic regression in two variables x_1,x_2 over $-1 \leq x_i \leq 1$ with a constant term.

(v) $\mathcal{U} = \{(1,x_1,x_2,x_3,x_1^2,x_2^2,x_3^2,x_1x_2,x_1x_3,x_2x_3)' : x_i = -1,0,1 : i = 1,2,3\}$
$-\{1,0,0,0,0,0,0,0,0,0)\}$ for which $J = 26$, $k = 10$,
$p_j^* \geqslant .01895$. This is the analogue in three variables of the space $\mathcal{U}$ in the previous example.

Finally in the case of A-optimality examples are iterations 6.4.3 in the following three cases; $\mathcal{U}$ as in (i) and $\mathcal{U}$ as in (iv) above and also $\mathcal{U} = \{(1,0,0)',(1,1,0)',(1,1,1)'\}$.

These results not withstanding it is of course likely that $p_j^{(r)}$ will be negative for some r if p_j^* is small. The following gives some minor indication of when negative iterates are likely to occur.

Consider the space $\mathcal{U}$ of example 6.4.2(iii) for which $k = 4$, namely $\mathcal{U} = \{(1,0,0,0)',(0,1,0,0)',(0,0,1,0)',(0,0,0,1)',(x,x,x,x)'\}$, $x \neq 0$. In the case of D-optimality and also $\phi(p) = -\operatorname{tr}\{M^{-t}(p)\}$, $\mathcal{U} = \operatorname{Sup}(p^*)$ if $x > 1/2$, but p_5^* will be small if x just exceeds $1/2$; for $x = .55,.54,.53,.52,.51$ respectively $p_5^* = .055,.045,.035,.024,.013$. Only for $x \geqslant .55$ do iterations 6.4.2 remain positive. In the cases $x = .54,.53,.52$, $p_5^{(1)} < 0$, while both $p_5^{(1)}$, $p_5^{(2)}$ are negative when $x = .51$. The case $x = .5$ is of course particularly interesting here, for then $p = (.25,.25,.25,.25,0)$ solves equations 6.4.5, 6.4.6 and also equation 6.4.4 when $\phi(p) = -\operatorname{tr}\{M^{-t}(p)\}$. Not surprisingly $p_5^{(r)}$ is always negative when using iteration 6.4.2 to solve equation 6.4.5.

Other examples of this type arise when, with $k = 4$, $\mathcal{V} = \mathcal{V}_1 = \{\underline{e}_1, \underline{e}_2, \underline{e}_3, \underline{e}_4, (.5,.5,.5,.5)', (-.5,.5,.5,.5)'\}$ and when $\mathcal{V} = \mathcal{V}_1 \cup \{(.5,.5,-.5,-.5)'\}$. In both these cases a solution to equation 6.4.4 puts $p^*_j = 1/4$, $j = 1,2,3,4$ and $p^*_j = 0$, $j \geq 5$, and in both cases, with $p_j^{(0)} = 1/J$, $p_j^{(r)}$ is negative for $j \geq 5$ under iteration 6.4.2.

Other instances when negative iterates are realised are in the case of a vertex whose optimal wieght is .037 and in the case of two points whose optimal weight is .002.

Before concluding this discussion of the first two issues with a final example, it is of relevance to the third issue to note here that in all of the above examples, including those when optimal weights are zero, we always had $p_j^{(r)} > -0.1$, in fact $p_j^{(r)} \geq -.08$.

The example we now wish to consider concerns the space of example 6.4.2 for which $J = 9$. For any λ, μ such that $\lambda + \mu = 1/4$, $p = (\lambda,\lambda,\lambda,\lambda,\mu,\mu,\mu,\mu,0)$ is a solution to 6.4.5 and hence there are an infinity of D-optimal designs, those for which $\lambda, \mu \geq 0$. However, starting from $p_j^{(0)} = 1/9$ to solve 6.4.5, iterations 6.4.2 behaved as follows. Firstly $p_q^{(r)}$ is negative for $r = 1,2$ and then is approximately zero for $r \geq 3$. The behaviour of the other components is distinctive. For $r = 1,\ldots,5$, $p_j^{(r)} < 0$, $j = 1,\ldots,4$, $p_j^{(r)} > 0$, $j = 5,\ldots,8$, while these inequalities are reversed for $r = 6,7,8,9$. The solution converged to takes $\lambda = .4563$, $\mu = -.2063$. That this is not a positive solution is disappointing, but also disquieting is that the iterations fluctuate markedly despite the fact that for $r \geq 3$, $p^{(r)}$ is a solution correct to at least three decimal places. For $r = 3,4,5,6,7,8,9$, $p^{(r)}$ is approximately the solution corresponding respectively to $\lambda = -.375, -.375, -.683, .651, .554, .456, .456$. Admittedly this example is a severely testing one.

We turn now to the third issue. This is of relevance because, as has already been suggested, we might consider using iterations 6.4.2, 6.4.3 when we are not entirely sure that $\mathcal{V}$ and $\text{Sup}(p^*)$ are the same. If, starting from $p_j^{(0)} = 1/J$, iterates adopt non-negligible negative values and persist in doing so it will be reasonable to assume that $\text{Sup}(p^*) \subset \mathcal{U}$. The results below suggest that such a conclusion could be reached at the earliest iterations.

The following four design spaces consist of four vertices in the case k = 3. In each instance the first three points form the D-optimal $\text{Sup}(p^*)$.

$$\mathcal{U} = \{(1,0,0)',(0,1,0)',(0,0,1)',(.3,.2,.1)'\}$$
$$\mathcal{U} = \{\quad " \quad , \quad " \quad , \quad " \quad ,(.6,.5,.4)'\}$$
$$\mathcal{U} = \{(4,3,2), \quad " \quad , \quad " \quad ,(1,0,0)'\}$$
$$\mathcal{U} = \{(7,3,2)', \quad " \quad , \quad " \quad , \quad " \quad \}$$

There are therefore no positive solutions to equation 6.4.5. In each case, with $p_j^{(0)} = 1/4$, iterations 6.4.3 turned negative immediately and remained so. The minimum of the four components of $p^{(1)}$ are respectively -.11,-.40,-.23,-.39 in the four examples. The corresponding minima in the case of $p^{(2)}$ are -.31.-.23,-.12,-1.08.

These results together with the observation that $p_j^{(r)} > -0.1$ in the examples on convergence to a positive solution suggest that a not unreasonable rule might be to conclude that $\text{Sup}(p^*) \subset \mathcal{U}$ if the inequality $p_j^{(r)} < -0.1$ is realised. This rule was in fact employed in the results reported in section 6.5.

To end this discussion it is of interest to report the solution to which iterates 6.4.3 converged in the above four examples. These were respectively

$$p^* = (-.239,-.131,-.041,1.410)$$

$$p^* = (.417,.395,.375,-.186)$$

$$p^* = (.342,.397,.365,-.105)$$

$$p^* = (-.259,-1.020,-.634,2.913)$$

Convergence then, from $p_j^{(0)} = 1/4$, is to a solution which allocates "weights" with a common sign to the members of $\text{Sup}(p^*)$, thus forcing the weight of the one non-member of that support to be opposite in sign in view of the constraint $\underline{1}'p = 1$.

In fact in all four examples $p^{(1)}$ shared the same sign structure as p^*. So also did $p^{(r)}$ for each r in the first and fourth cases.

Another instance of this phenomena occurs in the case of the space $\mathcal{U}$ of example 6.4.2(iv) for which $k = 3$, $J = 5 = (k+2)$. In this case the first four points form $Sup(p^*)$. There is again only one non-optimal support point. Two solutions to equation 6.4.5 were reported. One solution contained exactly two negative components, but from $p_j^{(0)} = 1/5$, iterations 6.4.3 converged to the other solution, namely $p = (.359,.213,.229,.767,-.569)$.

§6.5 A Composite Algorithm

In the previous sections of this chapter we, as previously said, have been considering algorithms suitable for sets of vertices which do not include large numbers of non-optimal-support points, some, as in the case of using adapted Newton Raphson in the solution of equations 6.4.1, only suitable when the set of vertices forms the full support of the optimum, while others, such as diagonalised versions of constrained steepest ascent, conjugate gradient or quasi Newton, will tolerate or operate fairly efficiently in the presence of a small number of non-optimal support vertices. Also the more complex of these schemes requires the inversion of a $J \times J$ matrix at least once, and even if this is not required at every iteration, a matrix of lower order has to be inverted every time a vertex is eliminated.

Clearly use of such procedures, i.e. no non-support tolerated or small non-support tolerated, will only be possible if we can identify a subset of the vertices which contains a subset, if not all, of $\mathrm{Sup}(p^*)$, and a small number if at all of non-optimal support points.

Fortunately it would appear from the empirical results of section 5.6 that, while the convergence of a vertex direction algorithm can be slow, it can quickly identify most if not all of $\mathrm{Sup}(p^*)$. It would be reasonable to conclude that the set, which we will denote by $\mathcal{U}_{(1)}$, of, for some T, those T vertices selected most often or given large weight after a suitable number of iterations will belong to $\mathrm{Sup}(p^*)$, and would also form the full support of the optimum on $\mathcal{U}_{(1)}$. On the other hand a set of vertices $\overline{\mathcal{U}}$ less rigorously chosen than $\mathcal{U}_{(1)}$ could contain most if not all of $\mathrm{Sup}(p^*)$ and a small number of non-optimal support points. One would want $\overline{\mathcal{U}}$ to include those vertices regularly selected, often and not so often, initial iterations excepted. One might also include the neighbouring vertices to these when the complete $\mathcal{U}_J = \{u_1, \ldots, u_J\}$ is a discretisation of a continuous set. The less rigorous the selection of $\overline{\mathcal{U}}$, the more we would expect the support of the optimum on $\overline{\mathcal{U}}$ to be a proper subset of $\overline{\mathcal{U}}$.

Suppose we select a $\overline{\mathcal{U}}$. Then in theory we apply one of the procedures that will tolerate a small number of non-optimal support points being prepared to put weights to zero according to the rules of section 4.3.3. We continue applying this procedure until the optimum $\bar{p}^*$ on $\bar{\mathcal{U}}$ is determined, or at some point when all non-support points of $\bar{p}^*$ have been identified we might switch to one of the more efficient algorithms. If we have chosen $\bar{\mathcal{U}}$ well then $\bar{p}^*$ should be the optimum for $\mathcal{U} = \mathcal{U}_J$. If however $\bar{p}^* \neq p^*$ how then do we proceed. We are then in a position which is more likely to occur if we opt for a set like $\mathcal{U}_{(1)}$ described above.

For generality however let $\mathcal{U}_{(1)}$ be an arbitrary subset of $\mathcal{U} = \mathcal{U}_J$. Suppose that we find the optimising $p^*_{(1)}$ on $\mathcal{U}_{(1)}$ by some technique or other. (If $\mathcal{U}_{(1)}$ is the set considered above we would use a 'no non-optimal-support tolerated' procedure such as adapted Newton Raphson for solving equations 6.4.1 or 6.4.4). We check for the optimality of $p^*_{(1)}$ on $\mathcal{U} = \mathcal{U}_J$. If $p^*_{(1)} \neq p^*$ how do we proceed?

This is a circumstance which we have already considered in several places including the discussion following lemma 5.5.1. We must adopt some other algorithmic rule to select a direction in which to move from $p^*_{(1)}$. The following suggestion is a variation on a forward vertex direction choice.

The suggestion is to improve $\mathcal{U}_{(1)}$ to the set

$$\mathcal{U}_{(2)} = \mathrm{Sup}\{p^*_{(1)}\} \cup \{u^{(j)}\} \quad , \quad F\{x(p^*_{(1)}), u^{(j)}\} > 0 \ , \quad u^{(j)} \in \mathcal{U}_J .$$

Now find $p^*_{(2)}$, the optimum on $\mathcal{U}_{(2)}$ and repeat.

This procedure generates a sequence of designs $p^*_{(t)}$, $t = 1,2,\ldots$ corresponding to subsets $\mathcal{U}_{(1)}, \mathcal{U}_{(2)}, \ldots, \mathcal{U}_{(t)}$ of $\mathcal{U} = \{u_1, \ldots, u_J\}$, the design $p^*_{(t)}$ being the optimum on $\mathcal{U}_{(t)}$. The subsets are defined by the relation

$$\mathcal{U}_{(t+1)} = \mathrm{Sup}\{p^*_{(t)}\} \cup \{u^{(j)}\} \quad , \quad u^{(j)} \in \{u_i \in \mathcal{U} : F\{x(p^*_{(t)}), u_i\} > 0\}$$

As a consequence the sequence, $\psi\{x(p^*_{(t)})\}$, $t = 1,2,\ldots,$ is monotonic nondecreasing. It would seem that $p^*_{(t)}$ should converge to p^*.

The essential features in passing from $\mathcal{U}_{(t)}$ to $\mathcal{U}_{(t+1)}$ is that non-support points of $p^*_{(t)}$ in $\mathcal{U}_{(t)}$ are eliminated and $\mathrm{Sup}\{p^*_{(t)}\}$

is then augmented by one other suitable vertex. We have an "eliminating/augmenting" procedure. The idea has the following geometrical motivation.

Assume that $\mathcal{U}_J = \{u_1, u_2, \ldots, u_J\}$ is a subset of real n dimensional space R^n, so that $x(p) \in R^n$. For any differentiable $x(p)$, let $R(p) = \{u \in R^n : F\{x(p), u\} \leq 0\}$. Since $\sum p_j F\{x(p), u_j\} = 0$, then typically the sets $R(p)$ and $\mathcal{U}_J$ intersect; some $u_j \in \mathcal{U}_J$ lie outside $R(p)$. One exception is $R(p^*)$ for then $F\{x(p^*), u_j\} \leq 0$ with equality if $u_j \in \mathrm{Sup}(p^*)$. Hence $\mathrm{Sup}(p^*)$ lies on the boundary of $R(p^*)$, and so $R(p^*)$ circumscribes $\mathcal{U}_J$. This suggests that $R(p^*)$ is in some sense a smallest set containing $\mathcal{U}_J$. This shall be seen more clearly in the case of D-optimality.

Consider now the set $R(\underset{\sim}{p}^*_{(t)})$ where $\underset{\sim}{p}^*_{(t)}$ is the optimum for some subset $\mathcal{U}_{(t)}$ of $\mathcal{U}_J$. If $u_j \in \mathrm{Sup}(\underset{\sim}{p}^*_t)$ then u_j will lie on the boundary of $R(\underset{\sim}{p}^*_{(t)})$, while u_j typically lies inside $R(\underset{\sim}{p}^*_{(t)})$ if $u_j \in \left\{\mathcal{U}_{(t)} - \mathrm{Sup}\{\underset{\sim}{p}^*_{(t)}\}\right\}$. However if $\underset{\sim}{p}^*_{(t)} \neq p^*$ then there must exist some $u_i \notin \mathcal{U}_{(t)}$ which lies outside $R(\underset{\sim}{p}^*_{(t)})$. Expansion of $R(\underset{\sim}{p}^*_{(t)})$ is necessary if it is to circumscribe $\mathcal{U}_J$.

Intuitively an improvement would be to expand $R(\underset{\sim}{p}^*_{(t)})$ to include on its boundary that u_j farthest from $R(\underset{\sim}{p}^*_{(t)})$, which one might interpret to be $u^{(1)}$. Such an improvement would be $R(\underset{\sim}{p}^*_{(t+1)})$ when $\mathcal{U}_{(t+1)} = \left\{\mathrm{Sup}(\underset{\sim}{p}^*_{(t)}) \cup \{u^{(1)}\}\right\}$. Clearly one might argue that an alternative to $u^{(1)}$ is farthest from $R(\underset{\sim}{p}^*_{(t)})$.

If still $\underset{\sim}{p}^*_{(t+1)} \neq p^*$ one repeats the argument.

This picture of the procedure may make it appear ponderous and inefficient. Certainly the latter will be the case if $\mathcal{U}_{(1)}$ is chosen arbitrarily. In particular if $\mathcal{U}_{(1)}$ contains many vertices which are not support points of p^*, then the sequence will be long with vertices being brought in and later discarded and possibly vice versa. Note that all vertices not in $\mathcal{U}_{(t)}$ would have to be considered in the selection of the augmenting vertex.

Suppose however as originally suggested that the initial $\mathcal{U}_{(1)}$ contains only support vertices of p^*, and in fact only those vertices allocated large weight by p^*. Then intuitively $R(\underline{p}^*_{(1)})$ must almost circumscribe $\mathcal{U}_J$. Hence the same must be true of $R(\underline{p}^*_{(t)})$, $t = 1,2,\ldots$. It must be that $R(\underline{p}^*_{(t)})$ is only a slight expansion of $R(\underline{p}^*_{(t-1)})$, the augmenting vertex not lying far outside $R(\underline{p}^*_{(t-1)})$. Given such a closeness between $R(\underline{p}^*_{(t-1)})$ and $R(\underline{p}^*_{(t)})$ it would seem that those vertices which lie on the boundary of the former will also almost certainly sit on that of the latter. That is, given the above initial choice of $\mathcal{U}_{(1)}$, one would have thought that almost always we would have $\mathrm{Sup}(\underline{p}^*_{(t)}) = \mathcal{U}_{(t)}$ for each t. At no stage then would support points of p^* be discarded to be brought in later (again) and no non-support vertices of p^* would at any stage be included. The procedure would successfully identify the remaining support points of $\mathrm{Sup}(p^*)$. It would be just an "augmenting" procedure or at least an "(eliminating)/Augmenting" procedure. Also one would expect similarities in the weightings under $\underline{p}^*_{(t)}$, $\underline{p}^*_{(t-1)}$, suggesting that $\underline{p}^{(0)}_{(t+1)}$ should be based on $\underline{p}^*_{(t)}$. We will see at the end of this chapter evidence to support the above conjecture in some examples arising in the design context.

It was in the design context that the above geometrical argument was first proposed, by Silvey and Titterington (1973). Then the set $R(p)$ can be restated to be $R(p) = \{v \in R^k : F\{M(p), vv'\} \le 0\}$, $M(p) = \sum p_j v_j v_j'$ and now the set $R(p)$ and the design space $\mathcal{V} = \mathcal{V}_J = \{v_1, v_2, \ldots, v_J\}$ intersect for $p \neq p^*$, while $R(p^*)$ just circumscribes $\mathcal{V}_J$. Some v_j lie inside $R(p)$, some lie outside and the above ideas carry over with u_j replaced by v_j, $\mathcal{U}_{(t)}$ by $\mathcal{V}_{(t)}$.

Silvey and Titterington had D-optimality in mind in which case $R(p) = \{v \in R^k : v' M^{-1}(p) v \le k\}$, an ellipsoid, centre the origin. A duality theorem of Sibson (1972) establishes that $R(p^*)$ is the smallest such ellipsoid containing $\mathcal{U}_J$. That is, suppose that the matrix N is such that the set $\{v \in R^k : v'Nv \leq k\}$ contains $\mathcal{U}_J$. Then the smallest such set is given by $N = M^{-1}(p^*)$.

It was the geometry really of this result which suggested the above "(eliminating)/Augmenting" procedure to Silvey and Titterington. They prove convergence of the procedure for this D-optimal case.

Note that in the case of design criteria depending on the matrix $AM^+(p)A'$, where A is of order $s \times k$, $\text{rank}\{(A)\} = s$, we know that the support of p* contains at least s vertices. Hence we must take $T \geq s$, and moreover $\mathcal{U}_{(1)}$ must contain at least s linearly independent v_j's. If $A = I$ then one advantage of taking $T = k$ is that explicit solutions for $\underset{\sim}{p}^*_{(1)}$ are available in the case of A-optimality and D-optimality, $\underset{\sim}{p}^*_{(1)}$ being $(1/k, \ldots, 1/k)$ in the latter case.

Wu (1976) also uses the procedure in his computations, but with some modifications. He does not go as far as computing $\underset{\sim}{p}^*_{(t)}$. Instead he opts for a vertex direction iteration if it becomes clear that he is close enough to $\underset{\sim}{p}^*_{(t)}$ to ascertain $\underset{\sim}{p}^*_{(t)} \neq p^*$. He thus takes a vertex direction step from an approximation $\underset{\sim}{\tilde{p}}_{(t)}$ to $\underset{\sim}{p}^*_{(t)}$. He does so if, for some $\varepsilon_0 > 0$

$$F_\phi\{p^{(r)}, p^{(r)} + [\bar{\varepsilon}_r(m^{(r)})]m^{(r)}\} < \varepsilon_0$$

where $p^{(r)} = \underset{\sim}{\tilde{p}}_{(t)}$ and $m^{(r)}$ is a constrained steepest ascent direction.

He appears to consider not only $u^{(1)}$ but also $u^{(4)}$ of 5.3.3. He is also prepared to settle for a u_j for which $F\{x(\underset{\sim}{\tilde{p}}_{(t)}), u_j\}$ is large if the computations in finding an optimal $u^{(i)}$ is time consuming. He considers several possibilities for making such a choice including choosing u_j such that $F\{x(\underset{\sim}{\tilde{p}}_{(t)}), u_j\} > \delta$ for some $\delta > 0$ or choosing u_j such that $F\{x(\underset{\sim}{\tilde{p}}_{(t)}), u_j\} > \beta\, F\{x(\underset{\sim}{\tilde{p}}_{(t)}), u^{(1)}\}$ where $0 < \beta < 1$

Finally he opts for an optimal step from $\underset{\sim}{\tilde{p}}_{(t)}$ to his chosen u_j, and he contemplates a number of standard algorithms for computing these.

Wu (1976, 1978b) also provides various convergence results relating to his resultant global algorithm for general criteria. He does not though, indulge in any discussion about the choice of the initial $\mathcal{U}_{(1)}$.

Which of the above two options are better, the approaching from below or forward method using the "(eliminating)/Augmenting" procedure based on a confident identification of a subset $\mathcal{U}_{(1)}$ of $\mathrm{Sup}(p^*)$, or the approaching from above or backward method based on a confident identification of a small subset $\overline{\mathcal{U}}$ of the vertices, but one which is big enough to contain $\mathrm{Sup}(p^*)$.

On balance the latter approach would seem better. From the outset the computations with the "(eliminating)/Augmenting" procedure are more complex, for there is no point in adopting this approach unless the more efficient algorithms are used to calculate $\underset{\sim}{p}^*_{(t)}$. In contrast, with the backward approach, one could start out using the computationally simpler but less efficient diagonalised versions of the steepest ascent or Newton methods, while clearly having available the option of switching to one of the more efficient algorithms.

There is one further argument in favour of the backward method in the design context. We know in the case of design criteria based on the matrix $AM^+(p)A'$ that the optimum can be nondifferentiable, having a support consisting of fewer than k vertices. However it would seem undesirable to invite nondifferentiability too early. The forward method could not in general proceed from an initial $\mathcal{U}_{(1)}$ containing $T < k$ points. While the optimum $\underset{\sim}{p}^*_{(1)}$ on $\mathcal{U}_{(1)}$ could be found by the above methods, if the criterion enjoys support differentiability, we would not in general be able to identify an augmenting vertex if it is suspected that $\underset{\sim}{p}^*_{(t)} \neq p^*$, for then $F\{M(\underset{\sim}{p}^*_{(t)}), v_j v_j'\} = -G(M_*, M_*)$, $M_* = M(\underset{\sim}{p}^*_{(t)})$, for all $v_j \notin L\{\mathrm{Sup}(\underset{\sim}{p}^*_{(t)})\}$.

The choice of the backward method in fact envisages a three stage procedure, namely opening with a vertex direction algorithm, switching to an intermediate algorithm and finally turning to a high-powered technique, the three methods satisfying the following.

(i) The vertex direction method as we have seen can be slow to converge but is computationally able to cope with large numbers of vertices. Most importantly though, it can quickly eliminate most of a large set of vertices as optimum support points.

(ii) The intermediate method should ideally be one that can cope with a small number of non-optimal-support points, and can take in its stride the further elimination of vertices when weights are formally put to zero using the rules of section 4.3.3. As a consequence it may not be high-powered, although it should be more efficient than vertex direction iterations.

(iii) The third stage method should be one that is suited to finding the optimum on a set of vertices which include no non-optimal-support points. It should be high-powered such as adaptations of Newton-Raphson techniques. Its efficiency would be impaired if applied to a set of vertices which did contain some non-optimal-support points.

If appropriate an added extra could be to collapse clusters of points before passing from the intermediate method to the final method.

We are therefore contemplating the following four-stage composite algorithm.

A COMPOSITE ALGORITHM

Stage	Technique
1	Vertex direction method
2	Intermediate method
3	Collapse clusters
4	High-powered method

Hence there will be no determined effort to compare different types of algorithms although some empirical comparisons are reported at the end of chapter 7.

We have already discussed vertex direction algorithms in detail. Also we have illustrated how to adapt standard high-powered iterations for corresponding unconstrained problems, when the only active constraint is $\sum p_i = 1$, as is envisaged at the third stage.

We have further suggested that some of the simpler steepest ascent, quasi Newton algorithms might make suitable intermediate stage methods. However there can be a disadvantage with these methods when a weight is set to zero. The inner product matrix A will require to be redefined and if it is not diagonal its inverse will require to be recalculated.

However minor these difficulties may be, in the next chapter we propose other possibilities for intermediate stage methods and general ideas for formulating these.

First we conclude this chapter with some results illustrating Silvey and Titterington's "(eliminating)/Augmenting" procedure in action, in the design context.

Recall that the principle of this technique is to select a subset $\mathcal{V}_{(1)}$ of the design space and to find $\underline{p}^*_{(1)}$, the optimum on $\mathcal{V}_{(1)}$. Then subsequently we calculate $\underline{p}^*_{(t)}$, the optimum on the set $\mathcal{V}_{(t)}$ with

$$\mathcal{V}_{(t+1)} = \mathrm{Sup}\{\underline{p}^*_{(t)}\} \cup \{v^{(j)}\} \quad , \quad v^{(j)} \in \{v_i \in \mathcal{V} : F\{M(\underline{p}^*_{(t)}), v_i v_i'\} > 0\}$$

This procedure was used to calculate D-optimal designs in several examples. In each case the initial set $\mathcal{V}_{(1)}$ consists of k linearly independent vertices so that $\underline{p}^*_{(1)}$ assigns weight 1/k. The augmenting vertex is taken to be $v^{(1)}$, that which maximises $F\{M(\underline{p}^*_{(t)}), vv'\}$ over $\mathcal{V}$.

We consider again the case of the discretised trigonometric regression design space

$$\mathcal{V} = \{v_{(x)} = (x, x^2, \sin 2\pi x, \cos 2\pi x)' \quad : x \in \mathcal{X}_d\}$$

$$\mathcal{X}_d = \{0, .01, .02, \ldots, .99, 1\},$$

for which $\mathrm{Sup}(p^*) = \{v_{(x)} : x = .08, .09, .38, .73, .74, 1\}$. If $\mathcal{V}_{(1)} = \{v_{(x)} : x = 0, .33, .67, 1\}$, the "eliminating/augmenting" procedure realises the sequence $\mathcal{V}_{(t)} = \{v_{(x)} : x \in \mathcal{X}_{(t)}\}$ for the sets $\mathcal{X}_{(t)}$

listed in Table 6.5.1.

The initial set $\mathcal{U}_{(1)}$ is not a good choice, containing as it does only one member of Sup(p^*). As a consequence the procedure does not perform efficiently. On nine occasions non-optimal-support points are the augmenting vertices. Subsequently they have to be eliminated as do the three non-optimal-support points in $\mathcal{U}_{(1)}$. In consequence the number of points in $\mathcal{U}_{(t)}$ fluctuates. The five elements of Sup(p^*) not in $\mathcal{U}_{(1)}$ are the 7th,10th,12th,13th,14th augmenting points.

Consider now the discretised polynomial regression design space.

$$\mathcal{U} = \{v_{(x)} = (1, x, x^2, \ldots, x^{k-1})' \ : \ x \in \mathcal{X}_d\},$$

$$\mathcal{X}_d = \{-1, -.99, -.98, \ldots, .99, 1\}.$$

In the case $k = 4$, $\text{Sup}(p^*) = \{v_{(x)} \ : \ x = \pm.44, \pm.45, \pm 1\}$

If $\mathcal{U}_{(1)} = \{v_{(x)} \ : \ x = \pm.33, \pm 1\}$ the procedure realises the sequence $\mathcal{U}_{(t)} = \{v_{(x)} \ : \ x \in \mathcal{X}_{(t)}\}$ for the sets $\mathcal{X}_{(t)}$ listed in Table 6.5.2.

Again $\mathcal{U}_{(1)}$ is not a wise choice although it contains two points of Sup(p^*). Twelve augmentations are necessary and there is the following unfortunate feature. The element of Sup(p^*) corresponding to $x = .44$ is the 7th augmenting vertex. However it is subsequently eliminated in the formulation of $\mathcal{U}_{(10)}$ and has to augment again at the last stage.

We note that in the calculation of $\underline{p}^*_{(t)}$, for each t, the dilated Newton Raphson iterates 6.4.3 were initially used in these two examples and in the remaining examples below. Of relevance to the above examples is that the test $p_j^{(r)} < \cdot 01$ successfully identified when Sup($\underline{p}^*_{(t)}$) was a strict subset of $\mathcal{U}_{(t)}$.

We have reported the above results to give an impression of how the procedure behaves, but its use was only contemplated assuming $\mathcal{U}_{(1)} \subset \text{Sup}(p^*)$.

When $\mathcal{U}_{(1)}$ consists of the k vertices in Sup(p^*) with the k largest optimal weights, then no non-optimal-support points are employed as augmenting vertices in the case of the trigonometric regression design space, and in the case of the polynomial regression design spaces for $k = 4,5,8$. The procedure successfully adds in those elements of Sup(p^*) which are not in $\mathcal{U}_{(1)}$ without any eliminations. In the cases of $k = 6,9$ of the polynomial example, respectively one and two non-members of Sup(p^*) are employed as augmenting points. In each case the points are immediate neighbours of elements of Sup(p^*). In the case of $k = 7$ Sup(p^*) contains exactly k vertices, there being no clusters, so that $\mathcal{U}_{(1)}$ = Sup(p^*) under the condition imposed on the choice of $\mathcal{U}_{(1)}$.

We conclude this discussion by recording that we identified this choice of $\mathcal{U}_{(1)}$ in each example by employing Wynn's vertex direction algorithm $V\{v^{(1)}, 1/(k+r+1)\}$ until its iterations had moved towards or "called" k vertices at least 10 times, presence in Sup($p^{(0)}$) being regarded as one "call". Sup($p^{(0)}$) consisted of the k vertices $v_{(x)}$ corresponding to the set of k equally spaced values of x which includes the end points of $\mathcal{X}_d$, and $p^{(0)}$ allocated equal weights to its support points.

TABLE 6.5.1

t	$\mathcal{X}_{(t)}$
1	$\{0,.33,.67,1\}$
2	$\{0,.13,.33,.67,1\}$
3	$\{0,.13,.33,.67,.77,1\}$
4	$\{0,.13,.43,.77,1\}$
5	$\{0,.06,.13,.43,.77,1\}$
6	$\{.06,.13,.35,.43,.77,1\}$
7	$\{.06,.35,.43,.71,.77,1\}$
8	$\{.06,.35,.43,.71,.74,1\}$
9	$\{.06,.10,.35,.43,.74,1\}$
10	$\{.06,.10,.39,.74,1\}$
11	$\{.08,.10,.39,.74,1\}$
12	$\{.08,.10,.37,.39,.74,1\}$
13	$\{.08,.37,.39,.73,.74,1\}$
14	$\{.08,.37,.38,.73,.74,1\}$
15	$\{.08,.09,.38,.73,.74,1\}$

TABLE 6.5.2

$\underline{t}$	$\underline{\mathcal{X}_{(t)}}$
1	$\{-1,-.33,.33,1\}$
2	$\{-1,-.33,.53,1\}$
3	$\{-1,-.49,-.33,.53,1\}$
4	$\{-1,-.49,.40,1\}$
5	$\{-1,-.49,.40,.46,1\}$
6	$\{-1,-.49,-.42,.46,1\}$
7	$\{-1,-.46,-.42,.46,1\}$
8	$\{-1,-.46,.44,1\}$
9	$\{-1,-.46,-.44,.44,1\}$
10	$\{-1,-.46,-.44,.45,1\}$
11	$\{-1,-.45,-.44,.45,1\}$
12	$\{-1,-.45,-.44,.44,.45,1\}$

CHAPTER 7

INTERMEDIATE ALGORITHMS

§7.1 A First Class Of Algorithm

§7.1.1 We now consider some ideas for formulating an algorithm for problem (P2) which is less complex computationally than steepest ascent or adapted Newton Raphson, but not as simple as a vertex direction algorithm. By less complex is mainly meant that inverses of J×J matrices at each iteration is not required.

Still the algorithm should aim to identify an optimal p^*, and we are still thinking of the approach which, for some $m^{(r)}$ or some $q^{(r)}$, derives iterate $p^{(r+1)}$ from $p^{(r)}$ according to the rule

$$p^{(r+1)} = \begin{cases} p^{(r)} + \alpha_r m^{(r)} & : \ \underline{1}' m^{(r)} = 0 \ , \ F\{p^{(r)}, p^{(r)} + m^{(r)}\} > 0 \\ (1-\alpha_r) p^{(r)} + \alpha_r q^{(r)} & : \ \underline{1}' q^{(r)} = 1 \ , \ F\{p^{(r)}, q^{(r)}\} > 0 \\ (1+\alpha_r) p^{(r)} - \alpha_r q^{(r)} & : \ \underline{1}' q^{(r)} = 1 \ , \ F\{p^{(r)}, q^{(r)}\} < 0 \quad (\alpha_r > 0) \end{cases}$$

We want a computationally simple formula for $m^{(r)}$, $q^{(r)}$.

Consider from 2.3.5 that, under the restriction $1'm^{(r)} = 0$, it is the case that

7.1.1 $$F_\phi\{p^{(r)}, p^{(r)} + m^{(r)}\} = \sum_{j=1}^{J} m_j^{(r)} F_\psi\{x(p^{(r)}), u_j\} .$$

Clearly, for the given values of $F\{x(p^{(r)}), u_j\}$, there will be many choices of $m_1^{(r)}, \ldots, m_J^{(r)}$ which will guarantee that the directional derivative here is positive. The formula 7.1.1 makes it fairly easy to identify suitable $m^{(r)}$. Our problem is more one of deciding which of these should we select. Clearly some will be 'better' than others, the 'best' being constrained steepest ascent, but we wish to exclude selection both of the latter unless it is simple, and also of vertex direction iterations.

In principle our choice must be arbitrary, though obviously one would try to determine it on the basis of some restricted or indirect optimality considerations, or choose it to have some intuitively desirable features. However such criteria will not usually be enjoyed by just one $m^{(r)}$ and must themselves be an arbitrary choice.

We consider now, and in the next section, some examples of reasonable $m^{(r)}$ or $q^{(r)}$.

Let

$$\mathcal{U} = \{u_1, \dots\dots, u_J\}$$
$$\mathcal{U}_r^+ = \{u_j : u_j \in \mathcal{U},\ F\{x(p^{(r)}), u_j\} > 0\}$$
$$\overline{\mathcal{U}}_r^+ = \{u_j : u_j \in \mathrm{Sup}(p^{(r)}),\ F\{x(p^{(r)}), u_j\} > 0\}$$
$$\overline{\mathcal{U}}_r^- = \{u_j : u_j \in \mathrm{Sup}(p^{(r)}),\ F\{x(p^{(r)}), u_j\} < 0\}.$$

Let $\mathcal{V}, \mathcal{V}_r^+, \overline{\mathcal{V}}_r^+, \overline{\mathcal{V}}_r^-$ denote similar terms in the design context.

§7.1.2 A first suggestion is to choose $m^{(r)}$ such that

7.1.2 $$m_j^{(r)} \begin{cases} \geq 0 & \text{if } u_j \in \overline{\mathcal{U}}_r^+ \\ \leq 0 & \text{if. } u_j \in \overline{\mathcal{U}}_r^- \\ = 0 & \text{else.} \end{cases}$$

In view of 7.1.1 this clearly guarantees that $F(p^{(r)};p^{(r)}+m^{(r)})$ is a sum of positive terms. Clearly many $m^{(r)}$ will satisfy 7.1.2. The suggestion has the motivation that, if $F\{x(p^{(r)}),u_t\} > 0$, it would appear that we wish to move towards or nearer to u_t, which is done by increasing the weight at u_t and vice versa if $F\{x(p^{(r)}),u_t\} < 0$.

A more natural conclusion from this picture goes further and suggests that the larger $|F\{x(p^{(r)}),u_t\}|$ the larger should be $m_t^{(r)}$; that is the larger should be the move, nearer or farther, from u_t as appropriate; one might though argue that $F\{x(p^{(r)}),u_t\}$ should be replaced by $F^A\{x(p^{(r)}),u_t\}$ in this argument.

The following are examples which fall into this class or which partially satisfy the above picture, and which in some instances have in principle been recommended in the design context.

7.1.3 $$m_j^{(r)} = \begin{cases} F\{x(p^{(r)}), u_j\}/D_r^+ & \text{if} \quad u_j \in \overline{\mathcal{U}}_r^+ \\ F\{x(p^{(r)}), u_j\}/D_r^- & \text{if} \quad u_j \in \overline{\mathcal{U}}_r^- \\ 0 & \quad \text{otherwise} \end{cases}$$

7.1.4 $$m_j^{(r)} = \begin{cases} F\{x(p^{(r)}), u_j\}/D_r^+ & \text{if} \quad u_j \in \overline{\mathcal{U}}_r^+ \\ -1 & \quad u_j = u^{(2)} \\ 0 & \quad \text{otherwise} \end{cases}$$

$$7.1.5 \quad m_j^{(r)} = \begin{cases} 1 & , \quad u_j = u^{(1)} \\ F\{x(p^{(r)}), u_j\}/D_r^- & \text{if} \quad u_j \in \overline{U}_r^- \\ 0 & \text{otherwise} \end{cases}$$

$$7.1.6 \quad m_j^{(r)} = \begin{cases} 1 & , \quad u_j = u^{(1)} \\ -1 & , \quad u_j = u^{(2)} \\ 0 & \text{otherwise} \end{cases}$$

where $D_r^+ = \sum_{u_j \in \overline{U}_r^+} \{F\{x(p^{(r)}), u_j\}\}$

$$D_r^- = -\sum_{u_j \in \overline{U}_r^-} \{F\{x(p^{(r)}), u_j\}\}$$

We make some comments on these.

(i) The last suggestion is in fact the bi-vertex direction considered in section 5.4, while suggestion 7.1.4 is a variation of an improvement suggested by St. John and Draper (1975) to Atwood's $V\{v^{(12)}, \alpha_r^*(v^{(12)})\}$ in the case of D-optimality. Under this vertex direction scheme the optimal step is either taken towards $v^{(1)}$ or away from $v^{(2)}$, whichever leads to the largest increase in $\psi(\cdot)$. If the latter option is the one which is adopted, i.e. $v^{(12)} = v^{(2)}$, then under $V\{v^{(12)}, \alpha_r^*(v^{(12)})\}$ the weight which is removed from $v^{(2)}$ is distributed among all of the remaining support points of $p^{(r)}$ in a manner proportional to $p_j^{(r)}, v_j \neq v^{(2)}$. St.J.&D. recommended that the removed weight should be allocated only to those $u_j \in \mathrm{Sup}(p^{(r)})$ such that $F\{x(p^{(r)}), u_j\} > 0$ and in a manner proportional to the values of these positive $F\{x(p^{(r)}), u_j\}$. This is then an iteration in the direction of 7.1.4 where the steplength is fixed by the value of $\alpha_r^*(v^{(12)})$.

While this may be as reasonable a manner as any for selecting a non-optimal steplength, yet it is not clear that the value obtained would compare with a steplength derived directly from $m^{(r)}$ by one of the methods discussed previously. Certainly there is no guarantee that St.J.&D.'s $\psi\{x(p^{(r+1)})\}$ will exceed $\psi\{x(p^{(r)})\}$.

(ii) Suggestion 7.1.5 is a complementary one to 7.1.4. It is a direction in which weight, added to a vertex, is removed only from those support vertices such that $F\{x(p^{(r)}),u_j\} < 0$, and in a manner proportional to the numerical value of these negative directional derivatives. It is an 'improvement' on a forward vertex direction iteration. Now we cannot necessarily let the (optimal) steplength, that the vertex direction would have taken towards $u^{(1)}$, determine the steplength in the direction $m^{(r)}$ as above, for we then may pass out of the feasible region $\mathcal{P}$

(iii) Suggestion 7.1.3 is clearly a fusion of 7.1.4 and 7.1.5 and would seem a more sensible choice, if one is going to adopt a multivertex direction 'improvement' on a vertex direction iteration. There is little computationally to choose between 7.1.3, 7.1.4, 7.1.5.

(iv) Atwood (1973) also suggests improvements to vertex direction iterations including the following special one which could be shown to define an $m^{(r)}$ similar to the above. In the case of D_A-optimality it is known that $p_j^* \leq 1/s$. If then $p_j^{(r)} > 1/s$, remove the weight in excess of $1/s$ from such u_j's and redistribute it among the other design points. He did not suggest any particular method of distribution but again it would be appropriate to distribute only to those u_j such that $F\{x(p^{(r)}), u_j\} > 0$, and in a manner such that no 'augmented' weight exceeds $1/s$. The case $s = k$ is D-optimality.

(v) Variations on 7.1.3, 7.1.4, 7.1.5, 7.1.6 would be obtained by replacing $F(\cdot,\cdot)$ by $F^A(\cdot,\cdot)$; by replacing $\overline{\mathcal{U}}_r^+$ by $\mathcal{U}_r^+$ or by $\{\overline{\mathcal{U}}_r^+\} \cup \{\mathcal{W}_r\}$ where $\mathcal{W}_r$ is a subset of $\mathcal{U}_r^+ - \overline{\mathcal{U}}_r^+$ and finally by replacing $u^{(1)}$, $u^{(2)}$ by other appropriate vertices.

(vi) The condition $\underline{1}'m^{(r)} = 0$ is ensured in 7.1.3,...,7.1.6, because $\sum_{u_j \in \overline{\mathcal{U}}_r^+} m_j^{(r)} = 1$, $\sum_{u_j \in \overline{\mathcal{U}}_r^-} m_j^{(r)} = -1$.
If we do replace $\overline{\mathcal{U}}_r^+$ by $\mathcal{U}_r^+$ it is of interest that 7.1.6 maximises $F\{p^{(r)}, p^{(r)}+m^{(r)}\}$ over $m^{(r)}$ satisfying these two conditions.

(vii) A general rule for devising $m^{(r)}$ satisfying 7.1.2 is given by

7.1.7 $$m_j^{(r)} = \begin{cases} g\{F\{x(p^{(r)}), u_j\}\}/g_r^+ & \text{if } u_j \in \overline{\mathcal{U}}_r^+ \\ g\{\quad \cdot\cdot \quad\}/g_r^- & \text{if } u_j \in \overline{\mathcal{U}}_r^- \\ 0 & \text{otherwise} \end{cases}$$

where $g(\cdot)$ is any function such that $g(x) > 0$ if $x > 0$, $g(x) < 0$ if $x < 0$, while $g_r^+ = \sum_{u_j \in \overline{\mathcal{U}}_r^+} g\{F\{x(p^{(r)}), u_j\}\}$

$$g_r^- = -\sum_{u_j \in \overline{\mathcal{U}}_r^-} g\{\quad \text{"} \quad\}$$

It would seem desirable to impose the restriction that $g(\cdot)$ be monotonic increasing or be at least nondecreasing. Clearly 7.1.3,.., 7.1.6 are particular cases of this.

Again one might extend $\overline{\mathcal{U}}_r^+$ or replace $F(\cdot,\cdot)$ by $F^A(\cdot,\cdot)$.

(viii) A choice of $q^{(r)}$ corresponding to 7.1.2 is given by $q^{(r)}$ such that $\underline{1}'q^{(r)} = 1$ and

7.1.8 $$q_j^{(r)} \begin{cases} \geq p_j^{(r)} & \text{if } u_j \in \overline{\mathcal{U}}_r^+ \\ \leq p_j^{(r)} & \text{if } u_j \in \overline{\mathcal{U}}_r^- \\ 0 & \text{otherwise} \end{cases}$$

Clearly $m^{(r)} = q^{(r)} - p^{(r)}$ satisfies 7.1.2 so that $F(p^{(r)}, q^{(r)}) > 0$ and one would take $p^{(r+1)} = (1-\alpha_r)p^{(r)} + \alpha_r q^{(r)}$.

§7.2 Further Classes; A Positive Covariance

§7.2.1 A second rule which will guarantee that 7.1.1 is positive is

Rule 7.2.1 Choose $m^{(r)}$ to satisfy

(i) $1'm^{(r)} = 0$, (ii) $m_j^{(r)} = 0$ if $F\{x(p^{(r)}), u_j\} < 0$ and $u_j \notin \mathrm{Sup}(p^{(r)})$,

and (iii) assign non-zero values to at least two of the remaining components of $m^{(r)}$ in such a way that if u_j has the n^{th} largest $F\{x(p^{(r)}), u_j\}$ among vertices corresponding to non-zero components of $m^{(r)}$, then $m_j^{(r)}$ is the n^{th} largest among these non-zero components of $m^{(r)}$; or is at least as large as (n-1) of these components.

That this ensures that 7.1.1 is at least nonnegative is seen as follows:

We have from 7.1.1
$$F\{p^{(r)}, p^{(r)} + m^{(r)}\} = \sum_{i=1}^{n} x_i y_i$$

where n is the number of non-zero components of $m^{(r)}$, x_j is the j^{th} largest of these, y_j is the j^{th} largest among the corresponding vertex directional derivatives so that

$$x_1 \leq x_2 \leq \cdots \leq x_n \quad , \quad y_1 \leq y_2 \leq \cdots \leq y_n .$$

Also $\sum_{i=1}^{n} x_i = \sum_{j=1}^{n} m_j^{(r)} = 0$

$$\therefore \quad F_{\phi}\{p^{(r)}, p^{(r)} + m^{(r)}\} = S_{xy}^{n}$$

where $S_{xy}^{(n)}$ is the sample covariance of the pairs $(x_1, y_1), \ldots, (x_n, y_n)$

i.e. for $\bar{x}_n = \left(\sum_{i=1}^{n} x_i\right)/n$, $\bar{y}_n$ etc.,

$$S_{xy}^{(n)} = \sum_{i=1}^{n} (x_i - \bar{x}_n)(y_i - \bar{y}_n) = \sum_{i=1}^{n} x_i y_i - n\bar{x}_n \bar{y}_n .$$

An elementary derivation establishes that

$$S_{xy}^{(t+1)} = S_{xy}^{(t)} + \{t/(t+1)\}(x_{t+1} - \bar{x}_t)(y_{t+1} - \bar{y}_t)$$
$$\geq S_{xy}^{(t)} \quad \text{if} \quad x_{t+1} \geq \bar{x}_t , \; y_{t+1} \geq \bar{y}_t$$

Hence for $x_1 \leq x_2, \ldots, \leq x_n$, $y_1 \leq y_2, \ldots, \leq y_n$, $S_{xy}^{(t+1)} \geq S_{xy}^{(t)}$ for $t = 2,\ldots,n-1$ and trivially $S_{xy}^{(1)} = 0$.

Part (iii) of the rule could equivalently be restated by replacing $F\{x(p^{(r)}),u_j\}$ by $G\{x(p^{(r)}),u_j\}$. This follows either from the fact that $F\{x(p^{(r)}),u_j\} = G\{x(p^{(r)}),u_j\} - G\{x(p^{(r)}),x(p^{(r)})\}$ or more powerfully from the fact that for any $m^{(r)}$

$$F_\phi\{p^{(r)}, p^{(r)}+m^{(r)}\} = \sum_{j=1}^{J} m_j^{(r)} G_\psi\{x(p^{(r)}), u_j\} = \sum_{j=1}^{J} m_j^{(r)} d_j$$

where $d_j = G_\psi\{x(p^{(r)}), u_j\} = \partial\phi/\partial p_j^{(r)}$, $\phi(p) = \psi\{x(p)\}$.

What the rule is advocating is a very particular scheme of ensuring what we must have in general for any direction $m^{(r)}$ in which we opt to move; namely that, in view of the requirement that $\underset{\sim}{1}'m^{(r)} = 0$, the sample covariance or correlation between the $m_j^{(r)}$ and the $F\{x(p^{(r)}),u_j\}$ or between the $m_j^{(r)}$ and the $G\{x(p^{(r)}),u_j\}$ must be positive.

Clearly directions 7.1.3,..,7.1.6 are particular instances in which the conditions of rule 7.2.1 are satisfied. However the rule suggests a less restricted $m^{(r)}$ than these, for it allows of an $m^{(r)}$ which could remove weight from a vertex u_j such that $F\{x(p^{(r)}),u_j\} > 0$ or vice versa. Such u_j though would typically have numerically small $F\{x(p^{(r)}),u_j\}$ and so we should be less convinced about the status of such vertices. An example that could do this is $m(I,\underset{\sim}{1}',d)$; that is

7.2.1 $$m_j^{(r)} = \begin{cases} d_j - \bar{d} & u_j \in \mathrm{Sup}(p^{(r)}) \\ 0 & \text{otherwise} \end{cases}$$

Generalisations of 7.2.1 which will also fall into this category are

7.2.2 $$m_j^{(r)} = \begin{cases} g(d_j) - \bar{g}_d & u_j \in Sup(p^{(r)}) \\ 0 & u_j \notin Sup(p^{(r)}) \end{cases}$$

and

7.2.3 $$m_j^{(r)} = \begin{cases} g(F_j) - \bar{g}_F & u_j \in Sup(p^{(r)}) \\ 0 & u_j \notin Sup(p^{(r)}) \end{cases}$$

where $F_j = F\{x(p^{(r)}),u_j\}$, $g(\cdot)$ is a nondecreasing function and $\bar{g}_d$, $\bar{g}_F$ are appropriate averages ensuring $\underline{1}'m^{(r)} = 0$. In the case $g(x) = x$ both reduce to 7.2.1.

§ 7.2.2 The following generalisation of Rule 7.2.1 is one to which a number of iterations enjoying alternative motivations conform.

Rule 7.2.2 Choose $m^{(r)} = Bz^{(r)}$ where B is a diagonal matrix, $B = \text{diag}\{b_1,\ldots,b_J\}$, $b_j > 0$ such that (i) $\underline{1}'m^{(r)} = 1'Bz^{(r)} = 0$, (ii) $z_j^{(r)} = 0$ if $F\{x(p^{(r)}),u_j\} < 0$ and $u_j \notin Sup(p^{(r)})$, and (iii) assign values to the components of $z^{(r)}$ as in (iii) of Rule 7.2.1.

This may not seem a particularly intuitive rule but again the motivation is that this guarantees positive covariance between the $m_j^{(r)}$ and $F\{x(p^{(r)}),u_j\}$.

We now have

$$F\{p^{(r)}, p^{(r)} + m^{(r)}\} = \sum_{i=1}^{n} b_i x_i y_i$$

where n is the number of non-zero components of $z^{(r)}$, x_j is the j^{th} largest component of $z^{(r)}$, y_j is the j^{th} largest among the corresponding vertex directional derivatives, so that $x_1 \leq x_2 \leq \ldots \leq x_n$, $y_1 \leq y_2 \leq \ldots \leq y_n$. Also $\sum b_j x_j = \sum b_j z_j^{(r)} = \sum m_j^{(r)} = 0$. Hence

$$F\{p^{(r)}, p^{(r)} + m^{(r)}\} = \left(\sum_{i=1}^{n} b_i\right)\sum_{j=1}^{n} w_j x_j y_j$$

where $w_j = b_j \Big/ \left(\sum_{i=1}^{n} b_i\right) > 0$ and $\sum_{i=1}^{n} w_i = 1$, $\sum w_i x_i = 0$.

Hence we have

$$F\{p^{(r)}, p^{(r)} + m^{(r)}\} = C(x,y),$$

where $C(x,y)$ is the covariance between the discrete random variables x,y with the joint probability distribution

$$p(x,y) = \begin{cases} w_i & , (x,y) = (x_i, y_i), \ i = 1, \dots, n \\ 0 & \text{otherwise} \end{cases}$$

For such a distribution

$$C(x,y) = \sum_{i=1}^{n} w_i x_i y_i - \left(\sum_{i=1}^{n} w_i x_i\right)\left(\sum_{i=1}^{n} w_i y_i\right)$$

Let $C^t(x,y)$ denote the covariance between x and y given that $x \in \{x_1,\dots,x_t\}$, $y \in \{y_1,\dots,y_t\}$; that is

$$C^t(x,y) = \sum_{i=1}^{t} \tilde{w}_i x_i y_i - E_t(x)E_t(y)$$

where $\tilde{w}_i = w_i \Big/ \sum_{j=1}^{t} w_j \quad , \quad i = 1, \dots, t$

and $E_t(x) = \sum_{i=1}^{t} \tilde{w}_i x_i \quad , \quad E_t(y) = \sum_{i=1}^{t} \tilde{w}_i y_i$.

are the corresponding conditional expectations.

Let also $Q = w_{t+1} \Big/ \sum_{i=1}^{t} w_i$

Then it is the case that

$$C^{(t+1)}(x,y) = (1+Q)^{-1}\left\{C^t(x,y) + [x_{t+1} - E_t(x)][y_{t+1} - E_t(y)]Q/(1+Q)\right\}$$

Hence $C^{(t+1)}(x,y) \geq 0$ if $C^t(x,y) \geq 0$ and $x_{t+1} \geq E_t(x)$, $y_{t+1} \geq E_t(y)$. Trivially $C^{(1)}(x,y) = 0$ and hence $C^{(t)}(x,y) \geq 0$, $t = 1,\dots,n$, if $x_1 \leq x_2 \leq, \dots, \leq x_n$, $y_1 \leq y_2 \leq, \dots, \leq y_n$ and clearly $C(x,y) = C^{(n)}(x,y)$.

We are contemplating then $m^{(r)}$'s of, say, the form

7.2.4 $$m_j^{(r)} = \begin{cases} b_j z_j^{(r)} & , u_j \in Sup(p^{(r)}) \\ 0 & u_j \notin \quad " \end{cases}$$

where $b_j > 0$ and $z_j^{(r)}$ 'increases' with $F_j = F\{x(p^{(r)}), u_j\}$ or with $d_j = G_j = G\{x(p^{(r)}), u_j\}$.

Examples include $m(A, \underline{1}', d)$ for positive definite $A = \text{diag}\{a_1, \dots, a_J\}$ which as we have already said provides a simple enough formula for $m^{(r)}$, namely

§7.3 Fixed Point Directions And Algorithms

§7.3.1 Conditions corresponding to Rules 7.2.1 and 7.2.2 for selection of a $q^{(r)}$ towards which to move and away from which to move from $p^{(r)}$ are respectively:

Rule 7.3.1a Choose $q^{(r)}$ such that (i) $\underset{\sim}{1}' q^{(r)} = 1$,
(ii) $q_j^{(r)} = 0$ if $F\{x(p^{(r)}), u_j\} < 0$, $u_j \notin \text{Sup}(p^{(r)})$
(iii) $m^{(r)} = q^{(r)} - p^{(r)}$ satisfies the remaining relevant conditions.

Rule 7.3.1b (i), (ii) as above and (iii) with $m^{(r)} = p^{(r)} - q^{(r)}$.

Clearly the suggested $m^{(r)}$'s of sections 7.2.1, 7.2.2 define $q^{(r)}$'s satisfying rule 7.3.1a. However the resultant $q^{(r)}$ may have negative components. This is not so with the following suggestion.

For some positive function $h(z,\delta)$ which, for fixed δ, is nondecreasing in the first argument z, take $q^{(r)}$ to be the vector $q^{(r)}\{h(z,\delta)\}$ whose components are

7.3.1
$$q_j^{(r)} = p_j^{(r)} h(z_j, \delta) \Big/ \sum_{i=1}^{J} p_i^{(r)} h(z_i, \delta)$$

where either $z_j = F_j = F_\psi\{x(p^{(r)}), u_j\}$

or $z_j = d_j = \partial\phi/\partial p_j^{(r)} = G_\psi\{x(p^{(r)}), u_j\}$.

If for fixed δ, $h(z,\delta)$ is nonincreasing in z then this defines a $q^{(r)} \in \mathcal{P}$ satisfying Rule 7.3.1b.

Clearly the advantage of the extra variable or free parameter δ is that 7.3.1 defines not just one but a whole range of possible $q^{(r)}$'s.

Taking the two cases together we have

7.3.2
$$m^{(r)} = \begin{cases} q^{(r)}\{h(z,\delta)\} - p^{(r)} & , \ \partial h/\partial z > 0 \\ p^{(r)} - q^{(r)}\{h(z,\delta)\} & , \ \partial h/\partial z < 0 \end{cases}$$

and so the resultant iterative rule is

7.3.3 $$p^{(r+1)} = \begin{cases} (1-\alpha_r)p^{(r)} + \alpha_r q^{(r)}\{h(z,\delta)\} & , \partial h/\partial z > 0 \\ (1+\alpha_r)p^{(r)} - \alpha_r q^{(r)}\{h(z,\delta)\} & , \partial h/\partial z < 0 \end{cases}$$

We note the following about such iterations

(i) An iteration can change the weights of only the current support.

(ii) In consequence $p^{(0)}$ must assign non-zero weight to all points of $\mathcal{U}$ or $\mathcal{V}$ under consideration; $\mathrm{Sup}(p^{(0)}) = \mathcal{U}$ or $\mathcal{V}$.

(iii) If z = d or z = F, 7.3.3 will be well defined at a $p^{(r)}$ at which only support differentiability is enjoyed.

(iv) If $p^{(r)} \in \mathcal{P}$, then $p^{(r)}$ is a fixed point of the mapping $p^{(r)} \longrightarrow q^{(r)}\{h(z,\delta)\}$ if the z_j, i.e. d_j or F_j, corresponding to the support points of $p^{(r)}$ are all equal. Hence p^* and all optima solving (P2) for subsets of $\mathcal{U}$ or $\mathcal{V}$ are such fixed points, and these will be the only ones at a given δ if $h(z,\delta)$ is strictly monotonic in z for that δ.

We might refer to the $m^{(r)}$ of 7.3.2 as a fixed point direction. We will denote an algorithm employing the iterative rule 7.3.3 by $FP\{h(z,\delta),\alpha_r\}$

We study examples of such a scheme in the remainder of this chapter and in chapters 8 and 9.

§7.3.2 In this section we examine some properties relating to $FP\{h(z,\delta),\alpha_r\}$ when, for z = d or z = F, $h(z,\delta) = \exp\{\delta g(z)\}$, where g(z) is nondecreasing in z. In particular we will see some nice properties relating to the free parameter δ. First some examples.

The resultant $q^{(r)} = q^{(r)}\{h(z,\delta)\}$ might be

7.3.4 $$q_j^{(r)} = p_j^{(r)}\exp\{\delta g(d_j)\} \Big/ \sum p_i^{(r)}\exp\{\delta g(d_i)\}$$

7.3.5 $$q_j^{(r)} = p_j^{(r)}\exp\{\delta g(F_j)\} \Big/ \sum p_i^{(r)}\exp\{\delta g(F_i)\}$$

7.3.6 $$q_j^{(r)} = p_j^{(r)}\exp\{\delta d_j\} \Big/ \sum p_i^{(r)}\exp\{\delta d_i\}$$

7.3.7 $$q_j^{(r)} = p_j^{(r)}\exp\{\delta F_j\} \Big/ \sum p_i^{(r)}\exp\{\delta F_i\}$$

The function $g(z)$ must of course be defined at all the possible values of d_j or F_j. The function $g(z) = \ln z$ will not satisfy this in respect of the F_j some of which must be negative at nonoptimal $p^{(r)}$. However, if a function $\phi(p)$ enjoys positive derivatives d_j, then 7.3.4 will be defined in this instance. The resultant $q^{(r)}$ is

7.3.8 $$q_j^{(r)} = p_j^{(r)} d_j^{\delta} \Big/ \sum p_i^{(r)} d_i^{\delta} \quad ,$$

which was a proposal of Silvey, Titterington and Torsney (1978).

Since $g(z)$ is nondecreasing then $\exp\{\delta g(z)\}$ is nondecreasing or nonincreasing according as $\delta > 0$, $\delta < 0$ and so

7.3.9 $$m^{(r)} = \begin{cases} q^{(r)}\{e^{\delta g(z)}\} - p^{(r)} & , \quad \delta > 0 \\ p^{(r)} - q^{(r)}\{e^{\delta g(z)}\} & , \quad \delta < 0 \end{cases}$$

while

7.3.10 $$p^{(r+1)} = \begin{cases} (1-\alpha_r)p^{(r)} + \alpha_r q^{(r)}\{e^{\delta g(z)}\} & , \quad \delta > 0 \\ (1+\alpha_r)p^{(r)} - \alpha_r q^{(r)}\{e^{\delta g(z)}\} & , \quad \delta < 0 . \end{cases}$$

If we take $\delta = 1$ in the case of 7.3.8 the resultant $m^{(r)}$ is $m\{A(p), \underset{\sim}{1}', d\}$, $A(p) = \{\mathrm{diag}\{p_1, \ldots, p_J\}\}^{-1}$.

Let $q\{e^{\delta g(z)}\}$ denote the vector with components

$$q_j\{e^{\delta g(z)}\} = p_j \exp\{\delta g(z_j)\} \Big/ \sum_{i=1}^{J} p_i \exp\{\delta g(z_i)\}$$

where $z_j = d_j$ or $z_j = F_j$, $d_j = \partial\phi/\partial p_j$, $\phi(p) = \psi\{x(p)\}$, i.e. $d_j = G_j = G_\psi\{x(p), u_j\}$ while $F_j = F_\psi\{x(p), u_j\}$.

We now enumerate some properties of $q\{e^{\delta g(z)}\}$ which are generalisations of results quoted in Silvey, Titterington and Torsney (1978) for the case $g(z) = \ln z$

(i) If $\delta = 0$ and $p \in \mathcal{P}$ then $q\{e^{\delta g(z)}\} = p$

(ii) If $p_j > 0$ then $q\{e^{\delta g(z)}\} \in \mathcal{P}$ whether or not $p \in \mathcal{P}$.

(iii) Let $\tilde{u}_{(1)}$, $\tilde{u}_{(2)}$ respectively maximise and minimise $F\{x(p), u_j\}$ over $\mathrm{Sup}(p^{(r)})$, $x(p) = \sum p_j u_j$. Let $\tilde{e}_{(j)}$ be the unit vertex such that $\tilde{u}'_{(j)} = x\{\tilde{e}_{(j)}\}$ $j = 1,2$ and let $x\{q[e^{\delta g(z)}]\} = \sum u_j q_j\{e^{\delta g(z)}\}$ Then

$$x\{q[e^{\delta g(z)}]\} \longrightarrow \begin{cases} \tilde{u}_{(1)} & \text{as} \quad \delta \longrightarrow +\infty \\ \tilde{u}_{(2)} & \text{as} \quad \delta \longrightarrow -\infty \end{cases}$$

equivalently $\quad q\{e^{\delta g(z)}\} \longrightarrow \begin{cases} \tilde{e}_{(1)} & \text{as } \delta \to +\infty \\ \tilde{e}_{(2)} & \text{as } \delta \to -\infty \end{cases}$

suggesting that $q\{e^{\delta g(z)}\}$ sketches out an 'arc' in $\mathcal{P}$.

Hence vertex direction iterations can be viewed as a special case of $FP\{q^{(r)}[e^{\delta g(z)}], \alpha_r\}$.

The following picture is suggested by the result

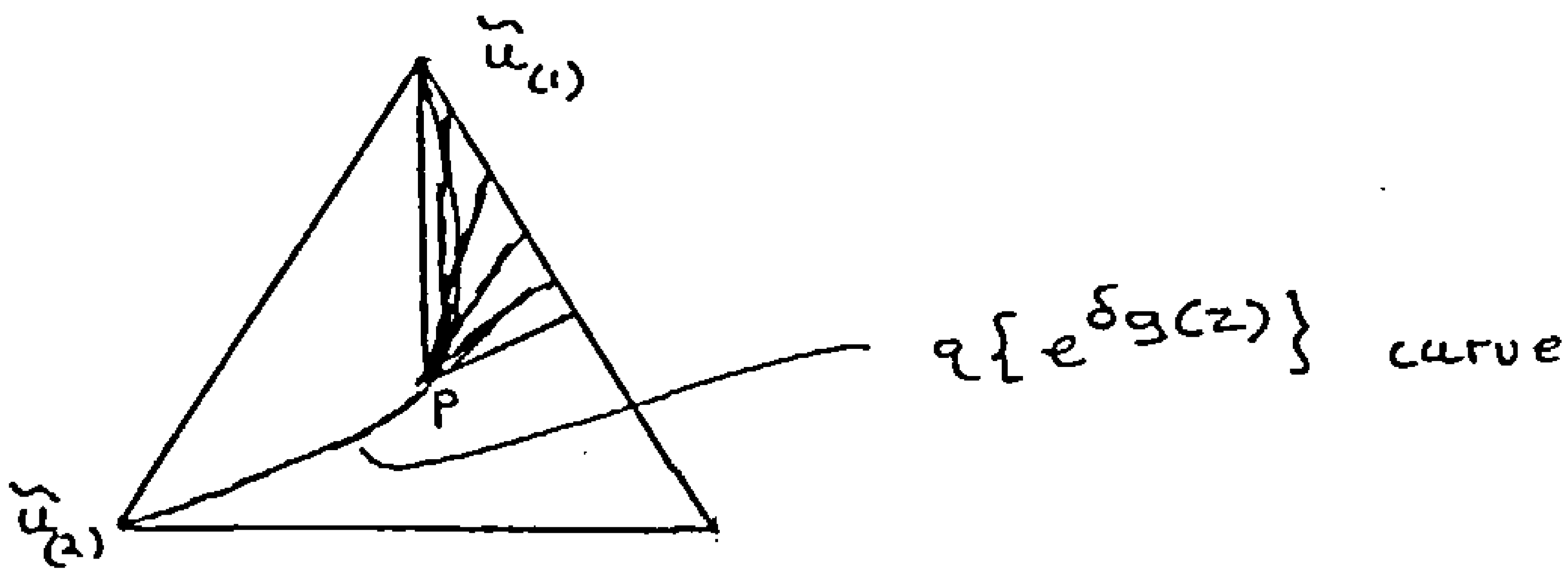


Figure 7.3.1

The shaded region is the set of points to which we could move from p by a step towards or away from $q\{e^{\delta g(z)}\}$ according as $\delta > 0$, $\delta < 0$.

(iv) We repeat again that a step towards or away from $q\{e^{\delta g(z)}\}$ will change the weights of only the current support but will be well defined at $p^{(r)}$ enjoying only support differentiability. We must have $\mathrm{Sup}\{p^{(0)}\} = \mathcal{U}$.

(v) $F_\phi\{p, q\{e^{\delta g(z)}\}\} \begin{cases} > 0 & , \delta > 0 \\ < 0 & , \delta < 0 \end{cases}$

(vi) The latter has already been established of course. However it would follow from the following seemingly respectable property by appealing to the fact that, in view of (i), it is the case that for $p \in \mathcal{P}$, $F_\phi\{p, q\{e^{\delta g(z)}\}\} = 0$ at $\delta = 0$. The result is that:

<u>Lemma 7.3.1</u>

If $d_j \geq 0$, $j = 1, \ldots, J$ then $F_\phi\{p, q\{e^{\delta g(z)}\}\}$ is nondecreasing in δ.

The condition that $\partial\phi/\partial p_j \geq 0$ is one that we have observed to be satisfied by the design criteria and the other examples of $\phi(p)$ considered in chapter 1.

<u>Proof</u> $$F_{\phi}\{p, q\{e^{\delta g(z)}\}\} = \frac{\sum p_j d_j \exp\{\delta g(z_j)\}}{\sum p_j \exp\{\delta g(z_j)\}} - \sum p_j d_j$$

$$= \frac{E_p\{d \exp\{\delta g(z)\}\}}{E_p\{\exp\{\delta g(z)\}\}} - E_p(d)$$

$$= f_1(\delta) - E_p(d),$$

where the expectations are with respect to the joint probability distribution

$$p(d, z) = \begin{cases} p_j, & (d, z) = (d_j, z_j) \quad j = 1, ---, J \\ 0, & \text{otherwise} \end{cases},$$

which implies the marginal distributions

$$p(d) = \begin{cases} p_j, & d = d_j \\ 0, & \text{else} \end{cases}, \quad p(z) = \begin{cases} p_j, & z = z_j \\ 0 & \text{else} \end{cases} \quad j = 1, ---, J$$

Now $$f_1'(\delta) = \frac{E_p\{d \exp[\delta g(z)]\}}{E_p\{\exp\{\delta g(z)\}\}} \left(f_2(1) - f_2(0)\right)$$

where $$f_2(t) = E_p\{d^t g(z) \exp\{\delta g(z)\}\} / E_p\{d^t \exp\{\delta g(z)\}\}.$$

Since $d_j \geqslant 0$, the factor formed by the ratio of the two moments is nonnegative so that we require to show that $f_2(1) \geqslant f_2(0)$.

Now

$$f_2'(t) = E_q\{g(z) \ln d\} - E_q\{g(z)\} E_q\{\ln d\}$$

where q is the joint probability distribution

$$q(d, z) = p(d, z) d^t \exp\{\delta g(z)\} / E_p\{d^t \exp\{\delta g(z)\}\}$$

Hence

$$f_2'(t) = \mathrm{Cov}_q\{g(z), \ln d\},$$

which in view of the results of section 7.2.2 is nonnegative for z = F or z = d. Clearly this result would be true for z = d, d continuous and $g(\cdot)$ nondecreasing. Shohat (1929) derives some relevant inequalities. □

The result of course does not contain any implication about a value of δ corresponding to a direction $q\{e^{\delta g(z)}\}$, towards or away from p, of steepest ascent, since $F_{\phi}\{p, q\{e^{\delta g(z)}\}\}$ is not normalised.

(vii) Assuming $p \in \mathcal{P}$ then in view of (i) and (v) it must be that

$$\phi\{q\{e^{\delta g(z)}\}\} \begin{cases} > \phi(p) \text{ for small positive } \delta \\ < \phi(p) \text{ for small negative } \delta \end{cases}$$

(viii) While for all $p \in \mathcal{P}$, $q\{e^{\delta g(z)}\} = p$ at $\delta = 0$, in the case $g(z)$ strictly monotonic and $\delta \neq 0$, $p \in \mathcal{P}$ is a fixed point of the mapping $p \longrightarrow q\{e^{\delta g(z)}\}$ if and only if the z_j, i.e. d_j or F_j, corresponding to the support points of p are all equal. Hence p^* and all optima solving (P2) for subsets of $\mathcal{U}$ are these fixed points.

§7.3.3 We now consider some points concerning the implementation of $FP\{q^{(r)}\{e^{\delta g(z)}\}, \alpha_r\}$

We must choose the value δ_r of δ at iteration r and of course α_r.

For simplicity the approach we consider is that chosen by Silvey, Titterington and Torsney (1978). This is to take δ_r small and positive and to choose $\alpha_r = 1$. The argument for this is that, in view of (vii) of section 7.3.2, this should produce a monotonic increasing sequence $\phi(p^{(r)})$, the iterates $p^{(r)}$ being

7.3.11 $$p^{(r+1)} = q^{(r)}\{\exp\{\delta_r g(z)\}\}$$

We are thereby using $FP\{e^{\delta_r g(z)}, 1\}$ which, in view of (viii) of section 7.3.11, is fully a fixed point algorithm. We have δ_r as it were replacing α_r.

In contrast if we take $\delta_r < 0$ then we must decide more formally on the value of the steplength of the consequent move away from $q^{(r)}\{\exp\{\delta_r g(z)\}\}$

How small should we choose positive δ_r?

In chapters 7 and 8, when taking $z = d$, $g(d) = h(d)$ and so, $h(d_j, \delta_r) = d_j^{\delta_r}$, and when considering a particular function $\phi(p)$, we postulate a value of δ_r which in some instances satisfies $\phi\{q^{(r)}\{e^{\delta_r g(z)}\}\} \geq \phi(p)$ and may do so more generally.

However in the absence of such analytic results the following are suggested.

(i) An ideal choice might seem to be to choose $\delta_r = \delta_r^*$ where δ_r^* maximises $\phi\{q^{(r)}\{\exp(\delta_r g(z))\}\}$ with respect to δ_r. Typically δ_r^* would need to be determined numerically.

(ii) Choose $\delta_r = \delta_r^{**}$ where δ_r^{**} maximises $F_\phi^A\{p^{(r)}, q^{(r)}\{e^{\delta_r g(z)}\}\}$ over $\delta_r > 0$. Again δ_r^{**} would need to be determined numerically, but these calculations may be less complex than for δ_r^*, particularly in the design context, since they do not require to evaluate the criterion.

Note that if we were to contemplate negative δ_r we might take $\delta_r = \delta_r^{***}$ where this maximises $\left| F_\phi^A\{p^{(r)}, q^{(r)}\{e^{\delta_r g(z)}\}\} \right|$.

(iii) Keep δ_r at a constant value while that maintains monotonicity and change its value when monotonicity fails. For example, take $\delta_r = \delta_0^*$ $r = 0,1,2,\ldots,t$, where t is such that monotonicity would fail at $r = t + 1$ if $\delta_{t+1} = \delta_0^*$. Instead choose $\delta_{t+1} = \delta_{t+1}^*$ and maintain δ_r at this value for $r \geq t+2$, changing it only to a current optimal value if its use were to induce failure of monotonicity, and repeat.

(iv) Choose $\delta_0, \delta_1, \delta_2, \ldots$ to be a predetermined sequence. Clearly this is a counterpart to arbitrary selection of steplengths.

Clearly one could propose other objective methods of selection of δ_r.

Silvey, Titterington and Torsney (1978) considered both (i) and (iii) referring to them respectively as adaptive and non-adaptive selections. This was for the case $g(d) = \ln(d)$, that is $h(d,\delta) = d^\delta$.

In these two instances the sequence $\phi(p^{(r)})$ is guaranteed to be nondecreasing and is bounded above if $\phi(p^*)$ is finite. Hence the iterates $p^{(r)}$ should converge unless these become undefined at a nondifferentiable point when support differentiability is also lacking.

It is not clear though that such convergence would be to p^*. The iterates may converge to a suboptimum. This may well happen if we are over-ready to set weights to zero. However the following suggests that convergence to p^* may not usually be impeded by lack of differentiability.

In section 3.2.5 it was established that for $c = (1,0)'$, the c-optimal design on the design space $\mathcal{V} = \{v_1, v_2\}$, $v_1 = (1,0)'$, $v_2 = (x_1, x_2)'$, $x_2 \neq 0$, is $p^* = (1,0)'$.
All observations should be taken at v_1.

Let $p = (p_1, p_2)'$ be any design such that $p_i \neq 0$, $i = 1,2$.
Then

$$M(p) = \begin{bmatrix} (p_1 + p_2 x_1^2) & p_2 x_1 x_2 \\ p_2 x_1 x_2 & p_2 x_2^2 \end{bmatrix} \quad , \quad M^{-1}(p) = \begin{bmatrix} p_2 x_2^2 & -p_2 x_1 x_2 \\ -p_2 x_1 x_2 & (p_1 + p_2 x_1^2) \end{bmatrix} \Big/ (p_1 p_2 x_2^2)$$

$$c' M^{-1}(p) c = c' M^{-1} v_1 = p_1^{-1} \quad , \quad c' M^{-1}(p) v_2 = 0$$

$$d_1 = p_1^{-2} \quad , \quad d_2 = 0$$

Hence, if $p_i^{(r)} \neq 0$, $i = 1,2$, we have

$$p_1^{(r+1)} = \frac{p_1^{(r)}\{\exp\{\delta_r g[1/(p_1^{(r)})^2]\}\}}{p_1^{(r)}\{\exp\{\delta_r g[1/(p_1^{(r)})^2]\}\} + p_2^{(r)}\{\exp\{\delta_r g(0)\}\}} \quad , \quad p_2^{(r+1)} = 1 - p_1^{(r+1)}.$$

Hence if $g(\cdot)$ is increasing so that $g[1/(p_1^{(r)})^2] > g(0)$, and if $\delta_r > 0$, then $p_1^{(r+1)} > p_1^{(r)}$. It would seem likely that $p_1^{(r)}$ should converge to 1.

In fact if we take $g(d) = \ln(d)$, so that
$p_j^{(r)} = p_j^{(r)} d_j^{\delta_r} / \{ p_1^{(r)} d_1^{\delta_r} + p_2^{(r)} d_2^{\delta_r} \}$ then clearly $p^{(1)} = p^*$.
The optimum is reached in one step.

This is of course a very simple example with no suboptima, but p^* is a one point design which one might argue enjoys the worst features of any design; nondifferentiability and degenerate support differentiability.

§7.4 Empirical Comparisons

§7.4.1 We now report some empirical results on the performance of some of the algorithms which we have discussed in this and previous chapters. The following six algorithms are in the main compared.

Let $d^{(r)}$ denote $\partial\phi/\partial p^{(r)}$.

I: Algorithm $V\{u^{(12)}, \alpha_r^*(u^{(12)})\}$, Atwood's (1973) vertex direction method. See equation 5.3.5.

II: Algorithm $S\{I, \alpha_r^*(m_*^{(r)})\}$; that is, gradient projection with $m_*^{(r)} = d^{(r)} - \bar{d}^{(r)}$. See section 6.2.3.

IIIa: Algorithm $C\{I, \beta_r, \alpha_r^*(m^{(r)})\}$ where $m^{(r)}$ is given by equations 6.3.7, 6.3.8 with $\beta_r = d^{(r+1)'}d^{(r+1)}/2d^{(r)'}d^{(r)}$.
This is a conjugate gradient projection scheme. See section 6.3.2.

IIIb: As in IIIa but with $\beta_r = d^{(r+1)'}d^{(r+1)}/20\,d^{(r)'}d^{(r)}$.

IV: As in IIIa but with $\beta_r = d^{(r+1)'}d^{(r+1)}/c_r\,d^{(r)'}d^{(r)}$,
where c_r takes the values 64, 32, 16, 8, 4, 2, 1 according as $F_\psi\{M(p^{(r)}), v^{(1)}v^{(1)'}\}$ is in $[0.001, 1/640)$, $[1/640, 1/320)$, $[1/320, 1/160)$, $[1/160, 1/80)$, $[1/80, 1/40)$, $[1/40, 1/20)$, $[1/20, \infty)$.

V: Algorithm $FP\{(d^{(r)})^{\delta_r^*}, 1\}$

VI: Algorithm $FP\{(d^{(r)})^{\delta_r}, 1\}$ where δ_r is chosen according to (iii) of section 7.3.2.

These were used to compute D-optimal and A-optimal designs in the following examples. The set $\mathcal{V}_J$ denotes the design space.

Example 1 $\mathcal{V}_1 = \{v_1, v_2, v_3, v_4\} = \{(1,-1,-1)', (1,-1,1)', (1,1,-1)', (1,2,2,)'\}$; $k = 3$, $J = 4$. This is Wynn's space.

Example 2 $\mathcal{U}_2 = \{v_1, v_2, v_3 v_4\}$ with v_1, v_2, v_3 as in example 1 but with $v_4 = (1,2,3)'$.

Example 3 $\mathcal{U}_3 = \{v_1, v_2, v_3, v_4\}$ with v_2, v_3, v_4 as in example 1 but with $v_1 = (1,-1,-2)'$.

Example 4 $\mathcal{U}_4 = \{(1,1,-1,-1)', (1,-1,1,-1)', (1,-1,-1,-1)'(1,2,2,-1)', (1,1,-1,1)', (1,-1.5,1,1)', (1,-1,-1,2)'\}$; $k = 4$, $J = 7$.

Example 5 $\mathcal{U}_5 = \mathcal{U}_4 \cup \{(1,-1,-1,2)'\}$; $k = 4$, $J = 8$.

This additional point has zero weight at the optimum for both D- and A-optimality.

Example 6 $\mathcal{U}_6 = \{(1,0,0,0,0,0)', (0,1,0,0,0,0)'(0,0,1,0,0,0)', (1/2,1/2,0,1/4,0,0)', (1/2,0,1/2,0,1/4,0)', (0,1/2,1/2,0,0,1/4)', (1/3,1/3,1/3,1/9,1/9,1/9)'\}$; $k = 6$, $J = 7$

Example 7 An example with $k = 4$, $J = 10$, the design space $\mathcal{U}_7$ consists of the four unit vectors plus 6 points inside the unit sphere. Both the D- and A-optimal designs assign weight $1/4$ to each of the unit vectors.

In examples 1,2,3,4,6 $\mathcal{U}_j = \mathrm{Sup}(p^*)$ for both D- and A-optimality, and in fact, excepting example 4, $\mathcal{U}_j$ is the support of the optimum design for some familiar types of regression model.

In examples 1,2,3 this could be $E(y) = \theta_1 + \theta_2 x_1 + \theta_3 x_2$, where $(1, x_1, x_2)' = v$ belongs to the quadrilateral with vertices v_1, v_2, v_3, v_4. An appeal to corollary 3.2.3.1 verifies this.

In example 6 the model is
$E(y) = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \theta_4 x_1 x_2 + \theta_5 x_1 x_3 + \theta_6 x_2 x_3$, where $x_1, x_2, x_3 \geq 0$, $x_1 + x_2 + x_3 = 1$. Kiefer (1975b) showed that an optimal design is supported at the barycentres of this simplex: $(1,0,0)', (0,1,0)', (0,0,1)', (1/2,1/2,0)', (1/2,0,1/2)', (0,1/2,1/2)', (1/3,1/3,1/3)'$.

§7.4.2 The results reported below in respect of algorithms I, II, IIIa, IIIb, IV were reported by Wu (1978a). Those in respect of algorithms V and VI were reported by Silvey, Titterington and Torsney (1978).

In all, except examples 1,2,3, $p_j^{(0)} = 1/J$.
In the exceptions Wu took $p^{(0)} = (1/3, 1/3, 1/3, 0)$ which is not feasible in the case of V and VI. Hence in order to make comparisons possible, $p^{(0)}$ was taken to be (.33,.33,.33,.01) in these two cases.

In view of the fact that $\mathcal{V}_i = \mathrm{Sup}(p^*)$ in examples 1,2,3, this means that the initial support has, at some stage, to be augmented by a vertex direction iteration in the case of II, IIIa, IIIb, IV. Wu employs one of the rules which were reported in section 6.5. For some $\varepsilon_0 > 0$, $0 < \beta \leq 1$, he augments the support at iteration r if $F_\phi\{p^{(r)}, p^{(r)} + \{\bar{\alpha}_r(m^{(r)})\} m^{(r)}\} < \varepsilon_0$, by choosing a v_j such that $F_\psi\{M(p^{(r)}), v_i v_i'\} > \beta F\{M(p^{(r)}), v^{(1)}v^{(1)\prime}\}$. He takes an optimal step in this vertex direction. He does not report his value of ε_0 or β, though there can be only one vertex, namely v_4, with which to augment $\mathrm{Sup}(p^{(0)})$.

In the case of algorithms V, VI weights were set to zero in accordance with the rules of section 4.2.3 taking $\varepsilon_1 = \varepsilon_2 = 0{\cdot}01$.

In table 7.4.1, for D-optimality, and in table 7.4.2, for A-optimality, are recorded, by example and algorithm, the number of iterations required to achieve $F_\psi\{M(p^{(r)}), v^{(1)}v^{(1)\prime}\} \leq .1, .01, .001$. Not all algorithms were used in all examples for the two criteria. The initial value selected for δ_0, namely δ_0^*, in the case of VI, is also given.

The same information is also recorded in Table 7.4.3 in respect of algorithm $FP(d^\delta, 1)$ for each of $\delta = .1, .2, \ldots, 2.1$ when used to calculate the D-optimum design in the case of example 1. For each δ, $p_j^{(0)} = 1/4$. Also recorded for $\delta = .5, .6, \ldots, 2.1$ is the number of iterations required to achieve $F_\psi\{M(p^{(r)}), v^{(1)}v^{(1)\prime}\} \leq 10^{-n}$, for n = 4,5,6. For $\delta \geq 2.2$ the algorithm does not converge.

There are also some further miscellaneous results to report.

Misc (i) Using algorithm FP(d,1) with $p_j^{(0)} = 1/J$ to compute D-optimal designs, the condition, $F_\psi\{M(p^{(r)}), v^{(1)}v^{(1)'}\} \leq .001$, was realised in one iteration in the cases

$$\mathcal{U}_8 = \{v_{(x)} = (1,x,x^2,x^3)' : x = \pm.44, \pm.45, \pm1\}$$
$$\mathcal{U}_9 = \{v_{(x)} = (1,x,x^2,x^3,x^4)' : x = 0, \pm.65, \pm.66, \pm1\}.$$

Misc (ii) Using algorithm FP$(d^{1/2},1)$ with $p_j^{(0)} = 1/4$ to compute the c-optimal design on $\mathcal{U}_1$ for $c = (0,1,1)'$, $F_\psi\{M(p^{(r)}), v^{(1)}v^{(1)'}\} \leq .001$ was achieved in 6 iterations; $p^* = (1/2,0,0,1/2)$ at which only support differentiability obtains.

Misc (iii) Wu reports that for examples 1 and 3 $V\{v^{(1)}, \alpha_r^*(v^{(1)})\}$ takes respectively 109 and 66 iterations to achieve $F_\psi\{M(p^{(r)}), v^{(1)}v^{(1)'}\} \leq .001$ in the case of D-optimality. In example 2 this is still not achieved by the 160th iteration, $F_\psi\{M(p^{(r)}), v^{(1)}v^{(1)'}\}$ being .00492 at this point.

Misc (iv) Wu also used $S\{A_6, \alpha_r^*(m_*^{(r)})\}$ which takes $m_*^{(r)} = m(A_6, \underline{1}, d^{(r)})$, $A_6 = \mathrm{diag}\{(v_1' v_1)^2, \ldots, (v_J' v_J)^2\}$ (See equation 6.2.8), i.e. $m_*^{(r)}$ is the normalised gradient projection direction, to compute the D-optimal designs in examples 1,2,3. The number of iterations required to achieve $F_\psi\{M(p^{(r)}), v^{(1)}v^{(1)'}\} \leq .001$ range from 6 to 14.

TABLE 7.4.1

D-OPTIMALITY

Algorithm / Example	I	II	IIIa	IIIb	IV	V	VI	Initial power
1	3,5,7	2,3,3	2,3,4	2,3,3	2,3,3	2,3,4	2,5,11	1.9
2	5,6,7	3,6,9	3,5,6	3,6,9	3,4,6	4,11,17	4,8,16	1.7
3	4,7,8	3,4,6	3,4,5	3,4,6	3,4,4	3,5,8	4,5,9	2.5*
4	12,37,64	5,15,27	4,9,15	5,15,25	4,5,15	5,33,78	4,29,82	1.3
5	6,59,>900	6,22,32	6,14,63	6,16,32	4,8,13	7,44,82	4,29,68	2.0
7						4,10,10	20,42,64	1.9

* In this example the power changed to 1.1 after three iterations.

TABLE 7.4.2

A-OPTIMALITY

Algorithm / Example	I	II	IIIa	IIIb	IV
4	16,35,51	8,18,28	4,12,22	7,17,28	5,11,19
5	6,63,141	4,28,58	3,29,55	4,28,58	5,27,57
6	19,24,30	12,16,21	10,14,19	8,11,14	9,11,16

TABLE 7.4.3

Entries are the number of iterations taken by $FP(d^{\delta}, 1)$ to achieve $F\{M(p^{(r)}), v^{(1)} v^{(1)'}\} \leq 10^{-n}$, when calculating the D-optimal design for example 1.

δ \ n	1	2	3	4	5	6
.1	19	81	162			
.2	9	40	80			
.3	6	27	53			
.4	5	20	39			
.5	4	16	31	47	63	83
.6	3	13	25	38	52	68
.7	2	11	21	32	44	58
.8	2	9	18	28	38	50
.9	2	8	16	24	33	43
1.0	1	7	14	22	29	38
1.1	2	6	13	19	26	34
1.2	2	6	11	17	23	31
1.3	2	5	10	16	21	28
1.4	2	5	9	14	19	25
1.5	2	4	9	13	18	23
1.6	2	5	8	12	16	21
1.7	3	7	11	15	19	23
1.8	3	7	13	19	23	31
1.9	3	11	19	27	33	43
2.0	5	19	29	41	55	70
2.1	9	39	71	95	123	161

§7.4.3 Discussion

(i) As Wu observes it is clear that IV the adaptive conjugate gradient method is fastest. Methods IIIa, IIIb, the two conjugate gradient methods, slightly improve over II, the gradient projection method.

(ii) Atwood's method I is slowest among I, II, IIIa, IIIb, IV but it is only initially slower than V and VI. It needs fewer iterations than the latter to achieve $F_{\psi}\{M(p^{(r)}), v^{(1)}v^{(1)'}\} \leq .001$.

Wu reports that in all of the examples in which it was used more of the iterations adopted a reverse vertex direction step. In example 4 with D-optimality out of the 64 iterations, 52 involve subtracting weight from only one vertex. The figure for example 6 with A-optimality is even more dramatic. None of the 30 iterations involve adding weight to only one vertex, an occurrence which can only be possible if $\mathrm{Sup}(p^{(0)}) = \mathrm{Sup}(p^*)$. In view of the fact that $\alpha_r^*(v^{(j)})$ can be evaluated explicitly in case of D- and A-optimality, Wu suggests the use of $V\{v^{(12)}, \alpha_r^*(v^{(12)})\}$ in these two cases for J of small or moderate size. This would extend to D_A-optimality and L_A-optimality. One might consider instead algorithm $V\{v^{(3)}, \alpha_r^*(v^{(3)})\}$ (see equation 5.3.1) since these two algorithms are identical in the case of D-optimality.

(iii) We consider now algorithms V and VI and also $FP(d^{\delta}, 1)$. There are two impressions here. On the one hand $FP(d,1)$ and V achieve some respectable results but more often than not all three are slow to converge. The ultimate convergence of $FP(d^{\delta}, 1)$ is slow for every δ in example 1.

We consider the first impression though. Some of these good performance are attributable to a specific result. This is that when the criterion is D-optimality and when $\mathrm{Sup}(p^*)$ consists of exactly k (linearly independent) vertices, algorithm $FP(d^{(r)}, 1)$ will pass exactly to the optimum, $p_j^* = 1/k$, in one step from an iterate $p^{(r)}$ for which $\mathrm{Sup}(p^{(r)}) = \mathrm{Sup}(p^*)$. See section 8.2.

The performance of V in example 6 is an illustration of this. At iteration 9 the last non-optimal-support point was eliminated and then the optimal $\delta^*_{10} = 1$ leads at once to the optimum.

Note also in Table 7.4.3 that FP(d,1) achieves $F_\psi\{M(p^{(1)}), v^{(1)}\ v^{(1)\prime}\} \le 0.1$ and also the performance of the same algorithm in the two spaces of Misc. (i).

In the latter two cases though (and also in the first case) we do not have $p^{(1)}$ exactly equal to p^*. However the following general approximation is suggested which clearly could be viewed as an extension of the above result.

The spaces $\mathcal{V}_7$ and $\mathcal{V}_8$ are respectively for $k = 3,4$ the set $\mathrm{Sup}(p^*)$ where p^* is the D-optimum design on the discretised polynomial regression design space

$$\mathcal{V} = \{v_{(x)} = (1,x,\ldots,x^{k-1})' \ :\ x \in \mathcal{X}_d\},\ \mathcal{X}_d = \{-1,-.99,\ldots,.99,1\}.$$

Now if $\mathcal{X}_d$ is replaced by its continuous analogue, the interval $[-1,1]$, the resultant optimising p^* has a support of k points, and p^* assigns these, weights $1/k$. Some of these k points however have been replaced by clusters of two points in $\mathrm{Sup}(p^*)$, and p^* assigns a total weight of $1/k$ to each cluster and to each other point. Interestingly it is the case, for both $\mathcal{V}_7$, $\mathcal{V}_8$, that $p^{(1)}$ assigned approximately weight $1/2k$ to each point in a cluster and weight $1/k$ to the remaining points. Similar results would appear to hold good in other regression models. In general when the space $\mathcal{V}$ is the support of the D-optimum design on a discretised one-variable regression design space, it seems likely that under FP(d,1), the iterate $p^{(1)}$ will be as above, and furthermore, that this design will be a rough approximation to p^* in that $p^{(1)}$ comes mildly close to attaining the conditions for an optimum. The components though of $p^{(1)}$ might not be a good reflection of those of p^*. In the case of $\mathcal{V}_8$ the symmetric $p^{(1)}$ assigns $x = 0,.65,.66.1$ the weights (.2000,.1000,.1000,.2000), while the design p^* which achieves $F_\psi\{M(p^*), v^{(1)}\ v^{(1)\prime}\} \le 10^{-6}$ assigns the comparable weights (.2000,.1153,.0847,.2000). In contrast for $\mathcal{V}_7$, $p^{(1)}$ assigns $x = .44,.45,1$ the weights (.1250,.1250,.2500) while the corresponding p^* assigns the respective weights

(.01910,.23090,.25000).

A more general result as shall be seen in section 8.5 is that $\phi(p^{(r+1)}) \geq \phi(p^{(r)})$ under FP($d^{(r)}$,1) in the case of D-optimality. The strength of this result is indicated by the fact that $\delta_0^* > 1$ in all examples. Also always $\delta_r^* > 1$. In addition FP(d^{δ} ,1) proved to be monotonic for all the powers δ in Table 7.4.3.

It is also the case that algorithm FP($d^{1/2}$,1) which, as reported in Misc. (ii), performed favourably in finding a c-optimal design for example 1, is monotonic in the case of L_A-optimality. A proof will be given in section 9. Furthermore this too can attain the optimum in one step in specific circumstances. See section 8.2.

However these results not withstanding, convergence of FP(d,1) , FP($d^{1/2}$,1) is typically slow, although Wu (1978a) expects the former to perform favourably with I (the gradient projection method) and also with normalised gradient projection. It may be however that initial performance of these and of FP(d^{δ} ,1) for δ near 1 will be comparable. Note from table 7.4.3 that, in example 1, the latter achieves $F_{\psi}\{M(p^{(r)}),\ v^{(1)}\ v^{(1)'}\} \leq 0.1$ in two iterations for δ = .7,.8,.9,1.1,...,1.6 and, as already mentioned, it achieves the same in one iteration when δ = 1, faster than any of the other algorithms.

Since FP(d^{δ} ,1) is slow to converge for every δ in example 1 it is not surprising that VI is also slow to converge. It is unlikely even that δ_0^* will be a power that would yield fastest convergence. The evidence of section 7.4.2 is that it is likely to prove too large a power to persist with. Note the dramatic reduction from 2.5 to 1.1 in the one example in which δ had to be changed.

However it is disappointing that V is not much of an improvement on VI. The reason for its slow convergence would seem to be that the optimal power settles into an oscillatory sequence. For instance in example 3 for D-optimality this sequence is $\{2.5,1.4,1.9,1.4,2.0,1.4,1.9,1.5\}$. Note that $\delta_r^* > 1$.

Other algorithms, such as steepest ascent, it is known exhibit similar behaviour and no doubt modifications of the parallel tangent and possibly of the conjugate gradient type would be improvements.

Clearly the optimal form of $FP\{(d^{(r)})^{\delta_r}, 1\}$ would take $\delta_r = \delta^*$ where δ^* achieves fastest convergence. Table 7.4.3 suggests a δ^* of 1.6. However we of course cannot know δ^*. A compromise might be to change at an appropriate point from V to $FP(d^{\delta}, 1)$ where δ is the average of the last few optimal δ_r's. This would be an improvement similar in spirit to parallel tanjent. In fact in the case of example 1, algorithm V when started from $p_j^{(0)} = 1/4$ realises the following sequence of approximate optimal powers 1.65, 1.55, 1.65, 1.65, 1.75, 1.65, 1.65, 1.75, 1.65 . Averages of say any sequence of three of these is just below 1.7 which is clearly close to δ^* above.

Naturally improvement would also be gained by taking optimal steplengths, be it $FP\{(d^{(r)})^{\delta_r}, \alpha_r^*\}$ or $FP\{(d^{(r)})^{\delta_r^*}, \alpha_r^*\}$ which we employ. For instance using $FP\{d^{(r)}, \alpha_r^*\}$ to calculate the D-optimum design in example 1, only eight iterations are required to achieve $F_\psi\{M(p^{(r)}), v^{(1)}v^{(1)'}\} \leq 0.00005$.

Such a remedy however removes from the simplicity of $FP(d^{\delta}, 1)$. It is to be noted that both V and VI enjoy good initial convergence, as does $FP(d^{\delta}, 1)$ for $\delta = .7, \ldots, 1.6$ in the case of example 1. They at first climb as quickly as any of the other methods. Arguably $FP(d, 1)$ is initially quickest of all the algorithms in example 1. This is good enough if we maintain the recommendation that these algorithms be used as intermediate methods. That is, for large $\mathcal{U}$ or large $\mathcal{V}$ they be used as a sequel to vertex direction techniques, a prelude to (adapted) Newton Raphson or Fletcher Powell methods.

CHAPTER 8

ON MONOTONICITY OF A FIXED POINT ALGORITHM FOR A GENERAL CRITERION

§8.1 A Monotonicity Conjectured

It is the concern of this and the next chapter to prove in some instances and to suggest in others that algorithm $FP\{d^{\delta},1\}$, for solution of problem (P1), generates a monotonic sequence $\phi(p^{(r)})$ for a particular value δ, the function $\phi(p)$ satisfying among other conditions, the basic requirement for use of this algorithm, that $d_j = \partial\phi/\partial p_j > 0$.

In particular in chapter 9 we will prove that monotonicity obtains for $\delta = 1/2$ in the case of L_A-optimality, while in this chapter we will see that $\delta = 1$ attains this in the case of D-optimality and in the case of some of the functions $\phi(p)$ of examples 1.1.1,..., 1.1.4.

In chapter 9 we will also present empirical evidence to suggest that monotonicity is achieved by $\delta = 1/(t+1)$ in the case of $\phi(p) = \psi_t\{M(p);A\}$, $\psi_t\{M;A\} = -tr(AM^{+}A')^{t}$. For the case A = I this was conjectured by Torsney (1977).

Meanwhile in this chapter some theoretical results will be derived which suggest, though do not prove, that $\delta = 1/(t+1)$ would guarantee monotonicity in the case of a function $\phi(p)$ enjoying a particular property of $\psi_t\{M(p);A\}$, namely a function $\phi(p)$ which is homogeneous of degree $(-t)$ as well as having the necessary positive derivatives.

Actually algorithm $FP\{h(z,\delta),\alpha_r\}$ owes it existence to the fact that it was proven, by Fellman (1974, theorem 3.1.5), that $\phi\{q\{d^{1/2}\}\} \geq \phi(p)$ for the particular case, c-optimality, of L_A-optimality. He is not though concerned with formulating an algorithm. However his result first lead Silvey, Titterington and Torsney (1976) to consider $FP\{d,1\}$ for D-optimality, in the belief, as proved justified, that this too would be monotonic. It was only thereafter that they proposed $FP\{(d^{(r)})^{\delta_r},1\}$ as outlined in the previous chapter.

Admittedly establishing that a particular value $\bar{\delta} > 0$ achieves monotonicity need not not be of great practical importance; $FP\{d^{\bar{\delta}}, 1\}$ may still converge painfully slowly. However intuitively it would follow, if $\phi(p)$ is concave, that $FP\{d^{\delta}, 1\}$ will also be monotonic at least for $0 < \delta \leq \bar{\delta}$. One would also expect the same property of $FP\{e^{\delta g(2)}, 1\}$. Such information could be potentially useful.

The justification for this assertion is the more general claim that, if $\phi(p)$ is concave, one would expect $\phi\{q\{e^{\delta g(2)}\}\}$ to be unimodal in δ or possibly just nondecreasing in δ. The properties (i),...,(viii) of section 7.3.2 lend credence to this suggestion. In particular one would expect unimodality in the case of $\phi_t(p) = -tr\{AM^+(p)A'\}^t$, since, in view of the fact that $q\{e^{\delta g(2)}\}$ converges to a unit vector as $\delta \to \pm\infty$, we will in general have $\phi_t\{q\{e^{\delta g(2)}\}\} \to -\infty$ as $\delta \to \infty$.

Whether or not it was useful to search for a proof of monotonicity in the instances listed, the quest lead to the discovery of several interesting results.

§8.2 Some Motivating Results

§8.2.1 Consider the effect of $\delta = 1/(t+1)$ and of other values of δ in the case of the following simple functions which were seen, in examples 3.3.1(i) and 3.3.1(ii), to be special cases of standard design criteria:

$$\phi_1 = \phi_1(p) = c\,p_1 p_2 \cdots p_k \quad , \quad c > 0, k \geq 2$$

$$\phi_2 = \phi_2(p) = -c\,(p_1 p_2 \cdots p_s)^{-1} , \quad c > 0, s \geq 2$$

$$\phi_3 = \phi_3(p) = -\sum_{i=1}^{k} p_i^{-t} \quad , \quad t > 0, k \geq 2$$

$$\phi_4 = \phi_4(p) = -\sum_{i=1}^{k} a_i p_i^{-t} \quad , \quad t > 0, k \geq 2, a_i > 0 .$$

In particular ϕ_1 is a special case of $\det\{M(p)\}$, ϕ_2 of $\{-\det\{AM^+(p)A'\}\}$ and ϕ_3, ϕ_4 of $\psi_t\{M(p), A\} = -\mathrm{tr}\{AM^+(p)A'\}^t$. These cases arise when J = k or J = s, J the number of points in the design space. All four functions are concave and homogeneous with positive derivatives. The degrees of homogeneity are respectively k, -s, -t, -t.

The "optimal designs" under ϕ_1, ϕ_2, ϕ_3, ϕ_4 take $p_j^* = (1/k), (1/s), (1/k), \left\{a_j^{1/(t+1)} / \left(\sum a_i^{1/(t+1)}\right)\right\}$ respectively.

Consider ϕ_1, ϕ_2 and ϕ_3.

$$\partial\phi_1/\partial p_j = \phi_1/p_j \quad ; \quad \partial\phi_2/\partial p_j = |\phi_2|/p_j \quad ; \quad \partial\phi_3/\partial p_j = t\,p_j^{-(t+1)}$$

Hence

$$\partial\phi_1/\partial p_j \propto p_j^{-1} \quad ; \quad \partial\phi_2/\partial p_j \propto p_j^{-1} \quad ; \quad \{\partial\phi_3/\partial p_j\}^{1/(t+1)} \propto p_j^{-1}$$

Thus

$$p_j(\partial\phi_1/\partial p_j)^n \propto p_j^{1-n} \quad ; \quad p_j(\partial\phi_2/\partial p_j)^n \propto p_j^{1-n} \quad ; \quad p_j\{\partial\phi_3/\partial p_j\}^{n/(t+1)} \propto p_j^{1-n} .$$

Let J denote the number of components in p for these three functions, so that J = k, s, k respectively for ϕ_1, ϕ_2, ϕ_3 and in consequence $p_j^* = 1/J$.

Suppose that algorithm $FP\{d^\delta, 1\}$ were employed with, in the case of ϕ_1 and ϕ_2, $\delta = n$ and, in the case of ϕ_3, $\delta = n/(t+1)$.

Let $p_j^{(0)} = p_j$ and let p_1 and p_m denote respectively the minimum and maximum among the p_j's, and assume the indices 1 and m to be unique.

We will have

$$p_j^{(r)} = p_j^{(1-n)^r} \Big/ \sum_{i=1}^{J} p_i^{(1-n)^r} ,$$

and the following is easily verified.

(i) If $n = 1$, $p^{(r)} = p^*$ for all r; that is, the optimum is attained in one step.

(ii) If $0 < n < 2$, $\operatorname{Lt}_{r\to\infty} p^{(r)} = p^*$.

(iii) If $n = 0$, $p^{(r)} = p$ for all r.

(iv) If $n = 2$, $p^{(2r)} = p$, $p^{(2r+1)} = \underline{p}^{-1}$ for all $r = 0,1,2,\ldots$, where $\underline{p}^{-1}$ is the vector whose j^{th} component is $\left\{p_j^{-1}\big/\left(\Sigma p_i^{-1}\right)\right\}$ and so the iterates $p^{(r)}$ oscillate between two values.

(v) If $n < 0$, $\operatorname{Lt}_{r\to\infty} p^{(r)} = \underline{e}_m$, the m^{th} unit vector.

(vi) If $n > 2$, $\operatorname{Lt}_{r\to\infty} p^{(2r)} = \underline{e}_m$, $\operatorname{Lt}_{r\to\infty} p^{(2r+1)} = \underline{e}_1$, so that $p^{(r)}$ eventually "oscillates" between $\underline{e}_m$ and $\underline{e}_1$.

Coinciding rather well with this latter result is the fact that algorithm FP$\{d^{\delta},1\}$ failed to converge for $\delta \geqslant 2.2$, as reported in section 7.4.2, when calculating the D-optimal design for the space $\mathcal{V} = \{(1,-1,-1)',(1,-1,1)',(1,1,-1)',(1,2,2)'\}$, a space for which $J = k+1$.

Simple modifications to (v) and (vii) hold if either the maximum or the minimum over j of p_j is not unique. We would obtain convergence to or oscillations between uniform designs on the relevant set of points.

The function ϕ_4 exhibits similar behaviour. In particular $\delta = 1/(t+1)$ attains the optimum in one step.

Other simple functions exhibit similar phenomena also. Consider

$$\phi_5 = \phi_5(p) = \sum_{i=1}^{k} p_i^t \quad , \quad t > 0$$

This is homogeneous of degree t, has positive derivatives and would be the form of $\phi_t(p) = tr\{M^t(p)\}$ on an orthonormal design space. It has an optimum at $p_i = 1/k$ which is a maximum for $0 < t \leq 1$ in which range ϕ_5 is concave, while it is a minimum for $t > 1$ in which range ϕ_5 is convex.

We have

$$\partial\phi_5/\partial p_j = t p_i^{(t-1)}$$

$$p_j \left(\partial\phi_5/\partial p_j\right)^{-1/(t-1)} \propto 1$$

Hence $\delta = -1/(t-1)$ attains the optimum in one step. Note that this power is positive or negative according as $t > 1$ or $t < 1$. In fact it is really an extension of the corresponding $\delta = 1/(t+1)$ for $\phi_3(p)$.

It is these instances of a particular power attaining the optimum in one step for these simple functions, which lends credence to the idea that the same power might yield monotonicity at least in the case of the design criteria, of which the simple functions are special cases.

§8.2.2 What law if any determines such a power? The functions considered were all homogeneous. The degree of homogeneity is clearly the determining factor in the case of functions ϕ_3, ϕ_4 and ϕ_5, but what of ϕ_1, ϕ_2 where $\delta = 1$ is the relevant power but the degree of homogeneity is k, -s respectively. The following would seem to be the governing rule.

Suppose that a function $\phi(p)$ is such that

8.2.1 $$\partial\phi/\partial p_j = c(p) f_j(p) \qquad \forall \; j = 1, ---, J ,$$

where the functions $f_j(p)$ do not have any further (natural) common factor depending on p.

Then it is only the functions $f_j(p)$ and not c(p) which will affect the value of $q\{d^\delta\}$ however we choose δ. Due to normalisation $q\{d^\delta\}$ does not depend on c(p).

Now homogeneity of $\phi(p)$ implies of course the same of $\partial\phi/\partial p_j$. Almost certainly both $c(p)$ and $f_j(p)$ will also be homogeneous. Let the common degree of homogeneity of the $f_j(p)$ be h. Then

8.2.2 $$\delta = -1/h$$

is a possible rule for identifying special powers like the above. It proposes a value of δ which renders $p_j[f_j(p)]^\delta$ a homogeneous function of degree zero, a sort of normalising effect. This would seem to be the defining feature of the values of δ which achieved p^* in one step.

As well as ϕ_1 and ϕ_2 there are other nondegenerate instances of functions $\phi(p)$ giving rise to 8.2.1; for example when $\phi(p)$ is a power $[g(p)]^n$ of a homogeneous function $g(p)$ and when $\phi(p)$ is a product of homogeneous functions. Considering the latter case suppose that

8.2.3 $$\phi(p) = c\prod_{i=1}^{m} g_i(p) ,$$

where c is a constant and $g_i(p)$ is a homogeneous function of degree h_i. Then a natural expression for $\partial\phi/\partial p_j$ is given by 8.2.1 where

$$c(p) = \phi(p)$$

and

8.2.4 $$f_j(p) = \sum_{i=1}^{m}\left[\{\partial g_i(p)/\partial p_j\}/g_i(p)\right] ,$$

which is homogeneous of degree h = -1 since $\partial g_i(p)/\partial p_j$ is homogeneous of degree (h_i-1). From 8.2.2 we obtain the suggestion $\delta = 1$, the same as for ϕ_1 and ϕ_2.

In fact both ϕ_1 and ϕ_2 are special cases of 8.2.3; $m = k$, $g_i(p) = p_i$ yields ϕ_1, $a = -c$, $m = s$, $g_i(p) = p_i^{-1}$ yields ϕ_2. So also are the functions $\phi(p)$ of examples 1.1.1, 1.1.2, 1.1.3 and 1.1.4 of chapter 1, the latter taken in the form of equation 1.1.1.

Actually ϕ_1 is a particular case of the multinomial model of example 1.1.1 and $\delta = 1$ also attains the optimum in one step for that model.

Further it has been shown by appealing to the theory of the EM algorithm of Dempster, Laird and Rubin (1977) that this same $\delta = 1$

achieves monotonicity in the case of the maximisations of examples 1.1.2 and 1.1.4. See Morgan and Titterington (1977) for the latter and Dempster et al (1977) for the former.

Finally we note that the monotonicity which, we have already said, obtains when using $\delta = 1$ in the case of D-optimality can be proved by showing that then, $FP\{d, 1\}$ is another EM algorithm, the criterion being taken in the form $\det\{M(p)\}$. This and further relationships between $FP\{d^{\delta}, 1\}$ and the EM algorithm will be pursued later.

For the moment we note that in these results we have evidence that the suggestion in 8.2.2 can produce a monotonic algorithm. As has been said we believe that a particular instance when this is so is the case of $\phi(p) = -\operatorname{tr}\left(AM^{+}(p)A'\right)^{t}$, $t > 0$. This is homogeneous of degree $(-t)$ and is neither a product nor power of homogeneous functions. It therefore most naturally satisfies 8.2.1 with $c(p) = 1$, $f_j(p) = \partial\phi/\partial p_j$ and hence, $h = -(t+1)$, so that 8.2.2 suggests $\delta = 1/(t+1)$. Note that this would suggest $\delta = 1$ at $t = 0$ which corresponds to D_A-optimality when $s \geq 2$, A being of order $s \times k$, and hence to D-optimality when $A = I$.

While it is inconceivable that recommendation 8.2.2 would produce monotonicity in any homogeneous function with positive derivatives, we can however, obtain a condition, with related results, which is sufficient for $\delta = 1/(t+1)$ to achieve monotonicity in the case of maximising a general function, which has positive derivatives, and which is homogeneous of degree $(-t)$, whether or not the function is a product or power of homogeneous functions.

Certainly for t large one would expect monotonicity to obtain whatever the function $\phi(p)$ for then $\delta = 1/t+1$ is small. Two examples when this choice would ignore recommendation 8.2.2 but would suggest a "smaller than necessary power" are the following:

(i) $\phi(p) = \phi_2(p)$; then $t = s$, $1/(t+1) = 1/(s+1)$

(ii) $\phi_6(p) = -\{\det M(p)\}^{-1}$; since $\det\{M(p)\}$ is a polynomial of degree k with positive coefficients (see theorem 1.1.2 of Fedorov (1972)) this is homogeneous of degree $-k$ so that $t = k$, $1/(t+1) = 1/(k+1)$.

In both cases $c(p) \neq 1$ and $\delta = 1$ is known to achieve monotonicity, and this is the power which rule 8.2.2 would select for $\phi_2(p)$ and for $\phi_6(p)$ when $\det\{M(p)\} = \phi_1(p)$. In fact $\delta = 1$ would attain the optimum in one step in the latter instance and in the case of $\phi_2(p)$, and it is in the light of this fact that to ignore rule 8.2.2 and to opt for $\delta = 1/(t+1)$ in these examples is to select too small a power. However one could conclude monotonicity of $\delta = 1/(t+1)$ from that of $\delta = 1$ if $\phi_1\{q\{d^\delta\}\}$ is unimodal in δ.

§8.2.3 We end this section with the following discussion.

Note that recommendation 8.2.2 may select a negative δ yielding a monotonic decreasing sequence $\phi(p^{(r)})$, as in the case of ϕ_5 when the positive degree of homogeneity t is in the range (0,1), and thereby, an algorithm which finds a minimum of a homogeneous function with positive derivatives.

Usually we wish to maximise concave functions such as the standard design criteria and minimise convex functions. Selection of a negative δ by rule 8.2.2 should reconcile with this as the following observations suggest.

First the basic definition of homogeneity that $\phi(cp) = c^t\phi(p)$ imposes concavity or convexity on each line running through the origin, for the function c^t is convex increasing, concave increasing, convex decreasing respectively in the ranges $t > 1$, $0 < t < 1$, $t < 0$.

The next lemma establishes some stronger relationships and makes use of the following two properties of a function $\phi(p)$ which is homogeneous of degree h:

(i) $\partial\phi/\partial p_j$ is homogeneous of degree (h-1)

(ii) $\sum p_j \partial\phi/\partial p_j = h\phi(p)$.

Lemma 8.2.1

Suppose that a function $\phi(p)$ does not change sign on the positive quadrant, that its partial derivatives have a common sign, and that $\phi(p)$ is either concave on the positive quadrant or is convex there. If further the function is homogeneous, then the degree h of homogeneity and the sign of $\phi(p)$ determines both the sign of the derivatives and which of the two properties concavity or convexity it has.

Proof Assume $p_i > 0$.

(i) Since $\sum p_i \partial\phi/\partial p_i = h\phi(p)$, $\partial\phi/\partial p_j$ has the same sign as $h\phi(p)$.

(ii) Similarly $\sum_j p_j \partial^2\phi/\partial p_i \partial p_j = (h-1)\partial\phi/\partial p_i$, since the derivatives of $\phi(p)$ are homogeneous of degree $(h-1)$. It follows that the particular quadratic form $p' H p = h(h-1)\phi(p)$, where H is the matrix of second derivatives of $\phi(p)$. Hence, restricting x to the positive quadrant, $x'Hx$ has the same sign as $h(h-1)\phi(p)$.

From these two results it follows that $\phi(p)$ is

(a) convex increasing if $h > 1$, $\phi(p) > 0$

(b) concave increasing if $0 < h < 1, \phi(\cdot) > 0$

(c) convex decreasing if $h < 0$, $\phi(\cdot) > 0$

(d) concave decreasing if $h > 1$, $\phi(\cdot) < 0$

(e) convex decreasing if $0 < h < 1$, $\phi(\cdot) < 0$

(f) concave increasing if $h < 0$, $\phi(\cdot) < 0$. □

Note that the general design criterion $\phi(p) = -tr(AM^{+}(p)A')^{t}$ falls into the last category with h = -t.

We have then a variety of functions here, but we do not propose to keep all of them in mind.

We will concentrate on the use of $FP\{d^{\delta},1\}$ with $\delta = 1/(t+1)$ in the case of solving problem (P1) for a function $\phi(p)$, which has positive derivatives, is homogeneous of degree $(-t)$ and, in particular examples, is typically concave.

§8.3 A Moment Inequality

The quest for a proof that $FP\{d^{\delta},1\}$ is monotonic for $\delta = 1/(t+1)$ in the circumstances outlined above has led to the discovery of some moment inequalities. With hindsight this is not particularly surprising, for one of the standard properties of a $\phi(p)$ which is homogeneous of degree h is that

8.3.1 $$\sum p_i \partial\phi/\partial p_i = h\phi(p)$$

If we restrict p to $\mathcal{P}$ this becomes

8.3.2 $$E_p\{f(p)\} = h\phi(p) ,$$

where f(p) is a discrete random variable with the probability distribution

$$Pr\{f(p) = \partial\phi/\partial p_j\} = p_j .$$

Further if $\partial\phi/\partial p_j > 0$ then f(p) is a positive random variable and so it is possible to define the discrete positive random variable $g(p) = \{f(p)\}^{\delta}$ whose probability distribution is then

$$Pr\{g(p) = [\partial\phi/\partial p_j]^{\delta}\} = p_j$$

satisfying

$$E_p\{[g(p)]^{-\delta}\} = h\phi(p)$$

As we have said, to study the behaviour of any homogeneous function in the positive quadrant, it would suffice to study its behaviour on the subset $\{p : \sum p_i = a\}$ for any positive constant a. Choosing a = 1 gains the benefit of results in the theory of moments and may lead to the suggestion of new results in that area.

In particular to prove that $\phi(\mu) \geq \phi(\lambda)$ where $\mu = p^{(r+1)}$, $\lambda = p^{(r)}$, as one would wish to do if monotonicity is to be established, is, in view of 8.3.2, to prove that

$$\begin{array}{lll} [E_\mu\{f(\mu)\}] & \geq [E_\lambda\{f(\lambda)\}] & \text{if } h > 0 \\ [\quad " \quad] & \leq [\quad " \quad] & \text{if } h < 0 , \end{array}$$

and $f(\mu), f(\lambda)$ will both be discrete positive random variables if $\partial\phi/\partial p_i > 0$ for all $p \in \mathcal{P}$.

The search for such a proof in the case of our conjectured monotonicity has led to the discovery of the following lemma.

Lemma 8.3.1

Let x,y be two positive random variables and let (a), (b), (c), (d), (e), (f) denote the following mutually exclusive and exhaustive conditions on two numbers r and n:

(a) : $0 < r < n$

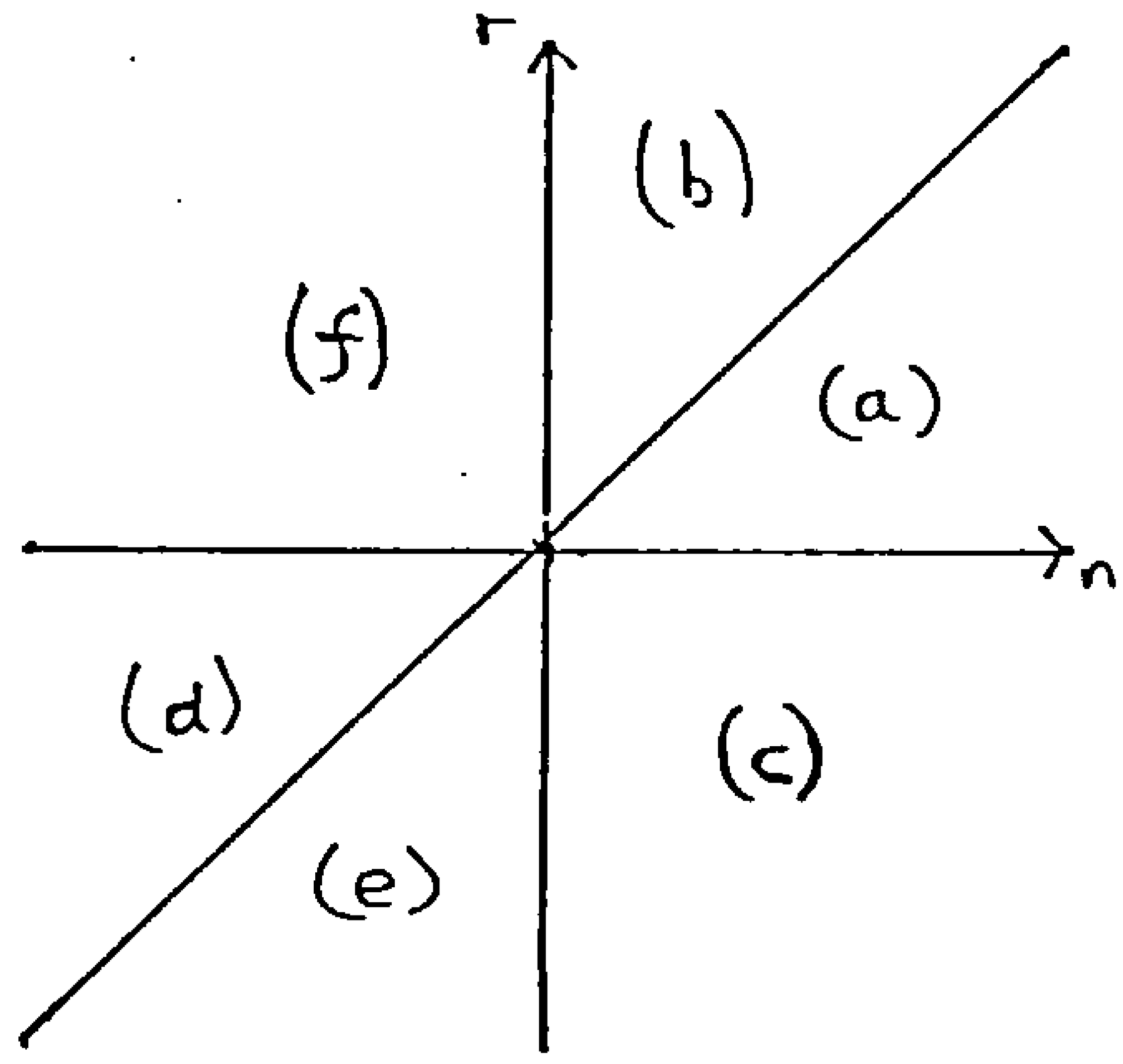
(b) : $r > n > 0$

(c) : $n > 0,\ r < 0$

(d) : $n < r < 0$

(e) : $r < n < 0$

(f) : $n < 0,\ r > 0$



The following is true

(i) If under (a) or (d) $E(y^n) \leq \{E(x^r)E(y^{n-r})\}$ then $E(x^n) \geqslant \{\quad " \quad\}$

(ii) If under (b) or (e) $E(y^n) \leq \{\quad " \quad\}$ then $E(x^n) \leq \{\quad " \quad\}$

(iii) If under (c) or (f) $E(y^n) \geqslant \{\quad " \quad\}$ then $E(x^n) \geqslant \{\quad " \quad\}$

This could be summarised in the form of a "multidimensional statement" with an obvious convention for its interpretation as follows:

If

$$E(y^n)\begin{cases} \leqslant E(x^r)E(y^{n-r}) & \text{under (a), (b), (d), (e)} \\ \geqslant E(x^r)E(y^{n-r}) & \text{under (c), (f)} \end{cases} \tag{8.3.4}$$

then

$$E(x^n)\begin{cases} \geqslant E(x^r)E(y^{n-r}) & \text{under (a), (c), (d), (f)} \\ \leqslant E(x^r)E(y^{n-r}) & \text{under (b), (e)} \end{cases} \tag{8.3.5}$$

Proof The result follows from two appeals to the result that for any positive (discrete) random variable z the function $M(t) = \{E(z^t)\}^{1/t}$ is nondecreasing in t.

More particularly we appeal to the consequence of this result that

8.3.6 $$\{E(x^{t_1})\}^{1/t_1} \geq \{E(x^{t_2})\}^{1/t_2}, \quad t_1 \geq t_2$$

This is sometimes referred to as Liapounov's inequality.

We proceed as follows: from 8.3.6

$$\{E(y^n)\}^{1/n} \begin{cases} \geq \{E(y^{(n-r)})\}^{1/(n-r)} & \text{under (a), (b), (f)} \\ \leq \quad '' & \quad '' \quad \text{(c), (d), (e)} \end{cases}$$

Hence

$$E(y^n) \begin{cases} \geq \{E(y^{(n-r)})\}^{n/(n-r)} & \text{under (a), (b), (d), (e)} \\ \leq \quad '' & \quad '' \quad \text{(c), (f)} \end{cases}$$

or

$$\{E(y^{(n-r)})\}^{n/(n-r)} \begin{cases} \leq E(y^n) & \text{under (a), (b), (d), (e)} \\ \geq \quad '' & \quad '' \quad \text{(c), (f)} \end{cases}$$

Therefore by assumption 8.3.4

$$\{E(y^{(n-r)})\}^{n/(n-r)} \begin{cases} \leq E(x^r)E(y^{(n-r)}) & \text{under (a), (b), (d), (e)} \\ \geq \quad '' \quad '' & \quad '' \quad \text{(c), (f)} \end{cases}$$

and hence

$$\{E(y^{(n-r)})\}^{r/(n-r)} \begin{cases} \leq E(x^r) & \text{under (a), (b), (d), (e)} \\ \geq \quad '' & \quad '' \quad \text{(c), (f)} \end{cases}$$

or

$$E(x^r) \begin{cases} \geq \{E(y^{(n-r)})\}^{r/(n-r)} & \text{under (a), (b), (d), (e)} \\ \leq \quad '' & \quad '' \quad \text{(c), (f)} \end{cases}$$

Hence

8.3.7 $$\{E(x^r)\}^{(n-r)/r} \begin{cases} \geq E(y^{(n-r)}) & \text{under (a), (c), (d), (f)} \\ \leq \quad '' & \quad '' \quad \text{(b), (e)} \end{cases}$$

Again appealing to 8.3.6 we have

$$\{E(x^n)\}^{1/n} \begin{cases} \geq \{E(x^r)\}^{1/r} & \text{under (a), (c), (e)} \\ \leq \quad '' & \quad '' \quad \text{(b), (d), (f)} \end{cases}$$

Hence

$$E(x^n) \begin{cases} \geq \{E(x^r)\}^{n/r} & \text{under (a), (c), (d), (f)} \\ \leq \quad '' & \quad '' \quad \text{(b), (e)} \end{cases}$$

that is

$$E(x^n) \begin{cases} \geq E(x^r)\{E(x^r)\}^{(n-r)/r} & \text{under (a), (c), (d), (f)} \\ \leq \quad '' & \text{'' (b), (e)} \end{cases}$$

and so from 8.3.7

$$E(x^n) \begin{cases} \geq E(x^r)E(y^{(n-r)}) & \text{under (a), (c), (d), (f)} \\ \leq \quad '' & \text{'' (b), (e)} \end{cases}$$

□

Corollary 8.3.2

If under (a) or (d) $E(y^n) \leq E(x^r)E(y^{(n-r)})$
then $E(x^n) \geq E(y^n)$. □

The result is obvious and states that we know the direction of the inequality between two moments if a further subsidiary inequality is satisfied. It is this particular result of which use is made in an attempt at proving the conjectured monotonicity.

§8.4 A Sufficient Condition, A Stationary Value And Related Results

§8.4.1 Let $\lambda = p^{(r)}$, $\tau = p^{(r+1)}$ where $p^{(r)}$, $p^{(r+1)}$ are successive iterates of algorithm $FP\{d^{\delta},1\}$, $\delta = 1/(t+1)$, in the solution of (P1). Thus

8.4.1 $$\tau_j = \lambda_j\{\partial\phi/\partial\lambda_j\}^{1/(t+1)} \Big/ \Big\{\sum_{i=1}^{J} \lambda_i\{\partial\phi/\partial\lambda_i\}^{1/(t+1)}\Big\}$$

If the algorithm is monotonic then

8.4.2 $$\phi(\tau) \geq \phi(\lambda).$$

Lemma 8.4.1

Assume that $\phi(p)$ has positive derivatives and, for $t > 0$, is homogeneous of degree $(-t)$. Let λ, τ be related by 8.4.1. Then a sufficient condition for the inequality $\phi(\tau) \geq \phi(\lambda)$ to be satisfied is the inequality.

8.4.3 $$\Phi(\lambda|\tau) \geq \Phi(\tau|\tau)\ ,$$

where

8.4.4 $$\Phi(\lambda|\rho) = \sum_{i=1}^{J} \lambda_i\{\partial\phi/\partial\lambda_i\}^{1/(t+1)} \cdot \{\partial\phi/\partial\rho_i\}^{t/(t+1)}.$$

Proof We appeal to corollary 8.3.2.

Let x and y be discrete positive random variables with the probability distributions.

$$\Pr\{x = \{\partial\phi/\partial\lambda_j\}^{1/(t+1)}\} = \lambda_j\ ;\ \Pr\{y = \{\partial\phi/\partial\tau_j\}^{1/(t+1)}\} = \tau_j$$

Then in view of the homogeneity of $\phi(\cdot)$,

$$\{E(x^{(t+1)})\} = -t\phi(\lambda)\quad ,\quad \{E(y^{(t+1)})\} = -t\phi(\tau)$$

Hence $\phi(\tau) \geq \phi(\lambda)$ if and only if

8.4.5 $$\{E(y^{(t+1)})\} \leq \{E(x^{(t+1)})\}$$

Take n,r in lemma 8.3.1 to be n = t+1, r = 1, in which case condition (a) is satisfied. Hence, from corollary 8.3.2, a sufficient condition for 8.4.5 is the inequality

$$E(x)E(y^t) \geq \{E(y^{(t+1)})\}$$

Now

$$\begin{aligned} E(x)E(y^t) &= \Big\{\sum \lambda_i\{\partial\phi/\partial\lambda_i\}^{1/(t+1)}\Big\}\Big\{\sum \tau_j\{\partial\phi/\partial\tau_j\}^{t/(t+1)}\Big\} \\ &= \sum \lambda_j\{\partial\phi/\partial\lambda_j\}^{1/(t+1)}\{\partial\phi/\partial\tau_j\}^{t/(t+1)} \\ &= \Phi(\lambda|\tau) \end{aligned}$$

The proof is complete when we observe that

$$\Phi(\tau|\tau) = \sum \tau_j\{\partial\phi/\partial\tau_j\} = -t\phi(\tau) = \{E(y^{(t+1)})\}. \quad \square$$

§8.4.2 The result then, suggests replacing the burden of establishing 8.4.2 by that of proving 8.4.3 for τ, λ related by 8.4.1.

In fact we can obtain some interesting and relevant results if we regard $\Phi(\lambda|\mu)$ as a function of λ for fixed arbitrary μ.

Lemma 8.4.2

Let $\phi(p)$ satisfy the conditions of lemma 8.4.1 and let $\Phi(\lambda|\mu)$ be defined by 8.4.4. Then

(i) $\Phi(\lambda|\mu)$ is homogeneous of degree zero in λ.

(ii) $\Phi(\mu|\mu) = -t\phi(\mu)$

(iii) $\Phi(\lambda|\mu)$ is stationary with respect to λ at $\lambda = \mu$.

Proof Part (ii) we have seen above and part (i) is a consequence of $\{\partial\phi/\partial\lambda_j\}$ being homogeneous of degree $-(t+1)$. Part (iii) is trivial also.

$$\Phi(\lambda|\mu) = \sum_i \lambda_i\{\partial\phi/\partial\lambda_i\}^{1/(t+1)}\cdot\{\partial\phi/\partial\mu_i\}^{t/(t+1)} .$$

Hence

$$\partial\Phi/\partial\lambda_j = \left\{\frac{\partial\phi}{\partial\lambda_j}\right\}^{1/(t+1)}\cdot\left\{\frac{\partial\phi}{\partial\mu_j}\right\}^{t/(t+1)} + \frac{1}{(t+1)}\sum_i \lambda_i\frac{\partial^2\phi}{\partial\lambda_i\partial\lambda_j}\left\{\frac{\partial\phi}{\partial\lambda_i}\right\}^{-t/(t+1)}\cdot\left\{\frac{\partial\phi}{\partial\mu_i}\right\}^{t/(t+1)}$$

$$\therefore \partial\Phi/\partial\lambda_j\Big|_{\lambda=\mu} = \partial\phi/\partial\mu_j + \{1/(t+1)\}\sum_i \mu_i\partial^2\phi/\partial\mu_i\partial\mu_j$$
$$= \partial\phi/\partial\mu_j - \partial\phi/\partial\mu_j = 0 .$$ □

Clearly this result lends credence to the tantalising prospect that $\Phi(\lambda|\mu)$ has a minimum at $\lambda = \mu$ whatever μ is, which of course would imply 8.4.3 and hence 8.4.2. We will in fact see an example of such a minimum in chapter 9, namely for $\phi_t(p) = -\mathrm{tr}\{AM^{t}(p)A'\}^{t}$, in the case t = 1, while for other t we will see empirical evidence in support of 8.4.3 for τ, λ satisfying 8.4.1, evidence which is wide ranging enough to suggest that such a minimum exists for other t. Of minor interest is that in the special case of the function ϕ_3 of section 8.2 we have trivially $\Phi(\lambda|\mu) = \Phi(\mu|\mu)$ for all $\lambda = (\lambda_1, \cdots, \lambda_J)$ such that $\lambda_j \geq 0$.

If μ were a minimum of $\Phi(\lambda|\mu)$ one might have thought that this would be due to convexity of $\Phi(\lambda|\mu)$, a property which in turn might derive from concavity of $\phi(p)$. If this is so it is not easy to prove, nor are the necessary second order conditions for μ to be a minimal turning point, easily verified.

However clearly such a minimum is unlikely to exist for all $\phi(p)$ satisfying the relevant conditions, nor is it a necessary condition for 8.4.3 to be satisfied by λ,τ related by 8.4.1. In the next subsection (subsection 8.4.3) we obtain further results which lend some support to the possibility that the latter might be true, while in section 8.5, when discussing the EM algorithm, we derive a condition which is sufficient for inequality 8.4.3 to be satisfied by λ,τ for $\tau=\mu$, where μ is any value, in which case μ does minimise $\Phi(\lambda|\mu)$ with respect to λ. Also we obtain a further stationary value corresponding to this additional sufficient condition.

§8.4.3 One approach of more practical value which might lead to establishing whether or not a particular power δ ($\delta = 1/(t+1)$) will render $FP\{d^{\delta},1\}$ monotonic, would be to attempt derivation of a range of values of δ yielding monotonicity. Assuming the argument at the end of section 8.1 to be valid, this would simply require, in the case of a concave $\phi(p)$, a positive upper limit on such positive values. Certainly we know from (vii) of section 7.3.2 that small positive δ must yield monotonicity. If such an upper limit exceeds $1/(t+1)$, our conjectured monotonicity is established. In fact we have already reported in section 7.4.3 such a range of values with such an upper limit. Monotonicity obtains for $\delta \leq 2.1$ when $FP\{d^{\delta},1\}$ is used to calculate the D-optimum design for the space $\mathcal{V}=\{(1,-1,-1)',(1,-1,1)',(1,1,-1)',(1,2,2)'\}$.

Clearly a study of the behaviour of $\phi\{q\{d^{\delta}\}\}$ is called for, but lemma 8.3.1 suggests also a comparison with a generalisation of the function $\Phi(\lambda|\mu)$.

Now let $\lambda= p^{(r)}$ and $\tau(m) = p^{(r+1)}$, where $p^{(r)},p^{(r+1)}$ are successive iterates of algorithm $FP\{d^{\delta},1\}$, $\delta = m/(t+1)$, in the solution of problem (P1).

Thus

8.4.6 $$\tau_j(m) = \lambda_j\{\partial\phi/\partial\lambda_j\}^{m/(t+1)}\Big/\{\sum\lambda_i\{\partial\phi/\partial\lambda_i\}^{m/(t+1)}\} .$$

Let x and y be discrete positive random variables with the probability distributions

$$\Pr\{x=\{\partial\phi/\partial\lambda_j\}^{1/(t+1)}\}=\lambda_j \quad ; \quad \Pr\{y=\{\partial\phi/\partial\tau_j(m)\}^{1/(t+1)}\}=\tau_j(m) .$$

Then as derived for the case m = 1 in the proof of lemma 8.4.1.

$$E(x^{(t+1)}) = -t\phi(\lambda) = f_0$$

$$E(y^{(t+1)}) = -t\phi\{\tau(m)\} = f_1(m)$$

$$E(x^m)E(y^{(t+1-m)}) = \sum_i \lambda_i\{\partial\phi/\partial\lambda_i\}^{m/(t+1)}\cdot\{\partial\phi/\partial\tau_i(m)\}^{(t+1-m)/(t+1)} = f_2(m) .$$

Now monotonicity for a given m requires $f_1(m) < f_0$. Relevant results are

(i) $f_1(0) = f_2(0) = f_2(t+1) = f_0$

(ii) $f_1'(0) = \{t/(t+1)\}f_2'(0)$

$$= -\{t/(t+1)\}\{\sum_j\lambda_j\{\partial\phi/\partial\lambda_j\}\{\ln\{\partial\phi/\partial\lambda_j\}\} - \{\sum\lambda_j\{\partial\phi/\partial\lambda_j\}\}\{\sum\lambda_j\ln\{\partial\phi/\partial\lambda_j\}\}$$

$$= -\{t/(t+1)\}\,\mathrm{Cov}_\lambda\{\partial\phi/\partial\lambda_i\, ,\ \ln\{\partial\phi/\partial\lambda\}\}$$

We have already encountered this covariance which is positive by the results of section 7.2.2. This would also follow from the convexity of $x\ln(x)$ and the concavity of $\ln(x)$

We already know that $f_1'(0)$ should be negative, so we have established the same of $f_2'(0)$.

(iii) From lemma 8.3.1 and corollary 8.3.2 it follows that, if, for any m satisfying $0 < m < (t+1)$, $f_1(m) \le f_2(m)$ then $f_1(m) \le f_2(m) \le f_0$.

(iv) From lemma 8.3.1 it follows that, if, for any m larger than $(t+1)$, $f_1(m) \le f_2(m)$, then $f_2(m) > f_0$.

Implications of these results are as follows:

(a) Since $f_1(0) = f_0$ and $f_1'(0) < 0$, then $f_1(m) < f_0$ for small positive m, a result which we already know.

(b) Since $f_1(0) = f_2(0) = f_0$ and $f_1'(0) < 0, f_2'(0) < 0$, both functions are decreasing, at zero, away from a common value, but, since $f_1'(0) = \{t/(t+1)\} f_2'(0)$, $f_1(m)$ is doing so at a slower rate than $f_2(m)$. For small m then, $f_1(m) \geq f_2(m)$. So, although we have $f_1(m) < f_0$ for small m, the condition, $f_1(m) \leq f_2(m)$, which is sufficient for $f_1(m) < f_0$, is not satisfied at such m.

(c) In view of (iii), $f_2(m)$ cannot exceed f_0 in the range $0 < m < (t+1)$ unless $f_1(m)$ does so. The curve $f_2(m)$ cannot cross the horizontal f_0 before $f_1(m)$ in that range, or stay above f_0 if $f_1(m)$ does not.

(d) The curve $f_2(m)$ does cross f_0 at $m = t+1$ assuming that this is not a maximal turning point of $f_2(\cdot)$.

Of interest is whether or not $f_1(m), f_2(m)$ cross in $(0, t+1)$, for we will then have a range of values of m for which the sufficiency condition $f_1(m) \leq f_2(m)$ for $f_1(m) \leq f_0$ will become operative.

Clearly the above results are not conclusive either way on this question.

There is the possibility that they do not cross, an occurrence which might seem more likely if $\phi\{q\{d^{\delta}\}\}$ is unimodal and converges to $-\infty$ as δ tends to ∞, as in the case of $\phi_t(p) = -\mathrm{tr}\{AM^{+}(p)A'\}^{t}$. Then $f_1(m)$ must have a minimal turning point and thereafter $f_1(m) \longrightarrow +\infty$ as $m \longrightarrow \infty$.

However at least $f_2(m)$ is not increasing at zero away from a common value with $f_1(m)$ and we have an undisputed requirement only on $f_2(m)$ to cross the horizontal f_0 in $(0, t+1)$, namely at $m = t+1$. This obviously increases the chances of the two curves crossing.

In the case of $\phi(p) = -\sum p_i^{-t}$ we have

$$f_1(m) = t\left\{\sum \lambda_j^{-t(1-m)}\right\}\left\{\sum \lambda_j^{(1-m)}\right\}^{t}$$

$$f_2(m) = t\left\{\sum \lambda_j^{-(t-m)(1-m)}\right\}\left\{\sum \lambda_j^{(1-m)}\right\}^{t+1-m},$$

and $f_1(1) = f_2(1) = t$.

This is the function ϕ_3 of section 8.2.

We have a crossing at the value of m = 1, which we saw in that section to achieve the optimum in one step, and the other results of that section suggest that $f_1(m)$ will not cross f_0 until after m = 2.

This function is also a particular case of $\phi_t(p) = -tr\{AM^+(p)A'\}^t$. As we have already said, we will see in chapter 9, that $f_1(1) < f_2(1)$ in the case t = 1, so that the functions must cross at some point prior to m = 1. It is reasonable to expect the same of $\phi_t(p)$ at least for t near to 1. We will encounter empirical evidence that it is true for any t.

For small t we might also expect $f_1(m)$ not to cross f_0 before m = 2. The empirical results reported in Table 7.4.1 for D-optimality, which corresponds to t = 0 here, lends weight to this claim.

There emerges three pictures from these discussions.

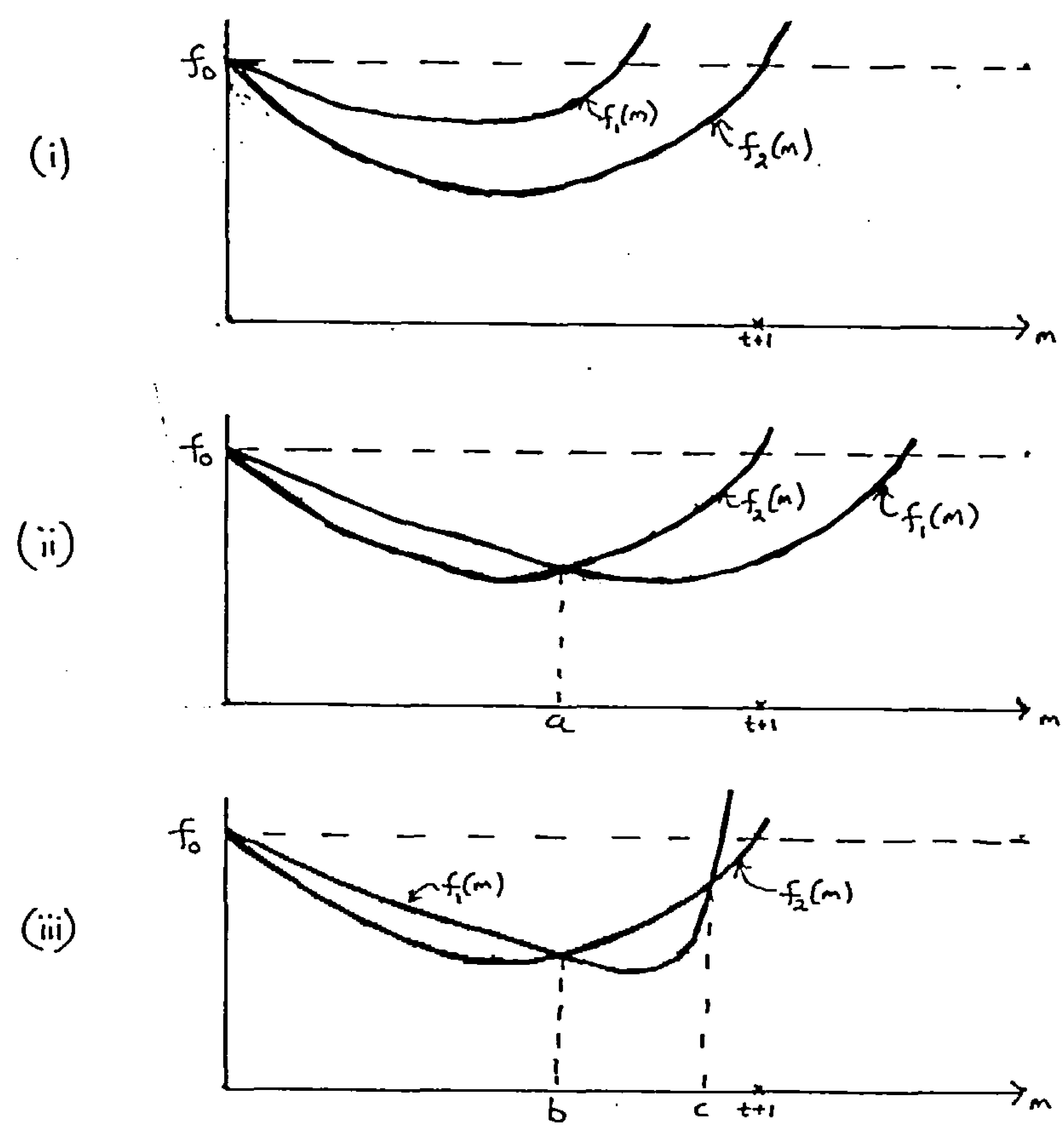


Figure 8.4.1

Graph (ii) might correspond to t small, graph (iii) to t large. Monotonicity obtains for $m = 1$ if $a < 1$ in graph (ii) or if $b < 1$, $c > 1$ in graph (iii).

§8.5 Relationship Between FP$\{d^{\delta},1\}$ And The EM Algorithm

§8.5.1 In this section a link between the EM algorithm and FP$\{d^{\delta},1\}$ is exposed, which provides further support that the latter is monotonic for $\delta = 1/(t+1)$ under the conditions previously postulated.

A monotonic iterative scheme itself, the EM algorithm was conceived or crystallised by Dempster, Laird and Rubin (1977) for contexts in which there is missing data.

Considering such a context let y denote data actually observed and let x denote what would have been the complete data. For complete generality Dempster et al express the relationship between y and x by stating that $y = y(x)$, for some many to one mapping $y(x)$ from x to y, a notation which can include instances when the missing data might be of an 'implicit' form. In many instances of course we would more naturally allocate an explicit representation to missing data, denoting it by z say, so that $x = (y,z)$.

Denote the likelihood of x by $f(x|\Theta)$. It is naturally desired to determine the maximum likelihood estimator of Θ with respect to the likelihood of y. The EM algorithm proceeds as follows.

Let $\Theta^{(r)}$ denote a current iterate. The next iterate is the end result of evaluating an expectation (E-step) and then solving a maximisation problem (M-step), namely:

E-step: compute $Q(\theta|\theta^{(r)}) = E\{\ln f(x|\theta)|y, \theta^{(r)}\}$

M-step: choose $\theta^{(r+1)}$ to be the value which solves $\max_{\theta} Q(\theta|\theta^{(r)})$.

The conditional expectation of $\ln f(x|\theta)$ given y which forms the E-step is thus evaluated on the assumption that within the conditional distribution of x given y, Θ is known to be $\Theta^{(r)}$, whereas Θ is left unspecified within $f(x|\cdot)$. The expectation is therefore a function of Θ and $\Theta^{(r)}$, and the maximisation forming the M-step has intuitive appeal in that it is a modified maximisation of the log-likelihood of the complete data x.

The following theorem due to Dempster et al implies that the algorithm is monotonic.

Theorem 8.5.1

Let $k(x|y,\theta)$ denote the conditional distribution of x given y and denote the marginal distribution of y by $g(y|\theta)$ so that

$$f(x|\theta) = k(x|y,\theta)g(y|\theta)$$

Let

$$H(\theta'|\theta) = E\{\ln k(x|y,\theta')|y,\theta\}$$
$$L(\theta) = \ln g(y|\theta)$$
$$Q(\theta'|\theta) = E\{\ln f(x|\theta')|y,\theta\} \; .$$

(In the above context we wish to maximise $g(y|\theta)$ or $L(\theta)$)

If $Q(\theta'|\theta) \geq Q(\theta|\theta)$ then $L(\theta') \geq L(\theta)$.

Proof

$$Q(\theta'|\theta) = E\{\ln k(x|y,\theta)|y,\theta\} + E\{\ln g(y|\theta')|y,\theta\}$$
$$= H(\theta'|\theta) + L(\theta')$$

so that

$$L(\theta') - L(\theta) = \{Q(\theta'|\theta) - Q(\theta|\theta)\} + \{H(\theta|\theta) - H(\theta'|\theta)\} .$$

Now

$$H(\theta'|\theta) - H(\theta|\theta) = E\left\{\ln\left\{\frac{k(x|y,\theta')}{k(x|y,\theta)}\right\}\Big| y,\theta\right\} .$$

Hence by Jensen's inequality (with equality if and only if $k(x|y,\theta') = k(x|y,\theta)$ a.e.)

$$H(\theta'|\theta) - H(\theta|\theta) \leq \ln\left\{E\{[k(x|y,\theta')/k(x|y,\theta)]|y,\theta\}\right\}$$
$$= \ln\left\{\int \frac{k(x|y,\theta')}{k(x|y,\theta)} . k(x|y,\theta)\, dx\right\}$$
$$= \ln(1) = 0$$

Hence the result. □

Monotonicity of the EM algorithm follows since, by choice $Q(\theta^{(r+1)}|\theta^{(r)}) \geq Q(\theta|\theta^{(r)})$ for any θ.

The theorem clearly establishes a stronger result than was necessary to establish the latter monotonicity. In consequence Dempster et al were prompted by it to formulate a Generalised EM algorithm or GEM. Corollary 8.5.2 readily follows from the theorem.

Defn 8.5.1

A GEM is an algorithm for which the recursion is $\theta^{(r+1)} = M(\theta^{(r)})$, where $M(\cdot)$ is some function such that $Q(M(\theta)|\theta) \geq Q(\theta|\theta)$.

Corollary 8.5.2

Let θ^* maximise $L(\theta)$. For every GEM the following is true.

(a) $L\{M(\theta^*)\} = L(\theta^*)$

(b) $Q\{M(\theta^*)|\theta^*\} = Q(\theta^*|\theta^*)$

(c) $k(x|y, M(\theta^*)) = k(x|y, \theta^*)$ a.e.

(d) $M(\theta^*) = \theta^*$ □

§8.5.2 We now consider a number of examples with the following features

(i) the function $\ln f(x|\theta)$ is linear in z so that the E-step simply consists of replacing the missing z by $E(z|y, \theta^{(r)})$.

(ii) the parameter θ is a probability vector p so that the likelihood $g(y|p)$ of the observed data is a function $\phi(p)$ and the maximum likelihood estimates $\hat{p}_{ML}$ is a p^* which solves (P1) for this function.

(iii) the EM algorithm reduces to $FP\{d, 1\}$ thereby establishing monotonicity of the latter.

Ex. 8.5.1 Consider first in a slightly modified form the example used by Dempster et al (1977) to introduce the EM algorithm. The example is based on data reported by Rao (1965 pp 368-369).

The complete data set would have been $(x_1, x_2, x_3, x_4, x_5)$, the cell counts in observations drawn from a 5-cell multinomial for which the probability vector is of the form

$$(1/2, p_1/4, p_2/4, p_2/4, p_1/4), \quad p_1, p_2 > 0 \quad , p_1 + p_2 = 1$$

so that

$$f(\underline{x}, p) = c(x)\ (1/2)^{x_1}(1/4)^{n-x_1}(p_1)^{x_2+x_5}(p_2)^{x_3+x_4},$$

where $n = x_1+x_2+x_3+x_4+x_5$.

Dempster et al expressed this in terms of π, where $\pi, (1-\pi) = p_1, p_2$.

The missing data was x_1 and x_2. Only $(y_1\ y_2\ y_3\ y_4) = (x_1+x_2, x_3, x_4, x_5)$ were observed for which the likelihood is

$$\phi(p) = g(y|p) = c_y(1/2 + p_1/4)^{y_1}(p_2/4)^{y_2+y_3}(p_1/4)^{y_4}$$

Clearly $f(\underline{x},p)$ is linear in $z = (x_1,x_2)$ so that the E-step would replace x_1,x_2 by

$$E(x_1|y) = E(x_1|y_1) = (y_1/2)\Big/\left(1/2 + p_1^{(r)}/4\right) = x_1^{(r)}$$

$$E(x_2|y) = E(x_2|y_1) = \left(y_1 p_1^{(r)}/4\right)\Big/\left(1/2 + p_1^{(r)}/4\right) = x_2^{(r)}.$$

The M-step fairly simply takes

$$p_1^{(r+1)} = \left(x_2^{(r)} + x_5\right)\Big/\left(x_2^{(r)} + x_3 + x_4 + x_5\right) \; ; \quad p_2^{(r+1)} = (x_3 + x_4)\Big/\left(x_2^{(r)} + x_3 + x_4 + x_5\right)$$

so that

$$p_1^{(r+1)} = p_1^{(r)}\left\{\left[y_1 + y_4 + 2y_4/p_1^{(r)}\right]\Big/\left[n p_1^{(r)} + 2(n - y_1)\right]\right\}$$

Hence

$$p_1^{(r+1)} = p_1^{(r)} \partial\phi/\partial p_1^{(r)} \Big/ \left\{\sum_{i=1}^{2} p_i^{(r)} \partial\phi/\partial p_i^{(r)}\right\} ,$$

for $\phi(p)$ has the positive derivatives

$$\partial\phi/\partial p_1 = \phi(p)\left[y_1 + y_4 + (2y_4/p_1)\right]\big/(2+p_1) \; ; \quad \partial\phi/\partial p_2 = \phi(p)\left[y_2 + y_3\right]/p_2 .$$

Hence $FP\{d, 1\}$ is monotonic in the case of the above $\phi(p)$ but note that this is not homogeneous. However we would obtain a homogeneous function by substituting '1' by $(p_1 + p_2)$ namely the function

$$\phi(p) = c_y\left[(p_1+p_2)/2 + p_1/4\right]^{y_1} (p_2/4)^{y_2+y_3} (p_1/4)^{y_4}$$

$$= c_y (1/4)^n (3p_1 + 2p_2)^{y_1} (p_2)^{y_2+y_3} (p_1)^{y_4} .$$

It would seem likely that $FP\{d,1\}$ would also be monotonic for this function, and note that, since it is a product of homogeneous functions, power $\delta = 1$ is what rule 8.2.2 would select as seen from 8.2.4.

Ex. 8.5.2 This is the problem referred to in Ex. 1.1.4.

The observed data y are the off-diagonal elements n_{ij} of a J×J contingency table. It is postulated that for $i \neq j$

$$p_{j|i} = p_j/(1-p_i) ,$$

where $p_{j/i}$ is the conditional probability of the $(i,j)^{th}$ cell, given that it is in row i and given $j \neq i$, while $p_1,\dots,p_J$ are unknown probabilities. Maximum likelihood estimates of the latter are desired. The likelihood of y conditional on the totals $\sum_{\substack{j=1\\ j\neq i}}^{J} n_{ij}$ is proportional to

$$\phi(p) = \prod_{\substack{i=1\\ i\neq j}}^{J} \prod_{j=1}^{J} \{p_j/(1-p_i)\}^{n_{ij}} = \prod_{j=1}^{J} p_j^{c_j} (1-p_j)^{-R_j}$$

where $C_j = \sum_{\substack{i=1 \\ i \neq j}}^{J} n_{ij}$, $R_i = \sum_{\substack{j=1 \\ j \neq i}}^{J} n_{ij}$

We have

$$\partial\phi/\partial p_j = \phi(p)\{C_j/p_j + R_j/(1-p_j)\}$$

so that

8.5.1 $$p_j\,\partial\phi/\partial p_j = \phi(p)\{C_j + p_j(R_j - C_j)\}/(1-p_j)$$

Clearly the missing data are the diagonal entries, so that $z = (n_{11},\ldots,n_{JJ})$ while $x = \{n_{ij}: i = 1,\ldots,J,\ j = 1,\ldots,J\}$.

The above conditional probabilities are implied by the following full independence model for the complete table:

$$f(x|\theta) = c_x \prod_{i=1}^{J}\prod_{j=1}^{J}(p_j q_i)^{n_{ij}}$$
$$= c_x \prod_{j=1}^{J} p_j^{n_{\cdot j}} \prod_{i=1}^{J} q_i^{n_{i\cdot}}$$

where $q_1,\ldots,q_J$ is a further set of probabilities, $n_{i\cdot} = \sum_{j=1}^{J} n_{ij}$, $n_{\cdot j} = \sum_{i=1}^{J} n_{ij}$, $\theta = (p,q)$.

The model states that cell probabilities satisfy

$$\Pr\{(i,j) = (r,s)\} = q_r p_s$$
$$\therefore\ \Pr\{(i,j) = (r,s) \mid r \neq s\} = q_r p_s \Big/ \Big[1 - \sum_{t=1}^{J} q_t p_t\Big] \quad , \quad r \neq s$$
$$\therefore\ \Pr\{j = s \mid i = r \neq s\} = p_s/(1-p_r) \implies p_{j|i} = p_j/(1-p_i).$$

Note however that the marginal distribution $g(y|\theta)$ determined by this $f(x|\theta)$ will not be the above $\phi(p)$.

For a complete data set we would have

$$\hat{p}_j = n_{\cdot j}/n \ ,\ \hat{q}_i = n_{i\cdot}/n \ ,\ n = \sum_{i=1}^{J}\sum_{j=1}^{J} n_{ij}$$

The M-step thus employs the iteration

8.5.2 $$p_j^{(r+1)} = \left(n_{jj}^{(r)} + C_j\right) \Big/ \left\{\sum_{t=1}^{J}\left(n_{tt}^{(r)} + C_t\right)\right\}$$

8.5.3 $$q_i^{(r+1)} = \left(n_{ii}^{(r)} + R_i\right) \Big/ \left\{\sum_{t=1}^{J}\left(n_{tt}^{(r)} + R_t\right)\right\}$$

Since $\ln f(x|\theta)$ is clearly linear in z we wish to replace n_{ii} by $E\{n_{ii}|y,\theta^{(r)}\}$ at the E-step. As Morgan and Titterington (1977) observe there are three options available to us. A formula for this expectation can be derived by first conditioning on the row sums or on the column sums or on the cell totals of the complete data table.

Employing the first option is to observe that

$$E\{n_{jj} \mid n_{j.}, \theta^{(r)}\} = n_{j.} \, p_j^{(r)}$$

Thus

$$E\{n_{jj} \mid y, \theta^{(r)}\} = p_j^{(r)}\{E\{n_{jj} \mid y, \theta^{(r)}\} + R_j\}$$

$$\Longrightarrow E\{n_{jj} \mid y, \theta^{(r)}\} = R_j \, p_j^{(r)}/(1 - p_j^{(r)})$$

This implies

8.5.4 $$n_{jj}^{(r)} = R_j \, p_j^{(r)}/(1 - p_j^{(r)})$$

Substituting into 8.5.2 we obtain

$$p_j^{(r+1)} = \frac{[C_j + p_j^{(r)}(R_j - C_j)]/(1 - p_j^{(r)})}{\sum_{i=1}^{I}\{[C_i + p_i^{(r)}(R_i - C_i)]/(1 - p_i^{(r)})\}} ,$$

which in view of 8.5.1 is to employ the rule

$$p_j^{(r+1)} = p_j^{(r)} \partial\phi/\partial p_j^{(r)} \Big/ \{\sum p_i^{(r)} \partial\phi/\partial p_i^{(r)}\} .$$

Thus FP$\{d,1\}$ is monotonic for the above $\phi(p)$. Again this is not homogeneous, but again we have a homogenised form of $\phi(p)$ in equation 1.1.1, which is a product of homogeneous functions.

If we adopted the second of the above options we would conclude that

8.5.5 $$n_{jj}^{(r)} = E\{n_{jj} \mid y, \theta^{(r)}\} = C_j \, q_j^{(r)}/(1 - q_j^{(r)})$$

The resultant EM iteration would reduce to FP$\{d,1\}$ for $\phi(q)$ with n_{ij} replaced by n_{ji}.

Adopting the third approach we obtain simultaneous linear equations for $E\{n_{tt} \mid y, \theta^{(r)}\}$.

First we note that

$$E\{n_{tt} \mid y, \theta^{(r)}\} = n \, p_t^{(r)} q_t^{(r)}$$

where $n = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij}$. Hence

$$E(n_{tt} \mid y, \theta^{(r)}) = p_t^{(r)} q_t^{(r)} \{\sum_{i=1}^{I} E(n_{ii} \mid y, \theta^{(r)}) + N\}$$

where $N = \sum_{j=1}^{J} C_j = \sum_{i=1}^{I} R_i = \sum_{\substack{i=1 \\ i \neq j}}^{I} \sum_{j=1}^{J} n_{ij}$

In consequence the solution is given by $\underset{\sim}{E} = N\{D^{-1} - \underset{\sim}{1}\underset{\sim}{1}'\}^{-1}\underset{\sim}{1}$,

$\underset{\sim}{E}$ being a $J\times1$ vector whose t^{th} component is $E\{n_{tt}|y, \theta^{(r)}\}$,
$D = \mathrm{diag}\{p_1^{(r)} q_1^{(r)}, \cdots\cdots, p_J^{(r)} q_J^{(r)}\}$ and $\underset{\sim}{1}$ is a $J\times1$ vector of one's.

Appealing to lemma 5.2.1 we have

$$\underset{\sim}{E} = N\{D + (D\underset{\sim}{1}\underset{\sim}{1}'D)/(1 - \underset{\sim}{1}'D\underset{\sim}{1})\}\underset{\sim}{1} \quad = \quad ND\underset{\sim}{1}\big/(1 - \underset{\sim}{1}'D\underset{\sim}{1}).$$

In consequence

8.5.6 $$n_{tt}^{(r)} = E(n_{tt}|y, \theta^{(r)}) = N p_t^{(r)} q_t^{(r)}\Big/\Big\{1 - \sum_{j=1}^{J} p_j^{(r)} q_j^{(r)}\Big\}$$

We will consider the resultant EM iteration in section 10.3.

<u>Ex. 8.5.3</u> Recall example 1.1.2. There the function

$$\phi(p) = \prod_{i=1}^{n}\Big\{\sum_{j=1}^{J} p_j f_j(y_i)\Big\}$$

was introduced. The functions $f_j(y)$ are component probability models, $\phi(p)$ is the likelihood of data $y_1, \ldots, y_n$ independently obtained from a mixture of these distribution, and maximum likelihood estimates of the mixing weights are desired.

Dempster et al (1977) analyse this problem in detail, the missing data being lack of knowledge of the component distribution from which each observation y_i comes. That is, we do not know the $J\times1$ vectors $\underset{\sim}{z}_1, \ldots, \underset{\sim}{z}_n$, where $\underset{\sim}{z}_i = \underset{\sim}{e}_j$, the j^{th} unit vector, if y_i comes from $f_j(y)$.

From the point of view of the EM algorithm we can let the component distributions depend on unknown parameters τ.

Thus we have $\theta = (p, \tau)$, $z = (\underset{\sim}{z}_1, \ldots, \underset{\sim}{z}_n)$

$x = (y_1, \ldots, y_n, \underset{\sim}{z}_1, \ldots, \underset{\sim}{z}_n)$

Now y_i, $\underset{\sim}{z}_i$ have the joint probability distribution.

$$p(y_i, z_i | p, \tau) = \begin{cases} p_j f_j(y_i|\tau), & z_i = e_j, \; j = 1, \ldots, J \\ 0 & \text{else} \end{cases}$$

from which it follows that

$$\Pr\{z_i = e_j \mid y_i, p, \tau\} = p_j f_j(y_i|\tau) \Big/ \left\{\sum_{t=1}^{J} p_t f_t(y_i|\tau)\right\}$$

and hence that

$$\underset{\sim}{E}_i(p,\tau) = E\{z_i \mid y_i, p, \tau\} = \left\{\sum_{s=1}^{J} p_s f_s(y_i|\tau)\, \underset{\sim}{e}_s\right\} \Big/ \left\{\sum_{t=1}^{J} p_t f_t(y_i|\tau)\right\}$$

Also

$$\ln p(y_i, z_i \mid p, \tau) = \begin{cases} \ln p_j + \ln f_j(y_i|\tau) & z_i = \underset{\sim}{e}_j\ ,\ j = 1, \dots, J \\ 0 & \text{otherwise} \end{cases}$$

Equivalently

$$\ln p(y_i, z_i \mid p, \tau) = z_i' \underset{\sim}{\ln p} + z_i' \underset{\sim}{f}(y_i|\tau) \quad , \quad z_i = \underset{\sim}{e}_1, \underset{\sim}{e}_2, \dots, \underset{\sim}{e}_J$$

where $\underset{\sim}{\ln p} = (\ln p_1, \ln p_2, \dots, \ln p_J)'$, $\underset{\sim}{f}(y_i|\tau) = (f_1(y_i|\tau), \dots, f_J(y_i|\tau))'$

Assuming independence between $(y_1, z_1), \dots, (y_n, z_n)$ it would follow that

$$\ln f(x|\tau, p) = \left(\sum_{i=1}^{n} z_i'\right) \underset{\sim}{\ln p} + \sum_{i=1}^{n} z_i' \underset{\sim}{f}(y_i|\tau)$$

and hence that

8.5.7 $$E\{\ln f(x|\tilde{\tau}, \tilde{p}) \mid y, \tau, p\} = [\underset{\sim}{a}'(p,\tau)]\, \underset{\sim}{\ln \tilde{p}} + \sum_{i=1}^{n} \underset{\sim}{E}_i'(p,\tau) \underset{\sim}{f}(y_i|\tilde{\tau}),$$

where $$\underset{\sim}{a}(p,\tau) = \sum_{i=1}^{n} E_i(p,\tau).$$

At the E-step we replace z_i by

$$z_i^{(r)} = \underset{\sim}{E}_i(p^{(r)}, \tau^{(r)}),$$

while at the M-step with $(\tau, p) = (\tau^{(r)}, p^{(r)})$ we would choose $(\tau^{(r+1)}, p^{(r+1)})$ to be the values $(\tilde{\tau}, \tilde{p})$ which maximise the right hand side of 8.5.7.

Thus $\tau^{(r+1)}$ will depend on the nature of $f(y|\tau)$, assuming that there are parameters τ. If not delete $\tau, \tau^{(r)}$ from the above.

Whichever of these two instances obtains we will have respectively,

8.5.8 $$p_j^{(r+1)} = a_j(p^{(r)}, \tau^{(r)}) \Big/ \left\{\sum_{i=1}^{J} a_i(p^{(r)}, \tau^{(r)})\right\}$$

8.5.9 $$p_j^{(r+1)} = a_j(p^{(r)}) \Big/ \left\{\sum_{i=1}^{J} a_i(p^{(r)})\right\}$$

where

$$a_j(p^{(r)}, \tau^{(r)}) = p_j^{(r)} \sum_{i=1}^{J} \left\{ f_j(y_i|\tau^{(r)}) \Big/ \sum_{t=1}^{J} p_t^{(r)} f_t(y_i|\tau^{(r)}) \right\}$$

$$= p_j^{(r)} b_j$$

and so
$$\sum_{i=1}^{n} a_i\left(p^{(r)}, \tau^{(r)}\right) = n \ .$$

This latter result is a consequence of the fact that this is the expected value of the sum of the components of the $\underline{z}_i$, which sum must be n.

It also has another explanation, for in the case of both 8.5.8, 8.5.9 with $\phi(p) = \prod_{i=1}^{n}\left\{\sum_{j=1}^{J} p_j f_j(y_i|\tau)\right.$ in the former case, we have
$$p_j^{(r+1)} = p_j^{(r)} \partial\phi/\partial p_j^{(r)} \Big/ \sum_{i=1}^{J} p_i^{(r)} \partial\phi/\partial p_i^{(r)}$$
with $\tau = \tau^{(r)}$ as appropriate. This is so because
$$\partial\phi/\partial p_j = \phi(p)\, b_j \ ,$$
and in view of the fact that $\phi(p)$ is a product of n functions, each linear in p, so that it is homogeneous of degree n, we have
$$\sum p_j \partial\phi/\partial p_j = n\phi(p) \ .$$

In the case when τ is absent, we again have the EM algorithm reducing to FP$\{d,1\}$.

§8.5.3 The latter monotonicity of FP$\{d,1\}$ could in fact have been established by an appeal to the result proved by Baum and Eagon (1967), for the relevant $\phi(p)$ is a homogeneous polynomial with positive coefficients. So also is det(M(p)) where $M(p) = \sum_{j=1}^{J} p_j v_j v_j'$, for
$$\det\{M(p)\} = \sum_{q \in S_k} \det\{M(q)\}$$
where S_k is the set of $\binom{J}{k}$ 'designs' q such that

(i) Sup(q) contains exactly k points.

(ii) $q = (q_1, \ldots, q_k)$ assigns the same weights to its support points as does p $\left(\text{hence } \sum_{i=1}^{k} q_i < 1\right)$

From 1.3.9 we have $M(q) = V_q D(q) V_q'$,

where V_q is the k*k matrix whose columns are the support points of q and where $D(q) = \text{diag}\{q_1, \ldots, q_k\}$. Hence

$$\det\{M(q)\} = \{\det(V_q)\}^2 q_1 q_2 \quad q_k ,$$

and so det $M(p)$ is a homogeneous polynomial of degree k.

Baum and Eagon proved explicitly for such a polynomial $\phi(p)$ of degree n, that

$$\phi(\tau(\lambda)) \geqslant \phi(\lambda)$$

where

$$\tau_j(\lambda) = \lambda_j(\partial\phi/\partial\lambda_j)/n\phi(\lambda) .$$

This establishes monotonicity of $FP\{d,1\}$ since the homogeneity of degree n implies that

$$\Sigma \lambda_i \partial\phi/\partial\lambda_i = n\phi(\lambda) .$$

This monotonicity in the case of D-optimality can also be proved in other ways. An appeal to another moment inequality which is a special case of a result of Kingman (1967) concerning Radon-Nikodym derivatives obtains the result; see Titterington (1976). The monotonicity can also be obtained by presenting the relevant $FP\{d,1\}$ algorithm as an EM algorithm for pseudo missing and observed data. Silvey (1977) established this.

However it is not natural to view det$\{M(p)\}$ as a function proportional to a likelihood, but such a device, and also Baum and Eagon's result, suggests that there might exist a more general result, of which theorem 8.5.1 would be a particular case, and whose application to a function $\phi(p)$ would have implications for $FP\{d,1\}$ and possibly $FP\{d^\delta,1\}$.

In fact Baum et al (1970) prove such a result, although, like Baum and Eagon, they were motivated to do so by a statistical problem, namely that of estimating the parameters of what they called a probabilistic function of a Markov chain. Dempster et al expressed this particular problem as a mixture problem in which the indicator variables $\underset{\sim}{z}_i$ were not independent and identically distributed, but followed a Markov chain.

The proof of Baum et al's result is outlined below for an appeal to it does have implications for $FP\{d^{\delta},1\}$. Further evidence supporting the conjectured inequality of earlier sections of this chapter can be obtained.

Using the notation of Dempster et al, Baum et al (1970) proceed as follows.

Let $\mu(\cdot)$ assign finite measure to a space $\mathcal{X}$. Let $f(x|\theta)$ be a positive function on a subset of Euclidean space. Let

$$P(\theta) = \int_{\mathcal{X}} f(x|\theta)\,d\mu(x) \;; \quad Q_B(\theta'|\theta) = \int_{\mathcal{X}} \ln\{f(x|\theta')\} f(x|\theta)\,d\mu(x) .$$

<u>Theorem 8.5.3</u>

If $Q_B(\theta'|\theta) \geq Q_B(\theta|\theta)$ then $P(\theta') \geq P(\theta)$.

<u>Proof</u>

$$\ln\{P(\theta')/P(\theta)\} = \ln\left\{\int_{\mathcal{X}}\{f(x|\theta')/P(\theta)\}\,d\mu(x)\right\}$$

$$= \ln\left\{\int_{\mathcal{X}}\left\{\frac{f(x|\theta')}{f(x|\theta)}\,\frac{f(x|\theta)}{P(\theta)}\right\}d\mu(x)\right\}$$

$$= \ln\left\{\int_{\mathcal{X}}\left\{\frac{f(x|\theta')}{f(x|\theta)}\,q(x|\theta)\right\}d\mu(x)\right\}$$

where $q(x|\theta) = f(x|\theta)/P(\theta)$.

Since $\int_{\mathcal{X}} q(x|\theta)\,d\mu(x) = 1$, an appeal to Jensen's inequality obtains that

$$\ln\{P(\theta')/P(\theta)\} \geq \int_{\mathcal{X}} \ln\left\{\frac{f(x|\theta')}{f(x|\theta)}\right\} q(x|\theta)\,d\mu(x)$$

$$= \{P(\theta)\}^{-1}\int_{\mathcal{X}}\{\ln f(x|\theta') - \ln f(x|\theta)\} f(x|\theta)\,d\mu(x)$$

$$= \{P(\theta)\}^{-1}\{Q_B(\theta'|\theta) - Q_B(\theta|\theta)\} .$$

This proves the theorem. □

Baum (1977) points out that theorem 8.5.1 follows as a corollary by observing that

$$L(\theta) = \ln P(\theta)$$

$$Q(\theta'|\theta) = Q_B(\theta'|\theta)/g(y|\theta)$$

$$d\mu(x) = \begin{cases} dx & x \in \mathcal{X}(y) \\ 0 & x \notin \mathcal{X}(y) \end{cases} ,$$

where $\mathcal{X}(y)$ is the set of x's which map onto y.

On balance then, this is a more general result than that of Dempster et al. Certainly the proof cannot naturally exploit any property that might result from the function $f(x|\theta)$ being a likelihood of data x with a component missing.

Baum et al (1970) also consider the algorithmic implications of their result. Like Dempster et al they recognise that the sequence $P(\theta^{(r)})$ would be nondecreasing where $\theta^{(r+1)} = M(\theta^{(r)})$ for some function $M(\theta)$ such that $Q_B(M(\theta),\theta) \geq Q_B(\theta|\theta)$. They also think of choosing $\theta^{(r+1)}$ to be the value of θ' which maximises $Q_B(\theta'|\theta^{(r)})$ and they show that when $\theta = p$, $P(\cdot) = \phi(\cdot)$, where $\phi(p)$ is a homogeneous polynomial with positive coefficients, that the resultant algorithm is $FP\{d,1\}$, thereby deriving Baum and Eagon's result as a corollary.

From theorem 8.5.3 we can also derive an inequality which is sufficient for the truth of

8.5.10 $$\Phi(\lambda|\mu) \geq \Phi(\mu|\mu).$$

where

$$\Phi(\lambda|\mu) = \sum \lambda_i (\partial\phi/\partial\lambda_i)^{1/(t+1)} \cdot (\partial\phi/\partial\mu_i)^{t/(t+1)},$$

and $\phi(p)$ is a homogeneous function of degree $-t$ with positive derivatives. This would prove monotonicity of $FP\{d^{\delta},1\}$, $\delta = 1/(t+1)$ in the case of such functions.

<u>Lemma 8.5.4</u>

For a function $\phi(p)$ as above let

$$f_i^{\mu}(\lambda) = \lambda_i (\partial\phi/\partial\lambda_i)^{1/(t+1)} \cdot (\partial\phi/\partial\mu_i)^{t/(t+1)}$$

$$\Phi(\lambda|\mu) = \sum_{i=1}^{J} f_i^{\mu}(\lambda)$$

$$Q_{\mu}(\lambda'|\lambda) = \sum_{i=1}^{J} f_i^{\mu}(\lambda) \ln f_i^{\mu}(\lambda')$$

If $Q_{\mu}(\lambda'|\lambda) \geq Q_{\mu}(\lambda|\lambda)$, then $\Phi(\lambda'|\mu) \geq \Phi(\lambda|\mu)$

Proof

The result follows from theorem 8.5.3 by taking

$$\mathcal{X} = \{1, 2, \text{-----}, J\} \quad , \quad \theta = \lambda ,$$
$$f(x|\lambda) = f_x^{\mu}(\lambda)$$

for then

$$P(\lambda) = \Phi(\lambda|\mu)$$
$$Q_B(\lambda'|\lambda) = Q_\mu(\lambda'|\lambda) \qquad \square$$

Corollary 8.5.5

A sufficient condition for inequality 8.5.10 to be true is clearly the inequality

8.5.11 . $$Q_\mu(\lambda|\mu) \geq Q_\mu(\mu|\mu) . \qquad \square$$

Clearly, as with inequality 8.5.10, it would seem inconceivable that 8.5.11 should be true at all λ,μ for any function $\phi(p)$ satisfying the necessary conditions, and we have no examples to quote. Instead the following result is reported which suggests that 8.5.11 may fairly often be true.

Lemma 8.5.6

Let $\phi(p)$ be a homogeneous function of degree $(-t)$ with positive derivatives and let $g(\lambda|\mu) = Q_\mu(\lambda|\mu)$. Then $g(\lambda|\mu)$, as a function of λ, has a stationary value at $\lambda = \mu$.

Proof

$$g(\lambda|\mu) = \sum_{i=1}^{J} f_i^{\mu}(\mu) \ln f_i^{\mu}(\lambda)$$

where

$$f_i^{\mu}(\lambda) = \lambda_i (\partial\phi/\partial\lambda_i)^{1/(t+1)} (\partial\phi/\partial\mu_i)^{t/(t+1)}$$

$$f_i^{\mu}(\mu) = \mu_i \partial\phi/\partial\mu_i$$

Hence

$$g(\lambda|\mu) = \sum_{i=1}^{J} (\mu_i \partial\phi/\partial\mu_i)\{\ln\lambda_i + (t+1)^{-1}\ln(\partial\phi/\partial\lambda_i) + [t/(t+1)](\partial\phi/\partial\mu_i)\}$$

Thus

$$\partial g/\partial\lambda_j = \mu_j(\partial\phi/\partial\mu_j)/\lambda_j + (t+1)^{-1}\sum_{i=1}^{J}\mu_i(\partial\phi/\partial\mu_i)(\partial^2\phi/\partial\lambda_i\partial\lambda_j)(\partial\phi/\partial\lambda_i)^{-1}$$

$$\therefore \partial g/\partial\lambda_j\big|_{\lambda=\mu} = \partial\phi/\partial\mu_j + (t+1)^{-1}\sum_{i=1}^{J}\mu_i(\partial^2\phi/\partial\mu_i\partial\mu_j)$$
$$= \partial\phi/\partial\mu_j + (t+1)^{-1}\{-(t+1)\partial\phi/\partial\mu_j\} = 0 \qquad \square$$

CHAPTER 9

ON MONOTONICITY OF A FIXED POINT ALGORITHM AND OTHER CONSIDERATIONS FOR A DESIGN CRITERION

§9.1 Introduction, Upper Bounds On Iterates And Optimum Weights

In this chapter we concentrate mainly on the design criterion

$$\phi_t(p|A) = \begin{cases} -\operatorname{tr}\{AM^{+}(p)A'\}^{t}, & t>0 \\ -\log_e \det\{AM^{+}(p)A'\}, & t=0 \end{cases}$$

If $A = \underset{\sim}{c}'$ put t = 1.

In section 9.3 we establish for the case t = 1 that algorithm $FP(d^{\delta},1)$, $\delta = 1/(t+1)$, is monotonic for this criterion, while in section 9.4 some empirical results, supporting this claim for general t, are reported. Some useful matrix results are established in section 9.2.

From section 3.1 relevant formulae are

$$d_j = \partial\phi_t/\partial p_j = \begin{cases} t\,Q_j(p,A,t), & p_j>0,\ t>0 \\ Q_j(p,A,0), & p_j>0,\ t=0, \end{cases}$$

where

9.1.1 $$\begin{cases} Q_j(p,A,t) = v_j' M^{+}(p)A'[AM^{+}(p)A']^{t-1}AM^{+}(p)v_j \\ Q_j(p,I,t) = v_j' M^{-(t+1)}(p)v_j \end{cases}$$

Hence, for t>0, the function of section 8.3 is

9.1.2 $$I(\lambda|p) = t\sum_{j=1}^{J}\lambda_j[Q_j(\lambda,A,t)]^{1/(t+1)}\cdot[Q_j(p,A,t)]^{t/(t+1)}.$$

We seek to establish, where possible the inequality

9.1.3 $$I(\lambda|p) \geq \Phi(p|p).$$

We can then appeal to the following consequences of lemma 8.4.1 to claim monotonicity.

Lemma 9.1.1

If $\sum\lambda_j[Q_j(\lambda,A,t)]^{1/(t+1)}\cdot[Q_j(p,A,t)]^{t/(t+1)} \geq \operatorname{tr}[AM^{+}(p)A']^{t}$,

then $\phi_t[\tau(\lambda)|A] \geq \phi_t(\lambda|A)$ where

9.1.4 $$\tau_j(\lambda) = \lambda_j[Q_j(\lambda,A,t)]^{1/(t+1)}\Big/\sum\lambda_i[Q_i(\lambda,A,t)]^{1/(t+1)}. \quad \square$$

This would prove monotonicity for all $t>0$ and would therefore imply monotonicity in the limit as $t\to 0$.

The following discussion also illustrates what would be another consequence of 9.1.3.

We have already reported that in the case of D_A-optimality, $p_j^* \leq 1/s$, when A is of order $s\times k$ and $\text{rank}\{(A)\} = s$, while we have shown, in section 3.3, that $p_j^* = 1/s$ or $p_j^* = 0$ when $\text{Sup}(p^*)$ contains s linearly independent points. This criterion corresponds to $t = 0$ above, and the case $A = I$, $s = k$ yields D-optimality.

Atwood (1973) and Sibson and Kenny (1975) proved this result for D-optimality. Atwood also proved it for D_s-optimality but using different **proofs** for the case $M(p^*)$ singular and non-singular. The following proof does not require to make such a distinction and merges and extends the approaches used by these two sets of authors.

Lemma 9.1.2

Let $M(p) = \sum p_j v_j v_j'$ and assume that the matrices $M(p)$ and $AM^+(p)A'$ might be singular. It is the case that

9.1.5 $\quad p_j v_j' M^+(p) v_j \leq 1$

9.1.6 $\quad v_j' M^+(p) A' [AM^+(p)A']^+ AM^+(p) v_j \leq v_j' M^+(p) v_j$

9.1.7 $\quad p_j v_j' M^+(p) A' [AM^+(p)A']^+ AM^+(p) v_j \leq 1$

Proof

(i)
$$\begin{aligned} v_j' M^+(p) v_j &= v_j' M^+(p) M(p) M^+(p) v_j \\ &= \sum_{i=1}^{J} p_i (v_j' M^+(p) v_i)^2 \\ &\geq p_j (v_j' M^+(p) v_j)^2. \end{aligned}$$

Hence 9.1.5.

Sibson and Kenny used this argument for $M(p)$ nonsingular.

(ii) Trivially 9.1.6 is true if $A = c'$, where c is $k\times 1$ and $c'M^+(p)c$ is 'singular', that is, $c'M^+(p)c = 0$, for then $[c'M^+(p)c]^+ = 0$.

For any A

$$v_j' M^+(p) A' [AM^+(p)A']^+ AM^+(p) v_j = v_j' M^{+/2}(p) B M^{+/2}(p) v_j$$

where

$$\begin{aligned} M^{+/2}(p) &= (M^+(p))^{1/2} \\ B &= M^{+/2}(p) A' [AM^+(p)A']^+ AM^{+/2}(p). \end{aligned}$$

The matrix B is idempotent, hence so is (I-B) which is therefore nonnegative definite.

Hence

$$v_j' M^{+/2}(p)[I - B] M^{+/2}(p) v_j \geq 0$$

$$\Longrightarrow \qquad v_j' M^{+/2}(p) B M^{+/2}(p) v_j \leq v_j' M^{+}(p) v_j .$$

Atwood used this result for the case of D_s-optimality as in corollary 9.1.2.1 below, but only for the case of $M(p^*)$ nonsingular, and of course $AM^{-1}(p^*)A'$ nonsingular.

(iii) Clearly 9.1.7 is an immediate consequence of 9.1.6 and 9.1.5. □

We can now draw the following conclusions.

Corollary 9.1.2.1

Let p^* be D_A-optimal. Then $p_j^* \leq 1/s$.

Proof The result is clearly true if $p_j^* = 0$. If $p_j^* > 0$ then $\varphi_j(p^*, A, 0) = s$. Hence from 9.1.7 $p_j^* s \leq 1$. This, of course, is uninformative if s = 1. □

Corollary 9.1.2.2

Let $\tau_j(\lambda)$ be defined as in equation 9.1.4. When t = 0 $\tau_j(\lambda) \leq 1/s$

Proof We have

$$\tau_j(\lambda) = \lambda_j \varphi_j(\lambda, A, 0) \Big/ \sum \lambda_i \varphi_i(\lambda, A, 0)$$
$$\leq \qquad 1 \Big/ \sum \lambda_i \varphi_i(\lambda, A, 0) ,$$

by 9.1.7. The result follows since

$$\sum \lambda_i \varphi_i(\lambda, A, 0) = \sum \lambda_i v_i' M^{+}(\lambda) A' [A M^{+}(\lambda) A']^{-1} A M^{+}(\lambda) v_i$$
$$= \mathrm{trace}\left[(AM^{+}(\lambda)A')^{-1} AM^{+}(\lambda) \{\sum \lambda_i v_i v_i'\} M^{+}(\lambda) A'\right]$$
$$= \mathrm{trace}\left[(AM^{+}(\lambda)A')^{-1} AM^{+}(\lambda) M(\lambda) M^{+}(\lambda) A'\right]$$
$$= \mathrm{trace}\left[(AM^{+}(\lambda)A')^{-1} AM^{+}(\lambda) A'\right] = \mathrm{trace}(I_s) = s .$$ □

Hence for D_A-optimality iterates under $FP\{d, 1\}$ confine themselves to a subset of $\mathcal{P}$ known to contain p^*, namely $\{p = (p_1, \cdots, p_J) : p_j \leq 1/s\}$.

It is possible that a similar result obtains for general t.

First we note the following point. In 9.1.5 and in 9.1.7 there is stated an upper bound on the values respectively of $p_j Q_j(p, I, 0)$, $p_j Q_j(p, A, 0)$. The bound can in fact be attained by any design p which assigns uniform weighting to s linearly independent v_j's $\left(\text{and for which } \mathcal{N}(M(\breve{p})) \subseteq \mathcal{N}(A)\right)$.

It was established in section 3.3.1 that such a design is D_A-optimal for its support. Thus when $\breve{p}_j = 1/s$, $Q_j(\breve{p}, A, 0) = s$, $\breve{p}_j Q_j(\breve{p}, A, 0) = 1$.

This suggests the following generalisation.

<u>Lemma 9.1.3</u>

Let $m = \max_{p \in \mathcal{P}} \left\{ p_j [Q_j(p, A, t)]^{1/(t+1)} \right\}$ and let p^* be maximal for $\phi_t(p|A)$, $\phi_t^* = \phi_t(p^*|A)$. Then

$$p_j^* \leq m / (-\phi_t^*)^{1/(t+1)} .$$

<u>Proof</u> At $p_j^* > 0$, $Q_j(p^*, A, t) = -\phi_t^*$. Hence $p_j^*(-\phi_t^*)^{1/(t+1)} \leq m$. □

Clearly this result is informative only if $m < (-\phi_t^*)^{1/(t+1)}$. It seems likely that this could be the case. Certainly

$$p_j [Q_j(p, A, t)]^{1/(t+1)} \leq \sum p_i [Q_i(p, A, t)]^{1/(t+1)}$$

$$\leq \left[\sum p_i Q_i(p, A, t) \right]^{1/(t+1)} = \left[-\phi(p|A) \right]^{1/(t+1)} .$$

Here we have used the fact that for a positive random variable x, $[E(x^r)]^{1/r}$ is nondecreasing in r, taking x to have the probability distribution $\Pr\left\{ x = [Q_j(p, A, t)]^{1/(t+1)} \right\} = p_j$

It may also be, that the bound m is attained by a design p which is optimal for a support of s linearly independent points. That is, that

$$m = \left[-\phi(\breve{p}|A) \right] \cdot \max_j \{p_j\} ,$$

the maximum being present since p would not assign uniform weighting to its support points.

We will not pursue these speculations and end by recording a further generalisation. Clearly we could have stated a more general result than that in lemma 9.1.3 by replacing $1/(t+1)$ by a general power δ, but for $\delta = 1/(t+1)$ we can prove the following result.

Lemma 9.1.4

Let m be as defined in lemma 9.1.3 and $\tau_j(\lambda)$ as defined in equation 9.1.4. If inequality 9.1.3 is true, then

$$\tau_j(\lambda) \le m/(-\phi_t^*)^{1/(t+1)}.$$

Proof Clearly $\tau_j(\lambda) \le m/\sum\lambda_i[\varphi_i(\lambda,A,t)]^{1/(t+1)}$

If 9.1.3 is true then, in particular,

$$\Phi(\lambda|p^*) \ge \Phi(p^*|p^*) = -t\phi_t^*,$$

that is

$$t\sum\lambda_j[\varphi_j(\lambda,A,t)]^{1/(t+1)}\cdot[\varphi_j(p^*,A,t)]^{t/(t+1)} \ge -t\phi_t^*,$$

$$\sum\lambda_j[\varphi_j(\lambda,A,t)]^{1/(t+1)}\cdot[\varphi_j(p^*,A,t)]^{t/(t+1)} \ge -\phi_t^*$$

Now $\varphi_j(p^*,A,t) \le -\phi_t^*$ with equality if $p_j^* > 0$.

Hence

$$\sum\lambda_j[\varphi_j(\lambda,A,t)]^{1/(t+1)}\cdot[\varphi_j(p^*,A,t)]^{t/(t+1)} \le \sum\lambda_j[\varphi_j(\lambda,A,t)]^{1/(t+1)}\cdot(-\phi_t^*)^{t/(t+1)}$$

and so

$$\sum\lambda_j[\varphi_j(\lambda,A,t)]^{1/(t+1)} \ge -\phi_t^*/(-\phi_t^*)^{t/(t+1)} = (-\phi_t^*)^{1/(t+1)}$$

Thus

$$m/\sum\lambda_j[\varphi_j(\lambda,A,t)]^{1/(t+1)} \le m/(-\phi_t^*)^{1/(t+1)}.$$

Hence the result. □

If inequality 9.1.3 is true, then iterates under $FP\{d^\delta,1\}$, $\delta = 1/(t+1)$, respect upper bounds known to be satisfied by p^*.

§9.2 Some Matrix Results

We now prove some matrix results. We shall make use of one of these later. The first is ancillary to the others.

Lemma 9.2.1

Let A and B be k×k symmetric matrices. Then A and B commute if and only if they share the same eigenvectors. Equivalently

9.2.1 $AB = BA$

if and only if

9.2.2 $A = PD_AP'$, $B = PD_BP'$

where P is an orthogonal matrix and D_A, D_B are diagonal matrices.

Proof Graybill (1969) in his theorem 12.2.12 proves this result.

That 9.2.2 implies 9.2.1 follows by the orthogonality of P and from the clear commutativity between diagonal matrices. The argument is

$$AB = PD_AP'PD_BP' = PD_AD_BP' = PD_BD_AP' = PD_BP'PD_AP' = BA .$$

A proof of the converse is simple if we assume that each eigenvalue of A has multiplicity 1. Assume x is the unique normalised eigenvector corresponding to λ. Then only multiples of x can satisfy the equation $Ax = \lambda x$.

Now

$$Ax = \lambda x$$

$$\Longrightarrow BAx = \lambda Bx$$

$$\Longrightarrow A(Bx) = \lambda(Bx) .$$

Thus Bx is an eigenvector of A corresponding to λ and so in view of the assumption of multiplicity 1 it must be that $Bx = \mu x$ for some μ. Hence x is an eigenvector of B.

The latter conclusion need not necessarily follow if λ has multiplicity larger than 1. Then there will be more than one orthogonal matrix P such that $A = PD_AP'$ or $P'D_AP = A$, where D_A is diagonal. It is then necessary to show that at least one of these orthogonal matrices will also give the diagonalisation $P'BP = D_B$, for D_B diagonal. Not all of them need do so. The proof is involved and not particularly informative and so we will not report it. □

A consequence of the lemma is that the matrices M^r, M^s commute and that M, M^{-1} commute for nonnegative definite M. A possibly less obvious case is that M, M^+ commute. More generally if nonnegative definite A,B commute then so do A^p, B^q.

We now prove a matrix result of which we shall make use.

Lemma 9.2.2

Let A,B be nonnegative definite k×k symmetric matrices and let x be a k×1 vector. Then

(i) $(x'A^2x)^{1/2}\cdot(x'B^2x)^{1/2} \geq x'ABx$.

If further A and B commute, then for $1/p + 1/q = 1$

(ii) $(x'A^px)^{1/p}\cdot(x'B^qx)^{1/q} \begin{cases} \geq x'ABx & \text{if } p > 1 \\ \leq x'ABx & \text{if } p < 1 \end{cases}$,

with the further qualification in the case $p<0$ that A,B be positive definite and that x is not an eigenvector of A or B.

Proof

(i) Let $y = Ax$, $z = Bx$ and by the Cauchy-Schwarz inequality

$$(x'A^2x)^{1/2}\cdot(x'B^2x)^{1/2} = (y'y)^{1/2}\cdot(z'z)^{1/2} \geq y'z = x'ABx\,.$$

(ii) The following proof appeals to 9.2.2 and then nicely appeals to the standard Holder inequality.

$$\begin{aligned}(x'A^px)^{1/p}\cdot(x'B^qx)^{1/q} &= (x'PD_A^pP'x)^{1/p}\cdot(x'PD_B^qP'x)^{1/q} \\ &= (y'D_A^py)^{1/p}\cdot(y'D_B^qy)^{1/q}\,, \quad y = P'x \\ &= \left(\sum_{i=1}^k y_i^2 d_{Ai}^p\right)^{1/p}\cdot\left(\sum_{i=1}^k y_i^2 d_{Bi}^q\right)^{1/q},\end{aligned}$$

where $y = (y_1,\ldots,y_k)'$, $D_A = \text{diag}\{d_{A1},\ldots,d_{Ak}\}$ etc.

Hence by the standard Holder inequality, which is applicable since $d_{Ai}, d_{Bi} \geq 0$, we have

$$(x'A^px)^{1/p}\cdot(x'B^qx)^{1/q} \begin{cases} \geq \left(\sum_{i=1}^k y_i^{2/p} d_{Ai}\, y_i^{2/q} d_{Bi}\right), & p > 1 \\ \leq \left(\quad '' \quad\right), & p < 1 \end{cases}$$

and the result follows since

$$\sum_{i=1}^{k} y_i^{2/p} d_{Ai}\, y_i^{2/q} d_{Bi} = \sum_{i=1}^{k} y_i^2 d_{Ai} d_{Bi} = y' D_A D_B y = x' P D_A P' P D_B P' x = x' A B x.$$

The extra conditions imposed in the case p<0 are necessary because the Holder inequality would then demand $y_i^{2/p} d_{Ai}\,,\, y_i^{2/q} d_{Bi} > 0$. □

Clearly this result is a slight generalisation of the Holder inequality, a matrix Holder inequality. Subsequent to obtaining the above proof it was found that Shisha and Mond (1967) also provide a proof of the result, but theirs is an indirect approach with part (ii) of the lemma following as a corollary and almost an afterthought of a much more general result on complementary inequalities. In particular they do not appeal to the standard Holder inequality.

Bechenbach and Bellman (1961, p.70) also prove a complementary inequality to part (i) of the lemma in the case A,B commuting. They derive an upper bound on $(x'A^2x)(x'B^2x)$ in that case.

Other inequalities are particular cases of the lemma. For example, for A positive definite,

$$(x'Ax)(x'A^{-1}x) \geq (x'x)^2$$

is an inequality referred to by Bechenbach and Bellman (1961, theorem 13, p.70).

A corresponding matrix version of the Minkowski inequality can be established in a similar way.

<u>Lemma 9.2.3</u>

Let A and B be nonnegative definite k×k matrices and let x be a k×1 vector. Then

(i) $$\left[x'(A+B)^2 x\right]^{1/2} \leq (x'A^2x)^{1/2} + (x'B^2x)^{1/2}.$$

If further A and B commute then

(ii) $$\left[x'(A+B)^p x\right]^{1/p} \begin{cases} \leq (x'A^p x)^{1/p} + (x'B^p x)^{1/p} & \text{if } p>1 \\ \geq (x'A^p x)^{1/p} + (x'B^p x)^{1/p} & \text{if } p<1, \end{cases}$$

with the further qualification in the case p<0, that A and B be positive definite and that x is not an eigenvector of A or B.

Proof

(i) Let $y = Ax$, $z = Bx$ and $y = (y_1,..,y_k)'$, $z = (z_1,..,z_k)'$.
Then appealing to the triangle inequality

$$\begin{aligned}\left[x'(A+B)^2x\right]^{1/2} &= \left[(y+z)'(y+z)\right]^{1/2} \\ &= \left[\sum_{i=1}^{k}(y_i+z_i)^2\right]^{1/2} \\ &\le \left(\sum_{i=1}^{k}y_i^2\right)^{1/2}+\left(\sum_{i=1}^{k}z_i^2\right)^{1/2} \\ &= (x'A^2x)^{1/2}+(x'B^2x)^{1/2} .\end{aligned}$$

(ii) From 9.2.2

$$A+B = P(D_A+D_B)P'$$
$$(A+B)^p = P(D_A+D_B)^pP' .$$

Hence

$$\begin{aligned}\left[x'(A+B)^px\right]^{1/p} &= \left[x'P(D_A+D_B)^pP'x\right]^{1/p} \\ &= \left[y'(D_A+D_B)^py\right]^{1/p} , \quad y = P'x \\ &= \left\{\sum_{i=1}^{k}y_i^2(d_{Ai}+d_{Bi})^p\right\}^{1/p} \\ &= \left\{\sum_{i=1}^{k}\left[y_i^{2/p}d_{Ai}+y_i^{2/p}d_{Bi}\right]^p\right\}^{1/p} .\end{aligned}$$

Thus by the Minkowski inequality, which is applicable since $d_{Ai}\ge 0$, $d_{Bi}\ge 0$, we have

$$\left[x'(A+B)^px\right]^{1/p} \begin{cases} \le \left[\left\{\sum_{i=1}^{k}y_i^2d_{Ai}^p\right\}^{1/p}+\left\{\sum_{i=1}^{k}y_i^2d_{Bi}^p\right\}^{1/p}\right] , & p>1 \\ \ge \left[\quad " \quad + \quad " \quad\right] , & p<1 .\end{cases}$$

The result follows since

$$\sum y_i^2d_{Ai}^p = y'D_A^py = x'PD_A^pP'x = x'A^px ,$$

and similarly for B. □

We close this section by recording an example of non-commuting matrices A,B which violate part (ii) of lemma 9.2.2. Take

$$A = \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} , \qquad B = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$$

These matrices are idempotent. Hence

$$x'A^p x = x'Ax = (1/2)(x_1+x_2)^2$$
$$x'B^q x = x'Bx = x_1^2$$
$$(x'A^p x)^{1/p}(x'B^q x)^{1/q} = 2^{-p}(x_1+x_2)^{2/p}x_1^{2/q}$$
$$x'ABx = x_1(x_1+x_2)/2$$

Taking $x_2 = -x_1/2$ we then have

$$(x'A^p x)^{1/p}\cdot(x'B^q x)^{1/q} = (2^{-3/p})x_1^2$$
$$x'ABx = 2^{-2}x_1^2$$

and

$$(2^{-3/p})\,x_1^2 < (2^{-2})x_1^2 \quad \text{for} \quad p < 3/2 .$$

§9.3 Analytic Results On Monotonicity

§9.3.1 We are now in a position to establish the following result.

Lemma 9.3.1

In the case of the L_A-optimality criterion the inequality $\Phi(\lambda|\mu) \geq \Phi(\mu|\mu)$ is true when $M(\lambda) \in \mathcal{M}_A$.

Proof The L_A-optimal criterion is

$$\phi(p) = -tr\{L[AM^+(p)A']\} ,$$

for L nonnegative definite, and of course the degree of homogeneity is $(-t)$ for $t = 1$. Its first derivatives satisfy

$$\begin{aligned} \partial\phi/\partial p_j &= v_j'M^+(p)A'LAM^+(p)v_j \quad , \quad p_j > 0 \\ &= v_j'M^+(p)C'CM^+(p)v_j \quad , \end{aligned}$$

the latter form due to the fact that L can be expressed in the form B'B, so that C = BA.

We therefore have

$$\begin{aligned} \Phi(\lambda|\mu) &= \sum\lambda_j[v_j'M^+(\lambda)C'CM^+(\lambda)v_j]^{1/2}\cdot[v_j'M^+(\mu)C'CM^+(\mu)v_j]^{1/2} \\ &= \sum\lambda_j(\underset{\sim}{y}_j'\underset{\sim}{y}_j)^{1/2}(\underset{\sim}{z}_j'\underset{\sim}{z}_j)^{1/2} \end{aligned}$$

where $\underset{\sim}{y}_j = CM^+(\lambda)v_j$, $\underset{\sim}{z}_j = CM^+(\mu)v_j$.

Appealing to the Cauchy-Schwarz inequality

$$\begin{aligned} \Phi(\lambda|\mu) &\geq \sum\lambda_j\underset{\sim}{y}_j'\underset{\sim}{z}_j \\ &= \sum\lambda_j v_j'M^+(\lambda)C'CM^+(\mu)v_j \\ &= tr\{C'CM^+(\mu)(\sum\lambda_j v_j v_j')M^+(\lambda)\} \\ &= tr\{C'CM^+(\mu)M(\lambda)M^+(\lambda)\} \\ &= tr\{A'LAM^+(\mu)M(\lambda)M^+(\lambda)\} \\ &= tr\{LAM^+(\mu)M(\lambda)M^+(\lambda)A'\} \\ &= tr\{LAM^+(\mu)A'\} = -\phi(\mu) . \end{aligned}$$

The last steps of the argument appeal to equation 1.4.4 to claim that $M(\lambda)M^+(\lambda)A' = A'$ if $M(\lambda) \in \mathcal{M}_A$.

The lemma follows since $\Phi(\mu|\mu) = -\phi(\mu)$. □

We have thus established the condition sufficient for the inequality $\phi(\tau) \geq \phi(\lambda)$ to be satisfied where

$$\tau_j = \lambda_j(\partial\phi/\partial\lambda_j)^{1/2}\Big/\Sigma\lambda_i(\partial\phi/\partial\lambda_i)^{1/2} ,$$

and hence that algorithm $FP\{d^{1/2},1\}$ is monotonic for L_A-optimality and in particular for the A-optimal criterion

$$\phi_1(\rho|I) = -\operatorname{tr}\{M^{-1}(\rho)\} .$$

The above proof is a generalisation of that used by Fellman (1974) to establish the same result for the particular case of c-optimality.

§9.3.2 Consider now the particular case of $\phi_t(\rho|A)$ when A = I, namely

$$\phi_t(\rho|I) = -\operatorname{tr}\{M^{-t}(\rho)\} .$$

We record some results in connection with possible monotonicity of $FP\{d^{\delta},1\}$ with $\delta = 1/(t+1)$ for this criterion.

Lemma 9.3.2

Assume $\phi(\rho) = \phi_t(\rho|I)$. Then if nonsingular $M(\lambda)$ and nonsingular $M(\mu)$ commute the inequality $\Phi(\lambda|\mu) \geq \Phi(\mu|\mu)$ is satisfied.

Proof We have

$$\begin{aligned}\Phi(\lambda|\mu) &= t\Sigma\lambda_j[v_j'M^{-(t+1)}(\lambda)v_j]^{1/(t+1)}\cdot[v_j'M^{-(t+1)}(\mu)v_j]^{t/(t+1)} \\ &= t\Sigma\lambda_j[v_j'\{M^{-1}(\lambda)\}^p v_j]^{1/p}\cdot[v_j'\{M^{-t}(\mu)\}^q v_j]^{1/q}\end{aligned}$$

where p = t+1, (1/p+1/q) = 1

By part (ii) of lemma 9.2.2

$$\begin{aligned}\Phi(\lambda|\mu) &\geq t\Sigma\lambda_j v_j'M^{-1}(\lambda)M^{-t}(\mu)v_j \\ &= t\cdot\operatorname{tr}\{M^{-t}(\mu)(\Sigma\lambda_j v_j v_j')M^{-1}(\lambda)\} \\ &= t\cdot\operatorname{tr}\{M^{-t}(\mu)M(\lambda)M^{-1}(\lambda)\} \\ &= t\cdot\operatorname{tr}\{M^{-t}(\mu)\} = \Phi(\mu|\mu)\end{aligned}$$ □

Of course this is not a particularly useful result since typically the assumed commutativity between $M(\lambda)$, $M(\mu)$ will not be justified. Only if the design space $\mathcal{V}$ were to consist of k orthonormal vectors would the latter hold for all λ, μ , for then

$$M(p) = VPV' \ ,$$

where $P = \text{diag}\{p_1,\dots,p_k\}$ and V is a $k \times k$ orthogonal matrix. The design points are then normalised eigenvectors common to all design matrices, in which case we have the necessary and sufficient condition for commutativity.

If $M(\lambda)$, $M(\mu)$ do not commute then a proof that $\Phi(\lambda|\mu) \geq \Phi(\mu|\mu)$, if that be true, cannot appeal to part (ii) of lemma 9.2.2 in respect of $\Phi(\lambda|\mu)$, as in the above proof, for it is not guaranteed that the result in part (ii) of that lemma will hold for $A = M^{-1}(\lambda)$, $B = M^{-t}(\mu)$, $x = v_j$, $p = t+1$.

However the following lemma offers the possibility that lemma 9.2.2 could still be instrumental in establishing the truth of $\Phi(\lambda|\mu) \geq \Phi(\mu|\mu)$.

<u>Lemma 9.3.3</u>

Let $\Phi(\lambda|\mu) = t\sum\lambda_j\left[v_j'M^{-(t+1)}(\lambda)v_j\right]^{1/(t+1)}\cdot\left[v_j'M^{-(t+1)}(\mu)v_j\right]^{t/(t+1)}$,

while for nonnegative definite N let

$$\psi(N) = -\text{tr}(N^t)$$

$$\Psi(\lambda,N) = t\sum\lambda_j\left[v_j'M^{-(t+1)}(\lambda)v_j\right]^{1/(t+1)}\cdot\left[v_j'N^{(t+1)}v_j\right]^{t/(t+1)}.$$

If (i) N commutes with $M(\lambda)$

(ii) $\psi(N) = \psi\{M^{-1}(\mu)\}$

(iii) $\Psi(\lambda,N) = \Phi(\lambda|\mu)$

then $$\Phi(\lambda|\mu) \geq \Phi(\mu|\mu) = t\cdot\text{tr}\left[M^{-t}(\mu)\right].$$

<u>Proof</u> The result is fairly clear. We will have $M^{-1}(\lambda)$ and N^t commuting, and so appealing to part (ii) of lemma 9.2.2 in exactly the same way as in the proof of lemma 9.3.2 we will have

$$\Psi(\lambda,N) \geq t\sum\lambda_j v_j'M^{-1}(\lambda)N^t v_j = t\cdot\text{tr}(N^t).$$

Hence the result in view of the assumed equalities in (ii) and (iii) and a clear implication is that no such matrix N can exist if the inequality $\Phi(\lambda|\mu) \geq \Phi(\mu|\mu)$ is not satisfied. □

Still the result may not seem of much practical value. However the problem of finding a suitable N if one exists takes on the following simple form.

Since $M(\lambda)$ and N are to commute we must have, for P an orthogonal matrix

$$M(\lambda) = P D_\lambda P' \quad , \quad N = PDP'$$

where D_λ , D are diagonal matrices whose entries are the respective eigenvalues of $M(\lambda)$, N. Standard results are

$$N^s = P D^s P'$$
$$tr(N^s) = tr(D^s) = \sum_{i=1}^{k} d_i^s$$
$$v_j' N^s v_j = u_j' D^s u_j = \sum_{i=1}^{k} u_{ji}^2 d_i^s$$

where $u_j = P'v_j = (u_{j1},..,u_{jk})'$

Hence we have the apparently simpler problem of establishing the existence of k numbers $d_1,..,d_k$ satisfying, for given matrices $M(\lambda)$, $M(\mu)$,

9.3.1 $$\sum d_i^t = tr\{M^{-t}(\mu)\}$$

and

9.3.2 $$t\sum_{j=1}^{J} \lambda_j \left[v_j' M^{-(t+1)}(\lambda) v_j\right]^{1/(t+1)} \cdot \left[\sum_{i=1}^{k} u_{ji}^2 d_i^{(t+1)}\right]^{-t/(t+1)} = \Phi(\lambda|\mu).$$

Clearly this imposes just two restrictions on $d_1,..,d_k$ suggesting that there should be more than one real solution when these exist, i.e. when in fact $\Phi(\lambda|\mu) \geq \Phi(\mu|\mu)$. One would have thought that this would have made it easier to establish the existence of such a solution, although admittedly we wish to do this, not by assuming, but as a means of verifying, the truth of the inequality $\Phi(\lambda|\mu) \geq \Phi(\mu|\mu)$ for the particular case $\mu = \tau(\lambda)$,

$$\tau_j(\lambda) = \lambda_j(\partial\phi/\partial\lambda_j)^{1/(t+1)} \Big/ \sum_{i=1}^{J} \lambda_i(\partial\phi/\partial\lambda_i)^{1/(t+1)}.$$

We have no analytic results to offer in this respect, only the empirical results of the following two examples in which solutions were found for the quantities $d_1,..,d_k$.

Ex. 9.3.2.1

Taking the design space to be

$$\mathcal{V} = \{(1,0)', (0,1)', (2,1)'\} = \{v_1, v_2, v_3\}$$

so that J = 3, k = 2, letting $\lambda = (1/4, 1/4, 1/2)$ and taking t = 2 we have

$$M(\lambda) = \frac{1}{4}\begin{bmatrix} 9 & 4 \\ 4 & 3 \end{bmatrix}$$

Hence $M(\lambda) = (PD_\lambda \cdot P')$, where othogonal P and diagonal D_λ are

$$P = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}, \quad D_\lambda = \frac{1}{4}\begin{bmatrix} 11 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$v_1' M^{-3}(\lambda) v_1 = 267(4/11)^3$$
$$v_2' M^{-3}(\lambda) v_2 = 1065(4/11)^3$$
$$v_3' M^{-3}(\lambda) v_3 = 5(4/11)^3$$

Finally $\tau(\lambda)$ of 9.1.4 is given by

$$\tau_1(\lambda) = (267)^{1/3} / \{(267)^{1/3} + (1065)^{1/3} + 2(5)^{1/3}\}$$
$$\tau_2(\lambda) = (1065)^{1/3} / \{ \quad '' \quad \}$$
$$\tau_3(\lambda) = 2(5)^{1/3} / \{ \quad '' \quad \}$$

Taking $\mu = \tau(\lambda)$, values d_1, d_2 which satisfy 9.3.1, 9.3.2 are given by

$$d_1^2 + d_2^2 = 5 \cdot 3264 = \operatorname{tr} M^{-2}(\mu).$$

and

$$2/11 \left\{ (267)^{1/3} \cdot \left[(1/5)(4d_1^3 + d_2^3)\right]^{2/3} + (1065)^{1/3} \cdot \left[(1/5)(d_1^3 + 4d_2^3)\right]^{2/3} + 2 \cdot (5)^{1/3} \cdot \left[5d_1^3\right]^{2/3} \right\} = 2(5.4706)$$

A solution is given by $d_1 = 0.8934$.

Ex.9.3.2.2

Wynn's design space is adopted here

$$\mathcal{V} = \{(1,-1,-1)', (1,1,-1)', (1,-1,1)', (1,2,2)'\} = \{v_1, v_2, v_3, v_4\}.$$

so that J = 4, k = 3. With $\lambda = (1/4, 1/4, 1/4, 1/4)$

$$M(\lambda) = (1/4)\begin{bmatrix} 4 & 1 & 1 \\ 1 & 7 & 3 \\ 1 & 3 & 7 \end{bmatrix}$$

so that

$$M(\lambda) = P D_\lambda P'$$

where orthogonal P and diagonal D_λ are with $c = \{2\sqrt{11}(\sqrt{11}-3)\}^{1/2}$

$$P = \begin{bmatrix} 0 & (\sqrt{11}-3)/c & (\sqrt{11}+3)/c \\ 1/\sqrt{2} & c^{-1} & -c^{-1} \\ -1/\sqrt{2} & c^{-1} & -c^{-1} \end{bmatrix} \quad D_\lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & (7+\sqrt{11})/4 & 0 \\ 0 & 0 & (7-\sqrt{11})/4 \end{bmatrix}$$

and taking t = 3

$$\begin{aligned} v_1' M^{-3}(\lambda) v_1 &= 1880\,(4/38)^3 \\ v_2' M^{-3}(\lambda) v_2 = v_3' M^{-3}(\lambda) v_3 &= 2762{\cdot}75\,(4/38)^3 \\ v_4' M^{-3}(\lambda) v_4 &= 584\,(4/38)^3 \end{aligned}$$

Thus $\tau(\lambda)$ of 9.1.4 is given by

$$\begin{aligned} \tau_1(\lambda) &= (1880)^{1/3}\Big/\left[(1880)^{1/3} + 2(2762{\cdot}75)^{1/3} + (584)^{1/3}\right] \\ \tau_2(\lambda) = \tau_3(\lambda) &= (2762{\cdot}75)^{1/3}\Big/\left[\quad " \quad\right] \\ \tau_4(\lambda) &= (584)^{1/3}\Big/\left[\quad " \quad\right] \end{aligned}$$

Taking $\rho = \tau(\lambda)$, values of d_1, d_2, d_3 which satisfy 9.3.1, 9.3.2 are given by

$d_1^2+d_2^2+d_3^2 = 2.0701386 = \mathrm{tr}\, M^{-2}(\rho)$

and

$$\begin{aligned} 2/38\Big\{ & (1880)^{1/3}\left[(1/(2\sqrt{11}))\{(3\sqrt{11}-1)d_2^3 + (3\sqrt{11}+1)d_3^3\}\right]^{2/3} \\ & + 2(2762.75)^{1/3}\left[(1/(2\sqrt{11}))\{4\sqrt{11}d_1^3 + (\sqrt{11}-3)d_2^3 + (\sqrt{11}+3)d_3^3\}\right]^{2/3} \\ & + (584)^{1/3}\left[(1/(2\sqrt{11}))\{(9\sqrt{11}+29)d_2^3 + (9\sqrt{11}-29)d_3^3\}\right]^{2/3}\Big\} \\ & = 2 \times 2.0784770 \end{aligned}$$

A solution is given approximately by $d_1 = .05$, $d_2 = 1$.

Note that if $\Phi(\lambda|\rho) \geq \Phi(\rho|\rho)$ in this particular context when we have $A = I$, then since λ is a probability vector, the inequality states that, while we may not have the inequality

$$\left[v_j' M^{-(t+1)}(\lambda) v_j\right]^{1/(t+1)} \cdot \left[v_j' M^{-(t+1)}(\rho) v_j\right]^{t/(t+1)} \geq v_j' M^{-1}(\lambda) M^{-t}(\rho) v_j$$

satisfied for every j, we do have it satisfied on 'average'.

§9.3.3 Recall that the general function of interest in this chapter is

$$\phi_t(p|A) = -\mathrm{tr}\{AM^+(p)A'\}^t$$

and that then

$$\Phi(\lambda|\mu) = \sum \lambda_j [\mathcal{Q}_j(\lambda,A,t)]^{1/(t+1)} \cdot [\mathcal{Q}_j(\mu,A,t)]^{t/(t+1)}$$

where

$$\mathcal{Q}_j(\lambda,A,t) = v_j' M^+(p)A'[AM^+(p)A']^{t-1} AM^+(p)v_j .$$

In the latter section 9.3.2 we considered only the case $A = I$, the reason for this being that $[\mathcal{Q}_j(\lambda,I,t)]^{1/(t+1)} \cdot [\mathcal{Q}_j(\mu,I,t)]^{t/(t+1)}$ then takes the form $(x'B^p x)^{1/p} \cdot (x'C^q x)^{1/t}$, and it was then natural to consider the possibility that lemma 9.2.2 might provide a means of establishing that $\Phi(\lambda|\mu) \geq \Phi(\mu|\mu)$. For general A however $[\mathcal{Q}_j(\lambda,A,t)]^{1/(t+1)} \cdot [\mathcal{Q}_j(\mu,A,t)]^{t/(t+1)}$ does not readily take on that form. Nevertheless we can still enlist lemma 9.2.2 in much the same way as does lemma 9.3.3.

<u>Lemma 9.3.4</u>

Assume that $M(\lambda)$ is nonsingular and let $\Phi(\lambda|\mu)$ be as above, while for nonnegative definite $(k \cdot k)$ N let

$$\psi(N) = -\mathrm{tr}\, N^t$$
$$\Phi(\lambda,N) = t \sum \lambda_j [v_j' M^{-(t+1)}(\lambda) v_j]^{1/(t+1)} \cdot [v_j' N^{(t+1)} v_j]^{t/(t+1)}$$

If

(i) N commutes with $M(\lambda)$

(ii) $\psi(N) = -\mathrm{tr}\{AM^{-1}(\mu)A'\}^t$

(iii) $\Phi(\lambda,N) = \Phi(\lambda|\mu)$

then $\Phi(\lambda|\mu) \geq \Phi(\mu|\mu) = t\cdot\mathrm{tr}\{AM^+(\mu)A'\}^t$.

The proof is clearly the same as for lemma 9.3.3. □

It is possibly less natural to consider using this approach for general A and there is the disadvantage that it can only lead to establishing the desired result for nonsingular $M(\lambda)$.

Of course it need not be necessary to establish that $\Phi(\lambda|\mu) \geq \Phi(\mu|\mu)$ for $\mu = \tau(\lambda)$, where

$$\tau_j(\lambda) = \lambda_j(\partial\phi/\partial\lambda_j)^{1/(t+1)} \Big/ \sum \lambda_i(\partial\phi/\partial\lambda_i)^{1/(t+1)},$$

by an approach that somehow or other makes use of lemma 9.2.2. However no further analytic results have been obtained. We believe though that the latter inequality is true and present empirical results in section 9.4 which support this belief.

§9.4 Empirical Results On Monotonicity

In the following sets of examples, algorithm $FP\{d^{\delta}, 1\}$, $\delta = 1/(t+1)$ is monotonic from $p^{(0)} = (1/J, \ldots, 1/J)$.

Firstly monotonicity obtains in the following instances when $A = I_k$.

(i) $\mathcal{U} = \{(1,-1,-1)', (1,-1,1)', (1,1,-1)', (1,2,2)'\}$ for $t = 0,1,2,\ldots,30$.

(ii) $\mathcal{U} = \{(1,2,3)', (4,3,2)', (3,5,4)'\}$ for $t = 0,1,2,3,4,\ldots,13$.

(iii) $\mathcal{U} = \{(1,0,0)', (1,1,0)', (1,1,1)'\}$ for $t = 0,1,2,3,\ldots,6$.

(iv) $\mathcal{U} = \{(1,-1,-1)', (1,1,-1)', (1,-1,-1)', (1,2,2)', (1,1,1)', (1,0,0)', (0,1,0)', (0,0,1)', (1,1,0)', (1,0,1)', (0,1,1)', (1,1,2)' (1,2,1)'\}$ for $t = 0,1,2,3,4,\ldots,12$.

(v) $\mathcal{U} = \{(1,0,0,0)', (0,1,0,0)', (0,0,1,0)', (0,0,0,1)', (.75,.75,0,0)' (.75,0,.75,0)', (.75,0,0,.75)', (0,.75,.75,0)', (0,.75,0,.75)' (0,0,.75,.75)\}$ for $t = 0,1,2,3,4$.

(vi) $\mathcal{U} = \{(1,0,0,0)', (0,1,0,0)', (0,0,1,0)', (0,0,0,1)', (.1,.2,.3,.4)', (.9,.2,.2,.1)', (.3,.3,.3,.3)', (.5,.7,.3,.2)', (.6,.4,.3,.1)', (.4,.4,.5,.5)'\}$ for $t = 0,1,2,3,4,5,6,7,8,9$.

(vii) $\mathcal{U} = \{(1,x,x^2,x^3)' : x = \pm.44, \pm.45, \pm1\}$ for $t = 0,1,2,3$.

(viii) $\mathcal{U} = \{(1,x,x^2,x^3,x^4)' : x = 0, \pm.65, \pm.66, \pm1\}$ for $t = 0,1,2$.

(ix) $\mathcal{U} = \{(1,x_1,x_2,x_1^2,x_1x_2,x_2^2)' : x_i = -1,0,1, i=1,2\}$ for $t = 0,1,2,3,4,\ldots,13$.

(x) $\mathcal{U} = \{(1,x_1,x_2,x_3,x_1^2,x_2^2,x_3^2,x_1x_2,x_1x_3,x_2x_3)' : x_i = -1,0,1, i=1,2,3\} - \{e_1\}$, for $t = 0,1,2,3,4$.

Of course in these cases monotonicity is known to obtain for $t = 0,1$. The next set of examples considers other choices of A in which case monotonicity has only been proved for $t = 1$.

For each of the cases $t = 0,1,2,3,4,5$ and for each of the three choices of A,

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix},$$

monotonicity obtains in the case of the two design spaces

(i) $\mathcal{V} = \{(1,1,-1)',(1,-1,1)',(1,1,-1)',(1,2,2)'\}$

(ii) $\mathcal{V} = \{(1,2,3,)',(3,4,2)',(5,4,3)'\}$.

Also monotonicity obtains for each of t = 0,1,2,3,4,5 in the case of

$\mathcal{V} = \{(1,0,0)',(0,1,0)',(0,0,1)'\}$ and $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix}$

Finally we report the following results concerning the truth or otherwise of inequality 9.1.3. In all the results the matrix A = I, in which case

$$\Phi(\lambda|p) = \sum \lambda_i [v_i' M^{-(t+1)}(\lambda) v_i]^{1/(t+1)} \cdot [v_i' M^{-(t+1)}(p) v_i]^{t/(t+1)}.$$

The inequality $\Phi(\lambda|p) \geq \Phi(p|p)$ has been proved true in the case t = 1. It is also true for $(\lambda, t) = (p^{(r)}, p^{(r+1)})$ for each iterate $p^{(r)}$ of $FP(d^{\delta}, 1)$, $\delta = 1/(t+1)$ in the following instances:

(i) in the case of $\mathcal{V} = \{(1,-1,-1)',(1,\pm1,\mp1)',(1,2,2)'\}$ for t = 2,3,4 both when $p_j^{(0)} = .25$ and when $p_4^{(0)} = .01$, $p_j^{(0)} = .33$, j = 1,2,3.

(ii) in the case $\mathcal{V} = \{(1,0)',(0,1)',(2,1)'\}$ for t = 2,3, when $p_1^{(0)} = p_2^{(0)} = .25$, $p_3^{(0)} = .5$.

Also in the case of this last space and when t = 2 and $p = (.5,.3,.2)$ the inequality, $\Phi(\lambda|p) \geq \Phi(p|p)$, is satisfied by all nondegenerate probability distributions λ.

§9.5 On Computing E_A-Optimal Designs

Some comments on the E_A-optimal criterion

$$\phi_\infty(p|A) = -\lambda_{max}\{AM^+(p)A'\}$$

are in order.

As was observed in chapter 3, this criterion lacks differentiability and support differentiability if the maximum eigenvalue of $AM^+(p)A'$ has a multiplicity in excess of 1, and this is the case whether or not $M(p)$ is singular.

The lack of even support differentiability means that it will not be easy to devise numerical techniques for finding p^*. Now even iterates of an algorithm which changes only the weights of a current support may be undefined.

However, as has been observed, the function

$$f_t(p|A) = -\left[-(1/s)\phi_t(p|A)\right]^{1/t}, \quad \phi_t(p|A) = -tr\{AM^+(p)A'\}^t,$$

which does enjoy support differentiability, is such that

$$\mathop{Lt}_{t\to\infty} f_t(p|A) = \phi_\infty(p|A).$$

This suggests that for t large, $p^*(t)$ should be a good approximation to $p^*(\infty)$, where $p^*(t), p^*(\infty)$ are respectively, maximising values for $\phi_t(p|A), \phi_\infty(p|A)$. It would seem that we should have $p^*(t) \to p^*(\infty)$ as $t \to \infty$. Certainly $f_t\{p^*(t)|A\} \geq f_t(p|A)$ for all p, so that we must have $\mathop{Lt}_{t\to\infty} f_t\{p^*(t)|A\} \geq \phi_\infty\{p^*(\infty)|A\}$.

The advantage of the above relationship is that the support differentiability of $\phi_t(p|A)$ means that we can, with a little care, evaluate $p^*(t)$ using the algortihms already considered. For large t though, convergence may be exceedingly slow, particularly in the case of $FP\{d^\delta, 1\}$, $\delta = 1/(t+1)$.

Consider the case A = I which corresponds to the E-optimal criterion

$$\phi_\infty(p|I) = -\lambda_{max}\{M^{-1}(p)\}.$$

Of relevance to $FP\{d^{\delta}, 1\}$ is the term

$$Q_j(p, I, t) = v_j' M^{-(t+1)}(p) v_j = v_j' Q D^{-(t+1)} Q' v_j ,$$

where Q is orthogonal, $D = \mathrm{diag}\{d_1, .., d_k\}$.

In general for each j

$$\underset{t\to\infty}{\mathrm{Lt}} \left[v_j' Q D^{-(t+1)} Q' v_j\right]^{1/(t+1)} = \max_j \{d_j^{-1}\} = -\phi_\infty(p|I).$$

For large t then, $p^{(r+1)}$ will be only a slight modification of $p^{(r)}$ under $FP\{d^{\delta}, 1\}, \delta = 1/(t+1)$. Use of $\delta > 1/(t+1)$ might alleviate this, or use of $FP\{d^{\delta}, \alpha^*\}$ may be essential. Similar comments will be true of other algortihms.

Note however that if v_j is orthogonal to the eigenvector corresponding to $\lambda_{max}\{M^{-1}(p)\}$ then the above limit does not hold. An extreme example of this would arise if the design space $\mathcal{V}$ were to consist of k orthogonal points $v_1, .., v_k$. Then

$$M(p) = VPV' = QDQ'$$

where $V = [v_1 .... v_k]$, $Q = V(V'V)^{-1/2}$, $D = (V'V)^{1/2}P(V'V)^{1/2}$; Q is orthogonal, D is diagonal.

In this instance the terms $Q_j(p, I, t)^{1/(t+1)}$ will not in general have a common limit as $t \to \infty$, which suggests that we could let t go fully to that limit in 9.1.4. However we are considering an example for which an explicit solution is available. Since $D = \mathrm{diag}\{v_1'v_1p_1, .., v_k'v_kp_k\}$ we have

$$-\lambda_{max}\{M^{-1}(p)\} = -\max_j\{(v_j'v_j p_j)^{-1}\}.$$

If p^* is to maximise this then p^* must solve $\min_p \max_j \left\{(v_j'v_jp_j)^{-1}\right\}$ and so solves $\max_p \min_j \left\{v_j'v_jp_j\right\}$. The solution is given by $p_j^* = (v_j'v_j)^{-1} \Big/ \sum_{i=1}^{k} (v_i'v_i)^{-1}$,

$$\lambda_{max}\{M^{-1}(p^*)\} = \sum_{i=1}^{k} (v_i'v_i)^{-1}.$$

Since $\mathrm{tr}\{M^{-t}(p)\} = \sum_{j=1}^{k} (v_j'v_jp_j)^{-t}$, we have

$$p_j^*(t) = (v_j'v_j)^{-t/(t+1)} \Big/ \sum_{i=1}^{k} (v_i'v_i)^{-t/(t+1)},$$

and so we do have $p^*(t) \to p^*(\infty) = p^*$ as $t \to \infty$.

In the case $v_1 = (2,-1,-1)', v_2 = (1,0,2)', v_3 = (6,15,-3)'$ $(p_1^*, p_2^*, p_3^*) = (.45, .54, .01)$, since $v_1'v_1 = 6, v_2'v_2 = 5, v_3'v_3 = 270$; $\sum_{i=1}^{3}(v_i'v_i)^{-1} = 10/27 = .37037$. Compare this $p^* = p^*(\infty)$ with the values of $p^*(t)$ in Table 9.5.1. There is evidence there that for relatively small t we have in $p^*(t)$ a reasonable approximation to $p^*(\infty)$.

Consider also Wynn's example, the design space being $\mathcal{V} = \{v_1, v_2, v_3, v_4\} = \{(1,-1,-1)', (1,-1,1)', (1,1,-1)', (1,2,2)'\}$. From symmetry considerations it is clear that $p_2^* = p_3^*$ under E-optimal p^*. It is then possible to obtain a simple explicit formula for $\lambda_{max}\{M^{-1}(p)\}$.

Under a design p of the form $p = (p_1, q, q, p_4)$, where of course $2q = 1-p_1-p_4$, we have

$$M(p) = \begin{bmatrix} 1 & (2p_4-p_1) & (2p_4-p_1) \\ (2p_4-p_1) & (1+3p_4) & (2p_1+5p_4-1) \\ (2p_4-p_1) & (2p_1+5p_4-1) & (1+3p_4) \end{bmatrix}$$

In consequence

$$\det[M(p)-\lambda I] = [2(1-p_1-p_4)-\lambda][\lambda^2-(2p_1+8p_4+1)\lambda+2(p_1+4p_4)-2(2p_4-p_1)^2]$$

Hence $M(p)$ has eigenvalues

$$\lambda_1(p) = 2(1-p_1-p_4)$$
$$\lambda_2(p) = \left\{(2p_1+8p_4+1) - \sqrt{(1/3)(6p_1-1)^2+(2/3)(12p_4-1)^2}\right\}/2$$
$$\lambda_3(p) = \left\{(\quad \text{"} \quad) + \sqrt{\quad \text{"} \quad}\right\}/2$$

Clearly then

$$\lambda_{max}\{M^{-1}(p)\} = \max\{\lambda_1^{-1}(p), \lambda_2^{-1}(p)\}.$$

Now

$$\min_p\{\lambda_{max}\{M^{-1}(p)\}\} = \min\{\lambda_2^{-1}(p)\} = 1 \quad \text{at} \quad p_1^* = 2\lambda, p_4^* = \lambda, \ \forall \lambda \in [1/12, 1/6].$$

It would appear that the values $p^*(t)$ in Table 9.5.2, which were obtained using $FP\{d^\delta, 1\}$, $\delta = 1/(t+1)$, (monotonically, as reported in the previous section), are approximately of this form for $t \geq 6$, while they are exactly so for $t \geq 26$.

TABLE 9.5.1

t	$p_1^*(t)$	$p_2^*(t)$	$p_3^*(t)$	$-f_t\{p^*(t)/I\}$
1	.44553	.48805	.066146	0.27988
2	.45284	.51137	.035793	0.31577
3	.45370	.52018	.026113	0.33147
4	.45359	.52482	.021582	0.34020
5	.45331	.52769	.018998	0.34574
6	.45302	.52964	.017341	0.34957
7	.45275	.53106	.016192	0.35237
8	.45252	.53213	.015350	0.35450
9	.45232	.53297	.014708	0.35619
10	.45214	.53365	.014202	0.35755
11	.45199	.53421	.013794	0.35867
12	.45186	.53468	.013457	0.35961
13	.45174	.53508	.103176	0.36041
14	.45164	.53542	.012936	0.36110

TABLE 9.5.2

t	$p_1^*(t)$	$j = 2,3$ $p_j^*(t)$	$p_4^*(t)$	$-f_t\{p^*(t)/I\}$
1	.19107	.31050	.18794	0.79767
2	.21054	.31745	.15456	0.82284
3	.21921	.32051	.13976	0.83914
4	.22366	.32235	.13164	0.85203
5	.22621	.32357	.12665	0.86317
6	.22771	.32448	.12334	0.87315
7	.22853	.32521	.12104	0.88223
8	.22901	.32580	.11939	0.89053
9	.22924	.32629	.11818	0.89810
10	.22931	.32672	.11726	0.90499
11	.22932	.32705	.11657	0.91125
12	.22928	.32734	.11604	0.91693
13	.22924	.32755	.11565	0.92207
14	.22920	.32772	.11563	0.92671
15	.22921	.32782	.11516	0.93092
16	.22926	.32785	.11503	0.93474
17	.22940	.32781	.11499	0.93820
18	.22966	.32765	.11504	0.94135
19	.23008	.32736	.11519	0.94422
20	.23072	.32691	.11546	0.94684
21	.23161	.32626	.11588	0.94924
22	.23286	.32534	.11647	0.95145
23	.23451	.32410	.11728	0.95348
24	.23666	.32250	.11835	0.95535
25	.23924	.32056	.11963	0.95709
26	.24214	.31839	.12107	0.95870
27	.24516	.31613	.12258	0.96019
28	.24816	.31388	.12408	0.96158
29	.25106	.31170	.12553	0.96288
30	.25380	.30965	12690	0.96409

CHAPTER 10

GENERALISATIONS OF PROBLEMS AND ALGORITHMS

§10.1 Introduction

Up until now we have considered primarily problems (P1) and (P2). One distinctive feature of these problems is that of maximising a function $\phi(p)$ subject to p being a probability vector. Also typically the function $\phi(p)$ was concave and certainly in all examples considered the function $\phi(p)$ had positive derivatives which made it possible to employ algorithm $FP\{d^{\delta}, 1\}$.

We now examine some generalisations both in problem and algorithm. In particular we consider three sources of such generalising:

(i) by relaxing the assumptions on $\phi(\cdot)$ including that of positive derivative:

(ii) by discussing a fusion of $FP\{d^{\delta}, 1\}$ and of algorithm $S\{A, \alpha\}$

(iii) by considering in more detail problem (P3), that of optimising a function $\Phi(\theta)$ subject to several linear constraints $C\theta = b$ and $\theta_j \geq 0$.

In short we consider respectively extensions via generalising $\phi(p)$, generalising $FP\{d^{\delta}, 1\}$ and generalising the constraints.

§10.2 Relaxing Assumptions On $\phi(p)$

§10.2.1 The primary concern of this section is to consider how we might proceed to solve (P1), when the function $\phi(p)$ does not enjoy the properties which have been enjoyed by the functions we have so far considered. In particular we have in mind that the function might not have positive derivatives everywhere. It may then be that $\phi(p)$ is more likely not to be concave. If so the recommendations below can only claim to find a constrained stationary value p^* of $\phi(p)$, subject to $p^* \in \mathcal{P}$. We will be assuming that $\phi(p)$ is differentiable.

The point about such a possibility is that algorithm $FP\{d^\delta, 1\}$ or $FP\{d^\delta, \alpha_r\}$ could then be undefined. Of course other algorithms would not be troubled by negative derivatives such as $FP\{h(d,\delta), 1\}$, and in particular $FP\{\exp[\delta g(d)], 1\}$. However there are some interesting empirical results, mainly concerning monotonicity, to report in respect of a modification of $FP\{d^\delta, 1\}$; namely $FP\{|d|^\delta, 1\}$.

The function $|d|^\delta$ does not enjoy one of the properties recommended for $h(d,\delta)$, namely that it should be nondecreasing in d. However a motivation for considering $|d|^\delta$ lies in the fact that, since p^* must still be a constrained stationary value, we must have

$$\partial\phi/\partial p_j^* \begin{cases} = \sum_{i=1}^{J} p_i^* \partial\phi/\partial p_i^* & , \text{ if } p_j^* > 0 \\ \leq \quad " & , \;\; " \;\; p_j^* = 0 \end{cases}$$

and so if $\phi(p)$ is homogeneous of degree $(-t)$ then we still have

$$\partial\phi/\partial p_j^* \begin{cases} = -t\phi(p^*) & , \text{ if } p_j^* > 0 \\ < \quad " & , \;\; " \;\; p_j = 0 \end{cases}$$

Hence at the optimum all derivatives corresponding to nonzero p_j^* share a common sign, assuming that these derivatives are nonzero.

Suppose that $p_j^* > 0$ for all j and that $\partial\phi/\partial p_j^* \neq 0$. Then, at p^*, all derivatives share a common sign, and it is to be expected that the same will be true at all p in a region about p^*. If an initial $p^{(0)}$ lies in this region and if $\phi(\cdot)$ is concave in the region, one might expect subsequent iterates to remain in the region under a reasonable algorithm. If this is so with $FP\{|d|^\delta, 1\}$, then this should enjoy properties which $FP\{d^\delta, 1\}$ enjoys in a comparable positive derivative context. In particular monoticity might obtain for specific values of δ.

Of course finding such an initial $p^{(0)}$ may not be easy, possibly requiring a search, and identification will be even more difficult if some optimal weights are zero. Note however that if, for some j, $\partial\phi/\partial p_j$ is consistently out of step in sign with other derivatives, this may suggest that $p_j^* = 0$.

Clearly it would be more sensible to employ at least initially some other algorithm such as $FP\{\exp[\delta s(d)], 1\}$. However it may still be that $FP\{|d|^\delta, 1\}$ would converge to p^* from $p^{(0)} = (1/J,\dots,1/J)$, and we have empirical results to report in this respect. In general though if $FP\{|d|^\delta, 1\}$ does converge, it is only guaranteed that it will converge to a point $\breve{p}$ which is such that the derivatives $\partial\phi/\partial\breve{p}_j$, corresponding to support points of p, are equal in numerical value. There may still be differences in sign among these derivatives. This will certainly be the case if $\Sigma\breve{p}_j\partial\phi/\partial\breve{p}_j = 0$, a condition which would hold if $\phi(\cdot)$ is homogeneous and $\phi(\breve{p}) = 0$, or if $\phi(p)$ is simply homogeneous of degree zero, as in the case of example 1.1.3. Neither $FP\{d^\delta, 1\}$ nor $FP\{|d|^\delta, 1\}$, in contrast with, say $FP\{\exp(\delta d), 1\}$, can converge to a point $\breve{p}$ such that $\partial\phi/\partial\breve{p}_j = 0$ for $\breve{p}_j > 0$, and hence neither algorithm can converge to a p^* if $\phi(p^*) = 0$ and $\phi(\cdot)$ is homogeneous, unless as contemplated in 8.2.1 we have

10.2.1 $$\partial\phi/\partial p_j = c(p)f_j(p),$$

and $c(\breve{p}) = 0$ but $f_j(\breve{p}) \neq 0$. The latter will almost certainly be the case if $c(p) = [\phi(p)]^m$ or if $c(p) = [-\phi(p)]^m$. If when $\breve{p}_j > 0$ the terms $|f_j(\breve{p})|$ share a common nonzero value, then convergence to $\breve{p}$ could obtain, since $FP\{|d|^\delta, 1\}$ would reduce to $FP\{|f|^\delta, 1\}$, where $f = (f_1(p),\dots,f_J(p))'$. We have empirical results to report on this specific point. The discussion has been proferred in anticipation of these results. They have been obtained in two examples.

§10.2.2 In this first example the function $\phi(p)$ is one which was used as a design criterion for a problem which arose in a chemical context. The following, three parameter, linear regression model described the relationship between the viscosity y and the concentration x of a chemical solution.

10.2.2 $$E(y) = \theta_1 x + \theta_2 x^{1/2} + \theta_3 x^2 \quad , \quad 0 < x \leq 0{\cdot}2$$

It was desired to estimate θ_3 as independently of θ_1 and θ_2 as possible.

There are two fairly obvious approaches to this problem, namely to design to reduce near to zero either covariances or correlations between appropriate least squares estimators. Adopting the covariance approach we would wish to solve problem (P1) for

10.2.3 $$\phi(p) = -(a'M^+(p)b)^2$$

since $|a'M^+(p)b|$ is the numerical covariance between $a'\hat{\theta}_{LS}$, $b'\hat{\theta}_{LS}$.

Clearly 10.2.3 is a generalisation of the c-optimal criterion. A further generalisation which springs to mind is

10.2.4 $$\phi(p) = -[tr\{AM^+(p)B\}]^2$$

where A,B are s×k matrices each of rank s.

It is not entirely clear what properties other than homogeneity of degree (-1) are possessed by 10.2.3. Certainly its derivatives which, at nonzero p_j, are

$$\partial\phi/\partial p_j = a'M^+(p)b\, a'M^+(p)v_j v_j' M^+(p)b \quad ,$$

need not always be positive. Note that they are of the form 10.2.1 with

$$c(p) = [-\phi(p)]^{1/2} \quad , \quad f_j(p) = a'M^+(p)v_j v_j' M^+(p)b .$$

In the case of c-optimality algorithm $FP\{d^{1/2}, 1\}$ is known to be monotonic and iterates under this scheme have the definition

$$p_j^{(r+1)} = \frac{p_j^{(r)}\left[(c'M^+(p^{(r)})v_j)^2\right]^{1/2}}{\sum p_i^{(r)}\left[(c'M^+(p^{(r)})v_i)^2\right]^{1/2}} = \frac{p_j^{(r)}\,|c'M^+(p^{(r)})v_j|}{\sum p_i^{(r)}\,|c'M^+(p^{(r)})v_i|} .$$

This suggests using algorithm $FP\{|f|^{1/2}, 1\}$ to solve problem (P1) for 10.2.1. This was tried both in the case of the above regression model, taking the discretised design space to be $x = .02(.02).20$ so that $J = 10$, and also in the case of Wynn's model

$$E(y) = \theta' v \quad , \quad v \in \mathcal{V} = \{v_1, v_2, v_3, v_4\} ,$$

where $v_1 = (1,-1,-1)'$, $v_2 = (1,-1,1)'$, $v_3 = (1,1,-1)'$, $v_4 = (1,2,2)'$.

In the case of the regression model three choices of the vector $\underline{a}$ were considered, namely $a = (1,0,0)'$, $a = (1,1,0)'$, $a = (-1,1,0)'$ while in each case $b = (0,0,1)'$. In each instance $p_j^{(0)} = (1/10)$.

In the case of Wynn's model two choices of the vector $\underline{a}$ were entertained, namely $a = (1,0,0)'$, $a = (-1,1,0)'$ while again $b = (0,0,1)'$ and $p^{(0)} = (1/4,1/4,1/4,1/4)$ for both cases.

In all cases the iterations were monotonic and convergent (though not to the optimum in one instance) and $a'M^{+}(p)b$ retained the same sign throughout. Convergence was again slow.

Consider the chemical example. In the case of all three choices of $\underline{a}$, the point of convergence seemed optimal, and in each of the three cases at least eight derivatives (the same eight derivatives) shared a common sign, the same sign at all iterates. The derivatives $\partial\phi/\partial p_3$, $\partial\phi/\partial p_9$ which correspond to $x = .06$ and $x = .18$ were sometimes out of step. Initially at $p^{(0)}$ all derivatives shared a common sign in each case, but, in the case of $a = (1,1,0)'$, $\partial\phi/\partial p_9$ changed sign to remain consistently out of step with the others, while this was the case with both $\partial\phi/\partial p_9$ and $\partial\phi/\partial p_3$ in the other two choices of $\underline{a}$. However in all three cases the point of convergence assigned zero weight to $x = .06$ and $x = .18$, and in fact their weights were put to zero in early iterations using the criteria of section 4.3. Convergence was always to a 3-point design, roughly $x = .02, .12, .2$ in the ratio 2:2:1.

Consider now the Wynn examples. These are interesting in that, for both choices of the vector $\underline{a}$, there exist designs p such that $a'M^{+}(p)b = 0$. For instance the design $p = (2/9,3/9,3/9,1/9)$

renders $M(p)$ diagonal so that $\mathrm{Cov}(\hat{\theta}_i, \hat{\theta}_j) = 0$, $i \neq j$, in which case $a'M^+(p)b = 0$ for any a,b, $a = (a_1,a_2,a_3)'$, $b = (b_1,b_2,b_3)'$, such that, for each j, at least one of a_j, b_j is zero. The latter is the case for the given choices of a,b above.

The two covariances $\mathrm{Cov}(\hat{\theta}_1, \hat{\theta}_3)$ and $\mathrm{Cov}(\hat{\theta}_2, \hat{\theta}_3)$ would also be put to zero by any design p such that

$$-p_1+p_2-p_3+2p_4 = 0$$

$$p_1-p_2-p_3+4p_4 = 0.$$

This is so because these are respectively the elements m_{13}, m_{23} of the matrix $M(p)$ as was reported in section 3.3.4. The matrix $M(p)$ is then of the form

$$M(p) = \begin{bmatrix} m_{11} & m_{12} & 0 \\ m_{12} & m_{22} & 0 \\ 0 & 0 & m_{33} \end{bmatrix},$$

and $M^{-1}(p)$ or $M^+(p)$ will also partition in this form. This guarantees that the above two covariances will be zero and hence so also will $a'M^+(p)b$ if $a = (1,0,0)'$, $b = (0,0,1)'$, or $a = (-1,1,0)'$, $b = (0,0,1)'$.

There are clearly many designs p satisfying the above two equations. These include the nonsingular design $p = (3/16,5/16,6/16,2/16)$ and the singular design $(1/2,1/2,0,0)$.

In the case of $a = (-1,1,0)'$, algorithm $FP\{|f|^{1/2}, 1\}$ converges to this latter design and happily both $a'\theta, b'\theta$ are estimable under it since $a = -(v_1+v_2)/2$, $b = (-v_1+v_2)/2$.

In the case of $a = (1,0,0)'$ the algorithm does not converge to any of the optimising p^*'s. The point of convergence is the design $p = (.19,.30,.35,.16)$ at which $|f_j| = .78427$, $f_j = a'M^{-1}(p)v_jv_j'M^{-1}(p)b$, with f_1 and f_3 being negative, while $a'M^{-1}(p)b = -0.051818$. We have already anticipated that this could happen. It is not clear that there are differences between these two examples which explain convergence to the optimum in one and not in the other. It may be a question of estimability. Under the design $p = (1/2,1/2,0,0)$, the terms f_j are $(f_1,f_2,f_3,f_4) = (-1/2,1/2,0,-1/2)$. Hence this p would be a fixed point of the algorithm. However $a'\theta$, for $a = (1,0,0)'$, is not estimable under this design. There may be no optimising design p^* which would be a fixed point of $FP\{|f|^{1/2}, 1\}$, and would offer estimability of $a'\theta, b'\theta$. We will not pursue this.

We note in conclusion that the alternative correlation based approach would seek to solve (P1) with $\phi(p) = \{a'M^+(p)b\}^2/\{a'M^+(p)a\, b'M^+(p)b\}$. This is a product of homogeneous functions and

$$\partial\phi/\partial p_j = \phi(p)\left[\frac{2a'M^+(p)v_j v_j'M^+(p)b}{a'M^+(p)b} - \frac{\{a'M^+(p)v_j\}^2}{a'M^+(p)a} - \frac{\{b'M^+(p)v_j\}^2}{b'M^+(p)b}\right]$$

$$= \phi(p) f_j(p).$$

Possibly $\delta = 1$ would render $FP\{|f|^\delta, 1\}$ monotonic.

§10.2.3 The second example exhibiting negative derivatives arises in a statistical context, the function $\phi(p)$ being a transformed multinomial likelihood. The problem also illustrates generalisations with respect to constraints and hence, details will be deferred until section 10.3.4. Save to say that it was desired to maximise a positive homogeneous function, whose derivatives could be negative, subject to the variables in its argument being positive and satisfying a linear equality constraint. At the optimum none of the derivatives could be negative since none of the variables could be zero. Starting from an initial $p^{(0)}$, which was particular to the problem, an appropriate generalisation of $FP\{|d|^\delta, 1\}$ converged monotonically to the optimum for $0 \le \delta \le 0.4$ and did not encounter negative derivatives.

§10.3 Generalising The Constraints

§10.3.1 As a problem exhibiting a generalisation of constraints, we consider problem (P3). We have already touched on this problem in chapter 6, viewing it there as a generalisation of (P1). Now we restate it as (P3): "Maximise a function $\psi(\theta)$, $\theta = (\theta_1, \dots, \theta_t)'$, subject to (i) $\theta_j \geq 0$, (ii) $C\theta = b$ where C is an $s \times t$ matrix of rank s".

As was said in section 1.1 this problem can define a further example of problem (P2), namely one in which the vertices $u_1, \dots, u_J$ would be those of the following intersections which belong to the constraint region; the intersections of the plane $C\theta = b$ with (t-s) of the regions $\theta_j = 0$. We would wish then to solve (P2) for $\phi(p) = \psi\{\theta(p)\}$, $\theta(p) = \sum p_j u_j$. Hence the reason for relabelling as $\psi(\cdot)$, the objective function $\Phi(\cdot)$ of (P3).

Discussion of conditions under which such a reduction of (P3) to (P2) is possible, is deferred to section 10.3.5. Prior to that, examples will be encountered which will illuminate that discussion.

Of general interest for the moment is that, as we shall see, other reductions of (P3) to either extensions of (P1) or to problems which are effectively (P1) can also sometimes be established.

The relevance of such links between (P1), (P2) and (P3) is that our primary concern is again in the formulation of algorithms for (P3). These links will suggest an immediate answer, namely any algorithm which has been formulated for (P2) applied to the appropriate reduction of (P3) to (P2) or (P1).

Of particular interest to us again is the use of $FP\{d^\delta, 1\}$, $d = d_1, \dots, d_J$, $d_j = \partial\phi/\partial p_j$. We note here the general result that if

$$\phi(p) = \psi\{\theta(p)\} \quad , \quad \theta(p) = \sum p_j u_j$$

then

$$\partial\phi/\partial p_j = u_j'[\partial\psi/\partial\theta] = G_\psi\{\theta(p), u_j\}$$

The iterative rule for p under $FP\{d^\delta, 1\}$ therefore transforms as follows

10.3.1 $$p_j^{(r+1)} = p_j^{(r)}\left[\partial\phi/\partial p_j^{(r)}\right]^{\delta} \Big/ \sum\left\{p_i^{(r)}\left[\partial\phi/\partial p_i^{(r)}\right]^{\delta}\right\}$$

$$= p_j^{(r)}\left[u_j'\partial\psi/\partial\theta^{(r)}\right]^{\delta} \Big/ \sum\left\{p_i^{(r)}\left[u_i'\partial\psi/\partial\theta^{(r)}\right]^{\delta}\right\}$$

Hence

$$\theta^{(r+1)} = \sum p_j^{(r+1)} u_j$$

10.3.2 $$\theta^{(r+1)} = \left[\sum p_j^{(r)}\left[u_j'\partial\psi/\partial\theta^{(r)}\right]^{\delta} u_j\right] \Big/ \left[\sum p_i^{(r)}\left[u_i'\partial\psi/\partial\theta^{(r)}\right]^{\delta}\right]$$

This is a partial restatement of 10.3.1 in terms of θ. In general it is not possible to substitute for each individual p_j in terms of θ since there need not be a unique p satisfying $\theta = \theta(p) = \sum p_j u_j$.

In practice of course it would not be of interest to derive 10.3.2. One would simply first find p^* and then evaluate $\theta^* = \theta(p^*)$.

It is of interest to do so here since the application of $FP\{d^{\delta}, 1\}$ to the solution of (P1) or (P2) for other reductions of (P3) to these problems, suggest different iterative rules for θ. We derive these in the next sections with special cases being studied first.

§<u>10.3.2</u> We consider first the simplest case to deal with. This has s = 1 and is the problem:

(P6) "Maximise $\psi(\theta)$ subject to (i) $\theta_j \geq 0$, (ii) $c'\theta = 1$, where $c = (c_1, .., c_t)'$ and $c_j > 0$."

The choice of b = 1 and the restriction $c_j > 0$ is without loss of partial generality since $a'\theta = b$ can be transformed to $c'\theta = 1$, $c_j > 0$ if $a_1, \ldots, a_t$, b are nonzero and enjoy a common sign. The possibilities of some negative c_j, of $c_j = 0$ for some j, and of b = 0 will be commented on later in a general context.

There is a fairly natural approach to reducing this particular version of (P3) to (P1). Let $p_j = c_j \theta_j$ and we wish to solve (P1) for $\phi(p) = \psi(p_1/c_1,\ldots,p_t/c_t)$. This problem is also identical to solving (P2) where the vertex u_j has one nonzero component, its j^{th} component, which is $(1/c_j)$.

If $\psi(\cdot)$ is homogeneous with positive derivatives, then $\phi(\cdot)$ also enjoys these properties. The positive property is clear from the formula $\partial\phi/\partial p_j = c_j^{-1}\partial\psi/\partial\theta_j$, and in view of the one to one correspondence between p and θ in $p_j = c_j \theta_j$, iteration 10.3.2 simplifies to

10.3.3 $$\theta_j^{(r+1)} = \theta_j^{(r)} c_j^{-\delta}\left[\partial\psi/\partial\theta_j^{(r)}\right]^{\delta} \Big/ \sum\left\{\theta_i^{(r)} c_i^{(1-\delta)}\left[\partial\psi/\partial\theta_i^{(r)}\right]^{\delta}\right\}$$

An initial $\theta^{(0)}$ corresponding to $p_j^{(0)} = 1/t$ is

10.3.4 $$\theta_j^{(0)} = 1/(tc_j).$$

Results of applying this iteration in an example satisfying the necessary conditions are reported in section 10.3.4.

§10.3.3 A second special case of problem (P3) takes b = 1 and takes C to be a matrix which has exactly one positive entry in each column, the remaining entries being zero. If the variables $\theta_1,\ldots,\theta_t$ are suitably labelled then C will have the form

$$C = \begin{bmatrix} XXXX000000 & & 000 \\ 0000XXX00 & & 000 \\ 00 \quad 0XXXXX0 & & 000 \\ \cdot \quad \cdot & & \\ & & 000 \\ 0 & & 0XXX \end{bmatrix}$$

where X denotes a positive value.

This means that the variables $\theta_1,\ldots,\theta_t$ are partitioned into s mutually exclusive and exhaustive subsets $H_1,\ldots,H_s$ and the s constraints in $C\theta = b$ comprise constraints of the type considered in the previous section, one applied to each subset H_i only.

Put this way there seems a much more obvious approach to numerical solution of (P3), namely to apply iteration 10.3.3 to each subset $Ⓗ_i$. This has the impression of viewing (P3) as requiring the simultaneous solution of s problems, each of which can be reduced to an example of (P1) or (P2). Certainly this would be the case if $\psi(\theta)$ were to factor into s functions, the m^{th} function depending on $Ⓗ_m$ only.

The constraints can be restated, as below, in terms of some coefficients $c_1,\ldots,c_t$ and two sets of labels $i_1 < i_2 <,\ldots,< i_s$ and $j_1 < j_2 <,\ldots,< j_s$ where $i_1 = 1$, $j_m = i_{m+1} - 1$ for $m = 1,\ldots,(s-1)$ and $j_s = t$. They become

$$\sum_{k=i_m}^{j_m} c_j\theta_j = 1 \quad , \quad m = 1,\ldots\ldots,s$$

Iterations under the above suggestion are then given by

10.3.5 $$\theta_\ell^{(r+1)} = \theta_\ell^{(r)} c_\ell^{-\delta}\left[\partial\psi/\partial\theta_\ell^{(r)}\right]^\delta \Big/ \sum_{k=i_m}^{j_m}\left\{\theta_k^{(r)} c_k^{1-\delta}\left[\partial\psi/\partial\theta_k^{(r)}\right]^\delta\right\}$$

for $m = 1,\ldots,s$; $\ell = i_m,\ldots,j_m$.

This iteration is clearly quite different from 10.3.2. The following example confirms this. Take $s = 2$, $t = 4$ and

$$C = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

so that the constraints reduce to $\theta_1 + \theta_2 = 1$, $\theta_3 + \theta_4 = 1$.

The full constraint region is then the set

$$S = \{(\theta_1,\theta_2,\theta_3,\theta_4) : \theta_j \geq 0 , \theta_1 + \theta_2 = 1, \ \theta_3 + \theta_4 = 1\} ,$$

which is a polygon with vertices

$$u_1 = (1,0,1,0)'$$
$$u_2 = (1,0,0,1)'$$
$$u_3 = (0,1,1,0)'$$
$$u_4 = (0,1,0,1)' ,$$

these being the intersections of $C\theta = \underline{1}$ with the two planes $\theta_i = 0$, $\theta_j = 0$ for $(i,j) = (2,4),(2,3),(1,4),(1,3)$ respectively.

That S is such a polygon follows from the fact that $\theta \in S$ satisfies $\theta = \sum_{j=1}^{4} p_j u_j$ for any probabilities $p_1,..,p_4$ satisfying

$$p_1 = (\theta_3 - \theta_2) + p_4 ,$$

$$p_2 = \theta_4 - p_4 , \quad p_3 = \theta_2 - p_4 ,$$

$$\max\{0, (\theta_2 + \theta_4) - 1\} \le p_4 \le \min\{\theta_2, \theta_4\}.$$

That $\{(\theta_2 + \theta_4) - 1\}$ is no larger than either of θ_2, θ_4 follows from the fact that $\theta_1 \ge 0$, $\theta_2 \ge 0$, $\theta_1 + \theta_2 = 1$ implies $(\theta_2 - 1) < 0$, and $\theta_3 \ge 0$, $\theta_4 \ge 0$, $\theta_3 + \theta_4 = 1$ implies $(\theta_4 - 1) < 0$.

Consider now that the iterative rules 10.3.2, 10.3.5 imply respectively the following formulae for $\theta_1^{(r+1)}$.

$$\theta_1^{(r+1)} = \frac{p_1^{(r)}\left[\partial\psi/\partial\theta_1^{(r)} + \partial\psi/\partial\theta_3^{(r)}\right]^{\delta} + p_2^{(r)}\left[\partial\psi/\partial\theta_1^{(r)} + \partial\psi/\partial\theta_4^{(r)}\right]^{\delta}}{\sum_{i=1}^{4} p_i^{(r)}\left[u_i' \, \partial\psi/\partial\theta^{(r)}\right]^{\delta}}$$

and

$$\theta_1^{(r+1)} = \theta_1^{(r)}\left[\partial\psi/\partial\theta_1^{(r)}\right]^{\delta} \Big/ \left\{\theta_1^{(r)}\left[\partial\psi/\partial\theta_1^{(r)}\right]^{\delta} + \theta_2^{(r)}\left[\partial\psi/\partial\theta_2^{(r)}\right]^{\delta}\right\}$$

On the whole 10.3.5 would seem the more natural iteration. In the following example it defines an EM algorithm when $\delta = 1$.

Recall example 8.5.2 which concerned square J×J contingency tables with missing diagonal entries. Suppose a probability model proposes independence for such a table. Denoting the entries by n_{ij}, the likelihood, conditional on $N = \sum\sum_{i \ne j} n_{ij}$, is given by

$$\begin{aligned} \psi(\underset{\sim}{\alpha}, \underset{\sim}{\beta}) &= g(\underset{\sim}{n} \mid \theta, N) \\ &\propto \prod_{j=1}^{J}\prod_{\substack{i=1 \\ i \ne j}}^{J} (\alpha_j \beta_i)^{n_{ij}} \Big/ \left[1 - \sum_{k=1}^{J} \alpha_k \beta_k\right]^{N} \\ &= \left(\prod_{j=1}^{J} \alpha_j^{C_j}\right)\left(\prod_{i=1}^{J} \beta_i^{R_i}\right) \Big/ \left[1 - \sum_{k=1}^{J} \alpha_k \beta_k\right]^{N} \end{aligned}$$

where

$$C_j = \sum_{\substack{i=1 \\ i \ne j}}^{J} n_{ij} , \quad R_i = \sum_{\substack{j=1 \\ j \ne i}}^{J} n_{ij} , \quad \theta = (\underset{\sim}{\alpha}, \underset{\sim}{\beta}).$$

Note

$$N = \sum_{j=1}^{J} C_j = \sum_{i=1}^{J} R_i .$$

To find maximum likelihood estimates of $\underline{\alpha}, \underline{\beta}$ we wish to maximise $\psi(\theta)$ with respect to θ, subject to

$$\alpha_j \geq 0, \ \beta_i \geq 0, \ \sum_{j=1}^{J} \alpha_j = 1, \ \sum_{i=1}^{J} \beta_i = 1 \ ,$$

that is, subject to (i) $\theta_j \geq 0$, (ii) $C\theta = 1$ where

$$C = \begin{bmatrix} \underline{1}' & \underline{0}' \\ \underline{0}' & \underline{1}' \end{bmatrix}$$

where $\underline{1}, \underline{0}$ are $J \times 1$ vectors of 1's and 0's respectively.

We have

$$\partial\psi/\partial\alpha_j = \psi(\underline{\alpha}, \underline{\beta}) \left\{ \frac{C_j}{\alpha_j} + \frac{N\beta_j}{(1 - \Sigma\alpha_k\beta_k)} \right\}$$

$$= \psi(\underline{\alpha}, \underline{\beta}) \left\{ \frac{N\alpha_j\beta_j + C_j[1 - \Sigma\alpha_k\beta_k]}{\alpha_j[1 - \Sigma\alpha_k\beta_k]} \right\}$$

$$\partial\psi/\partial\beta_i = \psi(\underline{\alpha}, \underline{\beta}) \left\{ \frac{N\beta_i\alpha_i + R_i[1 - \Sigma\alpha_k\beta_k]}{\beta_i[1 - \Sigma\alpha_k\beta_k]} \right\} .$$

Under 10.3.5 with $\delta = 1$ iterative rules are

$$\alpha_j^{(r+1)} = \left\{\alpha_j^{(r)} \partial\psi/\partial\alpha_j^{(r)}\right\} \Big/ \left\{\sum_{\ell=1}^{J} \alpha_\ell^{(r)} \partial\psi/\partial\alpha_\ell^{(r)}\right\}$$

$$\beta_i^{(r+1)} = \left\{\beta_i^{(r)} \partial\psi/\partial\beta_i^{(r)}\right\} \Big/ \left\{\sum_{\ell=1}^{J} \beta_\ell^{(r)} \partial\psi/\partial\beta_\ell^{(r)}\right\}$$

These simplify to

$$\alpha_j^{(r+1)} = \left\{N\alpha_j^{(r)}\beta_j^{(r)} + C_j\left[1 - \Sigma\alpha_k^{(r)}\beta_k^{(r)}\right]\right\} \Big/ \sum_{\ell=1}^{J} \left\{N\alpha_\ell^{(r)}\beta_\ell^{(r)} + C_\ell\left[1 - \Sigma\alpha_k^{(r)}\beta_k^{(r)}\right]\right\}$$

$$= \left\{N\alpha_j^{(r)}\beta_j^{(r)} + C_j\left[1 - \Sigma\alpha_k^{(r)}\beta_k^{(r)}\right]\right\} \Big/ N$$

$$\beta_i^{(r+1)} = \left\{N\alpha_i^{(r)}\beta_i^{(r)} + R_i\left[1 - \Sigma\alpha_k^{(r)}\beta_k^{(r)}\right]\right\} \Big/ N$$

Imbedding the problem in an EM algorithm context the appropriate model for a complete table is one which proposes full independence. The likelihood is then

$$f\left(\underline{n} \,\middle|\, \underline{\alpha}, \underline{\beta}, N + \sum_{k=1}^{J} n_{kk}\right) \propto \prod_{j=1}^{J} \prod_{i=1}^{J} (\alpha_j\beta_i)^{n_{ij}} .$$

Since the natural logarithm of this is linear in n_{ij}, the E-step replaces n_{kk} by its current expectation. This was shown to be given by equation 8.5.6, namely

$$n_{kk}^{(r)} = N\alpha_k^{(r)}\beta_k^{(r)} \Big/ \left[1 - \sum_{\ell=1}^{J} \alpha_\ell^{(r)}\beta_\ell^{(r)}\right]$$

The M-step uses the formula 8.5.2, 8.5.3, namely

$$\alpha_j^{(r+1)} = \left(n_{jj}^{(r)} + C_j\right) \Big/ \sum_{k=1}^{J}\left(n_{kk}^{(r)} + C_k\right)$$

$$\beta_i^{(r+1)} = \left(n_{ii}^{(r)} + R_i\right) \Big/ \sum_{k=1}^{J}\left(n_{kk}^{(r)} + R_k\right) ,$$

which trivially reduce to the above.

§10.3.4 We turn now to more general C satisfying the necessary conditions of problem (P3).

There does not seem any natural analogue to iteration 10.3.5 unless the constraints $C\theta = b$ can be transformed to constraints of the type considered in section 10.3.3.

The following however is an alternative approach which may reduce (P3) to an example of the special case (P6) considered in section 10.3.2.

Suppose without loss of generality that

$$b = \begin{pmatrix} \underline{1}_m \\ \underline{0}_{s-m} \end{pmatrix}$$

Denoting the i^{th} row of C by $\underline{d}_i$ the constraints $C\theta = b$ are then given by

$$\underline{d}_i'\theta = 1 \quad , \quad i = 1, \ldots\ldots, m$$
$$\underline{d}_i'\theta = 0 \quad , \quad i = (m+1), \quad , s .$$

These can be restated as

$$\underline{d}_1'\theta = 1 \quad , \quad \underline{d}_i'\theta = \underline{d}_1'\theta \quad , \quad i = 2, \cdots, m$$
$$\underline{d}_i'\theta = 0 \quad , \quad i = (m+1), \quad , s .$$

Finally these transform to

$$\underline{c}_1'\theta = 1$$
$$\underline{c}_i'\theta = 0 \quad , \quad i = 2, \cdots, m ,$$

where $\underline{c}_1 = \underline{d}_1$, $\underline{c}_i = \underline{d}_i - \underline{d}_1$ for $i=2,\ldots,m$, $\underline{c}_i = \underline{d}_i$ for $i = (m+1),\ldots s$.

The suggestion we now make is to delete the last $(s-1)$ constraints by substitution for, say, the last $(s-1)$ of the θ_j's, in terms of the remaining θ_i's. This would transform the problem to an example of the following problem.

(P7) "Maximise $\Psi(\theta)$ subject to (i) $\theta_j \geq 0$, $j=1,\ldots,T$; (ii) $c'\theta = 1$ where $c = (c_1,\ldots,c_r)'$; (iii) $L_i(\theta) \geq 0$, $i=1,\ldots,l$, where $L_i(\theta)$ is a linear combination of $\theta_1,\ldots,\theta_T$."

Clearly this is similar to problem (P6) of section 10.3.2 but differs from that in two respects; the constraints (iii) are an additional feature and there is no requirement that $c_j > 0$. The latter is not guaranteed when (P7) is the above reduction of (P3) (in which case $T = t-(s-1)$, $l = s-1$), nor is $\Psi(\cdot)$ guaranteed to have positive derivatives.

However suppose that in the original problem (P3) the parameters θ_j, $j = 1,\ldots,t$, are strictly positive at the optimum, as would be the case if they are probability parameters for a likelihood. Then with respect to (P7) the constraints (iii) will be slack at the optimum. (If it emerged that $L_i(\theta^*) = 0$, then we have a constraint of the form $c_i'\theta = 0$ which could be dealt with as above.) The optimum θ^* will almost certainly be the optimum subject to only constraints (i) and (ii), and, if $c_j > 0$, we effectively have an example of problem (P6) to solve.

The latter possibilities were realised in the following example which is the one to which reference was made in sections 10.2.3 and 10.3.2.

The example arises in a statistical context, the particular instance of problem (P3) being that of determining the maximum likelihood estimates of the cell probabilities of a square $k \times k$ contingency table, under a hypothesis of marginal homogeneity, assuming a k^2-cell multinomial model for the cell entries o_{ij}.

The likelihood therefore satisfies

$$L_o(\underline{q}) \propto \prod_{i=1}^{k}\prod_{j=1}^{k} q_{ij}^{o_{ij}} \quad , \quad \sum_{i=1}^{k}\sum_{j=1}^{k} q_{ij} = 1 \quad , \quad q_{ij} \geq 0 \; .$$

The hypothesis of marginal homogeneity states that

10.3.6 $$\sum_{j=1}^{k} q_{mj} = \sum_{i=1}^{k} q_{im} \quad , \quad m = 1, 2, \cdots, k \; ;$$

that is, there is equality between the m^{th} row sum probabilities and the m^{th} column sum probabilities.

The problem of course is to maximise $L_o(q)$ subject to these k constraints and to the constraint $\sum\sum q_{ij} = 1$, as well as $q_{ij} \geq 0$. Clearly this is an example of (P3) with $t = k^2$, $s = k+1$. However the resultant matrix C has rank equal to k. Given the constraint $\sum\sum q_{ij} = 1$, one of the constraints in 10.3.6 is unnecessary. The last say can be omitted and a further simplification derives from the fact that the parameters q_{mm} cancel out in 10.3.6. The hypothesis of marginal homogeneity makes no claim about these parameters. As a consequence their maximum likelihood estimates under the hypothesis and under the multinomial model are the same, namely o_{mm}/n , $n = \sum\sum o_{ij}$

We can therefore reduce the above problem to the following

"Maximise $$f(\underline{q}) = \prod_{i=1}^{k}\prod_{\substack{j=1 \\ i \neq j}}^{k} q_{ij}^{o_{ij}} \quad ,$$

subject to (i) $$q_{ij} > 0$$

(ii)(a) $$\sum_{i=1}^{k}\sum_{\substack{j=1 \\ i \neq j}}^{k} q_{ij} = \left[1 - \left(\sum_{m=1}^{k} o_{mm}\right)\Big/ n\right] = b > 0$$

(ii)(b) $$\sum_{\substack{j=1 \\ j \neq m}}^{k} q_{mj} - \sum_{\substack{i=1 \\ i \neq m}}^{k} q_{im} = 0 \quad , \quad m = 1, \cdots, (k-1) \; .$$ "

Again this is an example of (P3) with $t = k^2-k$, $s = k$ and now the matrix C has rank s.

Consider the case $k = 3$. Then the full constraint region is the set

$$S = \{(q_{12},q_{13},q_{21},q_{23},q_{31},q_{32}) \;:\; q_{ij} \geqslant 0 ,$$
$$q_{12}+q_{13}+q_{21}+q_{23}+q_{31}+q_{32} = b ,$$
$$q_{12}+q_{13}-q_{21}-q_{31} = 0 ,$$
$$q_{21}+q_{23}-q_{12}-q_{32} = 0 \}$$

This is a polygon with vertices

$$u_1 = (b/2,0,b/2,0,0,0)'$$

$$u_2 = (0,b/2,0,0,b/2,0)'$$

$$u_3 = (0,0,0,b/2,0,b/2)'$$

$$u_4 = (b/3,0,0,b/3,b/3,0)'$$

$$u_4 = (0,b/3,b/3,0,0,b/3)'$$

However employing the recommendation that we deal with the constraints (ii)(b) by substitution, relevant formulae can be

$$q_{31} = q_{12}+q_{13}-q_{21} , \quad q_{32} = q_{21}+q_{23}-q_{12} .$$

The function $f(q)$ then becomes

$$f(q_{12}, q_{13}, q_{21}, q_{23}) = q_{12}^{o_{12}} q_{13}^{o_{13}} q_{21}^{o_{21}} q_{23}^{o_{23}} (q_{12}+q_{13}-q_{21})^{o_{31}} (q_{21}+q_{23}-q_{12})^{o_{32}}$$

and also

$$\sum_{\substack{i=1\\ i\neq j}}^{3} \sum_{j=1}^{3} q_{ij} = q_{12} + 2q_{13} + q_{21} + 2q_{23} .$$

Replacing $(q_{12},q_{13},q_{21},q_{23})$ by $(\theta_1,\theta_2,\theta_3,\theta_4)$ we have the following final reduction of the original problem.

"Maximise $\quad \psi(\theta) = \theta_1^{o_{12}} \theta_2^{o_{13}} \theta_3^{o_{21}} \theta_4^{o_{23}} (\theta_1+\theta_2-\theta_3)^{o_{31}} (\theta_3+\theta_4-\theta_1)^{o_{32}}$.

subject to (i) $\theta_i \geqslant 0$, $i = 1,..,4$,

(ii) $c'\theta = 1$, $c = (1/b,2/b,1/b,2/b)'$,

(iii) $(\theta_1+\theta_2-\theta_3) > 0$, $(\theta_3+\theta_4-\theta_1) > 0$. "

Clearly this is an example of problem (P6) since $c_j > 0$, but with the extra constraints (iii). However the latter must be slack at the optimum since all optimum q^*_{ij} must be strictly positive. If an initial $\theta^{(0)}$ satisfies constraints (iii) then hopefully iterations under formula 10.3.3 would satisfy the same conditions.

That $\theta^{(0)}$ suggested by 10.3.4 is in this category, for here it suggests

$$\theta_1^{(0)} = \theta_3^{(0)} = b/4$$
$$\theta_2^{(0)} = \theta_4^{(0)} = b/8$$

Since $\theta_1^{(0)} = \theta_3^{(0)}$ we have

$$\theta_1^{(0)} + \theta_2^{(0)} - \theta_3^{(0)} = \theta_2^{(0)} > 0$$
$$\theta_3^{(0)} + \theta_4^{(0)} - \theta_1^{(0)} = \theta_4^{(0)} > 0$$

The following alternative $\theta^{(0)}$ also satisfies the equation $\theta_1^{(0)} = \theta_3^{(0)}$.

A special case of marginal homogeneity in a square contingency table is that of symmetry. The table conforms to a hypothesis of symmetry if $q_{ij} = q_{ji}$. Maximum likelihood estimates under such a hypothesis are given by $\tilde{q}_{ij} = (o_{ij} + o_{ji})/2n$.

Taking these as initial approximations $q_{ij}^{(0)}$ to the maximum likelihood estimates q^*_{ij} under marginal homogeneity suggests in this particular example

10.3.7
$$\theta_1^{(0)} = q_{12}^{(0)} = (o_{12} + o_{21})/2n$$
$$\theta_2^{(0)} = q_{13}^{(0)} = (o_{13} + o_{31})/2n$$
$$\theta_3^{(0)} = q_{21}^{(0)} = \theta_1^{(0)}$$
$$\theta_4^{(0)} = q_{23}^{(0)} = (o_{23} + o_{32})/2n$$

It is conceivable however that iterations under 10.3.2 would be undefined for derivatives are

$$\partial\psi/\partial\theta_1 = \psi(\theta)\left\{\frac{o_{12}}{\theta_1} + \frac{o_{31}}{\theta_1+\theta_2-\theta_3} - \frac{o_{32}}{\theta_3+\theta_4-\theta_1}\right\}$$

$$\partial\psi/\partial\theta_2 = \psi(\theta)\left\{\frac{o_{13}}{\theta_2} + \frac{o_{31}}{\theta_1+\theta_2-\theta_3}\right\}$$

$$\partial\psi/\partial\theta_3 = \psi(\theta)\left\{\frac{o_{21}}{\theta_3} - \frac{o_{31}}{\theta_1+\theta_2-\theta_3} + \frac{o_{32}}{\theta_3+\theta_4-\theta_1}\right\}$$

$$\partial\psi/\partial\theta_4 = \psi(\theta)\left\{\frac{o_{23}}{\theta_4} + \frac{o_{32}}{\theta_3+\theta_4-\theta_1}\right\}$$

These derivatives particularly the first and third could conceivably be negative; only these could be negative if constraints (i) and (iii) are satisfied. However $\psi(\theta)$ is a homogeneous function, a product of homogeneous functions. Also if constraints (i) and (iii) are satisfied it is a positive function. Hence, because $\theta_j^* > 0$, all derivatives are positive at the optimum since from theorem 2.5.6 they are given by

$$\partial\psi/\partial\theta_j^* = N\psi(\theta^*) ,$$

where $N = \sum_{\substack{i=1 \\ i \neq j}}^{3} \sum_{j=1}^{3} o_{ij}$ is the degree of homogeneity.

If then derivatives are positive at $\theta^{(0)}$, one would expect the same at subsequent $\theta^{(r)}$.

The latter proved to be the case for the following data set on the unaided distance vision of 7477 women aged 30 - 39, contained in Plackett (1974) (See pages 22 and 61).

Grade of Right Eye	Grade of Left Eye: Highest	Second	Other
Highest	1520	266	190
Second	234	1512	510
Other	153	444	2648

In fact Plackett had the data in the form of a 4×4 table, the "other" category being subdivided into a third and lowest grade.

The hypothesis of marginal homogeneity is a natural one to consider for these data, and Plackett obtained maximum likelihood estimates under it for the 4×4 table using a Newton Raphson technique applied to a lagrangian which preserved the constraints.

Note that the observed frequencies are large. As a consequence the likelihood $L_o(q)$ must be a relatively flat surface suggesting an ill conditioned hessian matrix. This may explain why some difficulty was experienced in trying to use a Newton Raphson technique which differed from that used by Plackett.

However no great difficulty was experienced in implementing iterations 10.3.3. Relevant details are as follows.

$$\psi(\theta) = \theta_1^{260} \cdot \theta_2^{190} \cdot \theta_3^{234} \cdot \theta_4^{510} \cdot (\theta_1 + \theta_2 - \theta_3)^{153} (\theta_3 + \theta_4 - \theta_1)^{444}$$

$$b = 1-(5680)/(7477) = 1797/7477$$

Initial approximation 10.3.7 was used.

No negative derivatives were incurred, the constraints (iii) were never violated and monotonicity was obtained for $\delta = 0.1, 0.2, 0.3, 0.4$.

The iterations converged to the point

$$(\theta_1^*, \theta_2^*, \theta_3^*, \theta_4^*) = (.03371, .02268, .03313, .06406),$$

which yields the following expected frequencies for the table.

1520	252.02	169.62
247.74	1512	479.00
173.90	474.72	2648

These would seem to be the desired optimum values. They are comparable to the expected frequencies Plackett obtained for the corresponding 4×4 table.

Again convergence was slow. Of interest is that in the case of the 3×3 table the case $\delta = 0.5$ just simply produces a decreasing sequence of $\psi(\theta^{(r)})$ values. It is not clear why this happens. It imposes a relatively small upper limit on the values of δ yielding monotonicity, considering that $\psi(\theta)$ is a product of homogeneous functions. Possibly larger values of δ would have achieved monotonicity had initial approximation 10.3.4 been employed.

§10.3.5 We now turn to a brief discussion of the conditions under which problem (P3) can be reduced to an example of (P1) or (P2). We also comment on some other points concerning (P3).

Of particular interest is to establish conditions under which the convex constraint region

$$S = \{\theta : \theta_j \geq 0 , \; j = 1, ----, t \; ; \; C\theta = b\} ,$$

assumed nonempty, is a polygon or simplex with a finite set of finite vertices.

We wish to ascertain when there belongs to S, at least two distinct intersections among the intersections of the plane $C\theta = b$ with (t-s) of the regions $\theta_j = 0$.

This will certainly not be the case if b = 0 for then the origin is the only such intersection.

It will also not be the case if C has a column of zeros, say the last. Then $C\theta = b$ imposes no constraints on θ_t and so S is infinite.

Suppose then that the latter is not the case and assume without loss of generality that

$$b = \begin{pmatrix} \underline{1}_m \\ \underline{0}_{s-m} \end{pmatrix}$$

Then a sufficient condition for vertices of the above type to exist would seem to be that at least one of the first m rows of C should contain only positive values. Suppose the first row, $\underline{c}_1'$, satisfies this. Then $\underline{c}_1'\theta = 1$ is a constraint of the type considered in problem (P6). Its intersection with the positive quadrant is the "triangle" with vertices $(1/c_j)\underline{e}_j$, where $\underline{e}_j$ is the j^{th} unit vector and $\underline{c}_1' = (c_1, \ldots, c_t)$. Since S is the intersection of the remaining linear constraints in $C\theta = b$ with this finite traingle, then S will have the desired structure, assuming S contains more than one element. It follows that this will also be the case when the matrix C does not have a positive row among the first m rows, but the constraints $C\theta=b$ can be transformed, by taking linear combinations, to constraints $D\theta=d$ which do satisfy the condition. The special case considered in section 10.3.3 is an example of this.

It may be that the latter is a necessary condition if S is to be a finite simplex, but we will not pursue this. There may be some general result which is of relevance. Possibly Davis (1952, 1953, 1954) is informative in this respect.

We consider some other points.

(i) Suppose b = 0 and we choose to deal with the constraints $C\theta=0$ by substitution as in section 10.3.4. We thereby transform the problem to one of maximising a function $\psi(\theta)$ subject to $\theta_j \geq 0$ plus some additional inequality constraints. If though it can be established that all constraints are slack at the optimum then we essentially have an unconstrained maximisation problem. However if the function $\psi(\theta)$ is homogeneous one would wish to employ the constraint $\underline{1}'\theta = 1$.

(ii) Suppose C does have at least one column consisting of zeros. Then problem (P3) is of the form.

(P8) "Maximise $\psi(\underline{\theta}_1, \underline{\theta}_2)$ subject to

(i) $\underline{\theta}_1 \geq \underline{0}$, $\underline{\theta}_2 \geq \underline{0}$

(ii) $D\underline{\theta}_1 = b$

where D has no zero columns.

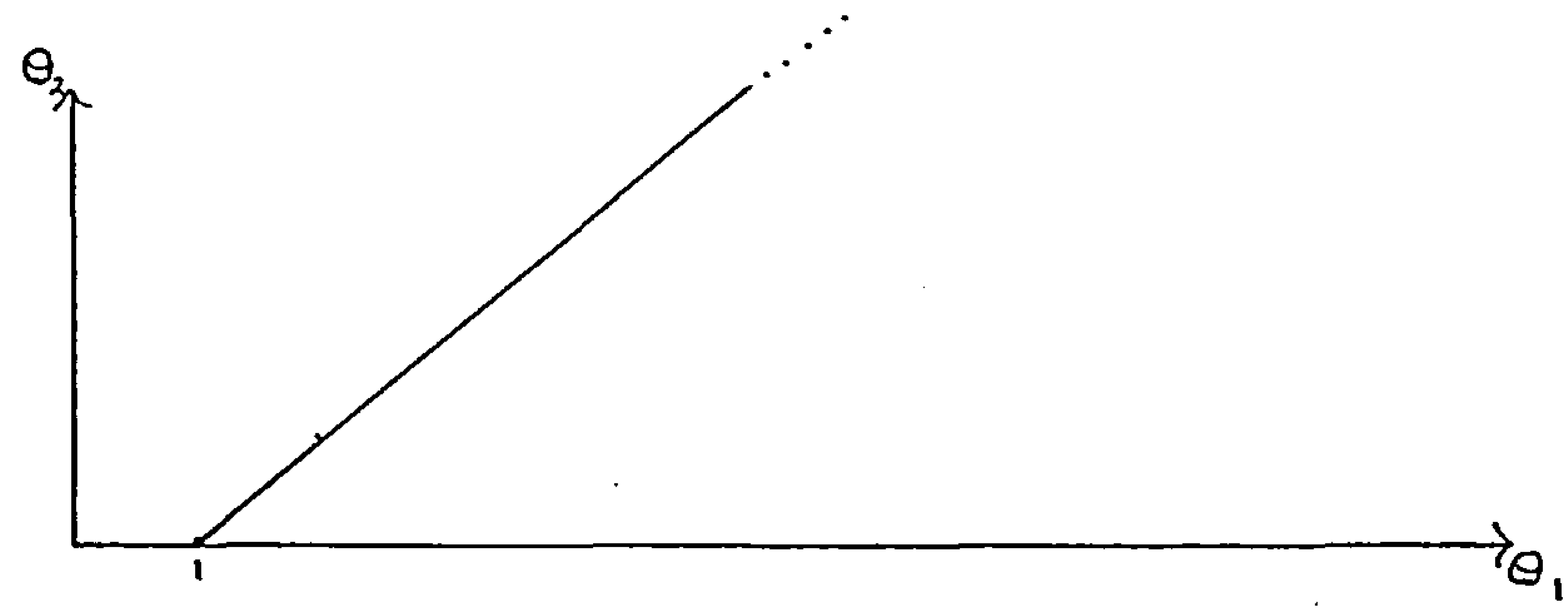
If the set $\{\underline{\theta}_1 : \underline{\theta}_1 \geq \underline{0}, D\theta_1 = b\}$ is a finite polygon , then numerical solution of (P8) could be achieved by combining algorithms of the type we have been considering applied to $\underline{\theta}_1$, with unconstrained iterations for $\underline{\theta}_2$, cycling between iterations for $\underline{\theta}_1$ and iterations for $\underline{\theta}_2$ perhaps.

(iii) Suppose the constraint region for problem (P3) is not a polygon. Then it must be infinite and in principle the numerical techniques which have been discussed in earlier chapters are not available to us. However S is still convex and it will be the case that we could view S as the limit of a finite polygon. We can therefore use theorem 2.5.6 to establish first order conditions for a differentiable finite optimum as the following example illustrates.

Consider the problem

"Maximise $\psi(\theta_1, \theta_2)$ subject to $\theta_1, \theta_2 \geq 0$ and $\theta_2 - \theta_1 = 1$"

The constraint region is those points in the positive quadrant which lie on the line $\theta_2 = \theta_1 + 1$



It is therefore given by $\underset{M\to\infty}{\mathrm{Lt}} \mathcal{P}(u_1, u_2(M))$,

where $\mathcal{P}(u_1, u_2(M))$ is the convex hull of the two points $u_1 = (1, 0)'$, $u_2 = (M, M-1)'$.

Any $\theta = (\theta_1, \theta_2)' \in \mathcal{P}(u_1, u_2(M))$ satisfies

$$\theta = p_1(M) u_1 + p_2(M) u_2(M)$$

where

$$p_1(M) = (M - \theta_1)/(M-1), \quad p_2(M) = (\theta_2 - 1)/(M-1) = \theta_1/(M-1).$$

Hence $p_1(M) \to 1$, $p_2(M) \to 0$ as $M \to \infty$.

For M large enough we will have $\theta^* \in \mathcal{P}(u_1, u_2(M))$, where θ^* is a finite maximum. If θ^* is differentiable then by theorem 2.5.6

$$F_\psi(\theta^*, u_1) = 0, \quad F_\psi(\theta^*, u_2(M)) = 0.$$

Equality will obtain in both cases since θ^* will not lie on the boundary of $\mathcal{P}(u_1, u_2(M))$ for large enough M.

Using the general results

$$F_\psi(\theta^*, u_j) = G_\psi(\theta^*, u_j) - G_\psi(\theta^*, \theta^*),$$
$$G_\psi(\theta^*, u_j) = u_j' \partial\psi/\partial\theta^*,$$
$$G_\psi(\theta^*, \theta^*) = \sum p_i^* G_\psi(\theta^*, u_i) = \sum p_i^* u_i' \partial\psi/\partial\theta^*,$$

we obtain

$$F_\psi\left(\theta^*, u_j\right) = M f_j(M) \quad ,$$

where

$$f_1(M) = \frac{[1 - p_1(M)]}{M}\frac{\partial\psi}{\partial\theta_1^*} - p_2(M)\left[\frac{\partial\psi}{\partial\theta_1^*} + \left(1 - \frac{1}{M}\right)\frac{\partial\psi}{\partial\theta_2^*}\right]$$

$$f_2(M) = [1 - p_2(M)]\left(\frac{\partial\psi}{\partial\theta_1^*} + \frac{\partial\psi}{\partial\theta_2^*}\right) - \frac{[1 - p_2(M)]}{M}\frac{\partial\psi}{\partial\theta_2^*} - \frac{p_1(M)}{M}\frac{\partial\psi}{\partial\theta_1^*} .$$

Hence for M large enough , θ^* must satisfy $f_j(M) = 0,\ j = 1,2$. Now let $M \to \infty$ and

$$f_1(M) \longrightarrow 0 \quad , \quad f_2(M) \longrightarrow \left(\partial\psi/\partial\theta_1^* + \partial\psi/\partial\theta_2^*\right) .$$

In fact $f_1(M) \to 0$ as $M \to \infty$ for any finite θ. This is a consequence of $F(\theta, u_2(M))$ depending on the distance between θ and $u_2(M)$.

The implication though from the second limit is clear. The first order conditions on θ^* is that

10.3.8 $$\partial\psi/\partial\theta_1^* + \partial\psi/\partial\theta_2^* = 0$$

This is also obtained by what is the more natural approach in this simple example, namely to substitute for say θ_2 in terms of θ_1. First order conditions are then given by

10.3.9 $$f'(\theta) = 0$$

where

$$f(\theta) = \psi\{\theta_1(\theta), \theta_2(\theta)\}$$

$$\theta_1(\theta) = \theta \quad , \quad \theta_2(\theta) = 1 - \theta$$

Since it follows that $f'(\theta) = \partial\psi/\partial\theta_1 + \partial\psi/\partial\theta_2$ we have that 10.3.8 and 10.3.9 are equivalent.

In more complex examples establishing first order conditions by using theorem 2.5.6 as above may be as simple an approach as any. A further consequence of the argument, that the solution θ^* to an example of (P3) with an unbounded constraint region, belongs to a finite polygon $\mathcal{P}(M)$ for large M, is the following. The maximum of $\psi(\cdot)$ over $\mathcal{P}(M)$ also occurs at θ^* and this is an example of (P2).

We could consider then using any algorithm which has been formulated for (P2) if we were convinced that Θ^* was an interior point of $\mathcal{D}(M)$.

§10.4 A Fusion Of Algorithms

In this section we extend some thoughts of Titterington (1977). He proposed a fusion of algorithms $FP\{d^{\delta}, 1\}$ and $S\{A, \alpha_r\}$ for problem (P1).

Recall that the constrained steepest ascent direction can be given by

10.4.1 $$\underset{\sim}{m}_1 = \{A^+ - (A^+\underline{1}\,\underline{1}'A^+)/(\underline{1}'A^+\underline{1})\}d \,.$$

Titterington suggests consideration of

10.4.2 $$\underset{\sim}{m}_2 = \{A^+ - (A^+\underline{1}\,\underline{1}'A^+)/(\underline{1}'A^+\underline{1})\}d^{\delta}$$

and he also considers

10.4.3 $$\underset{\sim}{m}_3 = \{A^+ - A^+C'(CA^+C')^+CA^+\}d^{\delta}$$

with problem (P3) in mind, although he assumes (CA^+C') to be non-singular.

Extending this approach one might consider

10.4.4 $$\underset{\sim}{m}_4 = \{A^+ - A^+C'(CA^+C')^+CA^+\}h(d, \delta) \,.$$

Such a direction of iteration is a fusion of algorithms $FP\{h(d,\delta), \alpha_r\}$ and $S\{A, \alpha_r\}$ in the case of $C = \underline{1}'$. It includes a wide class of algorithms. Denote by $Alg\{\underline{m}_i, \alpha\}$ the algorithm which takes a step α in the direction $\underline{m}_i$.

Titterington notes that $FP\{d^{\delta}, 1\}$ and $Alg\{\underset{\sim}{m}_2, \alpha\}$ are the same when $A^{-1} = diag\{p_1, \ldots, p_J\}$ and $\alpha = \left(\sum p_j d_j^{\delta}\right)^{-1}$. The same is true if $A^{-1} = \left(\sum p_j d_j^{\delta}\right) diag\{p_1, \ldots, p_J\}$ and $\alpha = 1$.
We have already seen it to be advantageous to allow A to depend on p. This illustrates that we might also allow it to depend on δ.

Titterington also notes that $Alg\{\underset{\sim}{m}_2, \alpha\}$ can for an appropriate choice of A, become a vertex direction algorithm if $\delta \to \infty$. We have already seen this to be the case with $FP\{d^{\delta}, 1\}$ and more generally with $FP\{\exp[\delta g(z)], 1\}$, $z = d$ or F.

It is not the intention to claim that any of the above directions is in any sense optimal. In the main though one would expect them to be reasonable directions in which to move. This should certainly be the case for δ near 1 in the case of $\underline{m}_2$, $\underline{m}_3$. However it is not in general guaranteed that $F\{p, p+\underline{m}_i\} = \underline{m}_i' d > 0$. For instance it is only guaranteed that $\underline{m}_2' d > 0$ for all $\delta > 0$ if A is nonnegative definite diagonal.

If it does emerge that $\underline{m}_i' d < 0$, there is, however, no great problem. We can take a step in the direction of $-\underline{m}_i$, in which case we are employing algorithm Alg$\{\pm\underline{m}_i, \alpha\}$.

Titterington in fact suggests, as a further generalisation of $\underline{m}_2$, taking A nonpositive definite and $\delta < 0$, in which case $\underline{m}_2' d$ is more likely to be negative.

§10.5 A Generalisation Of (P2)

We end as we began; namely by citing a generalisation of problem (P2). This illustrates a wider class of problems requiring solution of (P1). An example is given. The use of theorem 2.5.6 is seen again.

Consider first the design problem (P4). This was referred to as a particular explicit example of problem (P2). However regarding the design points $v_1,..,v_J$ as a basic set of vertices, we could view (P3) as seeking to solve (P2) not for the vertices $v_1,..,v_J$ but for a one to one transformation of these. A slightly more general formulation of (P2) is suggested by this, namely

(P9) "Solve (P1) for $\phi(p) = \psi\{x(p)\}$ where for a given set of vertices $\mathcal{V} = \{v_1,...,v_J\}$ and for a one-to-one function $G(\cdot)$, $x(p) = E\{G(v)\} = \sum_{j=1}^{J} p_j G(v_j)$."

The design problem, of course, is an example of this which takes $G(v) = vv'$.

This formulation suggests the following generalisation.

(P10) "Solve (P1) for $\phi(p) = \psi\{x_1(p), x_2(p),...,x_n(p)\}$ where for a given set of vertices $\mathcal{V} = \{v_1,..,v_J\}$ and for one-to-one functions $G_i(\cdot)$, $i = 1,...,n$, $x_i(p) = E\{G_i(v)\} = \sum_{j=1}^{J} p_j G_i(v_j)$."

This however could be viewed as a particular example of (P2); namely that which takes $u_j = \{G_1(v_j), G_2(v_j),...,G_n(v_j)\}$. This points out how to use theorem 2.5.6 to derive first order conditions for an optimising p^* for (P10). This is illustrated in the following example, considered by Silverman and Titterington (1980). They seek to solve (P1) for

$$\phi(p) = -\log_e \det\{M_o(p)\},$$

where

$$M_o(p) = \sum_{j=1}^{J} p_j(v_j - \bar{v})(v_j - \bar{v})', \quad \bar{v} = \sum p_j v_j .$$

Since

$$M_o(p) = \sum_{j=1}^{J} p_j v_j v_j' - \bar{v}\bar{v}' = M(p) - \bar{v}\bar{v}',$$

they seek to solve (P10) for

$$n = 2$$
$$\psi(X_1,x_2) = -\log_e \det\left[N\{(X_1,x_2)\}\right]$$
$$N\{(X_1,x_2)\} = X_1 - x_2x_2'$$

X_1 is a $k\times k$ matrix

x_2 is a $k\times 1$ vector

$$G_1(v) = vv'$$
$$G_2(v) = v$$
$$u_j = (v_jv_j', v_j) \quad .$$

With a view to deriving first order conditions for the optimising p^* we evaluate the directional derivative of the matrix $N\{(X_1,x_2)\}$ at X_1,x_2 in the direction of Y_1,y_2.

$$N\{(1-\epsilon)(X_1,x_2)+\epsilon(Y_1,y_2)\} = N\{[(1-\epsilon)X_1+\epsilon Y_1, (1-\epsilon)x_2+\epsilon y_2]\}$$
$$= (1-\epsilon)X_1+\epsilon Y_1 - [(1-\epsilon)x_2+\epsilon y_2][(1-\epsilon)x_2+\epsilon y_2]'$$
$$= X_1+\epsilon(Y_1-X_1) - [x_2+\epsilon(y_2-x_2)][x_2+\epsilon(y_2-x_2)]'$$
$$= X_1+\epsilon(Y_1-X_1) - \{x_2x_2'+\epsilon(y_2-x_2)x_2' + \epsilon x_2(y_2-x_2)' + \epsilon^2(y_2-x_2)(y_2-x_2)'\}$$

Hence

$$\left[N\{(1-\epsilon)(X_1,x_2)+\epsilon(Y_1,y_2)\} - N(X_1,x_2)\right]/\epsilon$$
$$= (Y_1-X_1) - (y_2-x_2)x_2' - x_2(y_2-x_2)' + \epsilon(y_2-x_2)(y_2-x_2)'$$
$$\longrightarrow (Y_1-X_1) - (y_2-x_2)x_2' - x_2(y_2-x_2)' \quad \text{as } \epsilon\longrightarrow 0\,.$$

Thus the function $N\{(X_1,x_2)\}$ has directional derivative

$$F_N\{(X_1,x_2),(Y_1,y_2)\} = Y_1 - X_1 - y_2x_2' - x_2y_2' + 2x_2x_2'$$
$$= (Y_1-y_2y_2') - (X_1-x_2x_2') + y_2y_2' - y_2x_2' - x_2y_2' + x_2x_2'$$
$$= N\{(Y_1,y_2)\} - N\{(X_1,x_2)\} + (y_2-x_2)(y_2-x_2)'\,.$$

Now using the result that $\psi(X_1,x_2)$ has directional derivative

$$F_\psi\{(X_1,x_2),(Y_1,y_2)\} = -\mathrm{tr}\{N^{-1}(X_1,x_2)F_N\{(X_1,x_2),(Y_1,y_2)\}\}$$

we obtain

$$F_\psi\{(X_1,x_2),(Y_1,y_2)\} = k - \mathrm{tr}\{N^{-1}(X_1,x_2)N(Y_1,y_2)\} - \mathrm{tr}\{[N^{-1}(X_1,x_2)](y_2-x_2)(y_2-x_2)'\}$$
$$= k - \mathrm{tr}\{N^{-1}(X_1,x_2)N(Y_1,y_2)\} - (y_2-x_2)'[N^{-1}(X_1,x_2)](y_2-x_2)\,.$$

Now take $(Y_1, y_2) = u_j$. Then

$$N(Y_1, y_2) = v_j v_j' - v_j v_j' = 0$$

Thus

$$F_\psi\{(X_1, x_2), u_j\} = k - (v_j - x_2)' N^{-1}(X_1, x_2)(v_j - x_2) .$$

$$F_\psi\{(M(p), \bar{v}), u_j\} = k - (v_j - \bar{v})'[M_0^{-1}(p)](v_j - \bar{v}) .$$

Appealing to theorem 2.5.6 we conclude that first order conditions for a maximum are

10.5.1 $$(v_j - \bar{v})'[M_0^{-1}(p)](v_j - \bar{v}) \begin{cases} = k & \text{if } p_j^* > 0 \\ \leq k & \text{"} \quad p_j^* = 0 \end{cases} .$$

Silverman and Titterington let this result be suggested by a duality theorem which establishes that the solution to the problem of finding a vector c and matrix N to maximise det(N) subject to

$$(v_j - c)' N (v_j - c) \leq k \quad , \quad j = 1, \ldots, J ,$$

is given by $C = \bar{v}$, $N = M_0^{-1}(p)$. The solution defines the ellipsoid of smallest content which contains $\mathcal{V} = \{v_1, \ldots, v_J\}$.

These first order conditions can also be obtained from theorem 2.5.6 by an alternative route. Titterington (1975) has shown that the solution p^* to problem (P1) is the same in the following two instances

$$\phi(p) = \phi_1(p) = -\log_e \det\{M_0(p)\}$$

and

$$\phi(p) = \phi_2(p) = -\log_e \det\{\breve{M}(p)\}$$

where

$$\breve{M}(p) = \sum p_j w_j w_j' \quad , \quad w_j' = (v_j', 1) .$$

The second problem is a straightforward D-optimal design problem in (k+1) dimensions. We already know from theorem 2.5.6 that first order conditions on p^* are

10.5.2 $$w_j' \breve{M}^{-1}(p^*) w_j \begin{cases} = (k+1) & \text{if } p_j^* > 0 \\ \leq (k+1) & \text{"} \quad p_j^* = 0 \end{cases} .$$

Conditions 10.5.1 and 10.5.2 are equivalent since

$$w_j' \breve{M}^{-1}(p) w_j = (v_j - \bar{v})'[M_0^{-1}(p)](v_j - \bar{v}) + 1 .$$

Titterington (1975) also shows the same p^* to be D_s-optimal for $s = k$ with respect to the design space $\mathcal{W} = \{w_1, \ldots, w_J\}$, i.e. p^* solves (P1) for $\phi(p) = \phi_3(p) = -\log_e \det\{A\tilde{M}^+(p)A'\}$, $A = [I_k : \underset{\sim}{0}]$. This is so because, under the transformation $w_j' = (v_j', 1)$, the above ellipsoid becomes a central cylinder, i.e. a cylinder with axis passing through the origin. Silvey and Titterington (1973) established that a D_s-optimal problem and the problem of choosing a nonnegative definite symmetric $s \times s$ matrix A and a matrix B, $s \times (k-s)$, to maximise $\log_e \det(A)$ subject to

$$(\underset{\sim}{v}_1 + B\underset{\sim}{v}_2)' A (\underset{\sim}{v}_1 + B\underset{\sim}{v}_2) \le s \quad , \quad v = (\underset{\sim}{v}_1', \underset{\sim}{v}_2')' \in \mathcal{V}$$

are duals with related solutions. The second problem is referred to as a thinnest central cylinder problem. The solution defines that smallest central cylinder, ellipsoidal in the first s components, which contains the design space $\mathcal{V}$.

These discussions illustrate that an optimisation problem can possibly be expressed as an example of (P10) in more than one way.

A generalisation of the original problem is to solve (P10) for

$$\psi(X_1, x_2) = -\mathrm{tr}\{A N^+(X_1, x_2) A'\}^t$$
$$N(X_1, x_2) = X_1 - x_2 x_2'$$

Certainly for $N(X_1, x_2)$ nonsingular it would appear that

$$F_\psi\{(X_1, x_2), (Y_1, y_2)\} =$$
$$= t\Big[\mathrm{tr}[AN^{-1}(X_1, x_2)A']^t - \mathrm{tr}\{[AN^{-1}(X_1, x_2)A']^{t-1} A[N^{-1}(X_1, x_2)N(Y_1, y_2)N^{-1}(X_1, x_2)]A'\}$$
$$- (y_2 - x_2)'[N^{-1}(X_1, x_2)]A'[AN^{-1}(X_1, x_2)A']^{t-1} A[N^{-1}(X_1, x_2)](y_2 - x_2).$$

Taking $(Y_1, y_2) = (v_j v_j', v_j) = u_j$, $(X_1, x_2) = (M(p), \bar{v})$,

$$(1/t) F_\psi\{(M(p), \bar{v}), u_j\} = \mathrm{tr}[AM_0^{-1}(p)A']^t - (v_j - \bar{v})' M_0^{-1}(p) A' [AM_0^{-1}(p)A']^{t-1} A M_0^{-1}(p)(v_j - \bar{v}).$$

Clearly any of the algorithms which have been considered for (P2) can be used to solve (P10). If there is more than one (P10) variant of a problem, then this may suggest more than one variant of a particular type of algorithm.

Silverman and Titterington employ the "Eliminating/augmenting" algorithm of Silvey and Titterington (1973), which we considered in section 6.5. This will be the same for the above two variants of (P10). It proved to be highly efficient in the examples they considered, for which J was large but k = 2.

Titterington (1976, 1978) also considered iterations under FP(d,1) for the two problems. These are $(p^{(r)}, p^{(r+1)}) = (\lambda, \tau)$ where, respectively for the problems,

10.5.3 $$\tau_j = \lambda_j [v_j - \bar{v}(\lambda)]' M_o^{-1}(\lambda) [v_j - \bar{v}(\lambda)] / k \quad , \quad \bar{v}(\lambda) = \sum \lambda_j v_j$$

and

10.5.4 $$\tau_j = \lambda_j w_j' \tilde{M}^{-1}(\lambda) w_j / (k+1) .$$

Note that the matrix $M_o(p)$ is nonnegative definite since

$$\begin{aligned} x' M_o(p) x &= \sum p_j (x' v_j)^2 - (x' \bar{v}(p))^2 \\ &= \sum p_j (x' v_j)^2 - \left(\sum p_j x' v_j \right)^2 = E(x'v)^2 - [E(x'v)]^2 \geq 0 . \end{aligned}$$

Hence under 10.5.3 $\tau_j > 0$ if $\lambda_j > 0$.

Note also that since $M_o(p) = M(p) - \bar{v}\bar{v}'$, lemma 5.2.1 could be used to express $M_o^{-1}(p)$ in terms of $M^{-1}(p)$. We have already considered extensively possibilities for a simple updating of $M^{-1}(p)$ as p changes from $p^{(r)}$ to $p^{(r+1)}$. These will have analogues in the case of $M_o(p)$.

Titterington reports that iterations under 10.5.3 were faster to converge than under 10.5.4. Monotonicity we know to obtain under 10.5.4. Titterington reports strong empirical evidence for it under 10.5.3 also.

Of interest finally is that with respect to $\phi(p) = \phi_1(p) = -\log_e \det\{\{M_o(p)\}\}$, formula 10.5.4 defines iterates under $FP\{h_1(d,1),1\}$, where $h_1(d, \delta) = (d+1)^{\delta}$, while, with respect to $\phi(p) = \phi_2(p) = -\log_e \det\{\{\widetilde{M}(p)\}\}$, formula 10.5.3 defines iterates under $FP\{h_2(d,1),1\}$, where $h_2(d, \delta) = (d-1)^{\delta}$.

REFERENCES

Andrews, D.F. et al (1972). Robust Estimates of Location. Princeton, N.J. Princeton University Press.

Ash, A. and Hedayat A. (1978). An introduction to design optimality with an overview of the literature. Commun. Stats. A, Vol. 7, 1295-1325.

Atkinson, A.C. and Cox, D.R. (1974). Planning experiments for discriminating between models. J.R.S.S.B., Vol. 36, 321-348.

Atwood, C.L. (1969). Optimal and efficient designs of experiments. Ann. Math. Stats., Vol. 40, 1570-1602.

Atwood, C.L. (1973). Sequences converging to D-optimal designs of experiments. Ann. Stats., Vol. 1, 342-352.

Atwood, C.L. (1976a). Convergent design sequences for sufficiently regular optimality criteria. Ann. Stats., Vol. 4, 1124-1138.

Atwood, C.L. (1976b). Computational considerations for convergence to an optimal design. Proc. 1976 Conf. on Information Sciences and Systems, Dept. of Elect. Eng., John Hopkins University, Baltimore, MD, 345-349.

Atwood, C.L. (1980). Convergent design sequences for sufficiently regular optimality criteria, II: singular case. Ann. Stats., Vol. 8, 894-912.

Baum, L.E. (1977). Contribution to discussion of a paper by Dempster, Laird and Rubin. J.R.S.S.B., Vol. 39, 28-29.

Baum. L.E. et al (1970). A maximisation technique occurring in the statistical analysis of probabilistic functions of Markov chains. Ann. Math. Stats., Vol. 41, 164-171.

Baum, L.E. and Eagon, J.A. (1967). An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. Bull. Amer. Math. Soc., Vol. 73, 360-363.

Beckenbach, E.F. and Bellman, R. (1961). Inequalities. Berlin: Springer Verlag.

Blumen, L., Kogan, M. and McCarthy, P.J. (1955). The Industrial Mobility of Labour as a Probability Process, Vol. 6 of Cornell Studies of Industrial and Labour Relations, New York: Cornell University.

Box, G.E.P. and Wilson, K.B. (1951). On the experimental attainment of optimum conditions. J.R.S.S.B., Vol, 13, 1-45.

Bradley, R.A. (1965). Another interpretation of a model for paired comparisons. Psychometrika, Vol. 30, 315-318.

Bradley, R.A. and Terry, M.E. (1952). The rank analysis of incomplete block designs I. The method of paired comparisons. Biometrika, Vol. 39, 324-345.

Brooks, R.J. (1972). A decision theory approach to optimal regression designs. Biometrika, Vol. 59, 563-571.

Brooks, R.J. (1974). On the choice of an experiment for prediction in linear regression. Biometrika, Vol. 61, 303-311.

Brooks, R.J. (1977). Optimal regression design for control in linear regression. Biometrika, Vol. 64, 319-325.

Chernoff, H. (1953). Locally optimal designs for estimating parameters. Ann. Math. Stats., Vol. 30, 755-770.

Clark, V. (1965). Choice of levels in polynomial regression with one or two variables. Technometrics, Vol. 7, 327-333.

Davidson, R.R. (1969). On a relationship between two representations of a model for paired comparisons. Biometrics, Vol. 25, 597-599.

Davies, M. (1974). Derivation of a directional derivative. Private communication.

Davis, C. (1952). The intersection of a linear subspace with the positive orthant. Mich. Math. J., Vol. 1, 163-168.

Davis, C. (1953). Remarks on a previous paper. Mich. Math. J., Vol. 2, 23-25

Davis, C. (1954). Theory of positive linear dependence. Amer. J. of Maths., Vol. 76, 732-745.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). J.R.S.S.B., Vol. 39, 1-38.

Ehrenfeld, S. (1955). On the efficiency of experimental designs. Ann. Math. Stats., Vol. 26, 247-255.

Elfving, G. (1952). Optimum allocation in linear regression theory. Ann. Math. Stats., Vol. 23, 255-262.

Elfving, G. (1955). Geometric allocation theory. Skand. Aktuarietidskr. Vol. 37, 170-190.

Elfving, G. (1959). Design of linear experiments. Cramer Festschrift Volume. New York: Wiley, 58-74.

Eplett, W.J.R. (1980). An influence curve for two sample rank tests. J.R.S.S.B., Vol. 42, 64-70.

Fan, K. (1959). On a theorem of Weyl concerning eigenvalues of linear transformations. Proc. Nat. Acad. Sci. USA, Vol. 35, 652-655.

Federer, W.T. and Ballam, L.N. (1972). Bibliography on Experiment and Treatment Design Pre 1968. Edinburgh: Oliver and Boyd.

Fedorov, V.V., (1969). Sequential methods for design of experiments in the study of the mechanism of a phenomenon. New Ideas in Experimental Design (V.V. Nalimov, Ed.). Moscow: Mauka.

Fedorov, V.V. (1972). Theory of Optimal Experiments. New York: Academic Press.

Fedorov, V.V. (1975). Optimal experimental designs for discriminating two rival regression models. A Survey of Statistical Design and Linear Models (J.N. Srivastava, Ed.). Amsterdam: North-Holland Publishing Co., 155-164.

Fedorov, V.V. (1978). Some extremal problems in designing discriminating experiments. Commun. Stats. A, Vol. 7, 1339-1345.

Fedorov, V.V. and Atkinson, A.C. (1975). The design of experiments for discriminating between two rival models. Biometrika, Vol. 62, 57-71.

Fedorov, V.V. and Malyutov, M.B. (1972). Optimal designs in regression experiments. M.O.S., Vol. 14, 237-324.

Fedorov, V.V. and Tukey, P. (1976). Construction of regularized and composite designs. Ann. Stats., To appear.

Fellman, J. (1974). On the allocation of linear observations. Comment. Phys. Math., Vol. 44, 27-78.

Ford, I. (1976). Ph.D. Thesis. University of Glasgow.

Frank M. and Wolfe P. (1956). An algorithm for quadratic programming. Naval Research Logistics Quarterly. Vol. 3, 95-110.

de la Garza, A. (1954). Spacing of information in polynomial regression. Ann. Math. Stats., Vol. 25, 123-130.

de la Garza, A. (1956). Quadratic extrapolation and a related test of hypothesis. J.A.S.A., Vol. 51, 644-649.

de la Garza, A. et al. (1955). Some minimum cost experimental procedures in quadratic regression. J.A.S.A., Vol. 50, 178-184.

Graybill, F.A. (1969). Introduction to Matrices with Applications in Statistics. Belmont, California: Wadsworth Pub. Co. Inc.

Gribek, P.R. and Kortanek, K.O. (1977). Equivalence theorems and cutting plane algorithms for a class of experimental design problems. SIAM J. Appl. Math., Vol. 32, 232-259.

Guest, P.G. (1958). The spacing of observations in polynomial regression. Ann. Math. Stats., Vol. 29, 294-299.

Hampel, F.R. (1968). Contributions to the theory of Robust Estimation. Ph.D. Thesis, University of California, Berkeley.

Hampel, F.R. (1971). A general qualitative definition of robustness. Ann Math. Stats., Vol. 42, 1887-1896.

Hill, P. (1976). Ph.D. Thesis, University of Glasgow.

Hoel, P.G. (1961a). Asymptotic efficiency in polynomial estimation. Ann. Math. Stats., Vol. 32, 1042-1047.

Hoel, P.G. (1961b). Some properties of optimal spacing in polynomial estimation. Ann. Inst. Stat. Math., Vol. 13, 1-8.

Hoel, P.G. (1965a). Minimax designs in two dimensional regression. Ann. Math. Stats., Vol. 36, 1097-1106.

Hoel, P.G. (1965b). Optimum designs for polynomial extrapolation. Ann. Math. Stats., Vol. 36, 1483-1493.

Hoel, P.G. and Levine, A. (1964). Optimal spacing and weighing in polynomial regression. Ann. Math. Stats., Vol. 35, 1553-1560.

Hotelling, H. (1944). Some improvements in weighing and other experimental techniques. Ann. Math. Stats., Vol. 15, 297-306.

Karlin, S. and Studden, W.J. (1966). Optimum experimental designs. Ann. Math. Stats., Vol. 37, 783-815.

Kelley, J.E. Jr. (1960). The cutting plane method for solving convex programs. J. Soc. Indust. Appl. Math. Vol. 8, 703-712.

Kiefer, J. (1958). On the randomized optimality and randomized non-optimality of symmetrical designs. Ann. Math. Stats., Vol. 29, 675-699.

Kiefer, J. (1959). Optimum experimental designs (with discussion). J.R.S.S.B., Vol. 21, 272-319.

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). Ann. Stats., Vol. 2, 849-879.

Kiefer, J. (1975a). Construction and optimality of generalised Youden designs. A Survey of Statistical Design and Linear Models (J.N. Srivastava, Ed.). Amsterdam: North-Holland Publishing Co., 333-353.

Kiefer, J. (1975b). Optimal design: variation in structure and performance under change of criterion. Biometrika, Vol. 62, 277-288.

Kiefer, J. (1978). Asymptotic approach to families of design problems. Commun. Stats. A, Vol. 7, 1347-1362.

Kiefer, J. and Wolfowitz, J. (1959). Optimum designs in regression problems. Ann. Math. Stats., Vol. 30, 271-294.

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. Canad. J. Math., Vol. 12, 363-366.

Kingman, J.F.C. (1967). An inequality involving Radon Nikodym derivatives. Proc. Camb. Phil. Soc., Vol. 63, 195-198.

Luenberger, D.G. (1973). Introduction to Linear and Nonlinear Programming. Reading, Mass.: Addison Wesley.

Mood, A.M. (1946). On Hotelling's weighing problem. Ann Math. Stats., Vol. 17, 432-446.

Morgan, B.J.T. and Titterington, D.M. (1977). A comparison of iterative methods for obtaining maximum likelihood estimates in contingency tables with a missing diagonal. Biometrika, Vol. 64, 265-269.

Murray, G.D. and Titterington, D.M. (1978). Estimation problems with data from a mixture. J.R.S.S.C., Vol. 27, 325-334.

Ortega, J.M. and Rheinboldt, W.C. (1970). Iterative Solution of Nonlinear Equations in Several Variables. New York: Academic Press.

Pazman, A. (1974a). A convergence theorem in the theory of D-optimum experimental designs. Ann. Math. Stats., Vol. 2, 216-218.

Pazman, A. (1974b). The ordering of experimental designs - a Hilbert space approach. Kybernetika (Prague), Vol. 10, 373-388.

Plackett, R.L. (1974). The Analysis of Categorical Data. London: Griffin Monograph.

Pukelsheim, F. (1979). On c-optimal design measures. Preprint.

Pukelsheim, F. (1980). On ϕ_p-optimal design measures: characterisations based on duality theory. Preprint.

Raghavaro, D. (1959). Some optimum weighing designs. Ann. Math. Stats., Vol. 30, 295-303.

Rao, C.R. (1946). On the most efficient designs in weighing. Sankhya, Vol. 7, 440.

Rao, C.R. (1965). Linear Statistical Inference and its Applications. New York: Wiley.

Rao, C.R. and Mitra, S.K. (1971). Generalised Inverse of Matrices and its Applications. New York: Wiley.

Rhode, C.A. (1965). Generalized inverses of partitioned matrices. J. Soc. Indust, Appl. Math., Vol. 13, 1033-1035.

Rockafeller, R.T. (1970). Convex Analysis. Princeton N.J.: Princeton University Press.

St. John, R.C. and Draper, N.R. (1975). D-optimality for regression designs: a review. Technometrics, Vol. 17, 15-23.

Searle, S.R. (1971). Linear Models. New York: Wiley.

Shisha, O. and Mond, B. (1967). Bounds on differences of means. Inequalities (O. Shisha, Ed.), Proceedings of a symposium held at Wright Paterson Air Force Base, Ohio. N.Y. and London: Academic Press, 293-307.

Shohat, J. (1929). Inequalities for moments of frequency functions and for various statistical constants. Biometrika, Vol. 21, 361-370.

Sibson, R. (1972). Contribution to discussion of a paper by H.P. Wynn. J.R.S.S.B., Vol. 34, 181-183.

Sibson, R. (1974a). D_A-optimality and duality. Progress in Statistics. Colloq. Math. Soc, Janos. Bolyai., Vol. 9, 677-692.

Sibson, R. (1974b). Cutting plane algorithms for D-optimal design. Unpublished typescript.

Sibson, R. and Kenny, A. (1975). Coefficients in D-optimal experimental design. J.R.S.S.B., Vol. 37, 288-292.

Silvey, S.D. (1969). Multicollinearity and imprecise estimation. J.R.S.S.B., Vol. 31, 539-552.

Silvey, S.D. (1972). Contribution to discussion of a paper by H.P. Wynn. J.R.S.S.B., Vol. 34, 174-175.

Silvey, S.D. (1974). Some aspects of the theory of optimal linear regression design with a general concave criterion function. Tech. Report 75, Dept. of Stats., Princeton Univ., Princeton N.J.

Silvey, S.D. (1977). Private communication.

Silvey, S.D. (1978). Optimal design measures with singular information matrices. Biometrika, Vol. 65, 553-559.

Silvey, S.D. (1980). Optimal Design. London: Chapman and Hall.

Silvey, S.D. and Titterington, D.M. (1973). A geometric approach to optimal design theory. Biometrika, Vol. 60, 21-32.

Silvey, S.D. and Titterington, D.M. (1974). A Lagrangian approach to optimal design. Biometrika, Vol. 61, 299-302.

Silvey, S.D., Titterington, D.M. and Torsney B. (1976). An algorithm for D-optimal designs on a finite design space. Unpublished typescript.

Silvey, S.D., Titterington, D.M. and Torsney B. (1978). An algorithm for optimal designs on a finite design space. Commun. Stats. A, Vol. 7, 1379-1389.

Silverman, B.W. and Titterington, D.M. (1980). Minimum covering ellipses. SIAM J. of Scientific and Statistical Computing. To appear.

Smith, K. (1918). On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance they give toward a proper choice of the distribution of observations. Biometrika, Vol. 12, 1-85.

Smith, A.F.M. and Makov, U.E. A quasi-Bayes sequential procedure for mixtures. J.R.S.S.B., Vol. 40, 106-112.

Studden, W.J. and VanArman, D.J. (1969). Admissable designs for polynomial spline regression. Ann. Math. Stats., Vol. 40, 1557-1569.

Titterington, D.M. (1975). Optimal design: some geometrical aspects of D-optimality. Biometrika, Vol. 62, 313-320.

Titterington, D.M. (1976). Algorithms for computing D-optimal designs on a finite space. Proc. 1976. Conf. on Information Sciences and Systems, Dept. of Elect. Eng., John Hopkins Univ., Baltimore, MD, 213-216.

Titterington, D.M. (1977). A survey of optimal design algorithms. Proceedings of the 10th European Meeting of Statisticians, Belgium, 64-65.

Titterington, D.M. (1978). Estimation of correlation coefficients by ellipsoidal trimming. J.R.S.S.C., Vol. 27, 227-234.

Torsney, B. (1977). Contribution to discussion of a paper by Dempster, Laird and Rubin. J.R.S.S.B., Vol. 39, 26-27.

Tsay, J.-Y. (1976a). Linear optimal experimental designs. Proc. 1976. Conf. on Information Sciences and Systems, Dept. of Elect. Eng., John Hopkins Univ., Baltimore, MD, 222-226.

Tsay, J.-Y. (1976b). On the sequential construction of D-optimal designs. J.A.S.A., Vol. 71, 671-674.

Tsay, J.-Y. (1977). A convergence theorem in L-optimal design theory. Ann. Stats., Vol. 5, 790-794.

Von Mises, R. (1947). On the asymptotic distribution of differentiable statistical functions. Ann. Math. Stats., Vol. 18, 309-348.

Vuchkov, I.N. (1977). A ridge-type procedure for design of experiments. Biometrika, Vol. 64, 147-150.

Wald, A. (1943). On the efficient design of statistical investigations. Ann. Math. Stats., Vol. 14, 134-140.

White, L.V. (1973). An extension of the general equivalence theorem to nonlinear models. Biometrika, Vol. 60, 345-348.

White, L.V. (1975). Ph.D. Thesis. Imperial College, London.

Whittle, P. (1971). <u>Optimization Under Constraints</u>. New York: Wiley-Interscience.

Whittle, P. (1973). Some general points in the theory of optimal experimental design. J.R.S.S.B., Vol. 35, 123-130.

Wolfe, P. (1961). Accelerating the cutting-plane method for nonlinear programming. J. Soc. Indust. Appl. Math., Vol. 9, 481-488.

Wu, C.-F. (1976). Contributions to optimization theory with applications to optimal design of experiments. Ph.D. Thesis, University of California, Berkeley.

Wu, C.-F. (1978a). Some iterative procedures for generating nonsingular optimal designs. Commun. Stats. A, Vol. 7, 1399-1412.

Wu, C.-F. (1978b). Some algorithmic aspects of the theory of optimal designs. Ann. Stats., Vol. 6, 1286-1301.

Wu, C.-F. and Wynn, H.P. (1978). The convergence of general step-length algorithms for regular optimum design criteria. Ann. Stats., Vol. 6, 1273-1285.

Wynn, H.P. (1970). The sequential generation of D-optimal experimental designs. Ann. Math. Stats., Vol. 41, 1655-1664.

Wynn, H.P. (1972). Results in the theory and construction of D-optimum experimental designs. (with discussion). J.R.S.S.B., Vol. 34, 133-147, 170-186.