REM WORKING PAPER SERIES

NEW APPROACH TO THE OPTIMIZATION PROBLEM

João Ferreira do Amaral

REM Working Paper 0171-2021

April 2021

REM – Research in Economics and Mathematics
Rua Miguel Lúpi 20,
1249-078 Lisboa,
Portugal

ISSN 2184-108X

REM – Research in Economics and Mathematics

Rua Miguel Lupi, 20
1249-078 LISBOA
Portugal

Telephone: +351 - 213 925 912
E-mail: rem@iseg.ulisboa.pt

https://rem.rc.iseg.ulisboa.pt/

A NEW APPROACH TO THE OPTIMIZATION PROBLEM

João Ferreira do Amaral[1]

Abstract

A new approach to the problem of optimization is developed using tools such as the concepts of aggregate and of combined functions. The solving of a simple problem of calculus of variations with inequality constraints illustrates the potentiality of this new method.

Keywords: optimization, calculus of variations, convexity, quasi-convexity

JEL classification: C61, C65

I Introduction

The problem of optimization that is described as the maximization of a functional defined on the space of real functions of real variables is often too simplified in its formulation.

The usual presentation of the problem includes a functional J : U→R where (U, ‖.‖) is the Banach space of real functions and a set V ⊂ U (Céa, 1971 p. 61).

The problem is presented as

Find sup J(u) for all the $u \in V \subset U$

This formulation may be enriched in order to contemplate, for example, the situation where the set V is determined by inequality constraints.

This paper presents a new formulation that for a particular situation (maximization and quasi-convex functionals) describes a simple process in order to solve a problem of calculus of variations with inequality constraints.

To develop the solution of the problem we need some auxiliary concepts. Actually it would be fair to say that the main intention of the paper is to develop some new (in the

[1] ISEG – Lisbon School of Economics and Management, Universidade de Lisboa; REM – Research in Economics and Mathematics. UECE – Research Unit on Complexity and Economics. e-mail: joaomfamaral@hotmail.com.
UECE/REM - ISEG, Universidade de Lisboa is financially supported by FCT (Fundação para a Ciência e a Tecnologia), Portugal. This article is part of the Strategic Project UIDB/05069/2020.

present context) concepts that can be useful for solving a maximization problems. This means that the development of the concepts takes a large proportion of the paper and that is why some patience is required from the reader before arriving to the real solving of problems. However we hope that the development of the concepts will be interesting in its own right.

Section II is devoted to the presentation of those auxiliary concepts. Section III solves a general problem of maximization and section IV is devoted to the solution of a specific problem of calculus of variations with inequality constraints.

II Preliminary concepts

1 Atomizable and non-atomizable functions

1.1 Atomizable functions

Let X be the space of real functions of real variable defined on a set A of real numbers and consider the set L of functions $F: X \to X$.

For each x of X let $y = F(x)$ be the value of F at x and consider the following function F^* $F^*: F(X) \times A \to R$ that to each y of $F(X)$ and each t of A associates the real number $y(t)$. We represent this number with the symbols $y(t) \equiv F^*(x, t) \equiv F(x)(t)$.

Before defining the concept of atomizable function we define the following relation between functions:

Definition 1 *(Equivalence). Let g:* $R^2 \to R$ *be a real function defined on* R^2*. We say that* $F \in L$ *and* $g \in C$ *where C is the set of real functions of* R^2 *are equivalent on A if* $F(x)(t) = g(x(t), t)$ *for each x of X defined on A and all the t of A.*

Remark 1. The function may depend on other functions y if these don't depend on the variable x, for instance if F is such that $F(x) = u.x^2$ where u is the function of L such that $u(t) = t$ for all t . For each t the value of the equivalent function in this case is g such that $g(x(t), t) = tx(t)^2$.

Remark 2. Obviously if the function F has an equivalent function on A this is unique.

We can now define atomizable function in the strong sense.

Definition 2 *(Atomizable function in the strong sense). A function F such that it has an equivalent function g is an atomizable function in the strong sense*.

This concept is a particular case of the concept we used in a previous paper (Amaral, 2007) that we called atomizable but that we now define as atomizable in the weak sense:

Definition 3 *(Atomizable function in the weak sense). F is atomizable in the weak sense if and only if* $F(x)(t^*) = F(z_{x(t^*)})(t^*)$ *for all the* $z_{x(t^*)}$ *such that* $z_{x(t^*)}(t^*) = x(t^*)$.

It is easy to see that an atomizable in the strong sense is atomizable in the weak sense:

Suppose that $F(x)$ is atomizable in the strong sense. It has an equivalent function, that is for each t^* of A we have $F(x)(t^*) = g(x(t^*), t^*) = g(z_{x(t^*)}(t^*), t^*)$ for all the $z_{x(t^*)}$ such that $z_{x(t^*)}(t^*) = x(t^*)$.

But as $F(z_{x(t^*)})(t^*) = g(z_{x(t^*)}(t^*), t^*)$ because F it is atomizable in the strong sense we have

$F(z_{x(t^*)})(t^*) = F(x)(t^*)$ for all the $z_{x(t^*)}$ such that $z_{x(t^*)}(t^*) = x(t^*)$ and F is atomizable in the weak sense.

However we can prove the equivalence of the two concepts, if the following axiom is valid.

Axiom *For each* $y = F(x)$ *and each t, the value y(t) depends strictly on the value of t and on the values of the function x at numbers t* of A, the same numbers for all the x*.

Theorem 1. *If the axiom is valid both concepts of atomizable functions coincide*.

Proof

We need to prove that if the axiom is valid a function atomizable in the weak sense is also atomizable in the strong sense.

Suppose that $F(x)(t^*) = F(z_{x(t^*)})(t^*)$ with $x(t^*) = z_{x(t^*)}(t^*)$ and that there exists a $t^{**} \neq t^*$ belonging to A such that

$$F(x)(t^*) = \theta(\{x(t^*), x(t^{**})\}, t^*) = F(z_{x(t^*)})(t^*) = \theta(\{x(t^*), z_{x(t^*)}(t^{**})\}, t^*)$$

But as the equality $F(x)(t^*) = F(z_{x(t^*)})(t^*)$ is valid for all the $z_{x(t^*)}$ such that $z_{x(t^*)} = x(t^*)$ then for $z = z_{x(t^*)} + h$ with any h such that $h(t) \neq 0$ for $t \neq t^*$ and $h(t^*) = 0$ we have,

$$F(z_{x(t^*)})(t^*) = \theta(\{(x(t^*), z_{x(t^*)}(t^{**}), t^*\}) = F(z)(t^*) = \theta(\{(x(t^*), z(t^{**}), t^*\})$$

As $z(t^{**}) \neq z_{x(t^*)}(t^{**})$, for all the $t^{**} \neq t^*$, this is not possible so that $F(z_{x(t^*)})(t^*)$ and consequently $F(x)(t^*)$ don't depend on any other value $x(t^{**})$ different from $x(t^*)$. Therefore $F(x)(t) = g(x(t), t)$. □

1.2 Non-atomizable functions

Non-atomizable functions and aggregates

Definition 4 *(Non-atomizable functions). Non-atomizable functions are the functions that verify the axiom and are not atomizable.*

Two important species of non-atomizable functions are respectively those that are based on a correspondence of sets and those based on correspondence of aggregates.

It is now necessary to define the concept of aggregate (later on in this section, p.16, more on aggregates). Given a function f defined on A we represent the aggregate of A under f as $f^*(A)$.

Consider a set A of real numbers t and let f be a real function that to each t of A associates an element $f(t)$ of the set R.

Definition 5 (A*ggregate). The aggregate $f^*(A)$ is the collection of elements $f(t)$, such that for this collection , if $f(t) = f(t^*)$ (according to the relation of equality of real numbers) with $t \neq t^*$, $f(t)$ and $f(t^*)$ are considered as distinct elements of the collection. For this reason an aggregate is not a set.*

Later on we will see examples of relations and operations on aggregates. For the time being it is sufficient to define the relation of belonging and inclusion.

Definition 6 *(Belonging). The element x belongs to the aggregate $f^*(A)$, $x \in f^*(A)$, if and only if there exists one and only one t of A such that $x = f(t)$.*

Definition 7 *(Inclusion). $f^*(C) \subset f^*(D)$ if and only if for every x belonging to $f^*(C)$, x belongs to $f^*(D)$.*

Remark 1. We use the same symbols of the relations of sets although the concept of aggregate and set do not coincide.

Remark 2. Obviously $f^*(C) \subset f^*(D)$ if and only if $C \subset D$, but for image-sets the double inclusion is not true. It is true that $C \subset D$ implies $f(C) \subset f(D)$, but the reciprocal is not always true.

With these concepts we may define disjoint aggregates, that is aggregates that have no common element (in the sense of element of an aggregate).

Obviously if C and D are disjoint sets so are the aggregates $f^*(C)$ and $f^*(D)$ and the reciprocal is also true (again, this is not always true for image sets).

Set functions, aggregate functions, mixed functions and non-atomizable functions

We may define real aggregate functions in a similar way as real set functions, that is as functions μ

$\mu: 2^{x^*(A)} \rightarrow R$ taking real values $\mu[x^*(C)]$.

If we have a function F defined on a set X of real functions of real variable and taking values in X that is

$F: \quad x \in X \rightarrow y \in X$

we may define the *aggregate function* μ with real values $\mu[F(x)^*(C)]$.

A more general concept is the concept of *mixed set/aggregate functions* that is X

$\mu: 2^{F(x)^*(A)} \times 2^A \rightarrow R$

with real values

$\mu[F(x)^*(C), D]$ with $C, D \in 2^A$.

Especially important is the particular case $C = D$.

Finally we can define mixed *aggregate/point functions* as

$\mu: 2^{F(x)^*(A)}\ x\ A \rightarrow R$ with real values $\mu[F(x)^*(C), t]$.

This allows us to define non-atomizable functions in terms of aggregates.

Definition 8 (*Non-atomizable functions in terms of aggregates). A non-atomizable function F : F(x) X A →R in terms of aggregates is a mixed aggregate/point function such that for each t of A we have*

$F(x)(t) = \mu[x^*(B(t)), t]$ *with* $t \in B(t)$, $B(t) \in 2^A$ *the same for every x*

The function μ is

$\mu: 2^{x^*(A)} \times A \rightarrow R$ and it is a mixed aggregate/point function defined for each $x^*(B(t))$ and each t of A.

Note that μ is based on a correspondence $t \rightarrow B(t)$, $B(t) \in 2^A$.

Definition 9 *(Non-atomizable functions in terms of sets). The definition is the same as Definition 8 replacing aggregate by set.*

These two definitions verify the conditions of the axiom on page 3 and the difference of both definitions from atomizable functions with values $F(x)(t) = g(x(t),t)$ is readily seen since, for a non-atomizable function, for each t it is determinant the set $x(B(t))$ or the aggregate $x^*(B(t))$ whereas for an atomizable function is determinant only the number $x(t)$. This of course is the justification for the name "atomizable functions". These functions, for each t have their values determined by each "atom" $x(t)$, something that is not the case for non-atomizable functions.

These are the basic cases of non-atomizable functions. More complex cases are those where we have the real values

$F(x)(t) = \mu[x^*(B(t)), x(B(t)), t]$

However these cases are not met again in this paper.

A definition that may be useful in some applications is the definition of continuity.

Definition 10 *(Continuity). A non-atomizable function F in terms of sets (the same for aggregates) such that* $F(x)(t) = \mu[x(B(t)),t]$ *is continuous in the non-empty and closed set A if and only if it is defined on all elements of A and*

$\lim_{t\rightarrow a} \mu[x(B(t)),t] = \mu[x(B(a)),a]$ *for each* $a \in A$.

We have the following theorem for non-atomizable functions in terms of sets.

Theorem 2. *Let x be continuous in A closed and non-empty. It is necessary for F to be continuous that the correspondence* $A \rightarrow 2^{x(A)}$ *is upper semi-continuous at each t of A.*

Proof

Suppose that $A \rightarrow 2^{x(A)}$ was not upper semi-continuous. Then we would have an $a^* \in A$ such that for a sequence $\{x(t_n)\}$ with $x(t_n) \in x(B(t_n))$ and $\lim x(t_n)_{\, t_n \rightarrow a^*} = x(a^*)$, $x(a^*) \notin x(B(a^*))$ so that $a^* \notin B(a^*)$. Then $F(x)$ would not be continuous in A because $\mu[x(B(t)),t]$ would not be defined for $t = a^*$.□

Remark. The theorem may be generalized to aggregates if the concept of upper semi-continuity is defined for correspondences of aggregates.

We will come back to aggregates later on in section. But before that it is necessary to define concepts of quasi-convexity and convexity.

2 Quasi-convexity and convexity

2.1 Atomizable functions

We begin by considering atomizable functions of two variables (x,y) both being real functions of real variable t.

Definition 11 *(Quasi-convexity). Let* C^2 *be the set of all the pairs of functions (u,v) that are convex linear combinations* $u = \lambda x_1 +(1- \lambda)x_2$ $v = \lambda y_1 +(1- \lambda)y_2$ *for all the* λ *of C such that* $0 \leq \lambda \leq 1$ *(that is* $0 \leq \lambda(t) \leq 1$ *for all the t of A), and all the* x_1, x_2, y_1, y_1 *belonging to C. A function* $f: C^2 \subset X^2 \rightarrow X$ *is quasi-convex in* C^2 *if and only if for each 4-uple* x_1, x_2, y_1 *e* y_2 *of functions of C, with* $x = \lambda x_1 +(1- \lambda)x_2$, $y = \lambda y_1 +(1- \lambda)y_2$ *and for all functions* λ *of X such that* $0 \leq \lambda \leq 1$ *we have*

$$f(x(t), y(t)) \leq \max \{f(x_1(t),y_1(t)), f(x_2(t),y_2(t))\}, \text{ for each } t \text{ de } A.$$

Remark 1. The function *v* is not necessarily the derivative of *u* although in a lot of cases this does happen.

Remark 2. A set that verifies the properties of C^2 (or for any other number of variables) is called a strong convex set or *s-convex set*. We could call it a convex set, as a generalization of the common concept of convex set. However most of the interesting properties of convex sets (for example those related to the concept of segment) do not apply to s-convex sets. That is why we prefer to designate them by a different name. Of course a s-convex set is a convex set.

Definition 11.a *(Convexity). A function $f\colon C^2 \rightarrow R$ defined on the s-convex set C^2 is convex if and only if under the same conditions of the previous definition we have*

$f(x(t),y(t)) \le \lambda(t)\, f(x_1(t),y_1(t))+(1-\lambda(t))\, f(x_2(t),y_2(t))$ *for each t of A.*

Remark 1. The usual definition of quasi-convex or convex function is a particular case of the antecedent when the $\lambda(t)$ are constant functions of t.

Remark 2. If a function with values $f(x,y)$ is quasi-convex or convex this does not mean that the corresponding composed function $F\colon R \rightarrow R$ with $F(t) \equiv f(x(t), y(t))$ is quasi-convex or convex.

For example, the function $f(x,y) \equiv x^2 + x'^2$ (where x' is the derivative of x) is convex but with $x(t) \equiv \log t,\ t > 0,\ F(t)=(\log t)^2+(1/t)^2$ F is not convex for sufficient large values of t.

We can prove the following theorem.

Theorem 3. *Let f be non-decreasing with x and non-decreasing with y in a s-convex set. If $x(t)$ e $y(t)$ are non-decreasing real functions of t defined on an interval A, $F(t)$ is quasi-convex.*

Remark. $f(x,y)$ is non-decreasing with x if and only if for all the $h \ge 0$ (that is such that $h(t) \ge 0$ for all the t of A) we have $f(x+h, y) \ge f(x,y)$

(that is, $f(x(t)+h(t), y(t)) \ge f(x(t), y(t))$ for all the t of A).

Proof

Consider t_1, t_2 , $t_1 < t_2$ and one t belonging to $[t_1, t_2]$.

Since x and y are non-decreasing we have $x(t) \le x(t_2)$ and $y(t) \le y(t_2)$ so that

$x \le x(t_2)$ and $y \le y(t_2)$ where $x(t_2)$ and $y(t_2)$ are constant functions, that is $x(t_2)(t) = x(t_2)$ for all the t.

As $f(x, y)$ is non-decreasing in x and y we have for every t

$f(x(t), y(t)) \leq f(x(t_2)(t), y(t_2)(t)) = f(x(t_2), y(t_2))$

That is, for every t of A, $F(t) \leq F(t_2) = max\{F(t_1), F(t_2)\}$ and $F(t)$ is quasi-convex.□

We have the following theorem concerning convex functions.

Theorem 4. *If f is convex and non-decreasing in x and y in a s-convex set and if x(t) and y(t) are convex functions, F(t) is a convex function in the interval A.*

Proof

Consider t_0 and t_1, a real number μ, $0 \leq \mu \leq 1$ and the corresponding t^* , $t^* = \mu t_0 + (1-\mu)t_1$

As $x(t)$ and $y(t)$ are convex in A we have

$x(t^*) \leq \mu x(t_0) + (1-\mu)x(t_1)$

$y(t^*) \leq \mu y(t_0) + (1-\mu)y(t_1)$

Therefore given the fact that f is non-decreasing in x and y we have for the following constant functions μ, $x(t_0)$, $x(t_1)$, $y(t_0)$ and $y(t_1)$,

$f(x(t^*),y(t^*)) \leq f(\mu x(t_0)(t^*) + (1-\mu)x(t_1)(t^*), \mu y(t_0)(t^*) + (1-\mu)y(t_1)(t^*))= f(\mu x(t_0) + (1-\mu)x(t_1), \mu y(t_0) + (1-\mu)y(t_1))$

Due to the convexity of f we have

$f(x(t^*),y(t^*)) \leq f(\mu x(t_0) + (1-\mu)x(t_1), \mu y(t_0) + (1-\mu)y(t_1)) \leq \mu f(x(t_0), y(t_0)) + (1-\mu) f(x(t_1), y(t_1))$

That is $F(t^*) \leq \mu F(t_0)+(1-\mu)F(t_1)$.□

Remark. In the particular case where $y = \varphi(x)$, is important not to confound the quasi-convexity of f that is defined by

$f(\lambda x_1 +(1-\lambda)x_2, \lambda\varphi (x_1) +(1-\lambda)\varphi(x_2)) \leq max \{f(x_1, \varphi (x_1)), f(x_2, \varphi (x_2))\}$

with the relation $f(\lambda x_1 +(1-\lambda)x_2, \varphi(\lambda x_1 +(1-\lambda)x_2)) \leq max\{f(x_1, \varphi (x_1), f(x_2, \varphi (x_2)\}$ which is something different.

Related to this we have the following theorem.

Theorem 5. *If f is quasi-convex and non-decreasing in* $\varphi(x)$*, where* φ *is convex we have* $f(x, \varphi(x)) \leq \max\{f(x_1, \varphi(x_1)), f(x_2, \varphi(x_2))\}$.

Proof

As φ is convex we have $\varphi(x) \leq \lambda \varphi(x_1) + (1-\lambda)\varphi(x_2)$ and since f is non-decreasing in φ we have for x_1 e x_2 and any $x = \lambda x_1 + (1-\lambda)x_2$,

$$f(x, \varphi(x)) \leq f(\lambda x_1 + (1-\lambda)x_2, \lambda \varphi(x_1) + (1-\lambda)\varphi(x_2))$$

But as f is quasi-convex

$$f(\lambda x_1 + (1-\lambda)x_2, \lambda\varphi(x_1) + (1-\lambda)\varphi(x_2)) \leq \max\{f(x_1, \varphi(x_1)), f(x_2, \varphi(x_2))\}$$

so that

$$f(x, \varphi(x)) \leq \max\{f(x_1, \varphi(x_1)), f(x_2, \varphi(x_2))\}. \square$$

Later on (section III) we consider a specific case where φ is not necessarily convex but where

$f(x, \varphi(x)) \leq f(x_1, \varphi(x_1)), f(x_2, \varphi(x_2))$ is still valid.

Let us look now to non-atomizable functions.

2.2 Quasi-convex non-atomizable functions in one variable

We define the concept for aggregates (for sets the definition would be analogous)

Definition 12 *(Quasi-convexity). Let* $F(x)(t) = \mu[x^*(B(t)),t]$ *be a non-atomizable function in terms of aggregates. Let* x_1*,* x_2 *be real functions belonging to a s-convex set C and* λ $(0 \leq \lambda \leq 1)$ *with* $x = \lambda x_1 (1-\lambda)x_2$ *.* $F(x)$ *defined on C is quasi-convex if and only if for each t of A*

$$F(x)(t) \leq \max\{F(x_1)(t), F(x_2)(t)\} \text{ that is}$$

$$\mu[x^*(B(t)),t] \leq \max\{\mu[x_1^*(B(t)),t], \mu[x_2^*(B(t)),t]\}.$$

This is the essential of quasi-convexity and convexity that is needed for the study of maximization in the context we have determined in the Introduction.

We introduce now another concept.

3 Combined functions-n (CF-n)

3.1.Definition. Equivalence 1 and 2

Let X be a set of real functions of a real variable and B a closed interval $B = [t_0\ t_1]$. Let $\{A_i\},\ i=1,\ldots n$ be a partition of B in n disjoint sets.

Definition 13 *(Combined functions-n and generated set).*

a) z is a CF-n (combined function-n) in A if and only if for a partition $\{A_i\}$*, and for each t of* A_i *we have* $z(t) = x_i(t)$*, where* x_i *are functions belonging to X*

b) The set of all the z is the set generated by X and it is represented by X(X)*

c) the natural number n is the degree of the CF.

It is possible to define a wide set of properties and operations for CF-n. For our purposes in this paper it is sufficient to develop the case CF-2, that is functions z defined on A such that for a partition $A = A_1 \cup A_2$ with A_1 and A_2 disjoint sets there exist x_1, x_2 of X such that $z(t) = x_1(t)$ for all the t of A_1 and $z(t) = x_2(t)$ for all the t of A_2.

We use the following evident notation to represent functions CF-2

$z \equiv (x_1, A_1, x_2, A_2)$

Remark. Until otherwise stated both the two sets of the partition are considered non-empty sets.

Definition 14 *(Equivalence-1, equivalence-2 and equality).*

The CF-2 function z is equivalent-1 to the function* $z \equiv (x_1, A_1, x_2, A_2)$ *if there is a partition* C_1, C_2 *of A such that*

$z^* \equiv (x_1, C_1, x_2, C_2)$.

The CF-2 function z is equivalent -2 to z if and only if*

$z \equiv (x_1, A_1, x_2, A_2)$

$z^* \equiv (y_1, A_1, y_2, A_2)$.

It is to verify that each of these relations is a relation of equivalence.

Now the definition of *equality*

For z = (x_1, A_1, x_2, A_2) and $w = (y_1, C_1, y_2, C_2)$

$w = z$ if and only if $A_i = C_i$ and $x_i = y_i$ for $i = 1,2$.

Remark. Note that we consider as two different CF-2 the functions (x, A_1, x, A_2) and (x, C_1, x, C_2) if $A_i \neq C_i$.

Evidently two CF-2 are equal if and only if they are equivalent-1 and equivalent-2.

Let us develop these concepts.

3.2 Functions equivalent -1

Sum and product

When we have two equivalent-1 CF-2 functions defined in the same set A we may define a sum and a product of those functions.

Definition 15 *(Sum). Let z and z* be two functions equivalents-1 with z defined on the partition* $A = A_1 \cup A_2$ *and z* defined on a different partition* $A = C_1 \cup C_2$ *, with*

$A_i \neq C_{i\pm 1}$. *The sum* $z^{**}=z \,╬\, z^*$ *is the function z** such that*

$$z^{**} \equiv (x_1, (A_1 \cap C_1) \cup (A_2 \cap C_2),\ x_2, (A_1 \cap C_2) \cup (A_2 \cap C_1)).$$

Note that as can be easily checked with $A_i \neq C_{i\pm 1}$, $\{(A_1 \cap C_1) \cup (A_2 \cap C_2), (A_1 \cap C_2) \cup (A_2 \cap C_1)\}$ is a partition of A in two non-empty sets so that $z^{**}$is equivalent-1 to z and z^*.

It is also easy to see that the operation is commutative and associative.

The following operation could be designated by multiplication but can be reduced to the previous as we will see.

Definition 15.a *(Product). Let z and z* be two equivalent-1functions with z defined on* $A = A_1 \cup A_2$ *and z* defined on a different partition* $A = C_1 \cup C_2$. *The product* $z^{**}= z \bullet z^*$ *is defined as*

$$z^{**} \equiv (x_1, (A_1 \cup C_1) \cap (A_2 \cup C_2),\ x_2, (A_1 \cup C_2) \cap (A_2 \cup C_1))$$

Note that again $\{(A_1 \cup C_1) \cap (A_2 \cup C_2), \; (A_1 \cup C_2) \cap (A_2 \cup C_1)\}$ *is a partition of A , so that z** is an equivalent-1 to z and z*.*

Symmetric and null elements

The symmetric element of $z = (x_1, A_1, x_2, A_2)$ is represented by $-z$ such that

$-z = (x_1, A_2, x_2, A_1)$

Obviously $-(-z) = z$ and $-(z ╬ z^*) = z ╬ (-z^*)$

On the other hand we have $z \bullet z^* = -(z ╬ z^*)$ as the following identities show

$z \bullet z^* \equiv (x_1, (A_1 \cup C_1) \cap (A_2 \cup C_2), \; x_2, (A_1 \cup C_2) \cap (A_2 \cup C_1)) =$

$= (x_1, (A_1 \cap A_2) \cup (A_1 \cap C_2) \cup (C_1 \cap A_2) \cup (C_1 \cap C_2) , \; x_2 \, (A_1 \cap A_2) \cup (A_1 \cap C_1) \cup$

$C_2 \cap A_2) \cup (C_2 \cap C_1)) = (x_1, (A_1 \cap C_2) \cup (A_2 \cup C_1), \; x_2 \, (A_1 \cup C_1) \cap (A_2 \cup C_2))=$

$= \; -(z ╬ z^*).$

Therefore as mentioned above the product reduces to a sum, after we define symmetric elements.

On the other hand as $-(z ╬ z^*) = z ╬ (-z^*)$ we can also write

$z \bullet z^* = z ╬ (-z^*)$

and this allows us to eliminate the operation of multiplication.

Till now we considered only non-empty sets. It is time now to introduce the possibility of empty sets. This can be done through the definition of null elements.

Null elements

For each pair x_1, x_2 there are two null elements defined respectively by

$0 = (x_1, A, x_2, \emptyset)$

and

$-0 = (x_1, \emptyset, x_2, A)$

Note that this is just an operational definition since it makes no sense to say that a function of a point is defined on the empty set, but is a concept that has the advantage of extending the definition of the sum to the case where $A_i = C_{i\pm 1}$.

We have

$0 ╬ z = z$

$-0 ╬ z = -z$

$z ╬ z = 0$

$-z ╬ z = -0$

If we define for a natural number m

$mz \equiv z ╬ z ╬ ... ╬ z$ with m terms,

we have $mz = 0$ *if* m is even and $mz = z$ if m is odd and other similar results for -0.

Unit elements

It is important to define unitary elements. The definition is such that there is a unitary element for each element t de A.

We define the unitary element 1 for the value of t as

$1(t) = (x_1, A-\{t\}, x_2, \{t\})$

Consider

$z \equiv (x_1, A_1, x_2, A_2)$,

if $t \in A_1$ we have

$z ╬ 1(t) = (x_1, [(A_1 \cap (A-\{t\})] \cup [(A_2 \cap \{t\})], x_2, [(A_1 \cap \{t\}) \cup (A_2 \cap (A-\{t\})] = (x_1, A_1-\{t\}, x_2, A_2 \cup \{t\})$

and *if* $t \in A_2$,

$z ╬ 1(t) = (x_1, A_1 \cup \{t\}, x_2, A_2 - \{t\})$.

In each of the cases this operation transfers one element of A from one of the sets to the other set.

For each t we have

z ╬ $-1(t) = -$ $(z$ ╬ $1(t))$

Let us now look at equivalent-2 functions.

3.3 Operations with CF-2 equivalent-2.Integration and aggregates

Generally when the functions are defined on the same partition, $\{A_1, A_2\}$ of a given set A , that is when we have equivalent-2 functions, if f represents a function of functions, *op* represents a binary operation and if

$z_* \equiv (x_1, A_1, x_2, A_2)$

$z^{**} \equiv (y_1, A_1, y_2, A_2)$

we have respectively

$f(z^*) \equiv (f(x_1), A_1, f(x_2), A_2)$

$z \equiv op(z^*, z^{**}) = (op(x_1,y_1), A_1, op(x_2,y_2), A_2)$

In what concerns integration if x_1 and x_2, are integrable functions of t we define

$\int_A z^* \, dt \equiv \int_{A_1} x_1 \, dt + \int_{A_2} x_2 \, dt.$

Aggregates and CF-2

For aggregates if $z \equiv (x_1, A_1, x_2, A_2)$ we have obviously

$z^*(A) = x_1^*(A_1) \cup x_2^*(A_2)$ and

$f(z)^*(A) = f(x_1)^*(A_1) \cup f(x_2)^*(A_2)$

We have also the following theorem

Theorem 6. *Let X be a s-convex set of functions defined on a set A. Then for each partition* $\{A_i\}$ *the set* $X^*(\{A_i\})$ *of all the equivalent-2 CF-2 functions is s-convex.*

Proof

Let $x^* = (x_1, A_1, x_2, A_2)$ and $y^*= (y_1, A_1, y_2, A_2)$ belonging to $X^*(\{A_i\})$ and $z^*= \lambda x^* + (1-\lambda)y^*$. Then as we have seen regarding binary operations

$z^* =(\lambda x_1+(1-\lambda)y_1, A_1, \lambda x_2+(1-\lambda)y_2, A_2)$

As X is a s-convex set, $\lambda x_1+(1-\lambda)y_1$ and $\lambda x_2+(1-\lambda)y_2$ belong to X so *that* z^* belongs to $X^*(\{A_i\})$.□

Distance between equivalent-2 functions

Metric issues are important since the space of real functions with real variables is a rich metric space. Based on this fact we can define a distance for equivalent-2 functions.

Consider a partition of A, $\{A_1, A_2\}$ and suppose that for each pair (x,y) of functions of the set X we define two distances on each set of the partition, designated respectively by d_{A1} and d_{A2}. The values $d_{A1}(x,y)$ and $d_{A2}(x,y)$ are dependent, respectively on A_1 and A_2.

Definition 16 *(Distance). We may define for each* $A = A_1 \cup A_2$, *where none of the sets is empty, for each pair of elements of* X^* *and for equivalent-2 functions w e z with* $w= (w_1, A_1, w_2, A_2)$ *and* $z =(z_1, A_1, x_2, A_2)$, *the distance between w and z on the set A as*

$d_A(w,z) \equiv d_{A1}(w_1,z_1) + d_{A2}(w_2,z_2)$.

Remark 1. It is easy to verify that $d_A(w,z)$ is indeed a distance. Note also that $d_{Ai}(x,y)$ is not the restriction of d_A to a subspace since $d_{Ai}(x,y) = 0$ does not imply $d_A(x,y) = 0$. We say that d_{Ai} is a *contraction* of d_A and d_A an *expansion* of d_{Ai}.

Remark 2. We assume that if $A_i = \{t\}$, $d_{Ai}(w_i,z_i) = d^*(w_i(t),z_i(t))$ where d^* is a distance defined on R.

It is now time to return to the concept of aggregate for further developments.

4 Aggregates and sets

Recall the concept of aggregate.

Consider a set A of elements t and a function f that to each t of A associates a certain element of the image set $f(A)$.

The aggregate $f(A)^*$ is the collection of elements $f(t)$ for each t of A, such that if $f(t) = f(t^*)$ when $t \neq t^*$, $f(t)$ and $f(t^*)$ are considered distinct elements of the collection.

4.1 Homologous aggregates and sets

It is useful to avoid possible confusions when we use the concepts of aggregates and sets.

Let A and B be two sets and $x: A \to B$ a function for which we calculate for each $C \subset A$ the set $x(C) \subset x(B)$ and the aggregate $x^*(C) \subset x^*(B)$.

Definition 17 *(Homologous). $x^*(C)$ is homologous to the set $x(C)$ and we represent this fact by $x^*(C) \leftrightarrow x(C)$ if for any t and t^* of C such that $t \neq t^*$ we have $x(t) \neq x(t^*)$. We use the convention $x^*(\emptyset) \leftrightarrow x(\emptyset)$ and $x^*(\{u\}) \leftrightarrow x(\{u\})$ for all the u of A.*

Let $\boldsymbol{F}_x \subset 2^A$ be the family of all the sets $C \subset A$ such that $x^*(C) \leftrightarrow x(C)$.

We have the following simple theorem

Theorem 7. *If $C \in \boldsymbol{F}_x$ then for any $C_1 \subset C$ we have $C_1 \in F_x$.*

and the corollary

Corollary *If $C \in \boldsymbol{F}_x$, then for any D of 2^A, $C \cap D$ and $C - D \in F_x$.*

Remark . Obviously there is always a function $H : 2^{x^*(A)} \to 2^{x(A)}$ such that $H(x^*(C)) = x(C)$ for each C of 2^A. But there is no inverse function unless the respective correspondence is restricted to the family $\boldsymbol{F}_x$.

The following theorem applies to real functions.

Theorem 8. *If x is a continuous real function of real variable defined on the closed interval A we have $x^*(A) \leftrightarrow x(A)$ if and only if x is strictly monotonous (increasing or decreasing).*

Proof

Suppose that $x^*(A) \leftrightarrow x(A)$.

Then if $t \neq t^*$, $x(t) \neq x(t^*)$. Suppose that $t < t^*$ and $x(t) < x(t^*)$ and that there were t^{**} and t^{***} with $t^{**} < t^{***}$ such that $x(t^{**}) > x(t^{***})$.

Consider the case $t < t^* < t^{**} < t^{***}$. If $x(t^{**}) > x(t^*)$ then $x(t^*) < x(t^{**})$ and $x(t^{***}) < x(t^{**})$ so that given the continuity of $x(t)$, for $\alpha > 0$ such that $x(t^{**}) - x(t^*) > \alpha$ and

$x(t^{**}) - x(t^{***}) > \alpha$ there would exist t^+ and t^{++}, $t^+ \neq t^{++}$ such that $x(t^+) = x(t^{++}) = x(t^{**}) - \alpha$, something that is not possible. The same argument with the necessary adaptations for $x(t^{**}) < x(t^*)$ and for any other case since there will always be t^+ that is not a extreme of the interval and a $\alpha > 0$ such that for two t^{++} and t^{+++}, $t^{++} \neq t^{+++}$ we have $x(t^{++}) = x(t^{+++}) = x(t^+) - \alpha$.

The reciprocal is obvious. Note that for the reciprocal to be true it is not necessary to suppose that x is continuous.□

This theorem is important because shows that this relation between homologous aggregates and sets happens in a very specific situation. For most of the cases there is a real risk of confounding the two concepts.

We continue now with the case of real functions.

4.2 The case of real functions of real variable

The particular case of aggregates that we are going to use is the case where $A \subset R$ and X is a set of real functions of real variable defined on A.

For each function x *of* X, $x(A)$ is the image set of A and $x^*(A)$ is the aggregate of all the values $x(t)$ for all the t of A.

The study of aggregates may proceed in three directions:

a) focus on a function x and study the aggregates $x^*(B)$ corresponding to sets $B \subset A$

b) focus on a set A and study the aggregates $x^*(A)$, $y^*(A)$,… for the corresponding functions of X.

c) consider simultaneously different aggregates and different functions

4.2 a) the same function and different sets and aggregates

Consider all the sets that are elements of the family 2^A.

We start by recalling the relation of belonging.

Belonging

An element r belongs to the aggregate $x^*(B)$ if and only if there is t of B such that $r = x(t)$.

Remark. Note that the elements r aren't plain real numbers. They are characterized not only by a value corresponding to a real number but also by a value of t, that creates a new order relation $\sqsubseteq$ (not the order of the set R) defined by $x(t) \sqsubseteq x(t^*)$ if and only it $t < t^*$. The ontology of aggregates and their elements is an interesting topic in its own right but it is not the object of the present paper.

Empty aggregate

$x^*(B)$ is empty if and only if B is the empty set and it is designated by $x^*(\emptyset)$.

Complement

The aggregate of all the elements $r = x(t^*)$ for all the t^* of A not belonging to B is the complement of $x^*(B)$ and it is designated by $\neg x^*(B)$ or by $x^*(A\text{-}B)$. Note that the same real number my belong to $x(B)$ and to $x(A\text{-}B)$ but this is by definition impossible for any element and the corresponding aggregates $x^*(B)$ and $x^*(A\text{-}B)$.

Inclusion *(see p.4)*

$x^*(B) \subset x^*(C)$ if and only if for every $x(t)$ belonging to $x^*(B)$ $x(t)$ belongs $x^*(C)$. It is easy to see that $x^*(B) \subset x^*(C)$ if and only if $B \subset C$.

Equality

$x^*(B) = x^*(C)$ if and only if $x^*(B) \subset x^*(C)$ and $x^*(C) \subset x^*(B)$.

Obviously $x^*(B) = x^*(C)$ if and only if $B = C$

Union and intersection

Given B and C we define:

Union of the two aggregates, $x^*(B) \cup x^*(C)$ is the aggregate of all the elements that belong to $x^*(B)$ or to $x^*(C)$.We have always $x^*(B) \cup x^*(C) = x^*(B \cup C)$ (this can be generalized for any number finite or infinite of sets).

Intersection of the two aggregates, $x^*(B) \cap x^*(C)$ is the aggregate of all the elements that belong to $x^*(B)$ and to $x^*(B)$. We have always $x^*(B) \cap x^*(C) = x^*(B \cap C)$ (again, this can be generalized for any number finite or infinite of sets).

Remark. Note that regarding intersection if $x(B)$ is the image set of B, then it is not always true that $x(B) \cap x(C) = x(B \cap C)$.

Obviously $x^*(B) \cap \neg x^*(B) = x^*(\emptyset)$.

Difference

We define $x^*(B) - x^*(C) \equiv x^*(B) \cap \neg x^*(C) = x^*(B - C)$.

Analysis of aggregate functions

Let $\mu: 2^{x^*(A)} \to R$ with real values $\mu[x^*(C)]$.

We can develop an analysis of real aggregate functions in a manner similar to the analysis of set functions. We exemplify briefly.

Let $\boldsymbol{F} \subset 2^A$ be a family of subsets of A and $x^*[\boldsymbol{F}]$ the family of the aggregates $x^*(B)$ for all the B of $\boldsymbol{F}$.

It is easy to verify that if $\boldsymbol{F}$ is a ring of sets (or a σ-ring) the same is true for the family $x^*[\boldsymbol{F}]$.

More generally if the family $\boldsymbol{F}$ is characterized by a property P such that $op\ (E_\alpha)$ belongs to $\boldsymbol{F}$ where $op(E_\alpha)$ is a set that is the result of an operation op on any number of sets E_α of $\boldsymbol{F}$ and if $op^*(x^*(E_\alpha)) = x^*(op(E_\alpha))$ where op^* is the corresponding operation for aggregates, then $x^*[\boldsymbol{F}]$ has the property P applied to aggregates.

Other properties could be defined and it is possible to define measures of aggregates. For instance if λ is a measure defined on the sets of $\boldsymbol{F}$ we may define a measure λ^* on the family $x^*(\boldsymbol{F})$ putting

$\lambda^*(x^*(C)) \equiv \lambda(C)$.

It easy to see that it is a measure, that is, non-negative σ-additive and such that $\lambda^*(x^*(\emptyset))=0$.

4.2b) The same set and different functions

We focus now on a given set and look at the aggregates $x^*(A)$, $y^*(A)$,…

Note that if $x^*(A) \neq y^*(A)$ we cannot have $x^*(A) \subset y^*(A)$ neither $y^*(A) \subset x^*(A)$ so that the relation of inclusion doesn't apply to this case. However we can define

Definition 18 *(partial order that is not the relation of inclusion).*

$x^*(A) \leq y^*(A)$ *if and only if for any* t *of* A, $x(t) \leq y(t)$.

Remark. In this definition since t is the same in the two members of the inequality it is the usual order defined for real numbers that matters and not the relation ∟ (p. 19).

Equality

$x^*(A) = y^*(A)$ if and only if $x(t) = y(t)$ for every t of A or equivalently,

if and only if $x^*(A) \leq y^*(A)$ and $y^*(A) \leq x^*(A)$.

Operations with real numbers

These operations are easily defined. Two examples are sufficient to illustrate the procedure

$x^*(A) + y^*(A) \equiv (x+y)^*(A)$.

Or , being λ a real number

$\lambda x^*(A) \equiv (\lambda x)^*(A)$.

We have now the following lemma that will be used later

Lemma. *If* $x(t) \leq y(t)$ *for all the values* $t \in A \subset R$ *then for all the functions* u *defined on* A *is necessary and sufficient for* u *to be* $u = \lambda x + (1-\lambda)y$ *for a certain* λ, $0 \leq \lambda \leq 1$ *that we have* $x^*(A) \leq u^*(A) \leq y^*(A)$.

Proof

If $u = \lambda x + (1-\lambda)y$ we have for every t $x(t) \leq u(t) \leq y(t)$ so that $x^*(A) \leq u^*(A) \leq y^*(A)$.

Reciprocally if for each t $x(t) \leq u(t) \leq y(t)$ we have $u(t) = x(t) + \lambda(t)[y(t) - x(t)]$ with

$0 \leq \lambda(t) \leq 1$. Then λ is the function that we were looking for.□

Remark. Similar results could be proved for other situations.

We have also the following theorem

Theorem 9. *Let u be defined on A and X be the set of all the functions x defined on A such that* $x^*(A) \leq u^*(A)$ $(u^*(A) \leq x^*(A))$. *Then X is a s-convex set*.

The proof is obvious.

Metric

We may define a metric between the pairs $x^*(A)$, $y^*(A)$... provided that there is a distance $d_A(x,y)$ defined for all the functions defined on *A*.

We put simply

$D(x^*(A),y^*(A)) \equiv d_A(x,y)$.

It easy to verify that is really a distance so that the family of all the aggregates $x^*(A)$, $y^*(A)$... can be easily be endowed with a useful structure of metric space. Note that this cannot be done for the family of image sets $x(A)$, $y(A)$,... because it may easily happen that $x(A) = y(A)$ with $x \neq y$.

A final important concept has to do with continuity of aggregates relative to mixed functions based on functions $F:X \rightarrow X$.

Definition 19 *(Continuity). Let* μ *be a real mixed aggregate/set function defined on the aggregate* $F(x)^*(B)$ *for all the x of X and on a given set* $B \subset A$. *The function* μ *is continuous at the function* $a \in X$ *if and only if for any* $\delta > 0$ *there exits* $\varepsilon > 0$ *such that for all the x such that* $\|x - a\| < \varepsilon$ *we have* $|\mu[F(x)^*(B)] - \mu[F(a)^*(B)]| < \delta$.

4.2 c) Different sets and different functions

Definition of equality

$x^*(A) = y^*(B)$ if an only if $x(t) = y(t)$ and $A = B$

Definition of partial order

If *A* is a non-empty set and $\{A_1, A_2\}$ is a partition of *A*, we have $x^*(A_1) \cup w^*(A_2) \leq y^*(A_1) \cup z^*(A_2)$ if and only if $x(t) \leq y(t)$ for all the *t* of A_1 and $w(t) \leq z(t)$ for all the *t* of A_2 .

4.2 d) Analysis of aggregate functions

We have already seen (p. 21) the possibility of defining measures for aggregates. The concept of additivity when the aggregates depend on real functions plays an important role in the analysis of real aggregate functions. We begin by the definition,

Definition 20 *(Additivity). A real function of aggregate μ is additive if for every x of X being $F(x)*(A)$ and $F(x)*(B)$ disjoint aggregates we have*

$$\mu[F(x)*(A) \cup F(x)*(B)] = \mu(F(x)*(A) + \mu(F(x)*(B)) .$$

Remark 1. An example is one where μ is the integral. In the case of a CF-2 (see p. 15) if $z^{**} \equiv (F(x) , A , F(x), B)$ we have $\int_{A\cup B} z^{**}dt = \int_A F(x)\, dt + \int_B F(x)dt$.

Remark 2 For each x we always have $\mu[F(x)*(A) \cup F(x)*(B)] = \mu[F(x)*(A \cup B)]$, so that if μ is additive,

$$\mu[F(x)*(A \cup B)] = \mu[F(x)*(A)] + \mu[F(x)*(B)].$$

Definition 21 *(Additivity fin terms of aggregates/sets). For a function μ that is mixed aggregate/set given A and B disjoint sets the property of additivity is defined as*

$$\mu[F(x)*(A) \cup F(x)*(B), A \cup B)] = \mu[F(x)*(A), A] + \mu[F(x)*(B), B] \text{ that is}$$

$$\mu[F(x)*(A \cup B), A \cup B] = \mu[F(x)*(A), A] + \mu[F(x)*(B), B].$$

Other properties may be important.

One example is sub-additivity. In the case where μ is a mixed aggregate/set function, μ is *sub-additive* if

$$\mu[F(x)*(A) \cup F(x)*(B), A \cup B)] \leq \mu[F(x)*(A), A] + \mu[F(x)*(B), B]$$

for any A and B not necessarily disjoint.

A more restricted sub-additivity may be defined for A and B disjoints only

Example of sub-additivity for disjoints aggregate/sets.

Consider F the identity function, that is $F(x) = x$ for every x of X and μ with values $\mu[x*(A), A)] = \alpha[x*(A), A)].\varphi[x*(A), A]$

where α and φ are additive functions such that their non-zero values have different signs, that is

$\alpha[x^*(A), A)].\varphi[x^*(B), B] \leq 0$

If A and B are disjoint we have

$\mu[x^*(A \cup B), A \cup B] = \alpha[x^*(A \cup B), A \cup B] .\varphi[x^*(A \cup B), A \cup B] =$

$= \alpha[x^*(A), A)].\varphi[x^*(A), A] + \alpha[x^*(A), A)].\varphi[x^*(B), B] + \alpha[x^*(B), B)].\varphi[x^*(A), A] + \alpha[x^*(B), B)].\varphi[(x^*(B), B]$

As α and φ have values with different signs, all the terms of the sum are non-positive so that

$\mu[x^*(A \cup B), A \cup B] \leq \mu[x^*(A), A] + \mu[x^*(B), B]$

so that μ is sub-additive for disjoint aggregates/sets.

4.2 e) Pseudo-monotony and quasi-convexity non-atomizable functions

It is possible to establish a characterization of quasi-convexity (see p. 10) for certain non-atomizable functions based upon the following definition of pseudo-monotony

Definition 22 *(Non-decreasing pseudo-monotony in terms of aggregates). The non-atomizable (in terms of aggregates) function* σ *is non-decreasing pseudo-monotonous if and only if given two aggregates* $x^*(E(t))$ *and* $y^*(E(t))$ *belonging to the domain of* σ, *such that* $x^*(E(t)) \leq y^*(E(t))$ *we have* $\sigma[x^*(E(t)), t] \leq \sigma[y^*(E(t)), t]$ *for each t of A.*

Remark. It is important not to confound pseudo-monotony with monotony of set or aggregate functions, for example $E \subset G \Rightarrow \mu(E) \leq \mu(G)$

We have the following important theorem for non-atomizable functions that provides a link between quasi-convexity and pseudo-monotony.

Theorem 10. *The real function F with values* $F(x)(t) \equiv \sigma[x^*(B(t)), t]$, $B(t) \subset A$, $t \in B(t)$, $x \in X$, *X s-convex is quasi-convex if* σ *is non-decreasing pseudo-monotonous in terms of aggregates in the set X.*

Proof

Note in the first place that the property of non-decreasing pseudo-monotony of σ, in terms of aggregates means that for each t of A and each pair of functions y_1 and y_2 of X such that $y^*_1(B(t)) \geq y^*_2(B(t))$ we have

$\sigma[y^*_1(B(t)),t] \equiv F(y_1)(t) \geq F(y_2)(t) \equiv \sigma[y^*_2(B(t)),t]$.

On the other hand for each t of A and for any x_1, x_2 of X, with $x = \lambda x_1 + (1-\lambda)x_2$ we have

$x(t) \leq max\{(x_1(t), x_2(t)\}$.

Let $A = A_1 \cup A_2$ where A_1 is the set of all the t of A such that $x_1(t) \geq x_2(t)$ and A_2 the set of all the t of A such that $x_2(t) > x_1(t)$.

Then, for each t of $A_1 \cap B(t)$ we have for any $x = \lambda x_1 + (1-\lambda)x_2$, $x(t) \leq x_1(t)$ and for each t of $A_2 \cap B(t)$, $x(t) \leq x_2(t)$.

Therefore, according to the definition of non-decreasing pseudo-monotony in terms of aggregates we have $x^*(A_1 \cap B(t)) \leq x^*_1(A_1 \cap B(t))$ for each t of $A_1 \cap B(t)$ and $x^*(A_2 \cap B(t)) \leq x_2^*(A_2 \cap B(t))$ for each t of $A_2 \cap B(t)$.

As σ is non-decreasing pseudo-monotonous in terms of aggregates, for each t of $A_1 \cap B(t)$ we have $F(x)(t) \leq F(x_1)(t)$ and for each t of $A_2 \cap B(t)$ we have $F(x)(t) \leq F(x_2)(t)$, so that for each t of $(A_1 \cap B(t)) \cup (A_2 \cap B(t)) = B(t)$ (since by assumption $B(t) \subset A$) we have $F(x)(t) \leq max\{(F(x_1)(t), F(x_2)(t)\}$.

But as we assume that for each t of A, $t \in B(t)$, we finally have for each t of A and each $x = \lambda x_1 + (1-\lambda)x_2$

$F(x)(t) \leq max\{(F(x_1)(t), F(x_2)(t)\}$.□

Pseudo-monotony for mixed aggregate/set functions

The definition of pseudo-monotony can be easily generalized for mixed aggregate/set functions:

Definition 23 *(Non-decreasing pseudo-monotony for mixed terms aggreate/set). The real mixed aggregate/set real function μ has the property of non-decreasing pseudo-monotony if and only if given two aggregates belonging to the domain of μ, $x^*(A)$ and $y^*(A)$ such that $x^*(A) \leq y^*(A)$ we have $\mu[x^*(A), A)] \leq \mu[y^*(A), A]$.*

Simplification of the notation

In what follows if we consider a mixed function with values $\mu[F(x)*(B), B]$ we simplify the notation putting $\mu[F(x)*(B)]$.

This ends the introduction of complementary concepts. We proceed now to section III that is to the formulation of the maximization problem under some specific conditions.

III Maximization

5 One variable

5.1 Solving the maximization problem

The entities

Consider the following entities:

a) Real functions of real variable t, continuous and differentiable, designated by x, y, z ,… with values $x(t), y(t), z(t)$ … defined on a set $A \subset R$

b) The set X of all the previous functions and another set $X*(n)(X)$ that is the set of all the CF-n functions generated by X.

c) A number m of constraints for each function x of X, each constraint represented by a proposition $m_i(x)$, being $X \cap M_i$ the set of functions of X for which $m_i(x)$ is true. That is $X \cap (\cap M_i)$, $i=1,...m$ - assumed to be non-empty – is the set of the x of X that verify all the constraints. To simplify the notation we designate this set by $X \cap M_i$.

d) The set $X \cap M_i$ and the corresponding set $X*(n)(X \cap M_i)$, that is the set of all the CF-n generated by the set $X \cap M_i$. It is assumed that $X \cap M_i$ is a s-convex set.

e) A function $F: X*(n)(X \cap M_i) \rightarrow X$, atomizable (we'll consider later on non-atomizable functions) and the real function $F(x) : A \rightarrow R$ with values $F(x)(t)$, defined for all the t of A.

f) A real mixed aggregate/set function based on real number sets designated by μ defined on the aggregate $F(x)*(B)$ and on the set B, for all the $B \in 2^A$ and all the x of $(X \cap M_i)$. As $2^A \subset 2^R$ is the family of all the subsets of A we have μ: $2^{F(x)*(A)} \times 2^A \rightarrow R$

Remark. The symbol $2^{F(x)*(A)}$ represents the family of all the aggregates $F(x)*(B)$ for each $B \subset A$. As we have seen at page 19 that $B \subset A$ is equivalent to $F(x)*(B) \subset F(x)*(A)$ the definition is coherent

We may now state the problem

The problem

Find $x^* \equiv (x^*_1, A_1, \ldots x^*_n, A_n) \in X^*(n)(X \cap M_i)$,such that:

$\Sigma_{i=1}^n \mu[F(x^*_i)^*(A_i)] = \max_{x \in X^*(n)(X \cap Mi)} \{\mu[F(x)^*(A)]\}$for any partition of A

and such that for all the $p < n$ $\max_{x \in X^*(p)(X \cap Mi)} \{\mu[F(x)^*(A)]\}$ may not exist.

Remark. This second condition means that n is the minimum degree of CF that is needed to guarantee that we find the maximum (as we'll see later on this degree depends in general on the number of restrictions $m_i(x)$).

Why this problem?

Usually the problems of maximization try to find the maximizing function in the set $(X \cap M_i)$ and not in the set $X^*(n)$ $(X \cap M_i)$. As sometimes this is not possible we try to find the maximizing function in the set $X^*(n)$ $(X \cap M_i)$ that is in some sense the "nearest" one to the set $(X \cap M_i)$. Nearest in the sense that at most $n -1$ equalities $x_i = x_j$ are necessary to obtain a x of $(X \cap M_i)$ from $x^* \equiv (x_1, A_1, \ldots x_n, A_n)$.

If we designate by X^+ the set of all the x^* of X^* such that $x_1 = x_2$ for all the partitions of A the proximity of the sets X^*-X^+ and X^+ may be easily attested by the fact that in a great number of cases the *separation* between the sets vanishes (not the Hausdorff distance but the separation S defined by $S(C,D) \equiv \inf\{_{x \in C, y \in D}\ d(x,y)\}$), for example if the set X is such that for each x of X exists a t of A such that there is a sequence y_n of elements of X distinct of x such that $\lim_{n\to\infty} y_n(t) = x(t)$.

Let us prove this for $n = 2$. If the condition applies we have $S(X^*\text{-}\ X^+, X^+) = 0$. Consider a sequence of functions z_n , belonging to X^*-X^+, $z_n = (x,\ A\text{-}\{t\}, y_n, \{t\})$ with $x \neq y_n$ and $\lim_{n\to\infty} y_n(t) = x(t)$ for some t. Consider x^* of X^+ $x^* = (x\ ,\ A\text{-}\{t\},\ x\ , \{t\})$. Using the concept of distance defined at page 16,

$d_A(z_n, x^*) = d_{A-\{t\}}(x, x) + d_{\{t\}}(x, y_n) = d(x(t), y_n(t))$ for every t , where the distance of the right member is the distance between real numbers. But for one of these t $inf\ \{d(x(t), y_n(t)\} = 0$ so that for the sequence $\{z_n\}$ of functions of $X^*\text{-}X^+$ $inf\ \{d_A(z_n, x^*)\} = 0$ that is , $S(X^*\text{-}X^+, X^+) = 0$.

Solving the problem for the simplest case, n=2 and m=1 (that is one restriction only)

For this case we have the following theorem

Theorem 11. If:

a) F is quasi-convex on the s-convex set $X \cap M_i$

b) μ *is defined on all the aggregates determined by subsets of A and is non-decreasing pseudo-monotonous and additive in terms of aggregates/sets*

c) there is only one restriction that is of the type $x_1 \leq x \leq x_2$

Then there are two disjoint sets A_1, A_2 *such that* $A_1 \cup A_2 = A$ *and for all the CF-2* x^* *of the set* $X^*(X \cap M_i)$ *we have*

$max_{x^* \in X^*(X \cap M_i)} \{\mu[F(x^*)^*(A)]\} = \mu[F(x_1)^*(A_1)] + \mu[F(x_2)^*(A_2)]$

Remark 1. The set *A* is considered constant (until we proceed later on with the sensitivity analysis)

Remark 2. Obviously for the set $X \cap M_i$ to be s-convex it is sufficient that *X* and each M_1 are s-convex since the intersection of s-convex sets is a s-convex set.

Proof of the theorem

There is only one restriction $x_1 \leq x \leq x_2$

so that for each *x* verifying the restriction we have

$x = \lambda x_1 + (1 - \lambda) x_2$

with $\lambda = (x_2 - x)/(x_2 - x_1)$ for every *t* such that $x_2(t) \neq x_1(t)$

As $F(x)$ is quasi-convex we have

$F(x) \leq max\ [F(x_1), F(x_2)]$

Let A_1 be the set of elements of A for which $F(x_1) > F(x_2)$ and $A_2 = A - A_1$ the set of elements of A for which $F(x_2) \geq F(x_1)$.

Then for every x of $X \cap M$ we have $F(x) \leq F(x_1)$ for all the elements of A_1 and $F(x) \leq F(x_2)$ for all the elements of A_2, that is in terms of aggregates

$F(x)*(A_1) \leq F(x_1)*(A_1)$ and $F(x)*(A_2) \leq F(x_2)*(A_2)$

As μ is non-decreasing pseudo-monotonous and additive in terms of aggregates and sets we have

$$\mu[F(x)*(A)\ ] = \mu[F(x)*(A_1) \cup F(x)*(A_1)\ ] = \mu[F(x)*(A_1)] + [F(x)*(A_1)] \leq \mu[F(x_1)*(A_1)] + \mu[F(x_2)*(A_2)]$$

Let us show now that for any other $x^{**} = (x_3, B_1, x_4, B_2)$ belonging to $X*(X \cap M)$ with $B_1 \cup B_2 = A$ the inequality is still verified.

We have

$$B_1 = (B_1 \cap A_2) \cup (B_1 \cap A_1)$$

$$B_2 = (B_2 \cap A_1) \cup (B_2 \cap A_2)$$

so that

$$F(x_3)*(B_1) = F(x_3)*\ (B_1 \cap A_2) \cup F(x_3)*\ (B_1 \cap A_1)$$

$$F(x_4)*(B_2) = F(x_4)*\ (B_2 \cap A_1) \cup F(x_4)*\ (B_2 \cap A_2)$$

and

$$\mu[F(x_3)*(B_1)] = \mu[F(x_3)*(B_1 \cap A_2)] + \mu[F(x_3)*(B_1 \cap A_1)] \leq \mu[F(x_1)*(B_1 \cap A_1)] + \mu[F(x_2)*(B_1 \cap A_2)]$$

$$\mu[F(x_4)*(B_2)] = \mu[F(x_4)*(B_2 \cap A_1)] + \mu[F(x_4)*(B_2 \cap A_2)] \leq \mu[F(x_1)*(B_2 \cap A_1)] + \mu[F(x_2)*(B_2 \cap A_2)]$$

Adding the two members

$$\mu[F(x_3)*(B_1)] + \mu[F(x_4)*(B_2)] \leq \{\mu[F(x_1)*(B_1 \cap A_1)] + \mu[F(x_1)*(B_2 \cap A_1)]\} + \{\mu[F(x_2)*(B_1 \cap A_2)] + \mu[F(x_2)*(B_2 \cap A_2)]\} = \mu[F(x_1)*(A_1)] + \mu[F(x_2)*(A_2)]$$

Therefore for any function of x^{**} of $X^*(X\cap M)$ we have $\mu[F(x^{**})^*(A)]\} \leq \mu[F(x_1)^*(A_1)] + \mu[F(x_2)^*(A_2)]$. As (x_1, A_1, x_2, A_2) belongs to $X^*(X\cap M)$ the theorem is proved.□

Remark 1. It is the fact that we can use properties of μ such as pseudo-monotony and additivity in terms of aggregates and sets that makes it helpful to use the concept of aggregate in maximization problems.

Remark 2. If instead of assuming that μ is additive we had assumed that it was sub-additive as the example of page 23 the inequality

$$\mu[F(x_3)^*(B_1)] + \mu[F(x_4)^*(B_2)] \leq \{\mu[F(x_1)^*(B_1\cap A_1)] + \mu[F(x_1)^*(B_2\cap A_1)]\} + \{\mu[F(x_2)^*(B_1\cap A_2)] + \mu[F(x_2)^*(B_2\cap A_2)]\},$$

would still be valid although we can't prove the theorem.

5.2 Sensitivity analysis of the values of the maximum

The intention of this section is to calculate the change of the maximum values when there are changes in the set A or in the set X^*.

We consider accordingly two types of sensitivity analysis.

First type: change in the set A

Suppose that set A changes to a new set B of real numbers. In most cases the set $(X\cap M_i)$ will also change so that we have to consider new sets $(X \cap M^*_i)$ and $X^{**}(X\cap M^*_i)$ respectively.

Let $\mu[F(x^{**})(A)]$ be the maximum $\mu[F(x^*_1)(A_1)] + \mu[F(x^*_2)(A\text{-}A_1)] = \max_{x^* \in X^*(X\cap M_i)} \{\mu[F(x)(A)]\}$ and $\mu[F(x^{++})(B)]$ the maximum $\mu[F(x^+_1)(B_1)] + \mu[F(x^+_2)(A\text{-}B_1)] = \max_{x^{**} \in X^{**}(X\cap M^*_i)} \{\mu[F(x)(B)]\}$.

we have the following theorem

Theorem 12. *With the assumptions of theorem relatively to F and* μ *and assuming that the functional J associated to* μ *is Gateaux-differentiable in the directions of the following expression we have*

$\mu[F(x^*_1)^*(A_1)] + \mu[F(x^*_2)^*(A_2)] - \mu[F(x^*_3)^*(B_1)] - \mu[F(x^*_4)^*(B_2)] = \mu[F(x^*_1)^*(A_1)] + \mu[F(x^*_2)^*(A_2)] - \mu[F(x^*_1)^*(B_1)] - \mu[F(x^*_2)^*(B_2)] +$

$D\mu[F(x^*_3+\theta_1(x^*_1-x^*_3), x^*_1-x^*_3)^*(A_1\cap B_1)] + D\mu[F(x^*_3+\theta_2(x^*_1-x^*_3), x^*_1-x^*_3)^*(B_1-A_1)]+$

$D\mu[F(x^*_4+\theta_1(x^*_2-x^*_4),x^*_2-x^*_4)^*(A_1\cap B_1)] + D\mu[F(x^*_4+\theta_2(x^*_2-x^*_4), x^*_2-x^*_4)^*(B_1-A_1)]$.

Remark 1. Functional J associated to μ is the functional $J: X\rightarrow R$ with $J(y) = \mu[F(y)^*(B)]$

Remark 2. $D\mu$ is the Gateaux differential of J. It is well known that if J is Gateaux-differentiable at the point y_1 in the direction $y_1 - y_2$ we have for a certain θ, $0 < \theta < 1$

$J(y_1)-J(y_2) = J'(y_2+\theta(y_1-y_2), y_1-y_2)$ that is in our notation

$\mu[F(y_1)^*(B)] - \mu[F(y_2)^*(B)] = D\mu[F(y_2+\theta(y_1-y_2), y_1-y_2)^*(B)]$.

Proof of the theorem

Consider the first difference $\mu[F(x^*_1)^*(A_1)] - \mu[F(x^*_3)^*(B_1)]$.

As μ is additive we have

$\mu[F(x^*_1)^*(A_1)] - \mu[F(x^*_3)^*(B_1)] = \mu[F(x^*_1)^*(A_1\cap B_1)] + \mu[F(x^*_1)^*(A_1-B_1)] - \mu[F(x^*_3)^*(B_1-A_1)] - \mu[F(x^*_3)^*(A_1\cap B_1)] =$

$\mu[F(x^*_1)^*(A_1\cap B_1)]+\mu[F(x^*_1)^*(A_1-B_1)] + \mu[F(x^*_1)^*(B_1-A_1)] - \mu[F(x^*_1)^*(B_1-A_1)] - \mu[F(x^*_3)^*(B_1-A_1)] - \mu[F(x^*_3)^*(A_1\cap B_1)] =$

$\mu[F(x^*_1)^*(A_1\cap B_1)]- \mu[F(x^*_3)^*(A_1\cap B_1)]+ \mu[F(x^*_1)^*(A_1-B_1)] - \mu[F(x^*_1)^*(B_1-A_1)] +$

$+ \mu[F(x^*_1)^*(B_1-A_1)] - \mu[F(x^*_3)^*(B_1-A_1)] =$

1) $\mu[F(x^*_1)^*(A_1-B_1)] - \mu[F(x^*_1)^*(B_1-A_1)] + D\mu[F(x^*_3+\theta_1(x^*_1-x^*_3), x^*_1-x^*_3)^*(A_1\cap B_1)]+ D\mu[F(x^*_3+\theta_2(x^*_1-x^*_3), x^*_1-x^*_3)(B_1-A_1)]$

But

$\mu[F(x^*_1)^*(A_1-B_1)] = \mu[F(x^*_1)^*(A_1)] - \mu[F(x^*_1)^*(A_1\cap B_1)]$

$\mu[F(x^*_1)^*(B_1-A_1)] = \mu[F(x^*_1)^*(B_1)] - \mu[F(x^*_1)^*(A_1\cap B_1)]$

So that subtracting both members of the two equalities we get

$\mu[F(x^*_1)^*(A_1\text{-}B_1)]$ - $\mu[F(x^*_1)^*(B_1\text{-}A_1)] = \mu[F(x^*_1)^*(A_1)]$ - $\mu[F(x^*_1)^*(B_1)]$

Substituting in 1) we obtain

$\mu[F(x^*_1)^*(A_1)]$- $\mu[F(x^*_3)^*(B_1)] = \mu[F(x^*_1)^*(A_1)]$ - $\mu[F(x^*_1)^*(B_1)]+ D\mu[F(x^*_3+\theta_1(x^*_1\text{-}x^*_3), x^*_1\text{-}x^*_3)^*(A_1\cap B_1)]+ D\mu[F(x^*_3+\theta_2(x^*_1\text{-}x^*_3), x^*_1\text{-}x^*_3)^*((B_1\text{-}A_1)]$.

For the second difference , $\mu[F(x^*_2)^*(A_2)]$- $\mu[F(x^*_4)^*(B_2)]$,

in the same way we get

$\mu[F(x^*_2)^*(A_2)]$- $\mu[F(x^*_4)^*(B_2)] = \mu[F(x^*_2)^*(A_2)]$- $\mu[F(x^*_2)^*(B_2)]+ D\mu[F(x^*_4+\theta_3(x^*_2\text{-}x^*_4), x^*_2\text{-}x^*_4)^*(A_2\cap B_2)]+ D\mu[F(x^*_4+\theta_4(x^*_2\text{-}x^*_4), x^*_2\text{-}x^*_4)^*((B_2\text{-}A_2)]$

and finally summing the two differences

$\Delta \equiv \mu[F(x^*_1)^*(A_1)]+ \mu[F(x^*_2)^*(A_2)]\text{-} \mu[F(x^*_3)^*(B_1)] - \mu[F(x^*_4)^*(B_2)] =$

$=\mu[F(x^*_1)^*(A_1)]\text{-}\mu[F(x^*_1)^*(B_1)]+D\mu[F(x^*_3+\theta_1(x^*_1\text{-}x^*_3),x^*_1\text{-}x^*_3)^*(A_1\cap B_1)]+ D\mu[F(x^*_3+\theta_2(x^*_1\text{-}x^*_3), x^*_1\text{-}x^*_3)^*(B_1\text{-}A_1)] +$

$+\mu[F(x^*_2)^*(A_2)]\text{-}\mu[F(x^*_2)^*(B_2)]+D\mu[F(x^*_4+\theta_3(x^*_2\text{-}x^*_4),x^*_2\text{-}x^*_4)^*(A_2\cap B_2)]+ D\mu[F(x^*_4+\theta_4(x^*_2\text{-}x^*_4), x^*_2\text{-}x^*_4)^*(B_2\text{-}A_2)] =$

$\mu[F(x^*_1)^*(A_1)]+ \mu[F(x^*_2)^*(A_2)]\text{-} \mu[F(x^*_1)^*(B_1)] - \mu[F(x^*_2)^*(B_2)] +$

$D\mu[F(x^*_3+\theta_1(x^*_1\text{-}x^*_3),x^*_1\text{-}x^*_3)^*(A_1\cap B_1)]+ D\mu[F(x^*_3+\theta_2(x^*_1\text{-}x^*_3), x^*_1\text{-}x^*_3)^*(B_1\text{-}A_1)]+$

$D\mu[F(x^*_4+\theta_3(x^*_2\text{-}x^*_4),x^*_2\text{-}x^*_4)^*(A_2\cap B_2)]+ D\mu[F(x^*_4+\theta_4(x^*_2\text{-}x^*_4), x^*_2\text{-}x^*_4)^*(B_2\text{-}A_2)]$,as we had to prove.□

When all the $D\mu$ are non-negative (non-positive) with at least one positive (negative) we have

$\Delta > (<) \mu[F(x^*_1)^*(A_1)]+ \mu[F(x^*_2)^*(A_2)]\text{-}\mu[F(x^*_1)^*(B_1)] - \mu[F(x^*_2)^*(B_2)$

When $A_1=B_1$ and $A_2=B_2$, that is when it is the set of constraints only that changes we have

$\Delta = D\mu[F(x^*_3+\theta_1(x^*_1\text{-}x^*_3),x^*_1\text{-}x^*_3)^*(A_1)]+D\mu[F(x^*_4+\theta_3(x^*_2\text{-}x^*_4),x^*_2\text{-}x^*_4)^*(A_2)]$.

Second type : transferring one element from A_1 to A_2

Let $x^{**}= (x_1,A_1,x_2,A_2)$ such that $\mu[F(x^{**})^*(A)] = max_{x^* \in X^*(X \cap M_i)} \{\mu[F(x^*)^*(A)]\}= \mu[F(x_1)^*(A_1)] + \mu[F(x_2)^*(A_2)]$

In may be important for certain cases determine if transferring one element t from A_1 to A_2 (or vice-versa) will change the value of the maximum.

We have $x^{***} = x^{**} ╬ 1(t)$

and we check if $\mu[F(x^{***})^*(A)] = \mu[F(x_1)^*(A_1-\{t\})] + \mu[F(x_2)^*(A_2 \cup \{t\})]$ is equal or less than $\mu[F(x_1)^*(A_1)] + \mu[F(x_2)^*(A_2)]$, that is if $\mu[F(x^{**} ╬ 1(t))^*(A)]$ is < or = to $\mu[F(x^{**})^*(A)]$ and the same for $1^*(t)$. This can be done also for an iteration of the operation

$x^{(n)} = x^{(n-1)} ╬ 1(t)$

Degree of approximation and economic decision

The analysis of the level of approximation can be done in two ways: one calculates the difference of the maximum to the value of a function equivalent-2 to the maximizing function and the second relatively to the proximity of the set X^*-X^+ to X^+ as we have exemplify at page 28.

Considering the first way, let x^{**} be the maximizing function and another CF-2 function equivalent-2 to x^{**}, $y = (y_1,A_1,y_2,A_2)$ satisfying the constraints. We obtain the difference

$\mu[F(x^{**})^*(A)] - \mu[F(y)^*(A)] = \mu[F(x_1)^*(A_1)] + \mu[F(x_2)^*(A_2)] - \mu[F(y_1)^*(A_1)] + \mu[F(y_2)^*(A_2)]$

and using the definition of distance given at page 16 we can compare the two quantities

$\{\mu[F(x^{**})^*(A)] - \mu[F(y)^*(A)]\}$ and $[(d_{A1}(x_1,y_1)+ d_{A2}(x_2,y_2)]$.

This can be a useful indicator for taking economic decisions when there is a benefit associated to the value of $\mu[F(x^{**})^*(A)] - \mu[F(y)^*(A)]\}$ and a cost associated to the distance between the two functions.

In what concerns the approximation of the solutions when calculated for functions y such that $y_1= y_2$ we can obtain some information in specific cases.

If $X \cap M_i$, is such that $x_1 \leq y \leq x_2$, which is a s-convex set we may write for any y, $y = \lambda x_1 + (1-\lambda)x_2$ with $0 \leq \lambda \leq 1$

so that

$$d_A(y, x^{**}) = d_{A1}(y, x_1) + d_{A2}(y, x_2) = \|(1-\lambda)(x_1-x_2)\|_{A1} + \|(\lambda(x_1-x_2)\|_{A2}$$

Even if x^{**} is such that $x_1 \neq x_2$ if we chose appropriate functions λ we can obtain better approximations of x^{**} to the functions of $X \cap M_i$. This is the case where the values $\lambda(t)$ are near 1 for the every t of A_1 and near 0 for every t of A_2 – provided of course that the properties of λ respect the conditions that allow x to belong $X \cap M_i$ (conditions of continuity or of differentiability, for example). It is the failure to verify these conditions that makes it impossible in most cases to have for a given x^{**} a y of $X \cap M_i$ such that $d_A(y, x^{**}) = 0$.

6 Generalization for two variables

We can generalize the problem of maximization assuming the existence of a function φ,

$\varphi : X \rightarrow X$ and of a function $F: X \times X \rightarrow X$.

The new problem is to find $(x^*_1, A_1, x^*_2, A - A_1)$ such that

$$\mu[F(x^*_1, \varphi(x^*_1))^*(A_1)] + \mu[F(x^*_2, \varphi(x^*_2))^*(A-A_1)] = max_{x \in X^*(X \cap Mi)}\{\mu[F(x, \varphi(x))^*(A)]\}$$

A very important particular case of this kind of problem is the one where X is the set of continuous and derivable real functions in $A \subset R$ and φ is such that for each $x \in X$, $\varphi(x)$ is the derivative of function x.

However in this more general formulation some additional assumptions are needed regarding the function F.

We have seen above (theorem 5 p. 10) that for φ convex and F quasi-convex we obtain $F(x, \varphi(x)) \leq max\{F(x_1, \varphi(x_1)), F(x_2, \varphi(x_2))\}$. This allows us to prove the following theorem.

Theorem 13. *If F is quasi-convex and non-decreasing monotonous in $\varphi(x)$ where φ is convex and if the assumptions of theorem 11 regarding μ apply, we have the result:*

$$\mu[F(x^*_1, \varphi(x^*_1))^*(A_1)] + \mu[F(x^*_2, \varphi(x^*_2))^*(A-A_1)] = max_{x \in X^*(X \cap Mi)}\{\mu[F(x, \varphi(x))^*(A)]\}.$$

Proof

As F is quasi-convex

$F(x, \varphi(x)) \leq \max \{F(x_1, \varphi(x_1)), F(x_2, \varphi(x_2))\}$

And from here the proof follows as in theorem 11.□

Another example is the one for each $\varphi(\lambda x_1+(1-\lambda)x_2) = \varphi(\lambda)(x_1 - x_2) + \lambda \varphi(x_1) + (1-\lambda) \varphi(x_2)$ (which is verified if φ is the operation of differentiation of x)

Theorem 14. I*f the set of constraints is given by* $x_1 \leq x \leq x_2$, *and other constraints that imply* $\varphi(\lambda) \leq 0$ ($\varphi(\lambda) \geq 0$) *for any possible* λ, *with* $0 \leq \lambda \leq 1$ *and if* $\varphi(\lambda x_1 + (1-\lambda)x_2) = \varphi(\lambda)(x_1 - x_2) + \lambda \varphi(x_1) + (1-\lambda) \varphi(x_2)$ *and F is quasi-convex and non-increasing (non-decreasing) in* φ *then if the assumptions of theorem 11 regarding* μ *apply we have* $\mu[F(x^*_1, \varphi(x^*_1))^*(A_1)] + \mu[F(x^*_2, \varphi(x^*_2))^*(A-A_1)] = \max_{x \in X^*(X \cap Mi)}\{\mu[F(x, \varphi(x))^*(A)]\}$.

Proof

For any possible $x = \lambda x_1 + (1-\lambda)x_2$ we have

$F(x, \varphi(x)) = F(x, \varphi(\lambda)(x_1 - x_2) + \lambda\varphi(x_1) + (1-\lambda)\varphi(x_2))$

Given the assumptions, $\varphi(\lambda)(x_1 - x_2) \geq 0$ and F is non-increasing in φ so that

$F(x, \varphi(x)) \leq F(x, \lambda\varphi(x_1) + (1-\lambda)\varphi(x_2))$

As F is quasi-convex we have

$F(x, \lambda\varphi(x_1) + (1-\lambda)\varphi(x_2)) \leq \max \{F(x_1, \varphi(x_1)), F(x_2, \varphi(x_2))\}$

so that $F(x, \varphi(x)) \leq \max \{F(x_1, \varphi(x_1)), F(x_2, \varphi(x_2))\}$ and the proof follows as in theorem 11.□

7 Maximization and non-atomizable functions

Theorem 11 is easily generalized for a quasi-convex non-atomizable function F and for a μ with the properties of the theorem. By theorem 10 if $F(x)(t) \equiv \sigma[x^*(B(t)),t]$ it is sufficient to have σ pseudo-monotonous non-decreasing to have $F(x)$ quasi-convex and the proof proceeds as in theorem 11.

IV A problem of calculus of variations with inequality constraints

The problem

Let f be a real function integrable over $A \equiv [t_0\ t_1]$ with arguments $x(t)$ and $x'(t)$, where $x(t)$ is a continuous real function with second order derivatives at all the elements of A.

Let X be the set of those functions.

Consider the problem:

Calculate $sup_X \int_A f(x(t)\ x'(t))dt$, with $A \equiv [t_0\ , t_1]$

subject to the restrictions $r_1 \leq g(x(t), x'(t)) \leq r_2$

$x(t_0) = x_0$

In terms of the previous notation we have $F : X^2 \rightarrow X$, where φ is the operation of derivation, the values of F are $F(x, x') = f(x, x')$ where the set $F(x, x')*(A)$ is the aggregate of all the elements $f(x(t), x'(t))$ for all the t of A and the mixed aggregate/set function μ such that $\mu[F(x, x')*(A)] = \int_A f(x(t)\ x'(t))\ dt$.

Note that this formulation applies only to atomizable functions F.

We have two sets M_i that is the set of all the x that verify $g(x(t), x'(t)) \leq r_2$ and the set of all the x that verify $r_1 \leq g(x(t), x'(t))$.

In the particular case that we solve the double inequality $r_1 \leq g(x(t), x'(t)) \leq r_2$ is given by

C1) $r_1 \leq x'(t) - mx(t) \leq r_2$

We join one more constraint given by

C2) $x''(t) - mx'(t) \geq 0$

So that there are in fact three sets M_i.

The set X^* is the set of all the CF-2 generated by the set X of continuous functions with second order derivatives.

The problem is :

Assumptions

Let $\cap M_i$, $i=1,2,3$ be the set of all the functions X that verify simultaneously the constraints

$r_1 \leq x'(t) - mx(t) \leq r_2$

$x''(t) - mx'(t) \geq 0$

and let $X^*[X(\cap M_i)]$ be the set of CF-2 generated by $X \cap M_i$.

Consider the function $f: X^2 \rightarrow X$ with values $f(x, x')$, such that

a) for each t the function f takes the value $f(x, x')(t) \equiv f[x(t), x'(t)]$

b) it is defined on $A \equiv [t_0 , t_1]$

c) is continuous in A

d) is quasi-convex on the set of all the linear combinations of functions of X

e) f is assumed non-increasing in x', that is $f(x, x') \geq f(x, x' + h)$ for any function h of X such that $h \geq 0$ (that is , $h(t) \geq 0$ for all the values t of A). ´

Problem

Find $x^*$of $X^*(X \cap M_i)$ such that

$\int_A f(x^*(t)\, x^{*\prime}(t))dt = \max_{x \in X^*(X \cap M_i)} \{ \int_A f(x(t)\, x'(t))\, dt\}$

where (see p. 15) $\int_A f(x^*(t)\, x^{*\prime}(t))dt$ represents the sum $\int_{A1} f(x_1(t)\, x_1'(t))dt + \int_{A2} f(x_2(t)\, x_2'(t))dt$ for $x^* \equiv (x_1, A_1, x_2, A_2)$.

The solution is given by the following theorem

Theorem 15. *Under the previous conditions the maximum of the integral* $\int_A f(x(t)\, x'(t))dt$ *on the generated set* $X^*(X \cap M_i)$ *is given* $\int_B f(x_1(t), x'_1(t))dt + \int_{A\, -B} f(x_2(t), x_2'(t))dt$ *where B is the set of all the t of A such that* $f(x_1(t), x'(t)) \leq f(x_2(t), x_2'(t))$ *for all the x of* $(X \cap M_i)$ *and* $x_1(t)$ *and* $x_2(t)$ *are respectively the solutions of the differential equations*

$x(t) - mx'(t) = r_1$ ($= r_2$), $m \neq 0$.

Remark. It is necessary to impose the condition for $F(t)$ being defined on all the elements of the set A because the fact that $x(t)$ and $x'(t)$ are defined on a set that contains A is not sufficient to ensure that F is defined on A. For example if $f(x, x'(t)) = \log x(t) + x'(t)$ with $x(t) = \log t$ and $A = [0.5, 2]$ the domain of $F(t)$ would not include the elements $t \leq 1$. We may have the inverse situation where the domain of F(t) includes the intersection of the domain of $x(t)$ and $x'(t)$. For example if $f(x, x') \equiv e^x + x'$ and $x(t) \equiv \log t$, the domain of $F(t)$ is the set of all the numbers different from 0 but the intersection of the domains if $x(t)$ and $x'(t)$ is the set of all the positive numbers.

Proof

Let us begin by characterizing the functions that belong to $\cap M_i$, $i=1,2,3$

The two restrictions *C1)* and *C2)* may be put in the form

$x(t) - mx'(t) = \theta(t)$, $r_1 \leq \theta(t) \leq r_2$,

where $\theta'(t) \geq 0$, that is θ is non-decreasing .

The solutions $x_1(t)$ ($x_2(t)$), are given by

$x(t) - mx(t) = r_1$ ($= r_2$) $m \neq 0$

that is ,

$x_1(t) = (x_0 + r_1/m)\, e^{m(t-t_0)} - r_1/m$

$x_2(t) = (x_0 + r_2/m)\, e^{m(t-t_0)} - r_2/m$

where $x_0 \equiv x(t_0)$.

It is easy to see that $x_1(t_0) = x_2(t_0)$ and $x_1(t) < x_2(t)$ (provided that $m \neq 0$) for all the $t \neq t_0$.

For any function x of $(X \cap M_i)$ we have

$x(t) - mx'(t) = \theta(t)$

and the solution is

$x(t) = e^{mt}[\int_{[t_0, t]} \theta(\tau) e^{-mt} d\tau + x_0 e^{-mt_0}]$

We can now prove the following lemma (see theorem 14 for a similar result).

Lemma 1. *Under the previous conditions and for each t of A we have*

$x(t) = \lambda(t)x_1(t) + (1-\lambda(t))x_2(t)$ *with* $0 \leq \lambda(t) \leq 1$ *and* $\lambda'(t) \leq 0$.

Proof

Obviously the equality is verified for $t = t_0$.

For any other t of A $t \neq t_0$ we can represent by λ(t) the quotient

1) $\lambda(t) = [x_2(t) - x(t)]/ [x_2(t) - x_1(t)] =$

$= [(r_2/m)e^{m(t-t_0)} - (r_2/m) - e^{mt} \int_{[t_0, t]} \theta(\tau)e^{-mt} d\tau]/ [(r_2 - r_1)(e^{m(t-t_0)} - 1)/m]$

As $r_1 \leq \theta(t) \leq r_2$ we have obviously $0 \leq \lambda(t) \leq 1$

Moreover we have $\lambda'(t) \leq 0$ for all the t. From 1) with a simple calculation we have

$\lambda(t) = [r_2 - \theta(t^*)] / (r_2 - r_1)$

where $\theta(t^*)$, with $t_0 \leq t^* \leq t$ is the mean point of the integral .

As we assume $\theta'(t) \geq 0$, $\theta(t^*)$ is non-decreasing so that $\lambda(t)$ is non-increasing, that is $\lambda'(t) \leq 0$ and this completes the proof .□

Let's us now prove a second lemma

Lemma 2. *If f is quasi-convex for any of X we have for each t of* $[t_0, t_1]$

$f(x(t), x(t)) \leq \max\{f(x_1(t), x'_1(t)), f(x_2(t), x_2'(t)\}$.

Proof

For each $t \neq t_0$

a) $x(t) = \lambda(t)x_1(t) + (1-\lambda(t))x_2(t)$

as f is quasi-convex, we have for each t

$f(x(t), \lambda(t)x_1'(t) + (1-\lambda(t))x'_2(t)) \leq \max\{f(x_1(t), x'_1(t)), f(x_2(t), x'_2(t)\}$

And from a) we can write for each t

$x'(t) = \lambda(t)x'_1(t) + (1-\lambda(t))x'_2(t) + \lambda'(t)(x_1(t) - x_2(t))$

As $x_1(t) \leq x_2(t)$ and by Lemma 1 $\lambda'(t) \leq 0$, we have for each t

$\lambda'(t)(x_1(t) - x_2(t)) \geq 0$

As f is non-increasing relatively to x', we have for each $t \neq t_0$

$f(x(t), x'(t)) \leq f(\lambda(t)x_1(t) + (1-\lambda(t))x_2(t), \lambda(t)x'_1(t) + (1-\lambda(t))x'_2(t)) \leq \max \{f(x_1(t), x'_1(t)), f(x_2(t), x'_2(t)\}$.

When $t = t_0$ we have $x_1(t_0) = x_2(t_0) = x(t_0)$ and obviously the inequality is verified.

Then for all the t of A

$f(x(t), x'(t)) \leq \max \{f(x_1(t), x'_1(t)), f(x_2(t), x'_2(t))\}$

as we had to prove.□

With these two lemmas we may prove

Theorem 16. *Under the conditions of the lemmas*

$\max \int_A f(x(t), x'(t))\, dt = \int_B f(x_2(t), x_2'(t))\, dt + \int_{A-B} f(x_1(t), x'_1(t))\, dt$

Where B is the set of all the t de $[t_0\ t_1]$ *such that* $f(x_1(t), x_1'(t)) \leq f(x_2(t), x_2'(t))$.

Proof

First let us verify that f is integrable over B (and therefore over A-B).

As we assume that $F(t)= f(x(t), x'(t))$ as a function of t is continuous in A (actually it would be enough for this purpose to assume that $F_1(t) = f(x_1(t), x'_1(t))$ and $F_2(t) = f(x_2(t), x'_2(t)))$ are continuous in A), $F_1(t)$ and $F_2(t)$ are continuous in B so that they are Borel-measurable and the set B of all the t such $F_1(t) < F_2(t)$ is Borel-measurable (König, 1997 pag 130). Then the integral on B and the integral on A exist.

We have now to prove that for any other $z = (x^*, C, x^{**}, A\text{-}C)$ of $X^*(X \cap M_i)$ we have $\int_C f(x^*(t), x^{*\prime}(t))\, dt + \int_{A-C} f(x^{**}(t), x^{**\prime}(t))\, dt \leq \int_B f(x_2(t), x_2'(t))\, dt + \int_{A-B} f(x_1(t), x_1'(t))\, dt$

If z belongs to X that is if $x^* = x^{**}$ the result is obvious due to lemma 2.

If $x^* \neq x^{**}$ let us consider the partition of A: $(B\cap C) \cup (B\cap A\text{-}C) \cup (A\text{-}B\cap C) \cup (A\text{-}B\cap A\text{-}C)$

Due to lemma 2 we have

$\int_{B\cap C} f(x^*(t), x^{*\prime}(t))\, dt + \int_{(A-B)\cap A-C} f(x^{**}(t), x^{**\prime}(t))\, dt \leq \int_{B\cap C} f(x_2(t), x_2'(t))\, dt + \int_{(A-B)\cap A-C} f(x_1(t), x_1'(t))\, dt$

$\int_{(A-B)\cap C} f(x^*(t), x^{*\prime}(t))\, dt + \int_{B\cap(A-C)} f(x^{**}(t), x^{**\prime}(t))\, dt \leq \int_{(B\cap(A-C)} f(x_2(t), x_2'(t))\, dt + \int_{(A-B)\cap C} f(x_1(t), x_1'(t))\, dt$

Summing the two inequalities we obtain the result that solves our problem □

The set of constraints

We have already mentioned (p. 28) that the minimum p of CF-p is generally dependent on the number of constraints.

Let us see an example with constraints similar to *C1) e C2).*

Suppose that there were two additional restrictions

C3) $r_2 \leq x(t) - nx'(t) \leq r_3 \quad n \neq 0,m$

C4) $x''(t) - nx(t) \geq 0$

Using the same process of the proof of theorem we have for $t \neq t_0$

c) $x(t) = (1-\lambda(t))x_1(t) + \lambda(t)x_2(t)$ with $0 \leq \lambda(t) \leq 1$, $\lambda'(t) \leq 0$, and $x_2(t) > x_1(t)$

d) $x(t) = (1-\mu(t))x_3(t) + \mu(t)x_4(t)$ with $0 \leq \mu(t) \leq 1$, $\mu'(t) \leq 0$ and $x_4(t) > x_3(t)$

So that $(1-\lambda(t))x_1(t) + \lambda(t)x_2(t) = (1-\mu(t))x_3(t) + \mu(t)x_4(t)$

and $\lambda(t) = [(x_4(t) - x_3(t))/ (x_2(t) - x_1(t))]\, \mu(t) + [(x_3(t) - x_1(t))/ (x_2(t) - x_1(t))]$

To have $\lambda(t) \leq 1$ is necessary that $[(x_3(t) - x_1(t))/ (x_2(t) - x_1(t))] \leq 1$ so that $x_2(t) \geq x_3(t)$.

Using the same process to solve in order to $\mu(t)$, we get $[(x_1(t) - x_3(t))/ (x_4(t) - x_3(t))] \leq 1$ and $x_4(t) \geq x_1(t)$ so that we have the following possibilities for any $t \in (t_0, t_1]$

$x_1 \leq x_3 \leq x_2 \leq x_4$, $x_1 \leq x_3 \leq x_4 \leq x_2$, $x_3 \leq x_1 < x_2 \leq x_4$ or $x_3 \leq x_1 \leq x_4 \leq x_2$.

The problems where one of this inequalities is not verified for one t, $t \in (t_0, t_1]$ are unsolvable because the set X is empty.

Note that as $m \neq n$ as times goes by the exponential function with exponent $max(m,n)$ prevails so that for t sufficiently large the problem has no solution. This may be important to explain some situations where there is a rupture at a given point in time.

If X is non-empty we have

$f(x_1(t), x'_1(t) \leq f(x_2(t), x'_2(t))$ for t of B $f(x_1(t), x'_1(t) > f(x_2(t), x'_2(t))$ for t of A-B

$f(x_3(t), x'_3(t) \leq f(x_4(t), x'_4(t))$ for t of C $f(x_3(t), x'_3(t) > f(x_4(t), x'_4(t))$ for t of A-C

For each t of $E_1 = (B \cap C)$ $f(x(t), x'(t)) \leq min\ (f(x_2(t), x'_2(t)), f(x_4(t), x'_4(t))$

For each t of $E_2 = B \cap (A\text{-}C)$ $f(x(t), x'(t)) \leq min\ (f(x_2(t), x'_2(t)), f(x_3(t), x'_3(t))$

For each t of $E_3 = (A\text{-}B) \cap C$ $f(x(t), x'(t)) \leq min\ (f(x_1(t), x'_1(t)), f(x_4(t), x'_4(t))$

For each t of $E_4 = (A\text{-}B) \cap (A\text{-}C)$ $f(x(t), x'(t)) \leq min\ (f(x_1(t), x'_1(t)), f(x_3(t), x'_3(t))$

On the other side for each E_i we have $E_i = F_{1i} \cup F_{2i}$ where $F_{1i} = \{t: f(x_j(t), x'_j(t)) \leq f(x_k(t), x'_k(t))\}$ and $F_{2i} = \{t: f(x_j(t), x'_j(t)) > f(x_k(t), x'_k(t))\}$.

Of course some of the F may be empty.

Therefore the maximum value will be

$J = \Sigma_{i=1}^{4} \int_{Ei} H_i\, dt$

where the H_i are the functions corresponding to the minima in the second members of the inequalities.,

That means that it makes sense in many cases with two set of constraints C_3 and C_4 to work with a set X^* of functions CF-4 because we have no guarantee that functions CF-p with $p < 4$ will solve the problem.

Conclusion

The main goal of this paper was to illustrate the use of some new (in this context) concepts for approaching the problem of maximization of one functional with inequality constraints. The method was successful but the problem solved was indeed a very simple one, dealing with quasi-convexity and maximization where ”boundary” solutions are to be expected. Further research should seek to determine if the method is useful for other situations.

References

Amaral, João Ferreira do [2007] *Generalized Dynamic Systems* WP 08/2007/DE/UECE, ISEG,ULisboa.

Céa, Jean [1971] *Optimisation - théorie et algorithmes*. Dunod. Paris

König, Heinz [1997] *Measure and Integration*. Springer

.