RAIRO. ANALYSE NUMÉRIQUE

P. J. LAURENT
C. CARASSO

An algorithm of successive minimization in convex programming

RAIRO. Analyse numérique, tome 12, nº 4 (1978), p. 377-400
<http://www.numdam.org/item?id=M2AN_1978__12_4_377_0>

R A I R O Analyse numérique/Numerical Analysis
(vol 12, n° 4, 1978, p 377 à 400)

AN ALGORITHM OF SUCCESSIVE MINIMIZATION IN CONVEX PROGRAMMING (*)

by P. J. LAURENT (1) and C. CARASSO (2)

Abstract — *A general exchange algorithm is given for the minimization of a convex function with equality and inequality constraints. It is a generalization of the Cheney-Goldstein algorithm, but following an idea given by Topfer, a finite sequence of sub-problems the dimension of which is decreasing, is considered at each iteration Given a positive number ε, under very general conditions, it is proved that the method, after a finite number of iterations, leads to an "ε-solution"*

In 1959, Cheney and Goldstein [6] (*see* also Goldstein [7]) proposed an algorithm for solving the problem of minimizing a convex function:

$$f(x) = \max_{t \in S} \left(\sum_{i=1}^{n} b_i(t)\, x_i - c(t) \right)$$

under the constraints:

$$\sum_{i=1}^{n} b_i(t)\, x_i \leqq c(t) \quad \text{for all } t \in U,$$

where S and U are two disjoint compact sets and $b_1, \ldots, b_n$, c are continuous real functions defined on $S \cup U$.

At each iteration ν of this algorithm, a polyhedral approximation of the problem is associated to a suitable subset A^ν consisting of $n+1$ points of $S \cup U$. Using the exchange theorem (Stiefel [11, 12, 13]; *see* also [8, 9]) a new element $t^\nu \in S \cup U$ is introduced: $A^{\nu+1} = (A^\nu \setminus t_0^\nu) \cup t^\nu$.

We propose here a new algorithm which is an extension of the Cheney-Goldstein algorithm for solving the same problem but under much weaker assumptions: the sets S and U are arbitrary and the mappings $b_1, \ldots, b_n$, c are

(*) Reçu décembre 1977
(1) Mathématiques appliquées I M A G Université Scientifique et Médicale de Grenoble
(2) U E R de Sciences, Université de Saint-Étienne, Saint-Étienne

only supposed to be bounded. Moreover, no Haar condition is introduced. At each iteration, we consider a sequence of nested minimization problems. The algorithm is based on an extension of the exchange theorem in which the exchanged quantities are not just a single point (*see* [3, 4]).

The idea of the algorithm is similar to the recursive method introduced by Töpfer [14], [15] (*see* also [3]) for problems of Tchebycheff best approximation.

In the case of a best approximation problem the algorithm becomes an extension of the Rémès algorithm (*see* [5]). For other applications, *see* [2].

1. PROBLEM AND ASSUMPTIONS

We denote by E the n-dimensional Euclidean space and by $\langle x, x' \rangle$ the usual inner-product of x and x' in E.

1.1. The minimization problem

We denote by L a finite set with l elements ($l < n$) and by S and U two arbitrary sets. Suppose that L, S and U have no common point and let $T = S \cup U$.

Let b and c be two bounded mappings from $L \cup T$ into E and $\mathbf{R}$ respectively (i. e., $b(T)$ and $c(T)$ are bounded).

We define the functionals f and g by:

$$f(x) = \operatorname*{Sup}_{t \in S} (\langle x, b(t) \rangle - c(t)),$$

$$g(x) = \operatorname*{Sup}_{t \in U} (\langle x, b(t) \rangle - c(t)).$$

It is easy to see that f and g are continuous convex functionals defined on E with values in $\mathbf{R}$.

We define the affine variety W by:

$$W = \{ x \in E \mid \langle x, b(t) \rangle = c(t),\ t \in L \}.$$

It is convenient to suppose that the $b(t)$, $t \in L$ are linearly independant and that they span a l-dimensional subspace:

$$V = \mathscr{L}(b(t) \mid t \in L).$$

Thus, the affine variety W is parallel to $V^{\perp}$, the orthogonal complement of V, and has the dimension $n - l$.

The problem (P) consists in minimizing $f(x)$ with x satisfying:

$$\langle x, b(t) \rangle = c(t) \qquad \text{for} \quad t \in L$$

and

$$\langle x, b(t)\rangle \leqq c(t) \quad \text{for} \quad t\in U.$$

i. e., $x\in W$ and $g(x)\leqq 0$.

Put:

$$\text{(P)} \quad \alpha = \inf_{\substack{x\in W\\ g(x)\leqq 0}} f(x)$$

and suppose that α is finite.

An element $\bar{x}\in E$ is called a *solution of* (P) if:

$$\bar{x}\in W, \qquad g(\bar{x})\leqq 0 \quad \text{and} \quad f(\bar{x})=\alpha.$$

An element $\tilde{x}\in E$ will be called an ε-*solution of* (P) (with $\varepsilon > 0$) if:

$$\tilde{x}\in W, \qquad g(\tilde{x})\leqq \varepsilon \quad \text{and} \quad f(\tilde{x})\leqq \alpha+\varepsilon.$$

For a given $\varepsilon > 0$ (arbitrarily small), the algorithm that we are going to describe, will give, after a finite number of iterations (depending on ε), an ε-solution of (P). The effective use of the method requires that for numbers $\hat{\varepsilon}$ satisfying $\eta\leqq\hat{\varepsilon}\leqq\varepsilon$ (where η is a positive number such that $\eta < \varepsilon/2^{n-l}$) and for any $x\in W$, it is possible to determine

$$s\in S \quad \text{such that} \quad \langle x, b(s)\rangle - c(s)\geqq f(x)-\hat{\varepsilon}$$

and

$$u\in U \text{ such that } \langle x, b(u)\rangle - c(u)\geqq g(x)-\hat{\varepsilon}.$$

The exact values of $f(x)$ and $g(x)$ are not directly used: only an upper bound in the calculation of the supremum is necessary.

1.2. Assumptions

We assume that:

(H1) there exist $\check{x}\in W$ and $\omega > 0$ such that $\langle \check{x}, b(t)\rangle - c(t)\leqq -\omega$, for all $t\in U$, (this implies the regularity of the constraints);

(H2) the set:

$$K=\{x\in E \mid \langle x, b(t)\rangle = 0,\ t\in L;\ \langle x, b(t)\rangle\leqq 0,\ t\in T\}$$

is a linear subspace.

Note that the preceding set K is equal to the recession cone of all non-empty level sets:

$$S_\lambda = \{ x \in W \,|\, f(x) \leqq \lambda,\ g(x) \leqq 0 \}.$$

Thus, the condition (H2) implies the existence of solutions for the probleme (P).

The condition (H2) is also equivalent to:

(H2′) $0 \in \operatorname{ri} \operatorname{co}(b(T)) + V$,

where ri co $(b(T))$ denotes the relative interior of the convex hull of $b(T)$.

As a consequence of (H2), there exist $\sigma \in \mathbf{R}$ and $\tau \in \mathbf{R}$ such that for all $x \in W$:

(M) $\displaystyle \operatorname*{Sup}_{t \in T} |\langle x, b(t) \rangle| \leqq \sigma \operatorname*{Sup}_{t \in T} \langle x, b(t) \rangle + \tau.$

1.3. Application to best approximation problems

The preceding formulation includes the general problem of best approximation in a finite dimensional subspace with equality and inequality constraints. In this case, the function to minimize is:

$$f(x) = \left\| \sum_{i=1}^{n} x_i y_i - y_0 \right\|,$$

where $y_0, y_1, \ldots, y_n$ are $n+1$ given elements of a normed linear space Y, the norm of which is denoted by $\|y\|$, for $y \in Y$. It is possible to find a subset $S \subset Y'$ (the topological dual of Y) such that f can be written in the following form:

$$f(x) = \operatorname*{Sup}_{y' \in S} (\langle x, b(y') \rangle - c(y')),$$

with

$$b(y') = [(y_1, y'), \ldots, (y_n, y')],$$

$$c(y') = (y_0, y'),$$

where (y, y') represents the value at y of the continuous linear functional $y' \in Y'$. For example, take for S the unit sphere of Y' or the set of its extremal points.

2. MINIMAL CONVEX SUPPORT (m. c. s.)

Subsequently, we will need the notion of minimal convex support of a linear subspace of E. This notion will be used not only relatively to V but also for other linear subspaces occuring in the algorithm.

Let $\mathscr{V}$ be a d-dimensional linear subspace of E spanned by the elements $b(t)$, $t \in D$ (not necessarely independant), where D is a finite subset of $L \cup T$.

2.1. Convex support of a linear subspace

A non-empty and finite subset $A \subset \bar{T}$ will be called a *convex support* of $\mathscr{V}$ if there exist coefficients $\rho(t) \geqq 0$, $t \in A$ satisfying $\sum_{t \in A} \rho(t) = 1$ such that:

$$\sum_{t \in A} \rho(t)\, b(t) \in \mathscr{V}$$

[i. e. if co $(b(A)) \cap \mathscr{V} \neq \emptyset$].

2.2. Minimal convex support of a linear subspace

A convex support A of $\mathscr{V}$ will be called *minimal* if there does not exist a convex support of $\mathscr{V}$ that is strictly included in A.

A subset $A = \{t_1, \ldots, t_{k+1}\}$ consisting of $k+1$ points of T is a minimal convex support (m. c. s.) of $\mathscr{V}$ if and only if:

(a) there exist *positive* coefficients $\rho(t)$, $t \in A$ satisfying $\sum_{t \in A} \rho(t) = 1$ such that $\sum_{t \in A} \rho(t) b(t) \in \mathscr{V}$;

(b) the subspace $\mathscr{L}(b(t) \mid t \in D \cup A)$ spanned by the $b(t)$, $t \in D \cup A$, has the dimension $d+k$.

Every convex support contains at least a m.c.s., and using Caratheodory's theorem, one shows that a m.c.s. contains at most $n-d+1$ elements.

2.3. Coefficients associated with a m.c.s.

One also proves that $A \subset T$ is a m. c. s. if and only if there exist *unique* positive coefficients $\rho_A(t)$, $t \in A$, satisfying:

$$\sum_{t \in A} \rho_A(t) b(t) \in \mathscr{V} \quad \text{and} \quad \sum_{t \in A} \rho_A(t) = 1.$$

These coefficients $\rho_A(t)$, $t \in A$, will be called the coefficients associated with the m. c. s. A. It will also be useful to introduce coefficients $\lambda_A(t)$, $t \in D$, such that:

$$\sum_{t \in A} \rho_A(t) b(t) + \sum_{t \in D} \lambda_A(t) b(t) = 0$$

[these $\lambda_A(t)$ are not necessarily unique].

2.4. Minimization associated with a m.c.s.

Let $A=\{t_1,\ldots,t_{k+1}\}$ be a m.c.s. of V, consisting of $k+1$ elements such that $A\cap S\neq\emptyset$. Put:

$$f_A(x)=\max_{t\in A\cap S}(\langle x, b(t)\rangle-c(t)),$$

$$g_A(x)=\max_{t\in A\cap U}(\langle x, b(t)\rangle-c(t))$$

[if $A\cap U=\emptyset$, then $g_A(x)\equiv-\infty$] and consider the problem (P_A) of minimizing $f_A(x)$ for x belonging to W and satisfying $g_A(x)\leqq 0$. Put:

$$(P_A)\quad \alpha_A=\min_{\substack{x\in W\\ g_A(x)\leqq 0}} f_A(x)\leqq\alpha.$$

We denote by W_A the set of solutions of (P_A), i.e. of elements $\bar{x}\in W$ satisfying $g_A(\bar{x})\leqq 0$, $\alpha_A=f_A(\bar{x})$.

Then we have the following result, the proof of which is simple:

THEOREM: *The amount* α_A *of* (P_A) *is given by:*

$$\alpha_A=\frac{1}{s_A}\sum_{t\in A}\rho_A(t)(\langle x_0, b(t)\rangle-c(t)),$$

(*where* x_0 *is an arbitrary element of* W *and* $s_A=\sum_{t\in A\cap S}\rho_A(t)$), *or by:*

$$\alpha_A=-\frac{1}{s_A}\Big(\sum_{t\in L}\lambda_A(t)c(t)+\sum_{t\in A}\rho_A(t)c(t)\Big).$$

The set of solutions in given by:

$$W_A=\{x\in W\,|\,\langle x, b(t)\rangle-c(t)+\delta(t)\alpha_A=\alpha_A,\ t\in A\},$$

where

$$\delta(t)=\begin{cases}0 & \text{if } t\in S,\\ 1 & \text{if } t\in U.\end{cases}$$

Thus, the set W_A is an affine variety which is parallel to $V_A^{\perp}$, with:

$$V_A=\mathscr{L}(b(t)\,|\,t\in L\cup A),$$

the dimension of which is equal to $l+k$. Therefore, the dimension of W_A is $n-(l+k)$ and the solution of (P_A) is unique when $k=n-l$.

For a given $\varepsilon > 0$, the algorithm that we are going to describe, will build a finite sequence $A^0, A^1, \ldots, A^\mu$ of m.c.s. of V and of associated solutions x^0, $x^1, \ldots, x^\mu$ ($x^\nu \in W_{A^\nu}$) such that the corresponding amounts $\alpha^0, \alpha^1, \ldots, \alpha^\mu$ ($\alpha^\nu = \alpha_{A^\nu}$) build a non-decreasing sequence such that:

$$x^\mu \in W, \qquad g(x^\mu) \leqq \varepsilon \quad \text{and} \quad f(x^\mu) \leqq \alpha^\mu + \varepsilon.$$

As $\alpha^\mu \leqq \alpha$, the element x^μ will be an ε-solution of (P).

3. A CONVERGENCE RESULT

The convergence result that we will state in this section corresponds, for the time being, to a theoretical algorithm, the use of which is not quite specified. We will need it, on the one hand, as a guide for the definition of the effective algorithm, on the other hand as a basis for proving its convergence.

Let A^ν, $\nu = 0, 1, \ldots$, be an infinite sequence of m.c.s. of V such that $A^\nu \cap S \neq \emptyset$, $\nu = 0, 1, \ldots$; and let $f^\nu = f_{A^\nu}$, $g^\nu = g_{A^\nu}$ be the associated polyhedral functionals (as in § 2.4.). Put:

$$\alpha^\nu = \min_{\substack{x \in W \\ g^\nu(x) \leqq 0}} f^\nu(x) \leqq \alpha$$

and denote by $W^\nu = W_{A^\nu}$ the set of solutions. Let $\rho^\nu(t) = \rho_{A^\nu}(t)$, $t \in A^\nu$, be the positive coefficients corresponding to A^ν (as in § 2.3.) and let $s^\nu = \sum_{t \in A^\nu \cap S} \rho^\nu(t)$.

Consider the functional h^ν defined by:

$$h^\nu(x) = \max(f(x); g(x) + \alpha^\nu)$$

$$= \sup_{t \in T} (\langle x, b(t) \rangle - c(t) + \delta(t)\alpha^\nu).$$

Finally, suppose that the set:

$$\hat{N} = \{\nu \in \mathbf{N} \,/\, A^{\nu+1} \neq A^\nu\},$$

is infinite and let $\hat{A}^\nu = A^{\nu+1} \setminus A^\nu$, for $\nu \in \hat{N}$.

Then, we have the following result:

3.1. THEOREM: *If, for all $\nu \in \hat{N}$, there exists $x^\nu \in W^\nu$ such that:*

$$h^\nu(x^\nu) - (\langle x^\nu, b(t) \rangle - c(t) + \delta(t)\alpha^\nu) \leqq \hat{\varepsilon},$$

for all $t \in \hat{A}^\nu$, *then we have:*

$$\alpha^{\nu+1} \geqq \alpha^\nu + \frac{1}{s^{\nu+1}} \Big(\sum_{t \in \hat{A}^\nu} \rho^{\nu+1}(t) \Big) \big(h^\nu(x) - \alpha^\nu - \hat{\varepsilon} \big)$$

(*this proves that* $\alpha^{\nu+1} > \alpha^\nu$, *as long as* $h^\nu(x^\nu) - \alpha^\nu - \hat{\varepsilon} > 0$) *and for any* $\varepsilon > \hat{\varepsilon}$, *there exists* $\mu \in \hat{N}$ *such that:*

$$h^\mu(x^\mu) - \alpha^\mu \leqq \varepsilon$$

[*what implies that* x^μ *is an* ε-*solution of* (P)].

Proof:

First part: By theorem 2.4, we have:

$$\alpha^{\nu+1} = \frac{1}{s^{\nu+1}} \sum_{t \in A^{\nu+1}} \rho^{\nu+1}(t) \, \big(\langle x^\nu, b(t) \rangle - c(t) \big).$$

For $t \in \hat{A}^\nu$, we have: $\langle x^\nu, b(t) \rangle - c(t) + \delta(t)\alpha^\nu \geqq h^\nu(x^\nu) - \hat{\varepsilon}$.
For $t \in A^{\nu+1} \cap A^\nu$, we have: $\langle x^\nu, b(t) \rangle - c(t) + \delta(t)\alpha^\nu = \alpha^\nu$.
Hence:

$$\begin{aligned} \alpha^{\nu+1} &\geqq \frac{1}{s^{\nu+1}} \sum_{t \in A^{\nu+1} \cap A^\nu} \rho^{\nu+1}(t) (\alpha^\nu - \delta(t)\alpha^\nu) \\ &\quad + \frac{1}{s^{\nu+1}} \sum_{t \in \hat{A}^\nu} \rho^{\nu+1}(t) (h^\nu(x^\nu) - \delta(t)\alpha^\nu - \hat{\varepsilon}) \\ &= \frac{1}{s^{\nu+1}} \sum_{t \in A^{\nu+1}} \rho^{\nu+1}(t) \, (\alpha^\nu - \delta(t)\alpha^\nu) \\ &\quad + \frac{1}{s^{\nu+1}} \sum_{t \in \hat{A}^\nu} \rho^{\nu+1}(t) \, (h^\nu(x^\nu) - \alpha^\nu - \hat{\varepsilon}) \end{aligned}$$

and finally:

$$\alpha^{\nu+1} \geqq \alpha^\nu + \frac{1}{s^{\nu+1}} \Big(\sum_{t \in \hat{A}^\nu} \rho^{\nu+1}(t) \Big) \, (h^\nu(x^\nu) - \alpha^\nu - \hat{\varepsilon}).$$

Second part: The following lemma is a property of the positive coefficients that are associated with a sequence of m.c.s. We will only use the fact that A^ν, $\nu = 0, 1, \ldots$, is a sequence of m.c.s. such that $\hat{N}$ is infinite.

LEMMA: *There exist an infinite subset* $\tilde{N}$ *in* $\hat{N}$ *and a bipartition of* A^ν *in* B^ν *and* $C^\nu \neq \emptyset$ *for* $\nu \in \hat{N}$ *such that:*

1° $\lim_{\nu \in \tilde{N}} \Big(\sum_{t \in B^\nu} \rho^\nu(t) \Big) = 0;$

2° *there exists* $m > 0$ *such that* $\rho^{\nu}(t) \geqq m$, *for all* $t \in C^{\nu}$ *and all* $\nu \in \tilde{N}$;

3° C^{ν} *is not included in* A^{ν^-}, *for all* $\nu \in \tilde{N}$, *where* ν^- *denotes the integer preceding* ν *in* $\tilde{N}$.

The proof of this lemma has been given in [4].

Third part: Suppose that we have $h^{\nu}(x^{\nu}) - \alpha^{\nu} > \varepsilon$, for all $\nu \in \hat{N}$ (with $\varepsilon > \hat{\varepsilon}$) and prove that this leads to a contradiction.

As a consequence of the first part, we then have:

$$\alpha^{\nu+1} \geqq \alpha^{\nu} + \frac{1}{s^{\nu+1}} (\sum_{t \in \hat{A}^{\nu}} \rho^{\nu+1}(t)\,(\varepsilon - \hat{\varepsilon}) \qquad \text{for all} \quad \nu \in \hat{N},$$

hence, α^{ν} is a non decreasing sequence such that $\alpha^{\nu} \leqq \alpha$.

Let $\tilde{\alpha}$ ($\tilde{\alpha} \leqq \alpha$) be its limit.

First we show that there exists a positive constant s such that

$$s^{\nu} = \sum_{t \in A^{\nu} \cap S} \rho^{\nu}(t) \geqq s > 0.$$

It is equivalent to prove that $C^{\nu} \cap S \neq \emptyset$ for all $\nu \in \tilde{N}$, sufficiently large.

By theorem 2.4, we have:

$$\langle x^{\nu}, b(t) \rangle - c(t) = \alpha^{\nu} - \delta(t)\,\alpha^{\nu} \qquad \text{for} \quad t \in A^{\nu}$$

and by (H1):

$$\langle \check{x}, b(t) \rangle - c(t) \leqq -\omega \qquad (\text{with } \omega > 0) \quad \text{for} \ \ t \in A^{\nu}.$$

The mappings b and c being bounded and α^{ν} converging to $\tilde{\alpha}$, it is possible to find a constant ξ such that:

$$|\langle x^{\nu} - \check{x}, b(t) \rangle| \leqq \xi \quad \text{for all} \ \ t \in A^{\nu} \text{ and all } \nu \in \tilde{N}.$$

Suppose that there exists an infinite subset $\approx{N} \subset \tilde{N}$ such that $C^{\nu} \cap S \neq \emptyset$, forall $\nu \in \underset{\approx}{N}$ and prove that this leads to a contradiction. Consider $\sum_{t \in A^{\nu}} \rho^{\nu}(t) \langle x^{\nu} - \check{x}, b(t) \rangle$ and decompose the sum according to B^{ν} and C^{ν}.

Letting $\eta^{\nu} = \sum_{t \in B^{\nu}} \rho^{\nu}(t)$, we have:

$$\Big| \sum_{t \in B^{\nu}} \rho^{\nu}(t) \langle x^{\nu} - \check{x}, b(t) \rangle \Big| \leqq \eta^{\nu} m,$$

with $\lim_{\nu \in \tilde{N}} \eta^{\nu} = 0$.

On the other hand, as $C^\nu \subset U$, we have $\langle x^\nu, b(t)\rangle - c(t) = 0$, for all $t \in C^\nu$, and letting $\theta^\nu = \sum_{t \in C^\nu} \rho^\nu(t)$, we obtain:

$$\sum_{t \in C^\nu} \rho^\nu(t) \langle x^\nu - \mathring{x}, b(t)\rangle \geqq \theta^\nu \omega \qquad \text{for all} \quad \nu \in \tilde{\tilde{N}},$$

with $\lim_{\nu \in \tilde{N}} \theta^\nu = 1$ and $\omega > 0$.

Thus, for $\nu \in \tilde{\tilde{N}}$ sufficiently large, we would have $\sum_{t \in A^\nu} \rho^\nu(t) \langle x^\nu - \mathring{x}, b(t)\rangle > 0$, what is in contradiction with the facts that A^ν is a m.c.s. of V and that $x^\nu - \mathring{x} \in V^\perp$.

Using the result of the second part, we see that for all $\nu \in \tilde{N}$, the set C^ν is not included in A^{ν^-}. Therefore, there exist elements of C^ν that have been introduced between the iteration ν^- and the iteration ν. Let $\hat{\nu} \in \hat{N}$ be the last iteration satisfying $\nu^- \leqq \hat{\nu} < \nu$ for which at least one element of C^ν *has been introduced*.

By theorem 2.4, we have:

$$\alpha^\nu = \frac{1}{s^\nu} \sum_{t \in A^\nu} \rho^\nu(t) \left(\langle x^{\hat{\nu}}, b(t)\rangle - c(t)\right)$$

with $s^\nu \geqq s > 0$.

We decompose again the sum according to B^ν and C^ν:

(α) *Sum corresponding to* B^ν.

As a consequence of (H2), [*see* (M) in § 1.2.], for all $t \in T$, we have:

$$|\langle x^{\hat{\nu}}, b(t)\rangle| \leqq \sigma \operatorname{Sup}_{t \in T} \langle x^{\hat{\nu}}, b(t)\rangle + \tau.$$

As the mapping c is bounded and the sequence α^ν is also bounded, we can find a constant χ such that:

$$|\langle x^{\hat{\nu}}, b(t)\rangle - c(t)| \leqq \sigma h^\nu(x^{\hat{\nu}}) + \chi,$$

for all $t \in T$. Hence, we have:

$$\left| \frac{1}{s^\nu} \sum_{t \in B^\nu} \rho^\nu(t) \left(\langle x^{\hat{\nu}}, b(t)\rangle - c(t)\right) \right| \leqq \frac{1}{s} \eta^\nu (\sigma h^\nu(x^{\hat{\nu}}) + \chi).$$

(β) *Sum corresponding to* C^ν.

For $t \in C^\nu \cap \hat{A}^{\hat{\nu}} \neq \emptyset$: $\langle x^{\hat{\nu}}, b(t)\rangle - c(t) \geqq h^{\hat{\nu}}(x^{\hat{\nu}}) - \delta(t)\alpha^{\hat{\nu}} - \hat{\varepsilon}$.

For $t \in C^\nu \setminus \hat{A}^{\hat{\nu}}$: $\langle x^{\hat{\nu}}, b(t)\rangle - c(t) = \alpha^{\hat{\nu}} - \delta(t)\alpha^{\hat{\nu}}$.

Hence, we have:

$$\frac{1}{s^\nu}\sum_{t\in C^\nu}\rho^\nu(t)\,(\langle x^{\hat\nu}, b(t)\rangle - c(t))$$

$$\geqq \frac{1}{s^\nu}\Big(\sum_{t\in C^\nu\cap \hat A^{\hat\nu}}\rho^\nu(t)\,(h^{\hat\nu}(x^{\hat\nu})-\delta(t)\,\alpha^{\hat\nu}-\hat\varepsilon)$$

$$+\sum_{t\in C^\nu\setminus \hat A^{\hat\nu}}\rho^\nu(t)(\alpha^{\hat\nu}-\delta(t)\alpha^{\hat\nu})\Big)$$

$$=\frac{1}{s^\nu}\Big(\sum_{t\in C^\nu\cap \hat A^{\hat\nu}}\rho^\nu(t)\Big)\,(h^{\hat\nu}(x^{\hat\nu})-\alpha^{\hat\nu}-\hat\varepsilon)+\frac{1}{s^\nu}\alpha^{\hat\nu}\Big(\sum_{t\in C^\nu}\rho^\nu(t)\,(1-\delta(t))\Big)$$

$$\geqq u^\nu\alpha^{\hat\nu}+m(h^{\hat\nu}(x^{\hat\nu})-\alpha^{\hat\nu}-\hat\varepsilon),$$

with $u^\nu=(1/s^\nu)\sum_{t\in C^\nu}\rho^\nu(t)\,(1-\delta(t))$ satisfying $\lim_{\nu\in\tilde N} u^\nu=1$.

Finally, joining (α) and (β) together, we obtain:

$$\alpha^\nu \geqq u^\nu\alpha^{\hat\nu}+m(h^\nu(x^{\hat\nu})-\alpha^{\hat\nu}-\hat\varepsilon)-\frac{\eta^\nu}{s}(\sigma h^{\hat\nu}(x^{\hat\nu})+\chi)$$

$$=u^\nu\alpha^{\hat\nu}+m(v^\nu h^{\hat\nu}(x^{\hat\nu})-\alpha^{\hat\nu}-\hat\varepsilon)-\eta^\nu\frac{\chi}{s},$$

with $v^\nu=1-(\eta^\nu\sigma/sm)$, satisfying $\lim_{\nu\in\tilde N} v^\nu=1$.

As α^ν converges towards $\tilde\alpha$, we deduce that:

$$\limsup_{\nu\in\tilde N}(v^\nu h^{\hat\nu}(x^{\hat\nu})-\alpha^{\hat\nu}-\hat\varepsilon)\leqq 0$$

hence:

$$\limsup_{\nu\in\tilde N} v^\nu h^{\hat\nu}(x^{\hat\nu})\leqq\tilde\alpha+\hat\varepsilon.$$

It is easy to prove that this inequality is contradictory with:

$$h^\nu(x^\nu)\geqq\alpha^\nu+\varepsilon \qquad \text{for all} \quad \nu\in\hat N,$$

in the cas where $\varepsilon>\hat\varepsilon$.

3.2. Remark: If we suppose that $A^0\cap S\neq\emptyset$, then, as long as we have $h^\nu(x^\nu)-\alpha^\nu-\hat\varepsilon>0$, then we have $A^{\nu+1}\cap S\neq\emptyset$:

As a matter of fact, if we would have $A^{\nu+1}\cap S=\emptyset$, this would mean that $A^{\nu+1}\cap A^\nu\subset U$ and that $\hat A^\nu\subset U$; hence

$$\langle x^\nu, b(t)\rangle - c(t)=0 \qquad \text{for all} \quad t\in A^{\nu+1}\cap A^\nu,$$

$$\langle x^\nu, b(t)\rangle - c(t)\geqq h^\nu(x^\nu)-\alpha^\nu-\hat\varepsilon>0 \qquad \text{for all} \quad t\in\hat A^\nu.$$

As we have, by (H1):

$\langle \check{x}, b(t) \rangle - c(t) < 0 \qquad \text{for all} \quad t \in A^{\nu+1},$

we would deduce that:

$$\sum_{t \in A^{\nu+1}} \langle x^{\nu} - \check{x}, b(t) \rangle > 0,$$

in contradiction with the facts that $A^{\nu+1}$ is a m. c. s. of V and that $\check{x}^{\nu} - \check{x} \in V^{\perp}$.

4. EXCHANGE THEOREM

The preceding convergence theorem shows that the sets $\hat{A}^{\nu}$ of new elements should be such that it is possible to exchange them with a subset C^{ν} of A^{ν} in such a way that $A^{\nu+1} = (A^{\nu} \setminus C^{\nu}) \cup \hat{A}^{\nu}$ is again a m. c. s. of V.

The next theorem shows how to operate this exchange. Subsequently we will have to do this operation, not only relatively to V but also for other linear subspaces occuring in the algorithm.

Let $\mathscr{V}$ be a d-dimensional linear subspace of E defined by:

$\mathscr{V} = \mathscr{L}(b(t) \mid t \in D)$

where D is a finite subset of $L \cup T$.

4.1. Exchange theorem

If A_0 is a m. c. s. of $\mathscr{V}$ and if A_1 is a m. c. s. of

$\mathscr{V}_0 = \mathscr{L}(b(t) \mid t \in D \cup A_0)$

then, there exists a bipartition of A_0 in B_0 and $C_0 \neq \emptyset$ such that:

$\tilde{A}_0 = B_0 \cup A_1$ *is a m. c. s. of* $\mathscr{V}$,

$\tilde{A}_1 = C_0$ *is a m. c. s. of*

$\tilde{\mathscr{V}}_0 = \mathscr{L}(b(t) \mid t \in D \cup \hat{A}_0).$

This theorem has been proved in [4]. It shows that it is possible to exchange with A_1 a non-empty part C_0 of A_0, in such a way that:

$\tilde{A}_0 = (A_0 \setminus C_0) \cup A_1$

is again a m. c. s. of $\mathscr{V}$.

4.2. Practice of the exchange

Denote by $\rho_0(t) > 0$, $t \in A_0$ and by $\lambda_0(t)$, $t \in D$ the coefficients associated with A_0, as in paragraph 2.3:

$$(a_0) \quad \sum_{t \in A_0} \rho_0(t) b(t) + \sum_{t \in D} \lambda_0(t) b(t) = 0,$$

$$\sum_{t \in A_0} \rho_0(t) = 1.$$

As A_1 is a m. c. s. of $\mathscr{V}_0$, we denote by $\rho_1(t) > 0$, $t \in A_1$ and by $\lambda_1(t)$, $t \in D \cup A_0$, the corresponding coefficients:

$$(a_1) \quad \sum_{t \in A_1} \rho_1(t) b(t) + \sum_{t \in D \cup A_0} \lambda_1(t) b(t) = 0,$$

$$\sum_{t \in A_1} \rho_1(t) b(t) = 0.$$

Substracting r-times the relation (a_0) from the relation (a_1), we obtain:

$$\sum_{t \in A_0} \rho_0(t) \left(\frac{\lambda_1(t)}{\rho_0(t)} - r \right) b(t) + \sum_{t \in A_1} \rho_1(t) b(t)$$

$$+ \sum_{t \in D} (\lambda_1(t) - r \lambda_0(t)) b(t) = 0.$$

If we choose $r = \min_{t \in A_0} (\lambda_1(t)/\rho_0(t))$, and we define:

$$C_0 = \left\{ t \in A_0 \,\middle|\, \frac{\lambda_1(t)}{\rho_0(t)} = r \right\} \quad \text{and} \quad B_0 = A_0 \setminus C_0$$

then the preceding relation becomes:

$$(\tilde{a}_0) \quad \sum_{t \in B_0 \cup A_1} \tilde{\rho}_0(t) b(t) + \sum_{t \in D} \tilde{\lambda}_0(t) b(t) = 0,$$

with:

$$\tilde{\rho}_0(t) = \begin{cases} \dfrac{1}{q} (\lambda_1(t) - r \rho_0(t)) & \text{if} \quad t \in B_0, \\ \dfrac{1}{q} \rho_1(t) & \text{if} \quad t \in A_1, \end{cases}$$

$$\tilde{\lambda}_0(t)=\frac{1}{q}\rho_1(t) \qquad \text{for} \quad t\in D,$$

$$q=\sum_{t\in B_0}(\lambda_1(t)-r\,\rho_0(t))+\sum_{t\in A_1}\rho_1(t).$$

The coefficients $\tilde{\rho}_0(t)$, $t\in\tilde{A}_0=B_0\cup A_1$ are positive with the sum equal to one.

Now the relation (a_0) can be written:

$$(\tilde{a}_1) \quad \sum_{t\in C_0}\tilde{\rho}_1(t)\,b(t)+\sum_{t\in D\cup B_0\cup A_1}\tilde{\lambda}_1(t)\,b(t)=0$$

with:

$$\tilde{\rho}_1(t)=\frac{1}{p}\rho_0(t) \qquad \text{for} \quad t\in C_0,$$

$$\tilde{\lambda}_1(t)=\begin{cases}\dfrac{1}{p}\lambda_0(t) & \text{if} \quad t\in D,\\[2mm] \dfrac{1}{p}\rho_0(t) & \text{if} \quad t\in B_0,\\[2mm] 0 & \text{if} \quad t\in A_1.\end{cases}$$

$$p=\sum_{t\in C_0}\rho_0(t).$$

The coefficients $\tilde{\rho}_1(t)$, $t\in\tilde{A}_1=C_0$ are positive with the sum equal to 1.

5. STRING OF M. C. S.

5.1. Successive minimization

The convergence theorem (§ 3) and the exchange theorem (§ 4) lead us to consider the following sub-problem:

$$(\text{SP}^{\nu}) \quad \beta^{\nu}=\operatorname*{Inf}_{x\in W^{\nu}} h^{\nu}(x)$$

with

$$h^{\nu}(x)=\max(f(x);\, g(x)+\alpha^{\nu})$$

$$=\operatorname*{Sup}_{t\in T}(\langle x,\, b(t),\rangle-c(t)+\delta(t)\,\alpha^{\nu}).$$

This sub-problem can be solved by the algorithm described in [4]:

Let A_1^ν be a m. c. s. of V, with $k_1^\nu+1$ elements,

$V^\nu = V_0^\nu = \mathscr{L}(b(t) \mid t \in L \cup A^\nu)$,

and denote by h_1^ν the polyhedral functional defined by:

$$h_1^\nu(x) = \max_{t \in A_1^\nu} (\langle x, b(t) \rangle - c(t) + \delta(t)\alpha^\nu).$$

We consider the minimization of $h_1^\nu(x)$ for $x \in W^\nu = W_0^\nu$.

Put:

$$\alpha_1^\nu = \min_{x \in W_0^\nu} h_1^\nu(x).$$

The set W_1^ν of solutions, can be written:

$$W_1^\nu = \{ x \in W_0^\nu \mid \langle x, b(t) \rangle - c(t) + \delta(t)\alpha^\nu = \alpha_1^\nu,\ t \in A_1^\nu \}.$$

It is an affine variety, which is parallel to $(V_1^\nu)^\perp$, where

$V_1^\nu = \mathscr{L}(b(t) \mid t \in L \cup A_0^\nu \cup A_1^\nu)$

is a $l + k_0^\nu + k_1^\nu$-dimensional linear subspace of E.

The same construction can be repeated relatively to V_1^ν: A_2^ν is a m. c. s. of V_1^ν, with $k_2^\nu+1$ elements, h_2^ν is the associated functional, α_2^ν the amount of its minimum on W_1^ν, W_2^ν the set of solutions, and $V_2^\nu = \mathscr{L}(b(t) \mid t \in L \cup A_0^\nu \cup A_1^\nu \cup A_2^\nu)$, the dimension of which is $l + \sum_{i=0}^{2} k_i^\nu$.

We continue this construction until we have $V_{m^\nu}^\nu = E$, hence $W_{m^\nu}^\nu$ is reduced to a single point.

5.2. String of m.c.s.

The preceding construction leads us to the notion of a string of m. c. s. (shortly "string"):

A finite sequence $\mathscr{C} = (A_0, \ldots, A_m)$ of subsets $A_i \subset T$ will be called a *string*, if, setting $V_{-1} = V$, we have:

A_i *is a m.c.s. of* V_{i-1},

$V_i = \mathscr{L}(b(t) \mid t \in L;\ t \in A_j, j = 0, \ldots, i)$,

$i = 1, \ldots, m$,

$V_m = E$.

If the subset A_j contains k_j+1 elements ($j=0,\ldots,m$), then the dimension of V_i is $l+\sum_{j=0}^{i} k_j$. Associated with each A_i of the string $\mathscr{C}$, we can define the coefficients $\rho_i(t)>0$, $t\in A_i$ and $\lambda_i(t)$, $t\in L\cup\left(\bigcup_{j=0}^{i-1} A_j\right)$ such that:

$$\sum_{t\in A_i} \rho_i(t)\,b(t)+\sum_{t\in L}\lambda_i(t)\,b(t)+\sum_{j=0}^{i-1}\sum_{t\in A_j}\lambda_i(t)\,b(t)=0,$$

$$\sum_{t\in A_i}\rho_i(t)=1.$$

5.3. Solution associated with a string

A string $\mathscr{C}=(A_0,\ldots,A_m)$ will be said *correct* if $A_0\cap S\neq\emptyset$. Put:

$$\left.\begin{aligned} f_i(x)&=\max_{t\in A_i\cap S}(\langle x, b(t)\rangle-c(t))\\ g_i(x)&=\max_{t\in A_i\cap U}(\langle x, b(t)\rangle-c(t))\end{aligned}\right\}\ i=0,\ldots,m.$$

We consider the sequence of successive minimization problems associated with a correct string $\mathscr{C}$:

$$\alpha_0=\min_{\substack{x\in W\\ g_0(x)\leqq 0}} f_0(x)$$

the set of solutions W_0 of which is an affine variety parallel to $V_0^{\perp}$, and

$$\alpha_i=\min_{x\in W_{i-1}}\max\,(f_i(x);\ g_i(x)+\alpha_0),$$

$i=1,\ldots,m$, the set of solutions W_i of which is an affine variety parallel to $V_i^{\perp}$.

As $V_m=E$, the affine variety W_m is reduced to a single point $x=x_{\mathscr{C}}$ that we will call the solution associated with the string $\mathscr{C}$. As $x=x_{\mathscr{C}}$ is a solution of the successive minimization problems, by theorem 2.4 above and theorem 2.3 of [4], it is characterized by the following conditions:

$$\begin{aligned}&\langle x, b(t)\rangle=c(t), && t\in L && (l \text{ conditions})\\ &\langle x, b(t)\rangle+\delta(t)\alpha_0-\alpha_i=c(t), && t\in A_i && (k_i+1 \text{ conditions}),\end{aligned}$$

$i=0,\ldots,m.$

We can use these $n+m+1$ linear equations for computing the $n+m+1$ unknown $x_1, \ldots, x_n, \alpha_0, \ldots, \alpha_m$. By construction, this linear algebraic system has a unique solution.

5.4. Exchange operation in a string

Let $\mathscr{C}=(A_0, \ldots, A_m)$ be a string and $(V_0, \ldots, V_m)$ the corresponding linear subspaces. We see that A_{j-1} is a m. c. s. of $\mathscr{V}=V_{j-2}$ and that A_j is a m. c. s. of the linear subspace:

$$\mathscr{V}_0=\mathscr{L}(b(t) \mid t \in D \cup A_{j-1})=V_{j-1},$$

with $D=L \cup \bigcup_{i=0}^{j-2} A_i$.

Thus we have the same situation as in theorem 4.1.

There exists a bipartition of A_{j-1} in B_{j-1} and $C_{j-1} \neq \emptyset$ such that, letting:

$$\tilde{A}_{j-1}=B_{j-1} \cup A_j \quad \text{and} \quad \tilde{A}_j=C_{j-1},$$

then $\tilde{\mathscr{C}}=(A_0, \ldots, \tilde{A}_{j-1}, \tilde{A}_j, \ldots, A_m)$ is again a string.

We will say that we have exchanged A_{j-1} and A_j in the string $\mathscr{C}$.

5.5. Regular string

A string $\mathscr{C}=(A_0, \ldots, A_m)$ will be said *regular* if each of the subsets A_i, $i=0, \ldots, m$, has at least two elements. Thus, if $\mathscr{C}$ is regular, the dimension of V_i is strictly greater than the dimension of V_{i-1} $(i=0, \ldots, m)$ and the integer m is necessarily smaller or equal to $n-l-1$.

If $\mathscr{C}$ is an arbitrary string, we obtain a regular string by taking away all the m. c. s. that are reduced to a single point. If A_0 is not reduced to a single point, this operation does not change the solution associated with the string as well as the amounts α_i corresponding to the remaining m. c. s. A_i.

6. ALGORITHM

If $\varepsilon > 0$ is the desired accuracy, let ε_i be positive numbers satisfying:

$$(\star) \quad \varepsilon_0=\varepsilon, \quad \varepsilon_{i+1}<\frac{\varepsilon_i}{2}, \qquad i=0, \ldots, n_0.$$

6.1. Description of the algorithm

Suppose that, at the iteration ν, we have a correct and regular string $\mathscr{C}^\nu=\{A_0^\nu, \ldots, A_{m^\nu}^\nu\}$, and denote by x^ν the associated solution and by $\alpha_0^\nu, \ldots, \alpha_{m^\nu}^\nu$ the corresponding amounts.

Determine $t^\nu \in T$ such that

$$h^\nu(x^\nu) - (\langle x^\nu, b(t^\nu)\rangle - c(t^\nu) + \delta(t^\nu)\alpha_0^\nu) \leqq \varepsilon_{m^\nu+1}$$

with

$$h^\nu(x) = \operatorname*{Sup}_{t \in T}(\langle x, b(t)\rangle - c(t) + \delta(t)\alpha_0^\nu)$$

and put:

$$A_{m^\nu+1}^\nu = \{t^\nu\},$$
$$\alpha_{m^\nu+1}^\nu = \langle x^\nu, b(t^\nu)\rangle - c(t^\nu) + \delta(t^\nu)\alpha_0^\nu.$$

We define the integer j^ν by:

$$j^\nu = \min(j \mid 0 \leqq j \leqq m^\nu + 1;\ \alpha_{m^\nu+1}^\nu + \varepsilon_{m^\nu+1} \leqq \alpha_j^\nu + \varepsilon_j)$$

[the fact that the above inequality is satisfied for a given integer j means that the corresponding sub-problem (*see* 5.1 and 5.3) has been sufficiently solved].

We will consider three cases according to the value of j^ν:

First case:

$$j^\nu = m^\nu + 1.$$

We introduce the new point t^ν in the string $\mathscr{C}^\nu$.

Using the exchange theorem, in the string:

$$(A_0^\nu, \ldots, A_{m^\nu}^\nu, A_{m^\nu+1}^\nu = \{t^\nu\})$$

we exchange $A_{m^\nu}^\nu$ and $A_{m^\nu+1}^\nu$. Then, we obtain:

either $\quad (A_0^\nu, \ldots, \tilde{A}_{m^\nu}^\nu, A_{m^\nu+1})$

in which $\tilde{A}_{m^\nu}^\nu$ contains t^ν but is not reduced to this single point,

or $\quad (A_0^\nu, \ldots, A_{m^\nu-1}^\nu, \{t^\nu\}, A_{m^\nu}^\nu)$,

what occurs in the case $b(t^\nu) \in V_{m^\nu-1}^\nu$.

In this latter case, we exchange again $A_{m^\nu-1}^\nu$ and $\{t^\nu\}$, and so on, until we finally obtain:

either $\quad \tilde{\mathscr{C}}^\nu = (A_0^\nu, \ldots, \tilde{A}_{i^\nu-1}^\nu, \tilde{A}_{i^\nu}^\nu, \ldots, A_{m^\nu}^\nu) \qquad (i^\nu \geqq 1)$,

in which $\tilde{A}_{i^\nu-1}^\nu$ contains t^ν but is not reduced to this single point,

or $\quad (\{t^\nu\}, A_0^\nu, \ldots, A_{m^\nu}^\nu)$.

But in this latter case, this means that x^ν is an $\varepsilon_{m^\nu+1}$-solution of (P), with $\varepsilon_{m^\nu-1}<\varepsilon$: As a matter of fact, $A^\nu_{m^\nu+1}=\{t^\nu\}$ is then a m.c.s. of V. By the remark 3.2. we will have $t^\nu \in S$ and thus:

$$\alpha^\nu_{m^\nu+1}=\langle x^\nu, b(t^\nu)\rangle - c(t^\nu)$$

will satisfy:

$$\alpha^\nu_{m^\nu+1} \leqq \alpha.$$

By the choice of t^ν, we have:

$$h^\nu(x^\nu) \leqq \alpha^\nu_{m^\nu+1}+\varepsilon_{m^\nu+1},$$

what implies that $h^\nu(x^\nu) \leqq \alpha+\varepsilon_{m^\nu+1}$, i.e. x^ν is a $\varepsilon_{m^\nu+1}$-solution of (P) and we stop the algorithm.

Second case:

$$1 \leqq j^\nu \leqq m^\nu.$$

Using again the exchange theorem, we then exchange $A^\nu_{j^\nu-1}$ and $A^\nu_{j^\nu}$ in the string $\mathscr{C}^\nu$ (*see* § 5.4). This leads to the new string:

$$\widetilde{\mathscr{C}}^\nu=(A^\nu_0, \ldots, \widetilde{A}^\nu_{j^\nu-1}, \widetilde{A}^\nu_{j^\nu}, \ldots, A^\nu_{m^\nu}).$$

Note that $\widetilde{A}^\nu_{j^\nu-1}$ cannot be reduced to a single point, for it contains $A^\nu_{j^\nu}$ and the string $\mathscr{C}^\nu$ has been supposed to be regular.

Third case:

$$j^\nu=0.$$

Then, we have:

$$h^\nu(x^\nu) \leqq \alpha^\nu_{m^\nu+1}+\varepsilon_{m^\nu+1} \leqq \alpha^\nu_0+\varepsilon_0$$

and this means that x^ν is an ε_0-solution of (P) and we stop the algorithm. In short, if we put:

$$k^\nu=\begin{cases} i^\nu & \text{if } j^\nu=m^\nu+1, \\ j^\nu & \text{if } 0 \leqq j^\nu \leqq m^\nu, \end{cases}$$

we see that we stop the computation (the accuracy ε being obtained) when $k^\nu=0$. In the other cases, the last exchange executed concerns the m.c.s. the indices of which are $k^\nu-1$ and k^ν. The m.c.s. A^ν_i, $i=0, \ldots, k^\nu-2$ are not modified.

It can happen that the new m.c.s. $\widetilde{A}^\nu_{k^\nu}$ is reduced to a single point. In that case, we suppress it in the string (*see* § 5.5). Thus, we obtain a new regular string:

$$\mathscr{C}^{\nu+1}=(A^{\nu+1}_0, \ldots, A^{\nu+1}_{m^{\nu+1}})$$

in which $m^{\nu+1}$ can be equal to $m^{\nu}-1$, m^{ν} or $m^{\nu}+1$.

By the remark 3.2, the string $\mathscr{C}^{\nu+1}$ is also correct ($A_0^{\nu+1} \cap S \neq \emptyset$).

6.2. Properties of the algorithm

Suppose that $k^{\nu} \geqq 1$. Then we have:

(*a*) $A_k^{\nu+1} = A_k^{\nu}$, $\alpha_k^{\nu+1} = \alpha_k^{\nu}$, $k=0, \ldots, k^{\nu}-2$.

(*b*) *For all* $t \in \hat{A}^{\nu} = A_{k^{\nu}-1}^{\nu+1} \setminus A_{k^{\nu}-1}^{\nu}$:

$$h^{\nu}(x^{\nu}) - (\langle x^{\nu}, b(t) \rangle - c(t) + \delta(t)\alpha_0^{\nu}) \leqq \varepsilon_{k^{\nu}}.$$

(*c*) $\alpha_{k^{\nu}-1}^{\nu+1} \geqq \alpha_{k^{\nu}-1}^{\nu} + \gamma^{\nu}(\varepsilon_{k^{\nu}-1} - \varepsilon_{k^{\nu}})$, where γ^{ν} is a positive number given by:

$$\gamma^{\nu} = \begin{cases} \dfrac{1}{s^{\nu+1}} \sum_{t \in \hat{A}^{\nu}} \rho_{k^{\nu}-1}^{\nu+1}(t) & \text{if } k^{\nu}=1, \\ \sum_{t \in \hat{A}^{\nu}} \rho_{k^{\nu}-1}^{\nu+1}(t) & \text{if } k^{\nu} \geqq 2. \end{cases}$$

Proof: The point (*a*) follows directly from the definition of the algorithm.

(*b*) We will consider two cases:

first case:

$j^{\nu} = m^{\nu}+1; \qquad k^{\nu} = i^{\nu}.$

As we exchange $A_{k^{\nu}-1}^{\nu}$ and $\{t^{\nu}\}$, we have: $A_{k^{\nu}-1}^{\nu+1} \setminus A_{k^{\nu}-1}^{\nu} = \{t^{\nu}\}$ and by the choice of t^{ν}, we have:

$$\langle x^{\nu}, b(t^{\nu}) \rangle - c(t^{\nu}) + \delta(t^{\nu})\alpha_0 \geqq h^{\nu}(x^{\nu}) - \varepsilon_{m^{\nu}+1} \geqq h^{\nu}(x^{\nu}) - \varepsilon_{k^{\nu}}.$$

second case:

$1 \leqq j \leqq m^{\nu}; \qquad k^{\nu} = j^{\nu}.$

We exchange $A_{k^{\nu}-1}^{\nu}$ and $A_{k^{\nu}}^{\nu}$. Thus we have $\hat{A}^{\nu} = A_{k^{\nu}}^{\nu}$. Now, for all $t \in A_{k^{\nu}}^{\nu}$, we have:

$$\langle x^{\nu}, b(t) \rangle - c(t) + \delta(t)\alpha_0^{\nu} = \alpha_{k^{\nu}}^{\nu}$$

and by definition of $k^{\nu} = j^{\nu}$:

$$\alpha_{k^{\nu}}^{\nu} + \varepsilon_{k^{\nu}} \geqq \alpha_{m^{\nu}+1}^{\nu} + \varepsilon_{m^{\nu}+1} \geqq h^{\nu}(x^{\nu}).$$

(*c*) The proof is similar to the first part of the proof of theorem 3.1. We will not give it here.

6.3. Starting the algorithm

Generally, we wish to start the algorithm with an initial string $\mathscr{C}^0$ consisting of a single m. c. s. A_0^0 of V (with exactly $n-l+1$ elements). The determination of A_0^0 can be difficult (even impossible). In the case, it is possible to modify the problem (P) without changing its amount and one part of its solutions in such a way that the determination of A_0^0 for the new problem is very easy.

Suppose we know $x_0 \in E$ and $r \in \mathbf{R}$ such that the problem (P) has at least a solution $\bar{x}$ satisfying: $\| \bar{x} - x_0 \| \leqq r$ and let $\theta \in \mathbf{R}$ be constant such that $\theta < \alpha$. Consider then the function:

$$z(x) = \eta \| x - x_0 \| + \theta$$

where $\eta > 0$ satisfies the condition $\theta + \eta r \leqq \alpha$, and the new minimization problem:

$$(\tilde{\mathrm{P}}) \quad \tilde{\alpha} = \inf_{\substack{x \in W \\ g(x) \leqq 0}} \tilde{f}(x)$$

where $\tilde{f}(x) = \max\,(f(x); z(x))$.

It is easy to prove that $\alpha = \tilde{\alpha}$ and that the set of solutions of $(\tilde{\mathrm{P}})$ is exactly equal to the set of solution $\bar{x}$ of (P) satisfying the condition:

$$\| \bar{x} - x_0 \| \leqq \frac{\alpha - \theta}{\eta}.$$

Note that the function $z(x)$ can be written

$$z(x) = \sup_{x' \in S'} (\langle x, \eta x' \rangle + \theta - \eta \langle x_0, x' \rangle)$$

where S' is the unit sphere of E. Thus, the function $\tilde{f}$ has the same form as f, replacing b and c by suitable extensions $\tilde{b}$ and $\tilde{c}$ to $S \cup S'$. It is easy to choose A_0^0 in S'.

7. CONVERGENCE OF THE ALGORITHM

Before proving the convergence, we need a theoretical convergence result which is very similar to theorem 3.1 but corresponds to the form of the sub-problems (*see* § 5.1).

7.1. An auxiliary convergence result

Suppose that $\mathscr{W}$ is an affine variety which is parallel to $\mathscr{V}^{\perp}$ (where $\mathscr{V}$ is defined as in paragraph 2) and consider the following minimization problem:

$$(\mathrm{SP}) \quad \beta = \inf_{x \in \mathscr{W}} h_0(x),$$

with

$$h_0(x) = \max(f(x); g(x)+\alpha_0)$$

$$= \sup_{t \in T}(\langle x, b(t)\rangle - c(t) + \delta(t)\alpha_0).$$

Let A^ν, $\nu = 0, 1, \ldots$, be an infinite sequence of m. c. s. of $\mathscr{V}$ and let f^ν and g^ν be the corresponding functionals (as in § 3). Put:

$$\alpha^\nu = \min_{x \in \mathscr{W}} \max(f^\nu(x); g^\nu(x)+\alpha_0) \leqq \beta$$

and denote by $\mathscr{W}^\nu$ the set of solutions.

If we suppose again that the set:

$$\hat{N} = \{\nu \in \mathbf{N} \mid A^{\nu+1} \neq A^\nu\}$$

is infinite (put $\hat{A} = A^{\nu+1} \setminus A^\nu$, for $\nu \in \hat{N}$) then we have the following result:

THEOREM: *If, for all* $\nu \in \hat{N}$, *there exists* $x^\nu \in \mathscr{W}^\nu$ *such that:*

$$h_0(x^\nu) - (\langle x^\nu, b(t)\rangle - c(t) + \delta(t)\alpha_0) \leqq \hat{\varepsilon}$$

for all $t \in \hat{A}^\nu$, *then for any* $\varepsilon > \hat{\varepsilon}$, *there exists* $\mu \in \hat{N}$ *such that:*

$$h_0(x^\mu) - \alpha^\mu < \varepsilon$$

[*this implies that* x^μ *is an* ε*-solution of* (SP)].

This result is in fact a particular case of theorem 6.1 in [4].

7.2. Convergence of the algorithm

THEOREM: *For an arbitrary positive number* ε, *the algorithm described in paragraph* 6, *after a finite number* μ *of iterations, leads to an element* $x^\mu \in W$ *which is an* ε*-solution of* (P).

More precisely, for a given accuracy $\varepsilon > 0$, there exists an integer μ (depending on ε) such that the element $x^\mu \in W$ and the first m. c. s. A_0^μ of the string $\mathscr{C}^\mu$ satisfy:

$$f(x^\mu) \leqq \alpha_0^\mu + \varepsilon \qquad \text{and} \qquad g(x^\mu) \leqq \varepsilon,$$

where $\alpha_0^\mu = \alpha_{A_0^\mu}$ is the corresponding amount (*see* § 2.4). As $\alpha_0^\mu \leqq \alpha$, this implies that x^μ is an ε-solution of (P).

Proof: We only have to prove that the algorithm stops, i. e. that there exists μ such that $k^\mu = 0$.

Suppose that we have $k^\nu \geqq 1$, for all ν and show that this leads to a contradiction:

Let $\tilde{k} = \liminf_{\nu \to \infty} k^\nu$, $(1 \leqq \tilde{k} \leqq n_0 + 1)$. There exists ν_0 such that for all $\nu \geqq \nu_0$, we have $k^\nu \geqq \tilde{k}$; and the set:

$$\tilde{N} = \{ \nu \in \mathbf{N} \mid \nu \geqq \nu_0,\ k^\nu = \tilde{k} \}$$

is infinite.

Hence, for $\nu \geqq \nu_0$, we have $A_k^\nu = A_k$ and $V_k^\nu = V_k$ (independant of ν) for $k = 0, \ldots, \tilde{k} - 2$.

By the definition of the algorithm, we have:

$$\overset{\nu}{\alpha}_{m^\nu+1} + \varepsilon_{m^\nu+1} > \overset{\nu}{\alpha}_{\tilde{k}-1} + \varepsilon_{\tilde{k}-1} \quad \text{for all } \nu \in \tilde{N}.$$

As $\tilde{k} \leqq m^\nu + 1$, we have $\varepsilon_{m^\nu+1} \leqq \varepsilon_{\tilde{k}}$, hence:

(i) $\overset{\nu}{\alpha}_{m^\nu+1} - \overset{\nu}{\alpha}_{\tilde{k}-1} > \varepsilon_{\tilde{k}-1} - \varepsilon_{\tilde{k}}$ for all $\nu \in \tilde{N}$.

Put $V_{\tilde{k}-2} = \mathscr{V}$. Thus $\overset{\nu}{A}_{\tilde{k}-1}$ is a m. c. s. of $\mathscr{V}$. For all $\nu \geqq \nu_0$, such that $\nu \notin \tilde{N}$, we have $\overset{\nu+1}{A}_{\tilde{k}-1} = \overset{\nu}{A}_{\tilde{k}-1}$ and for all $\nu \in \tilde{N}$, by 6.2 *b*, we have:

$$h^\nu(x^\nu) - [\langle x^\nu, b(t) \rangle - c(t) + \delta(t)\, \alpha_0^\nu] \leqq \varepsilon_{\tilde{k}}$$

for all $t \in \hat{A}^\nu = \overset{\nu+1}{A}_{\tilde{k}-1} \setminus \overset{\nu}{A}_{\tilde{k}-1}$.

The choice of the ε_i [*see* condition ($\star$) in § 6] implies that $\varepsilon_{\tilde{k}-1} - \varepsilon_{\tilde{k}} > \varepsilon_{\tilde{k}}$. Using theorem 3.1 in the case $\tilde{k} = 1$ and theorem 7.1 in the case $\tilde{k} > 1$ (with $\hat{\varepsilon} = \varepsilon_{\tilde{k}}$ and $\varepsilon = \varepsilon_{\tilde{k}-1} - \varepsilon_{\tilde{k}}$) there exists $\mu \in \tilde{N}$ such that:

$$h^\mu(x^\mu) - \overset{\mu}{\alpha}_{\tilde{k}-1} \leqq \varepsilon_{\tilde{k}-1} - \varepsilon_{\tilde{k}}.$$

As we have $\overset{\mu}{\alpha}_{m^\mu+1} \leqq h^\mu(x^\mu)$, we obtain:

(ii) $\overset{\mu}{\alpha}_{m^\mu+1} - \overset{\mu}{\alpha}_{\tilde{k}-1} \leqq \varepsilon_{\tilde{k}-1} - \varepsilon_{\tilde{k}}$.

The two inequalities (i) and (ii) are contradictory.

REFERENCES

1. C. Carasso, *L'algorithme d'échange en optimisation convexe*, Thèse, Grenoble, 1973.
2. C. Carasso, *Un algorithme de minimisation de fonctions convexes avec ou sans contraintes : l'algorithme d'échange*. 7th I.F.I.P. Conference on Optimization Techniques, Springer-Verlag, 1975, 2, pp. 268-282.

3 C CARASSO and P J LAURENT, *Un algorithme pour la minimisation d'une fonctionnelle convexe sur une variété affine* Séminaire d'Analyse numerique, Grenoble, 18 octobre 1973

4 C CARASSO and P J LAURENT, *Un algorithme de minimisation en chaîne en optimisation convexe* Séminaire d'Analyse numérique, Grenoble, 29 janvier 1976 S I A M J Control and Optimization, vol 16, No 2, 1978 pp 209-235

5 C CARASSO and P J LAURENT, *Un algorithme general pour l'approximation au sens de Tchebycheff de fonctions bornées sur un ensemble quelconque* Approximations-Kolloquium, Bonn, June 8-12, 1976, Lecture Notes in Math , No 556, Springer-Verlag, Berlin, 1976

6 E W CHENEY and A A GOLDSTEIN, *Newton's Method for Convex Programming and Tchebyscheff Approximation*, Num Math , vol 1, 1959, pp 253-268

7 A A GOLDSTEIN, *Constructive Real Analysis, Harper's Series in Modern Mathematics*, Harper and Row, 1967

8 P J LAURENT, *Approximation et Optimisation*, Hermann Paris, 1972

9 P J LAURENT, *Exchange Algorithm in Convex Analysis* Conference on Approximation theory, The Univ of Texas, Austin, 1973, Acad Press

10 R T ROCKAFELLAR, *Convex Analysis*, Princ Univ Press 1970

11 E L STIEFEL, *Uber diskrete und lineare Tschebyscheff-Approximationen*, Num Math , vol 1, 1959, pp 1-28

12 E L STIEFEL *Numerical Methods of Tschebycheff Approximation* In *On numerical Approximation*, R LANGER, Ed , Univ of Wisconsin, 1959, pp 217-232

13 E L STIEFEL, *Note on Jordan Elimination, Linear Programming and Tchebycheff Approximation*, Num Math , vol 2, 1960, pp 1-17

14 H J TOPFER, *Tschebyscheff-Approximation bei nicht erfullter Haarscher Bedingung*, Zeits fur angew Math und Mech , vol 45, 1965, T 81-T 82

15 H J TOPFER, *Tschebyscheff-Approximation und Austauschverfahren bei nicht erfullter Haarscher Bedingung* Tagung, Oberwolfach, 1965 I S N M 7, Birkhauser Verlag 1967, pp 71-89