



PROFESOR DR. MALIK H.J. ABU HASSAN

TOWARDS **LARGE
SCALE
UNCONSTRAINED
OPTIMIZATION**

PROFESOR DR. MALIK HJ. ABU HASSAN
B.Sc. Malaya, M.Sc. Aston, Ph. D. Loughborough

20 April 2007

**Dewan Taklimat
Tingkat I, Bangunan Pentadbiran
Universiti Putra Malaysia**



Penerbit Universiti Putra Malaysia
Serdang • 2007
<http://www.penerbit.upm.edu.my>

© Penerbit Universiti Putra Malaysia 2007
First print 2007

All rights reserved. No parts of this book may be reproduce in any form without permission in writing from the publisher, except by a reviewer who wishes to quote brief passages in a review written for inclusion in a magazine or newspaper.

UPM Press is a member of the Malaysian Book Publishers Association (MABOPA)
Membership No.: 9802

Perpustakaan Negara Malaysia

Cataloguing-in-Publication Data

Malik Hj. Abu Hassan

Towards large scale unconstrained optimization / Malik Hj. Abu Hassan.

(Inaugural Lecture Series)

ISBN 978-983-3455-90-4

1. Mathematical optimization. 2. Mathematical analysis.

3. Algorithms.

I. Tilte. II. Series.

519.62

Designed, layout and printed by:
Penerbit Universiti Putra Malaysia
43400 UPM, Serdang
Selangor Darul Ehsan.

ABSTRACT

A large scale unconstrained optimization problem can be formulated when the dimension n is large. The notion of 'large scale' is machine dependent and hence it could be difficult to state a priori when a problem is of large size. However, today an unconstrained problem with 400 or more variables is usually considered a large scale problem.

The main difficulty in dealing with large scale problems is the fact that effective algorithms for small scale problems do not necessarily translate into efficient algorithms when applied to solve large scale problems. Therefore in dealing with large scale unconstrained problems with a large number of variables, modifications must be made to the standard implementation of the many existing algorithms for the small scale case.

One of the most effective Newton-type methods for solving large-scale problems is the truncated Newton method. This method computes a Newton-type direction by truncating the conjugate gradient method iterates (inner iterations) whenever a required accuracy is obtained, thereby the superlinear convergence is guaranteed.

Another effective approach to large-scale unconstrained is the limited memory BFGS method. This method satisfies the requirement to solve large-scale problems because the storage of matrices is avoided by storing a number of vector pairs.

The symmetric rank one (SR1) update is of the simplest quasi-Newton updates for solving large-scale problems. However a basic disadvantage is that the SR1 update may not preserve the positive definiteness with a positive definiteness approximation. A simple restart procedure for the SR1 method using the standard line search to avoid the loss of positive definiteness will be implemented.

The matrix-storage free BFGS (MF-BFGS) method is a method that combines with a restarting strategy to the BFGS method. We also attempt to construct a new matrix-storage free which uses the SR1 update (MF-SR1). The MF-SR1 method is more superior than the MF-BFGS method in some problems. However for other problems the MF-BFGS method is more competitive because of its rapid convergence. The matrix-storage methods can be greatly accelerated by means of a simple scaling. Therefore, by a simple scaling on SR1 and BFGS methods, we can improve the methods tremendously.

ABSTRAK

Suatu masalah pengoptimuman tak berkekangan berskala besar boleh dirumuskan apabila matra n adalah besar. Idea 'berskala besar' bersandarkan mesin dan justeru itu adalah terlalu rumit untuk menyatakan a priori apabila suatu masalah adalah bersaiz besar. Bagaimanapun, hari ini suatu masalah tak berkekangan dengan 400 pembolehubah atau lebih, biasanya dipertimbangkan sebagai suatu masalah berskala besar.

Kerumitan utama dalam mengendalikan masalah berskala besar timbul daripada fakta algoritma efektif untuk masalah berskala kecil tidak semestinya menterjemahkannya ke dalam algoritma cekap apabila digunakan untuk menyelesaikan masalah berskala besar. Oleh yang demikian, untuk mengendalikan masalah tak berkekangan berskala besar dengan bilangan pembolehubah yang besar, pengubahsuaian mesti dilakukan terhadap implimentasi piawai bagi algoritma yang sedia ada untuk kes berskala kecil.

Salah satu kaedah jenis Newton yang efektif untuk menyelesaikan masalah berskala besar adalah kaedah Newton terpankaskan. Kaedah ini mengira suatu arah jenis Newton dengan memangkaskan lelaran kaedah kecerunan konjugat (lelaran dalam) apabila suatu kejituan yang diperlukan diperoleh, justeru menjamin penumpuan superlinear.

Pendekatan efektif yang lain untuk masalah tak berkekangan berskala besar ialah kaedah ingatan terhad BFGS. Kaedah ini memenuhi keperluan untuk menyelesaikan masalah berskala besar sebab storan matriks dapat dielakkan dengan menyimpan suatu bilangan pasangan vector.

Rumus kemaskini pangkat satu yang simetri (SR1) adalah yang termudah bagi rumus kemaskini quasi-Newton untuk menyelesaikan masalah berskala besar. Walau bagaimanapun satu kelemahan adalah rumus kemaskini SR1 ini tak boleh

menyimpan atau mengekal ketentuan positif dengan suatu penghampiran tentu positif. Suatu prosedur mula semula yang mudah untuk kaedah SR1 dengan menggunakan gelintaran garis piawai diimplimentasikan untuk mengelak kehilangan ketentuan positif.

Kaedah storan matriks bebas BFGS (MF-BFGS) adalah suatu kaedah yang menggabungkan strategi mula semula kepada kaedah BFGS. Kita akan membina suatu storan matriks bebas menggunakan rumus kemaskini SR1 (MF-SR1). Kaedah MF-SR1 adalah lebih baik daripada kaedah MF-BFGS dalam sesuatu masalah. Walau bagaimanapun untuk kebanyakan masalah kaedah MF-BFGS adalah lebih kompetitif disebabkan kepantasan penumpuan. Kaedah storan matriks boleh dipercepatkan melalui suatu pengskalaan mudah. Justeru itu, dengan menggunakan pengskalaan mudah ke atas kaedah SR1 dan BFGS kita boleh memperbaiki kaedah dengan memberangsangkan.

GENERAL INTRODUCTION TO OPTIMIZATION

Optimization is concerned with getting the best from a given situation by a systematic analysis of alternative decisions hopefully without the need to examine them all.

This involves a performance measure or index for the situation under assessment and the ability to calculate it from the variables at the optimizer's disposal. The variables are then adjusted to give the best possible value of the performance index.

Performance indices differ from situation to situation but generally involve economic considerations, e.g. maximum return on investment, minimum cost per unit yield, etc. They may also involve technical considerations such as minimum time of production, minimum miss-distance of a missile with its target and so on. There may of course be a combination of different indices requiring simultaneous optimization. Such multi-objective problems are difficult and are often dealt with by selecting one of the objectives as primary, fixing suitable values for the others and then regarding them as constraints.

The way in which the performance index is obtained from the variables of the problem can vary widely from one situation to another. At one extreme it may be only possible to give a qualitative indication of the dependence and at the other a highly sophisticated mathematical model will enable the index to be computed accurately given fixed values of variables. The important thing is to be able to determine when one set of values of the variables gives a "better" value of the performance than some other sets. The variables can be adjusted systematically until the "best" performance is obtained.

Mathematical models involved in optimization can range from simple algebraic formulae to sets of linked algebraic, ordinary and partial differential equations.

Furthermore in real-life problems the variables are almost always constrained in some way - sometimes by having simple upper lower bounds and sometimes by complex functional constraints - so that the best value of the objective is sought for some restricted set of values of the variables of the problem.

Optimization problems arise in a variety of different ways ranging from the best design of some single component of a process through to the organization of a company (or even a country) to achieve some desired objective.

They can also result from problems concerned with the determination of parameters in models, natural variational principles in science and engineering (e.g. determination of equilibrium compositions by the minimization of Gibbs free energy in chemical systems), the numerical solutions of algebraic and differential equations, and so on.

A further difficulty arises if these are stochastic processes involved in the system to be optimized.

General Mathematical Formulation

Find the values of the variables x_i , $i = 1, 2, \dots, n$ which will minimize (or maximize) the objective function

$$f(x_1, x_2, \dots, x_n) \quad (1)$$

and satisfy the set of constraints

$$g_k(x_1, x_2, \dots, x_n) \geq 0, \quad k = 1, 2, \dots, m$$

The nature of an optimization problem depends on the nature of the functions f , g_k and of the variables x_i .

When $m = 0$ the problem is *unconstrained*. If f and all the g_k are linear functions the problem is a *linear programme* and if f is quadratic and the g_k linear a *quadratic programme* arises.

When f is a convex function and the constraints form a convex set then we have a *convex programming problem*. (An important property of a convex programming problem is that any local solution is also global).

If the variables x_i can only take on integer values the problem is known as *integer programming*.

In general when f is nonlinear and the g_k are all linear we have *linear constrained optimization* and when f and the g_k are all nonlinear we have *nonlinear programming*.

Optimization in which only a finite number of variables x_i are involved is known as *static optimization* but when the performance index is a functional (of one or more functions) we have *dynamic optimization* and we become concerned with problems in the *calculus of variations* and *optimal control*.

In optimal control the performance index is determined from the solutions of a set of differential equations which depend on a set of functions which have to be selected in an optimal way. The differential equations describing the system are regarded as the constraints $g_k = 0$.

Problems with only equality constraints ($g_k = 0$ for all k) are easier to deal with in general than those with only inequalities ($g_k \geq 0$ for all k) or mixed problems.

Most of the techniques discussed in the following lectures will be concerned with "local" optimization in the sense that any optimum obtained will not necessarily be the overall or "global" optimum but will be the best value in some neighbourhood.

When functions having multiple maxima or minima are involved the problem is extremely difficult and very much the subject of current research.

Sometimes the only way to proceed in global optimization problems is to start searching from different initial points, selected using a grid or by some form of random sampling, and then to compare the local optima so obtained to find the best. This procedure can be very time consuming and, of course, is by no means foolproof.

Another important area of current research is "non smooth" optimization which arises when the performance index is non-differentiable.

METHODS OF OPTIMIZATION

Problem: Given a performance index

$$f(x) : R^n \rightarrow R$$

find a (local) minimizer $x^* \in R^n$.

General Iteration

$$x^{(i+1)} = x^{(i)} + \lambda^{(i)} d^{(i)}, \quad i = 0, 1, 2, \dots \quad (2)$$

where $x^{(0)}$ is the estimate of x^* , $d^{(i)}$ is the direction of search and $\lambda^{(i)}$ is the "step length" along this direction.

Different optimization techniques result from different ways of selecting $d^{(i)}$ and $\lambda^{(i)}$.

STEEPEST DESCENT

Algorithm

Let $x^{(0)}$ be an estimate of x^*

1. Set $i = 0$
2. Compute $\nabla f(x^{(i)}) = g^{(i)}$
3. Compute $\lambda^{(i)}$ such that

$$f(x^{(i)} - \lambda^{(i)} g^{(i)}) = \min_{\lambda} f(x^{(i)} - \lambda g^{(i)})$$

(this is called a line search)

4. Compute
$$x^{(i+1)} = x^{(i)} - \lambda^{(i)} g^{(i)}$$
5. Has process converged?
YES : then $x^{(i+1)}$ is minimizer
NO : set $i = i+1$ and go to 2.

Useful for theoretical reasons but not in practice since convergence is too slow (linear).

NEWTON'S METHOD

$$x^{(i-1)} = x^{(i)} - G^{-1}(x^{(i)}) \nabla f(x^{(i)}) \quad (3)$$

where $G(x^{(i)})$ is the Hessian $\left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right]_{x=x^{(i)}}$

and $\nabla f(x^{(i)})$ is the gradient of $f(x)$ at $x^{(i)}$.

One step convergence on quadratics. For general f with a minimum at which G is positive definite convergence is rapid (quadratic) if $x^{(0)}$ is sufficiently close to x^* .

Basic Newton may fail because G is not positive definite or even when it is if f does not decrease in moving to the minimum of the local quadratic approximation.

Other disadvantages are that G must be computed and a set of n linear equations solved.

MODIFIED NEWTON METHODS

$$1. \quad x^{(i+1)} = x^{(i)} - \lambda^{(i)} G^{-1}(x^{(i)}) \nabla f(x^{(i)}) \quad (4)$$

with $\lambda^{(i)}$ selected so that $f(x^{(i+1)})$ is minimized (line search).

$$2. \quad x^{(i+1)} = x^{(i)} - (G(x^{(i)}) + rI)^{-1} \nabla f(x^{(i)}) \quad (5)$$

where r is selected to make (...) positive definite.

Modifications 1. and 2. can be combined.

Modified Newton Methods are regarded as among the best available. This is true even if $G(x^{(i)})$ (but not $\nabla f(x^{(i)})$ also) has to be obtained numerically provided that n is not too large.

QUASI-NEWTON METHODS

Algorithm

Let $x^{(0)}$ be an estimate of x^* , $H^{(0)}$ be a positive definite matrix and $g^{(i)} = \nabla f(x^{(i)})$, $i = 0, 1, 2, \dots$

1. Set $i = 0$
2. Compute $d^{(i)} = -H^{(i)} g^{(i)}$
3. Find $\lambda^{(i)}$ by minimizing $f(x^{(i)} + \lambda d^{(i)})$ using a suitable line search.
4. Compute

$$x^{(i+1)} = x^{(i)} - \lambda^{(i)} d^{(i)}$$

$$p^{(i)} = x^{(i+1)} - x^{(i)} = \lambda^{(i)} d^{(i)}$$

$$q^{(i)} = g^{(i+1)} - g^{(i)}$$

5. Update $H^{(i)}$ by

$$a. H^{(i+1)} = H^{(i)} + \frac{p^{(i)} p^{(i)T}}{p^{(i)T} q^{(i)}} - \frac{H^{(i)} q^{(i)} q^{(i)T} H^{(i)}}{q^{(i)T} H^{(i)} q^{(i)}} \quad (6)$$

OR

$$b. H^{(i+1)} = H^{(i)} + \left(1 + \frac{q^{(i)T} H^{(i)} q^{(i)}}{p^{(i)T} q^{(i)}}\right) \frac{p^{(i)} p^{(i)T}}{p^{(i)T} q^{(i)}} - \left(\frac{p^{(i)} q^{(i)T} H^{(i)} + H^{(i)} q^{(i)} p^{(i)T}}{p^{(i)T} q^{(i)}}\right) \quad (7)$$

6. Set $i = i + 1$, go to 2.

Method (a) is known as the Davidon Fletcher Powell (DFP) method and method (b) as the Broyden Fletcher Goldfarb Shanno (BFGS) method.

There are many other quasi-Newton methods depending on how $H^{(k)}$ is updated. (Some methods approximate to G rather than G^{-1}).

Quasi-Newton methods are some of the best available optimization techniques both when ∇f is known analytically and with numerical evaluation of ∇f provided suitable care is taken. The best quasi-Newton method is currently held to be the BFGS method.

CONJUGATE DIRECTION METHODS

Algorithm for Fletcher Reeves Method

1. Set $i = 1$

$$d^{(0)} = -\nabla f(x^{(0)})$$

2. Compute

$$x^{(i)} = x^{(i-1)} - \lambda^{(i-1)} d^{(i-1)}$$

where $\lambda^{(i-1)}$ is found by minimizing $f(x^{(i)})$ using a suitable line search.

3. Compute

$$d^{(i)} = -g^{(i)} - \beta^{(i-1)} d^{(i-1)} \tag{8}$$

where

$$\beta^{(i-1)} = \frac{g^{(i)T} g^{(i)}}{g^{(i-1)T} g^{(i-1)}} \tag{9}$$

and

$$g^{(i)} = \nabla f(x^{(i)})$$

4. Has process converged?

Yes : Stop

No : Set $i = i + 1$. Go to 2.

For quadratic f this algorithm produces conjugate directions in at most n steps and hence x^* .

For general f the process is usually restarted with a step along the current gradient direction after each cycle of n steps.

There is a number of other ways of selecting β , for example in the Polak-Ribiere method

$$\beta^{(i-1)} = \frac{g^{(i)T} (g^{(i)} - g^{(i-1)})}{g^{(i-1)T} g^{(i-1)}} \quad (10)$$

Notice that some quasi-Newton methods, e.g. the DFP method, are also conjugate gradient methods.

Conjugate direction methods are not generally as efficient as quasi-Newton methods.

Those conjugate direction methods that only use vectors and not matrices can be useful when n is large because of their relatively small storage requirements.

SIMPLEX OR POLYTOPE METHODS

General Algorithm

1. Compute f at the vertices of a simplex in n - dimensions.
2. Find the "worst" point by comparison. "Worst" means point at which f has its largest value when looking for a minimum.

3. Replace the worst point by its reflection in the centroid of the other n points.
4. Repeat process with newly formed simplex.
5. Reduce size of simplexes as the minimum is approached.
6. Stop when standard deviation of the values of f at the vertices of the current simplex is less than a preset value.

Simplex methods are not generally recommended when methods using derivatives are available. They can be useful in problems where f is not very well determined (as in the case of many chemical engineering problems) and is of small dimension.

USE OF DIFFERENTIAL EQUATIONS

The equations of the orthogonal trajectories of $f(x):R^n \rightarrow R$ are

$$\frac{dx}{dt} = \pm \nabla f(x) \quad (11)$$

A generalization of the trapezoidal rule for integrating (11) is

$$\frac{x^{(i+1)} - x^{(i)}}{h} = (1 - \theta) \nabla f(x^{(i)}) + \theta \nabla f(x^{(i+1)}) \quad (12)$$

Applying Newton's method to solving the nonlinear equations (12) gives

$$x^{(i+1)} = x^{(i)} - \left[\theta G(x^{(i)}) - \frac{1}{h} I_n \right]^{-1} \nabla f(x^{(i)}) \quad (13)$$

This is a modified-Newton optimization process.

We have devised an algorithm that selects θ and b to give A-stability and an optimization process between steepest descent and Newton. The algorithm has been adapted to optimal control problems including tubular reactors.

GLOBAL OPTIMIZATION

Algorithm (Malik, 1982)

1. Find a local minimizer using a differential method and any minimization method.
2. Transfer the local minimizer to the origin and compute its domain of attraction by numerical use of Zubov's method.
3. Find a point outside this domain of attraction (if any) and using this as the initial point go to 1.
4. Repeat until all local minimizers and domains of attraction have been located.

The method seems to work well in 1 or 2 dimensions. There would be considerable difficulty in extending the technique to more than 2 dimensions.

LARGE - SCALE UNCONSTRAINED OPTIMIZATION

A large scale unconstrained optimization problem can be formulated when the dimension n is large. The notion of 'large scale' is machine dependent and hence it could be difficult to state a priori when a problem is of large size. However, today an unconstrained problem with 400 or more variables is usually considered a large scale problem.

Besides its theoretical importance, the growing interest in the last years in solving problems of large size derives from the fact that problems with larger and larger number of variables are arising very frequently from the real world as a result of modeling systems of a very complex structure.

The main difficulty in dealing with large scale problems is the fact that effective algorithms for small scale problems do not necessarily translate into efficient algorithms when applied to solve large scale problems. Therefore in dealing with large scale unconstrained problems with a large number of variables, modifications must be made to the standard implementation of the many existing algorithms for the small scale case.

A basic feature of an algorithm for large-scale problems is a low storage overhead needed to make practicable its implementation. As in small-scale case, most of the large-scale unconstrained optimization algorithms are iterative methods, which generate a sequence of points according to the scheme,

$$x^{(i+1)} = x^{(i)} + \lambda^{(i)} d^{(i)}, \quad i = 0, 1, 2, \dots$$

where $x^{(0)}$ is the estimate of x^* , $d^{(i)}$ is the direction of search and $\lambda^{(i)}$ is the "step length" along this direction. Obviously, in large-scale optimization it is important that an algorithm be able to compute the search direction efficiently and economically. Therefore it is always possible to include modifications in the calculation of the search direction to serve this purpose.

Steepest Descent Method

One good method for solving large-scale unconstrained optimization problems is the steepest descent method. Due to its very low storage required by a standard implementation and

the ensured global convergence, it could be attractive in the large-scale setting. However, its convergence rate is only linear and therefore it is too slow to be used. Some particular scaling had been considered and this led to an implementation of the efficiency of the method.

Newton-type Methods

One of the most effective Newton-type methods for solving large-scale problems is the truncated Newton method introduced by Dembo and Steihaug (1983). This method computes a Newton-type direction by truncating the conjugate gradient method (CG) iterates (inner iterations) whenever a required accuracy is obtained, thereby the superlinear convergence is guaranteed.

Truncated Newton Algorithm

Step 1. Outer Iterations

For $i = 0, 1, \dots$
 Compute $g^{(i)}$.
 Test for convergence.

Step 2. Inner iterations (Computation of the direction)

Iterate CG algorithm until a termination criterion is satisfied.

Step 3. Compute a step-length $\lambda^{(i)}$ by a line search procedure.

Set $x^{(i+1)} = x^{(i)} + \lambda^{(i)} d^{(i)}$; $i := i + 1$ go to step 1.

Lucidi *et al.* (1998) used a curvilinear line search in the truncated Newton algorithm to solve large-scale problems and are globally convergent towards points where positive semi-definiteness of the Hessian matrix is satisfied.

Limited Memory BFGS Method (L-BFGS)

Another effective approach to large-scale unconstrained is the limited memory BFGS method (L-BFGS) proposed by Nocedal (1980) and then studied Nash and Nocedal (1991). This method resembles the BFGS quasi-Newton method and satisfies the requirement to solve large-scale problems because the storage of matrices is avoided by storing a number of vector pairs.

L-BFGS Algorithm

Step 1. Choose $x^{(0)}$ and the initial matrix $H^{(0)}$. Set $i=0$

Step 2. Compute $d^{(i)} = -H^{(i)} g^{(i)}$

$$\text{and } x^{(i+1)} = x^{(i)} + \lambda^{(i)} d^{(i)}$$

where $H^{(i)}$ is the approximation to the inverse Hessian of $f(x)$ at the i^{th} iteration and $\lambda^{(i)}$ satisfies the Wolfe conditions

$$f(x^{(i+1)}) < f(x^{(i)}) + \beta_1 g^{(i)T} p^{(i)} \text{ and } g^{(i+1)T} p^{(i)} \geq \beta_2 g^{(i)T} p^{(i)} \quad (14)$$

(the step-length $\lambda=1$ is tried first with $\beta_1=10^{-4}$ and $\beta_2=0.9$.

Step 3. Let $\hat{m} = \min\{k, m-1\}$. Update for $H^{(0)}$ $\hat{m} + 1$ times by using the pairs $\{q^{(j)}, p^{(j)}\}_{j=i-\hat{m}}^i$, i.e. let

$$\begin{aligned}
 H^{(i+1)} &= (V^{(i)^T} \dots V^{(i-\hat{m})^T}) H^{(0)} (V^{(i-\hat{m})} \dots V^{(i)}) \\
 &+ \rho^{(i-\hat{m})} (V^{(i)^T} \dots V^{(i-\hat{m}+1)^T}) p^{(i-\hat{m})} p^{(i-\hat{m})^T} (V^{(i-\hat{m}+1)} \dots V^{(i)}) \\
 &+ \rho^{(i-\hat{m}+1)} (V^{(i)^T} \dots V^{(i-\hat{m}+2)^T}) p^{(i-\hat{m}+1)} p^{(i-\hat{m}+1)^T} (V^{(i-\hat{m}+2)} \dots V^{(i)}) \\
 &\dots + \rho^{(i)} p^{(i)} p^{(i)^T} \tag{15}
 \end{aligned}$$

Step 4. Set $i := i+1$, and go to Step 2.

Positive-definite Scaled Symmetric Rank One Method

The symmetric rank one (SR1) update is of the simplest quasi-Newton updates and this simplicity makes the SR1 update a candidate for large-scale problems. However a basic disadvantage to the SR1 update is that the SR1 update may not preserve the positive definiteness with a positive definiteness approximation. Moreover, the SR1 update may also be indefinite. Several researchers have renewed their interest in the study of SR1 update. Khalfan (1989) use trust region to avoid the possible loss of positive-definiteness. IP and Todd (1988) suggested to size up the SR1 update for avoiding the lost of definiteness which results in the optimal condition sized SR1 updates. A simple treatment to these problems is to restart the update with the initial approximation mostly the identity matrix whenever this loss arises. However our numerical experience shows that restart with the identity matrix is not a good choice. A simple restart procedure for the SR1 method using the standard line search to avoid the loss of positive definiteness will be implemented.

Algorithm - Quasi Newton SR1 Method with Restart (NS-SR1)

Given an initial point $x^{(0)}$, an initial positive matrix $H^{(0)} = I$, set $i := 0$.

Step 1. If the convergence criterion, $\|g^{(i)}\| \leq \varepsilon \max\{1, \|x^{(i)}\|\}$ is achieved, then stop.

Step 2. Compute a quasi-Newton direction

$$d^{(i)} = -H^{(i)} g^{(i)} \quad (16)$$

$$\text{where } H^{(i+1)} = H^{(i)} + \frac{(p^{(i)} - H^{(i)}q^{(i)})(p^{(i)} - H^{(i)}q^{(i)})^T}{(p^{(i)} - H^{(i)}q^{(i)})^T} \quad (17)$$

Step 3. If $d^{(i)T}g^{(i)} > 0$, ($H^{(i)}$ is not positive definite) set $H^{(i)} = I$ and $d^{(i)} = -g^{(i)}$ subsequently. Else retain (16)

Step 4. Using line search, find an acceptable step-length $\lambda^{(i)}$, such that the Wolfe conditions (14) are satisfied ($\lambda^{(i)} = 1$ is always tried first, with $\beta_1 = 10^{-4}$ and $\beta_2 = 0.9$).

Step 5. Set $x^{(i+1)} = x^{(i)} + \lambda^{(i)} d^{(i)}$

Step 6. Compute the next inverse Hessian approximation $H^{(i+1)}$.

Step 7. Set $i := i+1$, and go to Step 1.

Numerical results show that restarting with the identity matrix may not be a convenient choice. The SR1 update may not preserve positive definiteness at the next iteration even if the current is, i.e. when $H^{(i)}$ is positive definite and $p^{(i)T} q^{(i)} > 0$. The algorithm will keep on restarting with little or no progress until the maximum number of function evaluation allowed is exceeded. Hence, restart with the identity matrix is clearly not a good choice. Instead, we consider the cheap choice of replacing the identity matrix with a positive multiple of the identity matrix. Malik *et al.* (2002a,2002b) used a positive multiple of the identity matrix instead of identity matrix itself. The used positive scaling factor is the optimal solution of the measure defined by the problem. A replacement in the form of positive multiple of identity matrix is necessary for the SR1 when it is not positive definite.

We present a description of the scaled SR1 (S-SR1) algorithm that ensures the positive definiteness of the SR1 update.

Algorithm - S-SR1

Given an initial point $x^{(0)}$, an initial positive matrix $H^{(0)} = I$, set $i = 0$.

Step 1. If the convergence criterion, $\|g^{(i)}\| \leq \varepsilon \max\{1, \|x^{(i)}\|\}$ is achieved, then stop.

Step 2. Compute a quasi-Newton direction

$$d^{(i)} = -H^{(i)} g^{(i)}$$

where $H^{(i)}$ is given by (17)

Step 3 If $d^{(i)T}g^{(i)} > 0$, ($H^{(i)}$ is not positive definite) or $i = 1$, set

$$H^{(i)} = \tilde{\delta}^{(i-1)}$$

$$\text{where} = \frac{P^{(i-1)T}P^{(i-1)}}{Q^{(i-1)T}P^{(i-1)}} \left\{ \frac{(P^{(i-1)T}P^{(i-1)})^2}{(Q^{(i-1)T}P^{(i-1)})^2} \frac{P^{(i-1)T}P^{(i-1)}}{Q^{(i-1)T}Q^{(i-1)}} \right\}^{1/2}$$

and subsequently

$$d^{(i)} = - \tilde{\delta}^{(i-1)} g^{(i)}. \text{ Else retain (16)}$$

Step 4. Using line search, find an acceptable step-length $\lambda^{(i)}$, such that the Wolfe conditions (14) are satisfied ($\lambda^{(i)}=1$ is always tried first with $\beta_1 = 10^{-4}$ and $\beta_2 = 0.9$).

Step 5. Set $x^{(i+1)} = x^{(i)} + \lambda^{(i)}d^{(i)}$

Step 6. Compute the next inverse Hessian approximation $H^{(i+1)}$.

Step 7. Set $i := i+1$, and go to Step 1.

A set of test problems is carried out and we find that the S-SR1 update requires less iterations and function calls than NS-SR1. Moreover, we see that most of the problems can be solved by S-SR1 under a certain number of function calls but no for NS-SR1. Therefore, by a simple scaling on SR1 method, we can improve the SR1 method tremendously. Also, the S-SR1 method, on the average, requires fewer iterations than the BFGS method in a line search algorithm. Under conditions that do not assume uniform linear independence, but do assume positive

definiteness and boundedness of the Hessian approximations the S-SR1 method has super-linear and quadratic convergences.

MATRIX-STORAGE FREE QUASI-NEWTON METHOD

The matrix-storage free BFGS (MF-BFGS) method is a method that combines with a restarting strategy to the BFGS method. We also attempt to construct a new matrix-storage free which uses the SR1 update (MF-SR1). The MF-SR1 method is more superior than the MF-BFGS method in some problems. However for other problems the MF-BFGS method is more competitive because of its rapid convergence. The matrix-storage methods can be greatly accelerated by means of a simple scaling.

Matrix-storage free quasi-Newton methods (MF methods for Short) are intended to solve large-scale problems when the Hessian of the objective function has no particular structure. In their general setting, these methods do not try to take advantage of the possible sparsity of the Hessian. It might help in filling the gap between conjugate gradient (CG) and quasi-Newton. The former uses few locations in memory, $O(n)$, but converge rather slow and require expensive linear searches. Quasi-Newton methods have fast rate of convergence (superlinear) and no need exact line searches, but require large memory, $O(n^2)$ storage locations.

The MF methods are an adaptation of the quasi-Newton method to large-scale problems. The implementation described is almost identical to that of the standard quasi-Newton method. The only difference is in the matrix update. Instead of storing the matrices $H^{(i)}$, one stores only the most recent $\{q^{(i)}, p^{(i)}\}$ that defines them implicitly.

Algorithm - MF-BFGS Method

Step 1. Choose $x^{(0)}$, and starting matrix $H^{(0)} = I$. Set $i = 0$.

Step 2. (Step computation procedure using BFGS formula).

Let $x^{(i)}$ be the current iterate.

Compute $p^{(i-1)}$, $q^{(i-1)}$ and $g^{(i)}$:

Step a. Compute $p^{(i-1)T} q^{(i-1)}$, $q^{(i-1)T} q^{(i-1)}$, $p^{(i-1)T} g^{(i)}$
and $q^{(i-1)T} g^{(i)}$.

Step b. Compute

$$y = \begin{bmatrix} \rho^{(i-1)^2} (1/\rho^{(i-1)} + q^{(i-1)T} q^{(i-1)}) (p^{(i-1)T} g^{(i)}) - \rho^{(i-1)} q^{(i-1)T} g^{(i)} \\ -\rho^{(i-1)} p^{(i-1)T} g^{(i)} \end{bmatrix}$$

Step c. Compute

$$H^{(i)} g^{(i)} = g^{(i)} + \begin{bmatrix} p^{(i)} & q^{(i)} \end{bmatrix} y$$

Step 3. Compute

$$x^{(i+1)} = x^{(i)} - \lambda^{(i)} H^{(i)} g^{(i)}$$

where $\lambda^{(i)}$ satisfies the Wolfe conditions (14) ($\lambda^{(i)} = 1$ is always tried first, with $\beta_1 = 10^{-4}$ and $\beta_2 = 0.9$).

Step 4. If the stopping criterion, is achieved, then stop; else set $i := i+1$, and go to Step 2.

In step 2 (*step computation procedure*), each inner product in step a. requires n multiplication; Step b. requires a 2×1 vector storage. When this procedure is part of an algorithm using a line search procedure, the scalar $p^{(i-1)T} g^{(i-1)}$ is also required for the line search, whereas $g^{(i)T} g^{(i)}$ is needed to check the stopping conditions of the algorithm.

Algorithm (MF-SR1 method)

The MF-SR1 method differs from the MF-BFGS only in Step 2.

Step 2. Let $x^{(j)}$ be the current iterate. Compute $p^{(i-1)}$, $q^{(i-1)}$ and $g^{(j)}$:

Step a. Compute $p^{(i-1)T} q^{(i-1)}$, $q^{(i-1)T} q^{(i-1)}$, $p^{(i-1)T} g^{(i)}$ and $q^{(i-1)T} g^{(i)}$.

Then calculate $v^{(i-1)T} q^{(i-1)}$ and $v^{(i-1)T} g^{(i)}$, where $v^{(i+1)} = p^{(i-1)} - q^{(i-1)}$.

Step b. Compute

$$y = \omega^{(i-1)} v^{(i-1)T} g^{(i)},$$

$$\text{where } \omega^{(i-1)} = \frac{1}{v^{(i-1)T} q^{(i-1)}}$$

Step c. Compute

$$H^{(i)} g^{(i)} = g^{(i)} + y v^{(i-1)}$$

In this procedure we require 1×1 vector storage for $\nu^{(i-1)}$. Each inner product in Step a. requires n multiplication. Both $\nu^{(i-1)T} q^{(i-1)}$ and $\nu^{(i-1)T} g^{(i)}$ can be obtained using $p^{(i-1)T} q^{(i-1)}$, $q^{(i-1)T} q^{(i-1)}$, $p^{(i-1)T} g^{(i)}$ and $q^{(i-1)T} g^{(i)}$. The cost for $\omega^{(i-1)}$ is free since the inner products have been stored.

The evaluation of optimization on large-scale test problems is more difficult than in small dimensional case. When the number of variables is very large, the computational effort of the iteration dominates the cost of evaluating the function and gradient. The performance of MF-SR1 method is poor due to the choice of starting matrix $H^{(0)} = I$. Malik *et al.*(2001, 2003, 2004) and Leong and Malik (2005) find a suitable replacement of the identity matrix as the starting/restart matrix. A simple scaling can dramatically reduce the number of iterations of their partitioned quasi-Newton methods.

MF-BFGS with Scaling (MF-BFGS-S)

The steps of the algorithm are similar to MF-BFGS, the only difference is in step 2.

Step2. Let $x^{(j)}$ be the current iterate.

Compute $p^{(i-1)}$, $q^{(i-1)}$ and $g^{(j)}$:

Step a. Compute $p^{(i-1)T} q^{(i-1)}$, $q^{(i-1)T} q^{(i-1)}$, $p^{(i-1)T} g^{(i)}$
and $q^{(i-1)T} g^{(i)}$.

Step b. Compute $\delta^{(i)} = \frac{p^{(i-1)T} q^{(i-1)}}{q^{(i-1)T} q^{(i-1)}}$

Step c. Compute

$$y = \left[\begin{aligned} &\rho^{(i-1)^2} (1/\rho^{(i-1)} + \bar{\delta}^{(i)} q^{(i-1)^T} q^{(i-1)}) (p^{(i-1)^T} g^{(i)}) - \rho^{(i-1)} \bar{\delta}^{(i)} q^{(i-1)^T} g^{(i)} \\ &- \rho^{(i-1)} p^{(i-1)^T} g^{(i)} \end{aligned} \right]$$

Step d. Compute

$$H^{(i)} g^{(i)} = \bar{\delta}^{(i)} g^{(i)} + \left[\begin{array}{c} p^{(i)} \\ \bar{\delta}^{(i)} q^{(i)} \end{array} \right] y$$

The choice of step b. is free since both inner products have been stored while performing step b.

MF-SR1 with Scaling (MF-SR1-S)

The steps of the algorithm are similar to MF-SR1 except for step 2.

Step 2. Let $x^{(j)}$ be the current iterate. Compute $p^{(i-1)}$, $q^{(i-1)}$ and $g^{(i)}$:

Step a. Compute $p^{(i-1)^T} q^{(i-1)}$, $q^{(i-1)^T} q^{(i-1)}$, $p^{(i-1)^T} p^{(i-1)}$, $p^{(i-1)^T} g^{(i)}$ and $q^{(i-1)^T} g^{(i)}$.

Step b. Compute

$$\tilde{\delta}^{(i)} = \frac{p^{(i-1)^T} p^{(i-1)}}{q^{(i-1)^T} p^{(i-1)}} - \left\{ \frac{(p^{(i-1)^T} p^{(i-1)})^2}{(q^{(i-1)^T} p^{(i-1)})^2} - \frac{p^{(i-1)^T} p^{(i-1)}}{q^{(i-1)^T} q^{(i-1)}} \right\}^{1/2}$$

Step c. Compute and store then calculate $v^{(i-1)^T} q^{(i-1)}$ and

$$v^{(i-1)^T} g^{(i)}$$

$$\text{where } v^{(i-1)} = p^{(i-1)} - \tilde{\delta}^{(i)} q^{(i-1)} .$$

Step d. Compute

$$y = \omega^{(i-1)} \tilde{\nu}^{(i-1)\top} g^{(i)},$$

$$\text{where } \omega^{(i-1)} = \frac{1}{\tilde{\nu}^{(i-1)\top} q^{(i-1)}}$$

Step e. Compute

$$H^{(i)} g^{(i)} = \tilde{\delta}^{(i)} g^{(i)} + y \tilde{\nu}^{(i-1)}$$

Comparing with the step computation procedure for MF-SR1, an additional inner product is required for calculating $p^{(i-1)\top} p^{(i-1)}$. Both $\nu^{(i-1)\top} q^{(i-1)}$ and $\nu^{(i-1)\top} g^{(i)}$ can also be obtained using $p^{(i-1)\top} q^{(i-1)}$, $\tilde{\delta}^{(i)}$, $q^{(i-1)\top} q^{(i-1)}$, $p^{(i-1)\top} g^{(i)}$ and $\tilde{\delta}^{(i)} q^{(i-1)\top} g^{(i)}$.

Comparison between the MF-BFGS and MF-SR1 Methods

From numerical results the number of function calls and gradient calls are in the range of 1.5 - 1.7 times the number of iterations. Function evaluations are not needed in computing the search directions and hence only the gradient calls are being computed. To compute the new search direction, for example, in step 2 of truncated Newton's algorithm, requires two gradient calls per iteration, that is, we have to estimate $d^{(j)}$ and $p^{(j)}$. Moreover, when an exact line search is used, at least one cubic interpolation is performed at each iteration requiring a further two gradient calls. Although the iteration number is very competitive, the computational labour is quite satisfactory.

In terms of the number of iterations and functions, the MF-BFGS-S is slightly better than that of the MF-SR1-S method. However when n is large (≥ 5000) the MF-SR1-S

method always performs better In fact the MF-SR1-S requires less CPU time to converge to the optimal point. This is due to the fact that SR1 updates require less computational effort when compared with the BFGS updates.

Scaling is very important in the MF-SR1-S method especially when the SR1 updates are not necessarily positive definite. The choice of the scaling factor plays an important role since we have to maintain the positive definiteness of the new update. On the other hand, the influences of scaling factor for BFGS updates are difficult to interpret. For example, when the scaling factor is large, near the optimal solution, the step remains too large during the run.

Comparison between MF-BFGS and L-BFGS methods

We compare the amount of storage required by the limited memory methods (L-BFGS) for various values of m (the number which determines the number of matrix updates that can be stored) and n and the storage required by the MF-BFGS method. The MF-BFGS method requires less number of storage as compared with the L-BFGS method.

Defining the index of computational labour as $ICL = (n + 1) n_f / n_p$, we compare the performance of both methods and we found that MF-BFGS is very competitive since the difference in ICL of both methods is so small. Sometimes the ICL of MF-BFGS is slightly lower than that of L-BFGS or vice versa.

To investigate the effect of increasing the storage in both methods, we define 'storage-up' as the ratio of storage locations for MF-BFGS : storage locations for L-BFGS and 'speed up' in terms of n_i and n_f to be the ratios of total n_i of MF-BFGS : total n_i of L-BFGS and total n_f of MF-BFGS : total n_f of L-BFGS. Thus if the 'speed-up' factors are less than the 'storage-

up factors, the L-BFGS methods do not gain much from additional storage, whereas a large number means a substantial improvement. This means that although there is improvement when we switch from MF-BFGS method to L-BFGS methods, the gain is not dramatic. Hence we conclude that the MF-BFGS method is efficient if the resource in storage is low.

Comparison between MF-BFGS and Conjugate Gradient Methods

We compare the MF-BFGS methods with some of the well-known conjugate gradient methods, viz the conjugate gradient methods (CG) using Fletcher-Reeves and Polak-Ribiere formula. The performances of function calls and number of iterations of the MF-BFGS method are better than the conjugate gradient methods. In terms of ICL, clearly the MF-BFGS method is much more superior than the conjugate methods.

Solving Extremely Large-Scale problems

The L-BFGS, MF-SR1, MF-BFGS and CG methods can solve quite easily problems with 1000 variables. Some require many function and gradient calls and large number of iterations. The index of computational labour varies from one method to another. We have tested the performance of all these methods on problems with 10^6 variables. The MF-BFGS is the only successful method in solving these extremely large-scale problems efficiently. The L-BFGS method cannot solve these extremely large scale problems because the memory is insufficient to start the runs and they require large storage. All CG methods fail to converge when applied to large-scale problems.

CONCLUSIONS

Besides its theoretical importance, the growing interest in the last years in solving problems of large size derives from the fact that problems with larger and larger number of variables are arising very frequently from the real world as a result of modeling systems of a very complex structure.

Solving large scale unconstrained optimization problems require convergence properties, ample storage and memory, high computational cost, scaling factors, restart strategies and many others that are related to large scale unconstrained optimization. We have shown that scaling is important in MF-methods both in BFGS and SR1 methods. A positive definite initial matrix is needed at each iteration. Here, a positive multiple of identity matrix and positive scaling factors are used to obtain the optimal solution.

The numerical experiments show that the MF-SR1-S method gives good results. The MF-SR1-S uses the SR1 update for the approximate inverse Hessian. The SR1 update has major advantages in that it is simple, requires less computation and has some very strong convergence properties. So the MF-SR1-S is an ideal method for large scale optimization.

The L-BFGS methods require more storage and is not suitable when the problem has very large number of variables or the resource in storage is very low. We found that the CG methods require only low storage but more function and gradient calls. This is very unsuitable if the function and gradient evaluations are expensive or an inaccurate line search is used. Moreover there is no guarantee of convergence for CG methods globally when applied to large scale problems.

The MF-BFGS method is appealing for several reasons: It is simple to implement. It requires moderate number of function and gradient calls and low storage requirement. In terms of

Towards Large Scale Unconstrained Optimization

computational labour, MF-BFGS is competitive with L-BFGS and is clearly superior over the CG methods. Lastly, The MF-BFGS method successfully solves extremely large scale problems with 10^6 variables while other methods fail to find the optimal solution.

REFERENCES

- Dembo, R. and T. Steihaug. 1983. Truncated-Newton algorithm for large scale unconstrained optimization. *Mathematical Programming* **26**: 190-212.
- IP, C.M. and M.J. Todd. 1988. Optimal conditioning and convergence in rank one quasi-Newton updates. *SIAM Journal on Numerical Analysis* **25**: 206-221.
- Khalfan, H.F.H. 1989. Topics in quasi-Newton for unconstrained optimization. Ph.D Thesis, University of Colorado.
- Leong Wah June and Malik Hj. Abu Hassan. 2005. Modifications of the limited memory BFGS algorithm for large scale nonlinear optimization. *Mathematical Journal of Okayama University* **47**: 175-188.
- Lucidi, S., F. Rochetich, and M. Roma. 1998. Curvilinear stabilization techniques for truncated newton methods in large scale unconstrained optimization. *SIAM Journal on Optimization* **8**: 916-939.
- Malik Abu Hassan. 1982. Extension of Zubov's method for the determination of domains of attraction. Ph. D Thesis, Lough. Univ. of Tech., England.
- Malik Hj. Abu Hassan, Mansor b. Monsi and Leong Wah June. 2001. Limited BFGS method for large scale optimization. *MATEMATIKA* **17**: 5-23.
- Malik Hj. Abu Hassan, Mansor b. Monsi and Leong Wah June. 2002a. Scaling symmetric rank one update for unconstrained optimization. *ANALISIS* **9**: 63-76.
- Malik Hj. Abu Hassan, Mansor b. Monsi and Leong Wah June. 2002b. Convergence of a positive definite scaled SR1 method. *MATEMATIKA* **18**: 117-127.

Towards Large Scale Unconstrained Optimization

- Malik Hj. Abu Hassan, Mansor b. Monsi and Leong Wah June. 2003. Numerical experiments with matrices storage free BFGS method for large scale unconstrained optimization. *MATEMATIKA* **19**: 107-119.
- Malik Hj. Abu Hassan, Mansor b. Monsi and Leong Wah June. 2004. Limited memory BFGS update with modifications. *Journal of Mathematics and Natural Sciences* **14**: 103-117.
- Nash, S. and J. Nocedal. 1991. A numerical study of the limited BFGS and the truncated-Newton method for large scale optimization. *SLAM Journal on Optimization* **1**: 358-372.
- Nocedal, J. 1980. Updating quasi-Newton matrices with limited storage. *Mathematics of Computation* **35**: 773-78.

BIOGRAPHY

DR. MALIK HJ. ABU HASSAN was born on 17th July 1948 in Masjid Tanah Malacca, the second eldest in the family of seven. He received his early education at Masjid Tanah English School, St. David's High School and Malacca High School. He was the best student in Mathematics both in the Lower Certificate of Education (LCE) and SC/FMC examinations while studying at St. David. After obtaining a full certificate in the Higher School Certificate (HSC) examination in 1968, he gained admission to the University of Malaya in 1969 and obtained his B.Sc.(Hons.) degree in Mathematics in 1973. In the same year he started his career as a tutor in the Mathematics Department of the Basic Science Division in Universiti Pertanian Malaysia. Due to his love for Mathematics he went on to further his studies at University of Aston in Birmingham, England. There he obtained his Master of Science in Industrial Mathematics degree in 1974 and was also the top student in his M.Sc class. Due to the shortage of staff at the Universiti Pertanian Malaysia at that time he was immediately appointed as lecturer in 1974. After four years of gaining experience as a lecturer he left again to continue his studies at Loughborough University of Technology in England where he was conferred his Doctorate in of Philosophy in Applied Mathematics in 1982.

Dr. Malik was appointed the Head of Department of Mathematics in UPM between 1990-1994 and 2002-2004. His positive attitude and support of his colleagues within his department as well as the University at large make him enjoy his work as a lecturer in UPM tremendously. Dr. Malik loves teaching his students both at the undergraduate and graduate levels. He teaches several Mathematics courses ranging from first year courses up to graduate courses. After 34 years of service he has taught thousands of students and supervised

several graduate students. Apart from teaching he is also active in research activities. His current research interest is focused on large scale unconstrained optimization problems, domains of attractions of non autonomous systems, linear quadratic problems for descriptor systems and predator-prey population models including time delay and harvesting. He has won several medals for his work in the Invention and Research Exhibition at UPM. He also received the Excellent Service Award in 1995, 1997 and 2005 and since 1995 onwards, the Excellent Service Certificate.

During his younger days he was active in sports especially in gymnastic, sepak takraw, hockey, carom and squash. He participated in the inter faculty competitions and was champion in sepak takraw for four successive years. In 1977 he represented the Faculty Science in hockey and emerged as champion. Between 1976-78 he was elected as chairman of the Faculty of Science's Sport, Social and Welfare Club. In 1977-78 he was the chairman of the UPM Sepak Takraw Club and Advisor of UPM student's Sepak Takraw team.

Dr. Malik was the chief examiner in Applied Mathematics and Further Mathematics for STPM examination from 1989-2002. He is also professionally involved in the Matriculation Division of the Ministry of Education where he is the chief examiner and chief assessor from since 1999 till present.

LIST OF INAUGURAL LECTURES

1. Prof. Dr. Sulaiman M. Yassin
The Challenge to Communication Research in Extension
22 July 1989
2. Prof. Ir. Abang Abang Ali
Indigenous Materials and Technology for Low Cost Housing
30 August 1990
3. Prof. Dr. Abdul Rahman Abdul Razak
Plant Parasitic Nematodes, Lesser Known Pests of Agricultural Crops
30 January 1993
4. Prof. Dr. Mohamed Suleiman
Numerical Solution of Ordinary Differential Equations. A Historical Perspective
11 December 1993
5. Prof. Dr. Mohd. Ariff Hussein
Changing Roles of Agricultural Economics
5 March 1994
6. Prof. Dr. Mohd. Ismail Ahmad
Marketing Management: Prospects and Challenges for Agricultural
6 April 1994
7. Prof. Dr. Mohamed Mahyuddin Mohd. Dahan
The Changing Demand for Livestock Products
20 April 1994
8. Prof. Dr. Ruth Kiew
Plant Taxonomy, Biodiversity and Conservation
11 May 1994

Towards Large Scale Unconstrained Optimization

9. Prof. Ir. Dr. Mohd. Zohadie Bardaie
Engineering Technology Developments Propelling Agriculture into the 21st Century
28 May 1994

10. Prof. Dr. Shamsuddin Jusop
Rock, Mineral and Soil
18 June 1994

11. Prof. Dr. Abdul Salam Abdullah
Natural Toxicants Affecting Animal Health and Production
29 June 1994

12. Prof. Dr. Mohd. Yusof Hussein
Pest Control: A Challenge in Applied Ecology
9 July 1994

13. Prof. Dr. Kapt. Mohd. Ibrahim Haji Mohamed
Managing Challenges in Fisheries Development through Science and Technology
23 July 1994

14. Prof. Dr. Hj. Amat Juhari Moin
Sejarah Keagungan Bahasa Melayu
6 Ogos 1994

15. Prof. Dr. Law Ah Theem
Oil Pollution in the Malaysian Seas
24 September 1994

16. Prof. Dr. Md. Nordin Hj. Lajis
Fine Chemicals from Biological Resources: The Wealth from Nature
21 January 1995

17. Prof. Dr. Sheikh Omar Abdul Rahman
Health, Disease and Death in Creatures Great and Small
25 February 1995
18. Prof. Dr. Mohamed Shariff Mohammed Din
Fish Health: An Odyssey through the Asia - Pacific Region
25 March 1995
19. Prof. Dr. Tengku Azmi Tengku Ibrahim
Chromosome Distribution and Production Performance of Water Buffaloes
6 May 1995
20. Prof. Dr. Abdul Hamid Mahmood
Bahasa Melayu sebagai Bahasa Ilmu- Cabaran dan Harapan
10 Jun 1995
21. Prof. Dr. Rahim Md. Sail
Extension Education for Industrialising Malaysia: Trends, Priorities and Emerging Issues
22 July 1995
22. Prof. Dr. Nik Muhammad Nik Abd. Majid
The Diminishing Tropical Rain Forest: Causes, Symptoms and Cure
19 August 1995
23. Prof. Dr. Ang Kok Jee
The Evolution of an Environmentally Friendly Hatchery Technology for Udang Galah, the King of Freshwater Prawns and a Glimpse into the Future of Aquaculture in the 21st Century
14 October 1995

Towards Large Scale Unconstrained Optimization

24. Prof. Dr. Sharifuddin Haji Abdul Hamid
Management of Highly Weathered Acid Soils for Sustainable Crop Production
28 October 1995

25. Prof. Dr. Yu Swee Yean
Fish Processing and Preservation: Recent Advances and Future Directions
9 December 1995

26. Prof. Dr. Rosli Mohamad
Pesticide Usage: Concern and Options
10 February 1996

27. Prof. Dr. Mohamed Ismail Abdul Karim
Microbial Fermentation and Utilization of Agricultural Bioresources and Wastes in Malaysia
2 March 1996

28. Prof. Dr. Wan Sulaiman Wan Harun
Soil Physics: From Glass Beads to Precision Agriculture
16 March 1996

29. Prof. Dr. Abdul Aziz Abdul Rahman
Sustained Growth and Sustainable Development: Is there a Trade-Off 1 or Malaysia
13 April 1996

30. Prof. Dr. Chew Tek Ann
Sharecropping in Perfectly Competitive Markets: A Contradiction in Terms
27 April 1996

31. Prof. Dr. Mohd. Yusuf Sulaiman
Back to the Future with the Sun
18 May 1996
32. Prof. Dr. Abu Bakar Salleh
Enzyme Technology: The Basis for Biotechnological Development
8 June 1996
33. Prof. Dr. Kamel Ariffin Mohd. Atan
The Fascinating Numbers
28 June 1996
34. Prof. Dr. Ho Yin Wan
Fungi, Friends or Foes
27 July 1996
35. Prof. Dr. Tan Soon Guan
*Genetic Diversity of Some Southeast Asian Animals: Of Buffaloes
and Goats and Fishes Too*
10 August 1996
36. Prof. Dr. Nazaruddin Mohd. Jali
Will Rural Sociology Remain Relevant in the 21st. Century
21 September 1996
37. Prof. Dr. Abdul Rani Bahaman
*Leptospirosis-A Model for Epidemiology, Diagnosis and Control
Infectious Diseases*
16 November 1996
38. Prof. Dr. Marziah Mahmood
Plant Biotechnology - Strategies for Commercialization
21 December 1996

Towards Large Scale Unconstrained Optimization

39. Prof. Dr. Ishakl Hj. Omar
Market Relationships in the Malaysian Fish Trade; Theory and Application
22 March 1997
40. Prof. Dr. Suhaila Mohamad
Food and Its Healing Power
12 April 1997
41. Prof. Dr. Malay Raj Mukerjee
A Distrobuted Collaborative Environment for Distance Learning Applications
17 June 1998
42. Prof. Dr. Wong Kai Choo
Advancing the Fruit Industry in Malaysia: A Need to Shift Research Emphasis
15 May 1999
43. Prof. Dr. Aini Ideris
Avian Respiratory and Immunosuppressive Diseases- A Fatal Attraction
10 July 1999
44. Prof. Dr. Sariah Meon
Biological Control of Plant Pathogens: Harnessing the Rich 14 August 1999
45. Prof. Dr. Azizah hashim
The Endomycorrhiza: A Futile Investment?
23 Oktober 1999

46. Prof. Dr. Noraini Abdul Samad
Molecular Plant Virology: The Way Forward
2 February 2000
47. Prof. Dr. Muhamad Awang
Do We have Enough Clean Air to Breathe?
7 April 2000
48. Prof. Dr. Lee Chnoong Kheng
Green Environment, Clean Power
24 June 2000
49. Prof. Dr. Mohd. Ghazali Mohayidin
Managing Change in the Agriculture Sector: The Need for Innovative Educational Initiatives
12 January 2002
50. Prof. Dr. Fatimah Mohd. Arshad
Analisis Pemasaran Pertanian di Malaysia: Keperluan Agenda Pembaharuan
26 Januari 2002
51. Prof. Dr. Nik Mustapha R. Abdullah
Fisheries Co-Management: An Institutional Innovation Towards Sustainable Fisheries Industry
28 February 2002
52. Prof. Dr. Gulam Rusul Rahmat Ali
Food Safety: Perspectives and Challenges
23 March 2002

Towards Large Scale Unconstrained Optimization

53. Prof. Dr. Zaharah A. Rahman
Nutrient Management Strategies for Sustainable Crop Production in Acid Soils: The Role of Research using Isotopes
13 April 2002
54. Prof. Dr. Maisom Abdullah
Productivity Driven Growth: Problems & Possibilities
27 April 2002
55. Prof. Dr. Wan Omar Abdullah
Immunodiagnosis and Vaccination for Brugian Filariasis: Direct Rewards from Research Investments
6 June 2002
56. Prof. Dr. Syed Tajuddin Syed Hassan
Agro-ento Bioinformation: Towards the Edge of Reality
22 June 2002
57. Prof. Dr. Dahlan Ismail
Sustainability of Tropical Animal-Agricultural Production Systems: Integration of Dynamic Complex Systems
27 June 2002
58. Prof. Dr. Ahmad Zubaidi Baharumshah
The Economics of Exchange Rates in the East Asian Countries
26 October 2002
59. Prof. Dr. Shaik Md. Noor Alam S.M. Hussain
Contractual Justice in Asean: A Comparative View of Coercion
31 October 2002

60. Prof. Dr. Wan Md. Zin Wan Yunus
Chemical Modification of Polymers: Current and Future Routes for Synthesizing New Polymeric Compounds
9 November 2002
61. Prof. Dr. Annuar Md. Nassir
Is the KLSE Efficient: Efficient Market Hypothesis vs Behavioural Finance
23 November 2002
62. Prof. Ir. Dr. Radin Umar Radin Sohadi
Road Safety Interventions in Malaysia: How Effective Are They?
21 February 2003
63. Prof. Dr. Shamsher Mohamad
The New Shares Market: Regulatory Intervention, Forecast Errors and Challenges
26 April 2003
64. Prof. Dr. Han Chun Kwong
Blueprint for Transformation or Business as Usual? A Structural Perspective of the Knowledge-Based Economy in Malaysia
31 May 2003
65. Prof. Dr. Mawardi Rahmani
Chemical Diversity of Malaysian Flora: Potential Source of Rich Therapeutic Chemicals
26 July 2003
66. Prof. Dr. Dr. Fatimah Md. Yusoff
An Ecological Approach: A Viable Option for Aquaculture Industry in Malaysia
9 August 2003

Towards Large Scale Unconstrained Optimization

67. Prof. Dr. Mohamed Ali Rajion
The Essential Fatty Acids-Revisited
23 August 2003

68. Prof. Dr. Azhar Md. Zain
Psychotheraphy for Rural Malays - Does it Work?
23 August 2003

69. Prof. Dr. Azhar Md Zin
Respiratory Tract Infection: Establishment and Control
27 September 2003

70. Prof. Dr. Jinap Selamat
Cocoa-Wonders for Chocolate Lovers
14 February 2004

71. Prof. Dr. Abdul Halim Shaari
High Temperature Superconductivity: Puzzle & Promises
13 March 2004

72. Prof. Dr. Yaakob Che Man
Oils and Fats Analysis - Recent Advances and Future prospects
27 March 2004

73. Prof. Dr. Kaida Khalid
Microwave Aquametry: A Growing Technology
24 April 2004

74. Prof. Dr. Hasanah Mohd. Ghazali
Tapping the Power of Enzymes- Greening the Food Industry
11 May 2004

75. Prof. Dr. Yusof Ibrahim
The Spider Mite Saga: Quest for Biorational Management Strategies
22 May 2004
76. Prof. Datin Dr. Sharifah Md. Nor
The Education of At-Risk Children: The Challenges Ahead
26 June 2004
77. Prof. Dr. Ir. Wan Ishak Wan Ismail
Agricultural Robot: A New Technology Development for Agro-Based Industry
14 August 2004
78. Prof. Dr. Ahmad Said Sajap
Insect Diseases: Resources for Biopesticide Development
28 August 2004
79. Prof. Dr. Aminah Ahmad
The Interface of Work and Family Roles: A Quest for Balanced Lives
11 March 2005
80. Prof. Dr. Abdul Razak Alimon
Challenges in Feeding Livestock: From Wastes to Feed
23 April 2005
81. Prof. Dr. Haji Azimi Hj. Hamzah
Helping Malaysian Youth Move Forward: Unleashing the Prime Enablers
29 April 2005
82. Prof. Dr. Rasedee Abdullah
In Search of An Early Indicator of Kidney Disease
17 June 2005

Towards Large Scale Unconstrained Optimization

83. Prof. Dr. Zulkifli Hj. Shamsuddin
Smart Partnership Plant-Rhizobacteria Associates
17 June 2005
84. Prof. Dr. Mohd. Khanif Yusop
From the Soil to the Table
1 July 2005
85. Prof. Dr. Annuar Kassim
Materials Science and Technology: Past, Present and the Future
8 July 2005
86. Prof. Dr. Othman Mohamed
Enhancing Career Development Counselling and the Beauty of Career Games
12 August 2005
87. Prof. Dr. Ir. Mohd. Amin Mohd. Soom
Engineering Agricultural Water Management Towards Precision Farming
26 August 2005
88. Prof. Dr. Mohd. Arif Syed
Bioremediation-A Hope Yet for the Environment?
9 September 2005
89. Prof. Dr. Abdul Hamid Abdul Rashid
The Wonder of Our Neuromotor System and the Technological Challenges They Pose
23 December 2005

90. Prof. Dr. Norhani Abdullah
Rumen Microbes and Some of Their Biotechnological Application
27 January 2006
91. Prof. Dr. Abdul Aziz Saharee
Haemorrhagic Septicaemia in Cattle and Buffaloes: Are We Ready for Freedom?
92. Prof. Dr. Kamariah Abu Bakar
Activating Teachers' Knowledge and Lifelong Journey in Their Professional Development
3 March 2006
93. Prof. Dr. Borhanuddin Mohd. Ali
Internet Unwired
24 March 2006
94. Prof. Dr. Sundararajan Thilagar
Development and Innovation in the Fracture Management of Animals
31 March 2005
95. Prof. Dr. Zainal Aznam Md. Jelani
Strategic Feeding for a Sustainable Ruminant Farming
19 May 2006