UN ALGORITHME DE MINIMISATION DE FONCTIONS CONVEXES AVEC OU SANS CONTRAINTES " L'ALGORITHME D'ÉCHANGES"

C.CARASSO

Université de Saint-Etienne

23, rue du Dr Paul Michelon

42100 SAINT-ETIENNE FRANCE

Résumé.

On étudie un algorithme permettant de trouver le minimum d'une forme linéaire sur une intersection finie ou infinie de demi-espaces. Cet algorithme s'applique à une large classe de problèmes d'optimisation convexe (programmation mathématique, théorie de l'approximation,...)

I. INTRODUCTION

On s'interesse à la résolution d'un problème d'optimisation dans $\mathbb{R}^n$ du type :

(1-1) (P) $\alpha = \text{Inf}[(x|z) \mid x \in C \cap V_t]$

avec $C = \{x \in \mathbb{R}^n \mid c(s) \leq (x|a(s)) \quad \forall s \in S\}$, dans lequel z est un élément fixe de $\mathbb{R}^n$, S est un compact quelconque, c et a deux applications continues de S dans respectivement $\mathbb{R}$ et $\mathbb{R}^n$. L'ensemble V_t désigne le translaté d'un sous espace vectoriel V de dimension p de $\mathbb{R}^n$. On supposera que V_t est de la forme :

$$V_t = \{x \in \mathbb{R}^n \mid (x|u_i) = c_i \quad i = p+1,\dots,n\}.$$

La plupart des problèmes d'optimisation convexe peuvent se mettre sous la forme (P). Nous verrons au paragraphe 3 des exemples de problème mis sous cette forme. La forme (P) n'est pas la plus générale; on peut, en utilisant systématiquement la dualité et l'analyse convexe, étudier le problème (P) en se plaçant dans un espace vectoriel topologique localement convexe X en dualité avec X'. Pour une étude de ce type nous renvoyons à [3] et [7].

Nous étudions un algorithme appelé "algorithme d'échange" qui permet, moyennant des hypothèses sur (P) et sur le déroulement de l'algorithme, de construire une suite d'éléments x^k de V_t tels que :

$$\lim_{k\to\infty}(x^k|z) = \alpha \text{ et } \max_{s\in S}[c(s)-(x^k|a(s))] \leq \eta^k \qquad \eta^k>0$$

et $\lim_{k\to\infty} \eta^k = 0$.

Lorsqu'on s'interesse à la recherche du meilleur approximant dans un espace vectoriel normé, l' "algorithme d'échange" est l'algorithme de REMES généralisé ([6]). La convergence est alors démontrée sans faire l'hypothèse classique (et forte !) de "condition de HAAR vérifiée" (voir [1]).

II. DESCRIPTION DE L'ALGORITHME D'ECHANGE

On suppose que α est fini et que z n'appartient pas à $\mathcal{L}(u_{p+1},\dots,u_n) = V$.

On fera les hypothèses suivantes :

H1 <u>Il existe $x^* \in V_t$ tel que $c(s) < (x^* | a(s))$ pour tout $s \in S$</u>

H2 <u>Il existe $u_1,\dots,u_p \in a(S)$ et $\rho_1,\dots,\rho_n \in \mathbb{R}$ avec $\rho_i > o$ $i=1,\dots,p$ tels que</u> :

i) $\mathcal{L}(u_1,\dots,u_n) = \mathbb{R}^n$

ii) $\sum_{i=1}^{n} \rho_i u_i = z$

En notant $\mathcal{L}(u_1,\dots,u_n)$ le sous-espace vectoriel engendré par $u_1,\dots,u_n$.

<u>Passage de l'itération k à l'itération k+1</u>

Soient β et ε des scalaires positifs donnés et "petits".

On suppose qu'à l'itération k on dispose de p éléments $u_1,\dots,u_p \in a(S)$, de p éléments $c_1,\dots,c_p \in \mathbb{R}$ tels que $u_i = a(s_i)$ et $c_i = c(s_i)$ $(i=1,\dots,p)$ et de n coefficients $\rho_1,\dots,\rho_n \in \mathbb{R}$ tels que :

$$\sum_{i=1}^{n} \rho_i u_i = z \qquad \text{avec } \rho_i > o \quad i=1,\dots,p \tag{2-1}$$

et d'une précision η $(o \leq \eta)$.

Pour passer à l'itération k+1 on franchit les étapes suivantes :

étape 1 : Déterminer $x \in \mathbb{R}^n$ tel que :

$$(x|u_i) = c_i \qquad i=1,\dots,n \tag{2-2}$$

étape 2 : Faire $\eta := \eta/2$

étape 3 : Calculer $\text{APPUI}(x,\eta) = (d,\bar{u},\bar{c}) \in \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}$

étape 4 : si $d<\beta$ aller à l'étape 10 sinon aller à l'étape 5.

étape 5 : Déterminer $\rho' = (\rho'_1,\dots,\rho'_n) \in \mathbb{R}^n$ tel que

$$\sum_{i=1}^{n} \rho'_i u_i = -\bar{u} \tag{2-3}$$

étape 6 : Déterminer $\alpha_o \in \mathbb{R}$ et $\hat{i} \in \{1,\dots,p\}$ tels que :

$$\alpha_o = \frac{\rho'_{\hat{i}}}{\rho_{\hat{i}}} = \min_{i=1,\dots,p} \frac{\rho'_i}{\rho_i} \tag{2-4}$$

étape 7 : Echanger $u_{\hat{i}}$ avec $\bar{u}$ $(u_{\hat{i}} := \bar{u})$ et $c_{\hat{i}}$ avec $\bar{c}$ $(c_{\hat{i}} := \bar{c})$

étape 8 : Poser

(2-5) $\qquad \rho_i := \rho_i - \dfrac{\rho_i'}{\alpha_o} \qquad i = 1,\ldots,n \quad i \neq \hat{\imath}$

$\qquad \rho_{\hat{\imath}} := -\dfrac{1}{\alpha_o}$.

Si $\rho_i > o$ pour $i=1,\ldots,p$ aller à l'étape 1 sinon aller à l'étape 9.

étape 9 : L'algorithme n'est pas "itératif".

étape 10: Si $\beta+\eta > \varepsilon$ aller à l'étape 2 sinon x est solution du problème

$$\mathrm{Inf}\left[(x|z) \mid c(s) \leqq (x|a(s) + \varepsilon \quad \forall s \in S \quad x \in V_t\right]$$

L'application APPUI qui intervient à l'étape 3 est une application qui, à $x \in \mathbb{R}^n$ et $\eta \in \mathbb{R}$ associe l'élément $(d,\bar{u},\bar{c})$ $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$ tel que si l'on pose :

$$\bar{d} = \underset{s \in S}{\mathrm{Max}}\left[c(s)-(x|a(s))\right] \quad \text{on ait :}$$

$$o \leqq d = \bar{c}-(x|\bar{u}) \quad \text{et}$$

$$\bar{d}-\eta \leqq d \leqq \bar{d}$$

REMARQUES.

1°/Si on note k le numéro de l'itération on a :

$$u_i = a(s_i^k) \quad i=p+1,\ldots,n \ ; \quad c_i=c(s_i^k) \ ; \ \bar{u}=c(\hat{s}^k)$$

$$\bar{c} = c(\hat{s}^k) \quad \text{et } \rho_i=\rho_i^k \qquad (s_i^k \in S \quad \text{et} \quad \hat{s}^k \in S).$$

2°/ La matrice du système linéaire (2-3) est la transposée de la matrice du système linéaire (2-2). On peut donc numériquement utiliser (2-2) pour résoudre (2-3). De même, à chaque passage à l'étape 1, le système linéaire à résoudre ne diffère du précédent que par une ligne; il n'est donc en général pas utile de recommencer intégralement sa résolution.

3°/ L'élément η qui intervient dans APPUI à l'étape 3 représente la précision avec laquelle est approché $\underset{s \in S}{\mathrm{Max}}\left[c(s)-(x|a(s))\right]$. Dans de nombreux problèmes on peut prendre $\eta = o$.

4°/ Après passage à l'étape 8 on a :

(2-6) $\qquad \displaystyle\sum_{\substack{i=1 \\ i \neq \hat{\imath}}}^{n} \rho_i u_i + \rho_{\hat{\imath}} \bar{u} = z \quad \text{ou} \quad \sum_{\substack{i=1 \\ i \neq \hat{\imath}}}^{p} \rho_i^{k+1} a(s_i^k) + \sum_{i=p+1}^{n} \rho_i^{k+1} u_i + \rho_{\hat{\imath}}^{k+1} a(\hat{s}^k) = z$

5°/ Les éléments a,c et S du problème (P) n'interviennent qu'à l'étape 3 dans APPUI.

6°/ Pour permettre de calculer aisément les éléments de départ (itération o) vérifiant (2-1) on peut résoudre le problème :

(P') $\qquad \alpha = \mathrm{Min}\left[(x|z) \mid c(s) \leqq (x|a(s)) \quad s \in S, \ \|x\| \leqq r \text{ et } x \in V_t\right]$

où r est suffisamment grand pour que (P) et (P') aient la même solution.

En remarquant que $\|x\| \leqq r$ peut s'écrire sous la forme :

$$-r \leqq (x|t) \quad \forall t \in S' = \{x \in \mathbb{R}^n \mid \|x\| \leqq 1\}$$

le problème (P') s'écrit :

(P') $\qquad \alpha = \mathrm{Min}[(x|z) \mid c'(s) \leqq (x|a'(s)) \quad \forall s \in S \cup S' \quad \text{et} \quad x \in V_t]$

avec

$$a'(s) = \begin{cases} a(s) & \text{si } s \in S \\ s & \text{si } s \in S' \end{cases} \qquad c'(s) = \begin{cases} c(s) & \text{si } s \in S \\ r & \text{si } s \in S \end{cases}$$

L'application APPUI du problème (P') étant l'application qui, à $x \in \mathbb{R}^n$, associe $(d', \bar{u}', \bar{c}') \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$ définis par :

$$d' = \mathrm{Max}\,[d, \|x\| - r] \qquad (\bar{u}', c') = \begin{cases} (\bar{u}, \bar{c}) & \text{si } d' = d \\ (-\frac{x}{\|x\|}, -r) & \text{si } d' = \|x\| - r \end{cases}$$

les éléments d, $\bar{u}$ et $\bar{c}$ étant définis par l'application APPUI du problème (P).
Les éléments de départ pour (P') sont pris sous la forme $u_i = \varepsilon_i s_i$ $(i=1,\dots,p)$ avec $s_i \in S'$ et ε_i=signe ρ_i $(i=1,\dots,p)$

III. EXEMPLES D'APPLICATIONS

Nous donnons deux exemples classiques d'applications; pour d'autres exemples voir [4].

3-1 MINIMISATION D'UNE FONCTION CONVEXE DERIVABLE AVEC CONTRAINTES DERIVABLES.

On considère le problème :

(P) $\qquad \alpha = \mathrm{Inf}[g(t^*, x) \mid g(t,x) \leqq o \quad \forall t \in K \quad \text{et} \quad x \in W_t]$

où les applications $x \longmapsto g(t,x)$ de $\mathbb{R}^{n-1}$ dans $\mathbb{R}$ sont, pour tout $t \in K \cup \{t^*\}$ convexes et dérivables; K est un compact et on a :

$$W_t = \{x \in \mathbb{R}^{n-1} \mid (x|u_i') = c_i \qquad i=p+1,\dots,n\}$$

On suppose qu'il existe $x_0 \in \mathbb{R}^{n-1}$ tel que l'ensemble :

$$L = \{x \in \mathbb{R}^{n-1} \mid g(t^*, x) < g(t^*, x_0) \quad g(t,x) \leqq o\}$$ soit non vide et compact

Pour tout $x \in L$ on a ([8] p.242) :

$$g(t,x) = \underset{y \in L}{\mathrm{Max}}[\, g(t,y) + (g_x'(t,y) \mid x-y)]$$

où $g_x'(t,y)$ désigne le gradient en y de $x \mapsto g(t,x)$. On suppose $(t,y) \longmapsto g_x'(t,y)$ continue.

Le problème (P) peut donc s'écrire :

$$\alpha = \mathrm{Inf}[((x,x_n) \mid (o,1)) \mid g(t,y) - (g_x'(t,y) \mid y) \leqq ((x,x_n) \mid (-g_x'(t,y), \delta_t))$$
$$\forall (t,y) \in (K \cup \{t^*\}) \times L \quad (x,x_n) \in V_t]$$

avec $\quad V_t = \{(x,x_n) \in \mathbb{R}^{n-1} \times \mathbb{R} \mid (x,x_n)|u_i) = c_i \quad i=p+1,\dots,n\}$

où $u_i = (u_i',o)$ et $\delta_t = \begin{cases} 1 & \text{si } t = t^* \\ o & \text{sinon} \end{cases}$

On retrouve la forme (1-1). L'application APPUI de ce problème fait correspondre à $(x,x_n) \in \mathbb{R}^{n-1} \times \mathbb{R}$ et $o \leqq \eta$ l'élément $(d,\bar{u},\bar{c}) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$ défini par :

$$\bar{u} = (-g_x'(\tilde{t},x),\ \delta_{\tilde{t}})$$

$$\bar{c} = g(\tilde{t},x) - (g_x'(\tilde{t},x)|x)$$

$$o \leqq d = g(\tilde{t},x) - \delta_{\tilde{t}} x_n$$

avec $\quad \bar{d} - \eta \leqq d \leqq \bar{d} \quad$ où $\quad \bar{d} = \underset{t \in K \cup \{t^*\}}{\mathrm{Max}} [g(t,x) - \delta_t x_n]$

REMARQUE :

Si K est fini, on a aisément $d=\bar{d}$; on peut donc prendre $\eta=o$. S'il n'y a pas de contrainte il suffit de faire $K = \phi$. Si les fonctions $g(t,.)$ sont linéaires et l'ensemble K fini l'algorithme d'échange est l'algorithme du simplexe appliqué au problème dual de (P).

Les hypothèses H1 et H2 s'écrivent :

H1 Il existe $x^* \in W_t$ tel que $g(t,x^*) < o$ pour tout t de K.

H2 Il existe $(t_1,y_1),\dots,(t_p,y_p) \in (K \cup \{t^*\}) \times L$

et $\rho_1,\dots,\rho_n \in \mathbb{R}$ avec $\rho_i > o$ $i=1,\dots,p$ tels que :

i) $\displaystyle\sum_{i=1}^{p} \rho_i g_x'(t_i,y_i) + \sum_{i=p+1}^{n} \rho_i u_i' = o$ et $\displaystyle\sum_{i=1}^{p} \rho_i \delta_{t_i} = 1$

ii) $\mathbb{R}^n = \mathcal{L}((-g_x'(t_i,y_i),\ \delta_{t_i})\ i=1,\dots,p;\ (u_i',o)\ i=p+1,\dots,n)$

3-2 MEILLEUR APPROXIMANT DANS UN CONVEXE D'UN ESPACE NORMÉ.

On considère le problème :

(P) $$\alpha = \mathrm{Inf}\left[\|f_o - \sum_{i=1}^{n-1} x_i f_i\| \mid g(t,x) \leqq o \quad \forall t \in K \text{ et } x \in W_t\right]$$

où $\|.\|$ représente la norme d'un espace vectoriel E de dual topologique E'; $f_i \in E$ $i=o,\dots,n-1$; K est compact, $x \mapsto g(t,x)$ est dérivable et on a :

$$W_t = \{x \in \mathbb{R}^{n-1} \mid (x|u_i') = c_i \quad i=p+1,\dots,n\}$$

En remarquant que ([9]) : $\quad \|g\| = \underset{\ell \in S'}{\mathrm{Max}} <g,\ell> \quad \forall g \in E$

où $S' = \{\ell \in E' \mid \|\ell\| \leq 1\}$ $\quad <g,\ell> = \ell(g)$, le problème (P) s'écrit :

$$\alpha = \mathrm{Inf}[((x,x_n)|(o,1)) \mid <f_o,\ell> \leqq ((x,x_n)|(<f_1,\ell>,\dots,<f_{n-1},\ell>,1))\ \forall \ell \in S'$$

$$g(t,y)-(g'_x(t,y)|y) \leq ((x,x_n)|(-g'_x(t,y),o)) \ \forall (t,y) \in K\times L$$
$$\text{et} \quad (x,x_n) \in V_t]$$

où L est un ensemble supposé non vide et compact de la forme :

$$L = \{x \in \mathbb{R}^{n-1} \mid \|f_o - \sum_{i=1}^{n-1} x_i f_i\| < \|f_o - \sum_{i=1}^{n-1} x_i^o f_i\| \quad \text{et} \quad g(t,x) \leqq o \quad \forall t \in K\}$$

L'application APPUI, de ce problème qui a la forme (1-1), fait correspondre à $(x,x_n) \in \mathbb{R}^{n-1}\times\mathbb{R}$ et $o \leqq \eta$ (à partir des éléments

$$d_1 = \|f - \sum_{i=1}^{n-1} x_i f_i\| - x_n \quad , \qquad d_2 = \underset{t\in K}{\text{Max}}\ g(t,x),$$

$$\bar{d} = \text{Max}(d_1,d_2),$$

$$\bar{\ell} \in S' \quad \text{tel que} : d_1+x_n-\eta \leqq \langle f-\sum_{i=1}^{n-1} x_i f_i, \bar{\ell}\rangle \leqq d_1+x_n \quad \text{et} \quad \tilde{t} \in K$$

tel que $d_2-\eta \leqq g(\tilde{t},x) \leqq d_2$) l'élément $(d,\bar{u},\bar{c}) \in \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}$ tel que :

$$\bar{d}-\eta \leqq d \leqq \bar{d}$$

$$\bar{u} = \begin{cases} (\langle f_1,\bar{\ell}\rangle,\ldots,\langle f_{n-1},\bar{\ell}\rangle,1) \\ (-g'_x(\tilde{t},x),o) \end{cases} \qquad \text{et} \quad \bar{c} = \begin{cases} \langle f_o,\bar{\ell}\rangle & \text{si } \bar{d}=d_1 \\ g(\tilde{t},x)-(g'_x(\tilde{t},x)|x) & \text{si } \bar{d}=d_2 \end{cases}$$

REMARQUE :

Dans C(K) muni de la norme du max on retrouve l'algorithme de REMES, algorithme généralisé par LAURENT dans le cas d'un espace normé ([6]).

IV. CONVERGENCE DE L'ALGORITHME D'ECHANGE

THEOREME DE CARACTERISATION.

Si H1 est vérifiée alors $\bar{x} \in C \cap V_t$ est solution de (P) si et seulement si, il existe k ($1 \leqslant k \leqslant p$) éléments $s_i \in S$, **k** coefficients $\rho_i > o$ tels que :

i) $c(s_i) = (\bar{x}|a(s_i)) \quad i=1,\ldots,k$

ii) $\bar{\ell} = \sum_{i=1}^{p} \rho_i a(s_i) \in V^\perp + z$

DEMONSTRATION:

On a $\alpha = \underset{x\in\mathbb{R}^n}{\text{Inf}} f(x)$ avec $f(x)=(x|z)+\chi_C(x)+\chi_{V_t}(x)$. L'élément $\bar{x}$ est solution de (P) si et seulement si $o \in \partial f(\bar{x})$. Les fonctionnellles χ_C et χ_{V_t} étant d'après H1 finies et continues en x^* on a :

$$\partial f(\bar{x}) = \{z\} + \partial\chi_C(\bar{x}) + \partial\chi_{V_t}(\bar{x})$$

On a $C = \{x \in \mathbb{R}^n \mid r(x) \leqq o\}$ avec $r(x) = \underset{s\in S}{\mathrm{Max}}[c(s)-(x|a(s))]$;

donc ([6] p.386 et p.355) :

$\partial\chi_C(\bar{x}) = \mathrm{cc}\,\partial r(\bar{x})$ et $\partial r(\bar{x}) = \overline{\mathbf{co}}\; a(F(\bar{x}))$ avec

$F(\bar{x}) = \{s \in S \mid c(s) = (\bar{x}|a(s))\}$ donc, $F(\bar{x})$ étant compact :

$\partial\chi_C(\bar{x}) = -\mathrm{cc}\; a(F(\bar{x}))$. On a aussi $\partial\chi_V(\bar{x}) = V^{\perp}$. La relation $o \in \partial f(\bar{x})$ est donc équivalente à l'existence d'un élément $\bar{\ell} \in \mathrm{cc}\; a(F(\bar{x})) \cap (V^{\perp}+z)$ qui s'écrit aussi sous la forme ii) en appliquant le théorème de CARATHEODORY ([6] p.74).

Q.E.D.

THEOREME DE DUALITE (**voir** [4] pour la démonstration).

Si H1 et H2 sont vérifiées alors :

i) Il existe $\bar{x} \in C \cap V_t$ tel que $(\bar{x}|z) = \alpha$

ii) $$\alpha = (\tilde{x}|z) + \mathrm{Max}\left[\sum_{i=1}^{k} \rho_i c(s_i) - (\tilde{x}|\sum_{i=1}^{k}\rho_i a(s_i)) \,\middle|\, 1 \leqq k \leqq p \quad \rho_i > o;\ s \in S \quad \sum_{i=1}^{k}\rho_i a(s_i) \in V^{\perp} + z\right]$$

où $\tilde{x}$ est un élément fixe de V_t et $V^{\perp} = \mathcal{L}(u_{p+1},\dots,u_n)$.

On suppose vérifiées les hypothèses H1 et H2.

LEMME 4-1 :

A chaque itération k on a $\rho_i \geqq o$ $i=1,\dots,p$ et $\mathbb{R}^n = \mathcal{L}(u_i \quad i=1,\dots,n)$.

DEMONSTRATION :

1°/ Montrons que dans (2-4) on a $\alpha_o < o$. Si on avait $\alpha_o \geqq o$ on aurait $\rho_i' \geqq o$ $i=1,\dots,p$ et en multipliant scalairement (2-3) par $x^k \in V_t$ on aurait :

$$\sum_{i=1}^{p} \rho_i' c(s_i^k) + (x^k|a(\hat{s}^k)) = o$$

(en posant $x=x^k$, $u_i=a(s_i^k)$, $c_i=c(s_i^k)$ $\bar{u}=a(\hat{s}^k)$, $\bar{c}=c(\hat{s}^k)$, $\bar{d}=\bar{d}^k$ les éléments obtenus à l'étape k).

Par définition on a :

$$o < c(\hat{s}^k) - (x^k|a(\hat{s}^k)) = \bar{d}^k - \beta^k$$

d'où

$$\sum_{j=1}^{p} \rho_j' c(\hat{s}_j^k) + c(\hat{s}^k) = \bar{d}^k - \beta^k$$

Mais d'après H1 :

$$o = \sum_{j=1}^{p} \rho_j'(x^*|a(s_j^k))+(x^*|a(\hat{s}^k)) > \sum_{j=1}^{p} \rho_j' c(s_j^k)+c(\hat{s}^k) = \bar{d}^k - \beta^k$$

ce qui est impossible, donc $\alpha_o < o$.

2°/ Supposons qu'à l'instant k on ait :

(4-1) $\quad \mathbb{R}^n = \mathcal{L}(u_i;\ i=1,\dots,n) \quad$ et $\quad \sum_{i=1}^{n} \rho_i u_i = z$

Montrons que ces relations sont encore vérifiées à l'itération k+1. Raisonnons par l'absurde. Soit $x \neq o$ tel que $x \in V^{\perp} = \mathcal{L}(u_{p+1},\dots,u_n)$ et $x = \sum_{i=1}^{p} \mu_i a(s_i^{k+1})$

on a aussi : $x = \sum_{\substack{i=1 \\ i \neq \hat{\imath}}}^{p} \mu_i a(s_i^k) + \mu_{\hat{\imath}^k} a(\hat{s}^k) \in V^{\perp}$ d'où :

$$a(\hat{s}^k) = v + \sum_{\substack{i=1 \\ i \neq \hat{\imath}^k}}^{p} \alpha_i a(s_i^k)$$ avec $v \in V^{\perp}$. En portant cette valeur dans (2-6)

il vient : $\sum_{i=1}^{p} (\rho_i^{k+1} + \alpha_i \rho_{\hat{\imath}^k}^{k+1}) u_i + \sum_{i=p+1}^{n} \rho_i^{k+1} u_i = z$ ce qui est contradictoire avec (4-1).

Q.E.D.

On supposera que dans le déroulement de l'algorithme d'échange décrit au paragraphe 2 on ne passe jamais par l'étape 9 soit :

<u>H3 Quelle que soit l'itération k on a $\rho_i > o \quad i=1,\dots,p$</u>

On dit alors que l'algorithme est <u>itératif</u>.

Remarquons que, dans le problème (1-1), on peut théoriquement se passer de la contrainte $x \in V_t$. On a en effet le lemme suivant dont la démonstration est évidente :

LEMME 4-2

Soit dans $\mathbb{R}^p$ le problème suivant :

(4-2) $\quad \alpha - (\tilde{x}|z) = \mathrm{Inf}[(x'|z') \mid c'(s) \leq (x'|a'(s)) \quad \forall s \in S]$

où, si $u_1,\dots,u_p$ désigne une base de $V = \{x \in \mathbb{R}^n \mid (x|u_i) = o \quad i=p+1,\dots,n\}$ et $\tilde{x}$ un élément fixe de V_t, on a :

$$z' = ((u_1|z),\dots,(u_p|z)) \in \mathbb{R}^p$$
$$a'(s) = ((u_1|a(s)),\dots,(u_p|a(s))) \in \mathbb{R}^p$$
$$c'(s) = c(s) - (\tilde{x}|a(s))$$

L'algorithme d'échange et les hypothèses H1,H2 et H3 des problèmes (1-1) et (4-2) sont équivalents.

On a plus précisément $x^k = \sum_{i=1}^{p} x_i'^k u_i + x$ où x^k est l'élément de V_t obtenu à la $k^{\text{ième}}$-itération sur (1-1) et $x'^k = (x_1'^k,\dots,x_p'^k)$ l'élément de $\mathbb{R}^p$ obtenu à la $k^{\text{ième}}$-

itération sur (4-2).

On supposera dans ce qui suit que l'algorithme d'échanges avec les hypothèses H1, H2 et H3 es appliqué au problème (1-1) avec $V_t = \mathbb{R}^n$. (on a donc p=n).

Comme dans la démonstration du lemme 4-1 on utilisera un indice k pour indiquer le numéro de l'itération.

On pose :

$$(4\text{-}3) \qquad \alpha^k = \sum_{i=1}^{n} \rho_i^k c(s_i^k) = (x^k | z)$$

LEMME 4-3 :

La suite $\{\alpha^k\}$ est croissante et bornée supérieurement par $(x^* | z)$.

On a : $\alpha^{k+1} - \alpha^k = \rho_{\hat{\imath}^k}^{k+1} \left[c(\hat{s}^k) - (x^k | a(\hat{s}^k))\right]$

DEMONSTRATION :

On a par construction :

$$c(s_j^k) = (x^k | a(s_j^k)) \quad j=1,\dots,n \quad \text{et} \quad c(s_j^{k+1}) = (x^k | a(s_j^{k+1})) \quad j=1,\dots,n \quad j \neq \hat{\imath}^k$$

On a donc :

$$\alpha^{k+1} = (x^k \mid \sum_{i=1}^{n} \rho_i^{k+1} a(s_i^{k+1})) - (x^k | \rho_{\hat{\imath}^k}^{k+1} a(s_{\hat{\imath}^k}^{k+1})) + \rho_{\hat{\imath}^k}^{k+1} c(s_{\hat{\imath}^k}^{k+1})$$

De (2-6), (4-3) et $\hat{s}^k = s_{\hat{\imath}^k}^{k+1}$ on tire :

$$\alpha^{k+1} = \alpha^k + \rho_{\hat{\imath}^k}^{k+1} \left[c(\hat{s}^k) - (x^k | a(\hat{s}^k))\right].$$

L'hypothèse H3 entraîne $\rho_{\hat{\imath}^k}^{k+1} > o$ donc $\{\alpha^k\}$ est croissante. De l'hypothèse H1 on tire :

$$\alpha^k = \sum_{j=1}^{n} \rho_j^k \, c(s_j^k) < (x^* | \sum_{j=1}^{n} \rho_j^k \, a(s_j^k)) = (x^* | z)$$

Q.E.D.

LEMME 4-4 :

S'il existe s>o tel que $o < s \leq \rho_{\hat{\imath}^k}^{k+1}$ alors :

$$\lim_{t \to \infty} d^k = o \quad \text{et} \quad \lim_{k \to \infty} \alpha^k = \alpha$$

DEMONSTRATION :

1°/ Du lemme 4-3 et de la définition de d^k on tire :

$$\alpha^{k+1} - \alpha^k \geqq \rho_{\hat{\imath}^k}^{k+1} (\bar{d}^k - \eta^k)$$

d'où :

$$0 \leqq \bar{d}^k \leqq \frac{1}{s} (\alpha^{k+1} - \alpha^k) + \eta^k$$

D'après 4-3 et l'étape 2 de l'algorithme on a $\{\alpha^k\}$ qui converge et $\lim\limits_{k\to\infty} \eta^k = 0$ d'où $\lim\limits_{k\to\infty} \bar{d}^k = \lim\limits_{k\to\infty} d^k = 0$

2°/ On pose $\tilde{\alpha} = \lim\limits_{k\to\infty} \alpha^k$; de $\alpha^k \leqq \alpha$ on tire $\tilde{\alpha} \leqq \alpha$.

Supposons que l'on ait $\tilde{\alpha} < \alpha$, d'après le théorème de caractérisation on a:

$$\alpha = \sum_{i=1}^{n} \rho_i c(s_i) \qquad \text{avec} \quad \sum_{i=1}^{n} \rho_i\, a(s_i) \in V^{\perp} + z \qquad \rho_i \geqq 0 \qquad i=1,\dots,n$$

Posons $\alpha - \tilde{\alpha} = \mu > 0$. Quel que soit $k \in \mathbb{N}$ il existe $\bar{s} \in S$ tel que

$$c(\bar{s}) - (x^k | a(\bar{s})) \geqq \mu \Big(\sum_{i=1}^{n} \rho_i\Big)^{-1} = \beta$$

(Sinon pour un indice $\bar{k}$ on aurait $c(s) - (x^{\bar{k}} | a(s))$ $\forall s \in S$

et par suite $\sum\limits_{i=1}^{n} \rho_i c(s_i) - (x^{\bar{k}}\ z) = \alpha - \alpha^{\bar{k}} < \mu)$

On a alors : :

$$\bar{d}^k \geqq c(\bar{s}) - (x^k | a(\bar{s})) \geqq \mu \Big(\sum_{j=1}^{n} \rho_j\Big)^{-1} \qquad \text{on ne pourrait donc avoir } \lim_{k\to\infty} \bar{d}^k = 0$$

Q.E.D.

LEMME 4-5 :

Si on note A l'enveloppe convexe de $z_U - a(S)$, il existe une constante $w > 0$ telle que :

$$A_U - A \subset wA$$

DEMONSTRATION :

D'après H2 il existe $\rho_j > 0$ et $s_j \in S$ $(j=1,\dots,n)$ tel que

$$z = \sum_{i=1}^{n} \rho_i a(s_i)$$

Par suite :

$$0 = \frac{1}{K} z + \sum_{i=1}^{n} \frac{\rho_i}{K} (-a(s_i)) \qquad \text{avec} \quad K = 1 + \sum_{i=1}^{n} \rho_i > 0$$

et

$$0 = \frac{1}{K}(-z) + \sum_{i=1}^{n} \frac{\rho_i}{K} a(s_i)$$

donc o appartient à l'intérieur de $A_U - A$. On note $B(0,\varepsilon)$ une boule de centre o et de rayon ε contenue dans $A_U - A$; S étant compact, $A_U - A$ est borné, il existe donc $\mu > 0$ tel que $A_U - A \subset B(0,\mu)$.

On a donc $B(0,\mu) = \frac{\mu}{\varepsilon} B(0,\varepsilon) \subset \frac{\mu}{\varepsilon} A$. En posant $w = \frac{\mu}{\varepsilon}$ on a le résultat.

Q.E.D.

THEOREME 4-1 :

Les hypothèses H1,H2 et H3 étant vérifiées on a :

i) $\lim\limits_{k\to\infty} (x^k|z) = \alpha$

ii) il existe une sous-suite $\{x^{\psi(k)}\}$ de $\{x^k\}$ telle que :

$$\lim_{k\to\infty} \ \max_{s\in S} \ [c(s)-(x^{\psi(k)}|a(s))] = o$$

DEMONSTRATION :

Si pour tout i la suite $\{\rho_i^k\}_{k\in\mathbb{N}}$ est bornée inférieurement par un scalaire strictement positif on a d'après le lemme 4-4 :

$$\lim_{k\to\infty} \alpha^k = \lim(x^k|z) = \alpha \quad \text{et} \lim_{k\to\infty} d^k = o \quad \text{ce qui par définition de } d^k$$

entraîne $\lim\limits_{k\to\infty} \bar{d}^k = o$ et démontre le théorème.

Supposons maintenant qu'il existe $i_1 \in \{1,\ldots,n\}$ tel que $\{\rho_{i_1}^k\}$ ne soit pas bornée inférieurement par un scalaire strictement positif. On peut alors extraire une sous-suite $\phi_1 : \mathbb{N} \to \mathbb{N}$ telle que : $\lim\limits_{k\to\infty} \rho_{i_1}^{\phi_1(k)} = o$. S'il existe $i_2 \in\{1,\ldots,n\} - i_1$ tel que $\lim \rho_{i_2}^{\phi_1(k)} = o$ on extrait de $\{\rho_{i_2}^{\phi_1(k)}\}$ une sous-suite définie par $\phi_2 : \mathbb{N}\to\mathbb{N}$ telle que $\lim\limits_{k\to\infty} \rho_{i_2}^{\phi_2(k)} = o$. On continue l'opération jusqu'à ce qu'on ait une sous-suite définie par $\phi_q : \mathbb{N} \to \mathbb{N}$ telle que :

1°/ $\lim\limits_{k\to\infty} \rho_i^{\phi_q(k)} = o$ pour tout $i \in I_o = \{i_1,\ldots,i_q\}$

2°/ il existe $\mu>o$ tel que pour tout $i\in I_1 = \{1,\ldots,n\} - I_o$ on ait $\mu<\rho_i^{\phi_q(k)}$.

Remarquons que I_1 est non vide car, de la relation $z = \sum\limits_{i=1}^{n} \rho_i^k a(s_i^k)$ on tire :

$$o < \frac{||z||}{\max\limits_{s\in S}|a(s)|} \leq \sum_{i=1}^{n} \rho_i^k$$

Démontrons le lemme technique :

LEMME 4-6 :

Il est impossible que, pour k fixé, l'on ait pour tout k'>k

$$\{s_i^{\phi_q(k')} \mid i \in I_1\} \subset \{s_i^{\phi_q(k)} \mid i=1,\ldots,n\}$$

DEMONSTRATION :

D'après l'hypothèse (H1) on a : $\max\limits_{s\in S}[c(s)-(x^*|a(s))] = d<o$ d'où :

$$\sum_{i=1}^{n} \rho_i^k\, c(s_i^k) - (x^* \,|\, \sum_{i=1}^{n} \rho_i^k\, a(s_i^k)) \leqq d \sum_{i=1}^{n} \rho_i^k < o$$

Par suite, en tenant compte de (2-6) :

$$\sum_{i=1}^{n} \rho_i^k \leqq \frac{1}{|d|} (|\tilde{\alpha}| + \|x^*\| . \|z\|) \qquad \text{avec} \quad \tilde{\alpha} = \lim_{k\to\infty} \alpha^k$$

La suite $\{\rho_i^k\}_{k\in\mathbb{N}}$ est donc bornée pour tout $i \in \{1,\ldots,n\}$. On peut donc extraire de $\{\rho_{i_{q+1}}^{\phi_q(k)}\}_{k\in\mathbb{N}}$ pour $i_{q+1} \in I_1$ une sous-suite définie par $\{\rho_{i_{q+1}}^{\phi_{q+1}(k)}\}_{k\in\mathbb{N}}$ convergeant vers $\rho_{i_{q+1}} > o$; on peut de même extraire de $\{\rho_{i_{q+2}}^{\phi_{q+1}(k)}\}_{k\in\mathbb{N}}$ $(i_{q+2} \in I_1)$ une sous-suite définie par $\phi_{q+2} : \mathbb{N} \longrightarrow \mathbb{N}$ telle que :

$$\lim_{k\to\infty} \rho_{i_{q+2}}^{\phi_{q+2}(k)} = \rho_{i_{q+2}} > o.$$

En opérant ainsi sur tous les indices de I_1, on dispose d'une sous suite définie par $\phi \; : \mathbb{N} \longrightarrow \mathbb{N}$ telle que :

$$\lim_{k\to\infty} \rho^{\phi(k)} = o \quad \text{pour } i \in I_o$$

et

$$\lim_{k\to\infty} \rho^{\phi(k)} = \rho_i \quad \text{pour } i \in I_1 = \{1,\ldots\ldots,n\} - I_o \quad \text{avec} \quad o < \mu \leqq \rho_i$$

Si le lemme n'est pas vérifié on a :

$$z = \sum_{i\in I_o} \rho_i^{\phi(k')}\, a(s_i^{\phi(k')}) + \sum_{i\in I_1} \rho_i^{\phi(k')}\, a(s_i^{\phi(k)}) \qquad \text{pour tout } k' > k.$$

En faisant tendre k' vers l'infini on aurait $z = \sum_{i\in I_1} \rho_i\, a(s_i^{\phi(k)})$ ce qui est impossible car d'après H3 on a $z = \sum_{i=1}^{n} \rho_i^{\phi(k)} a(s_i^{\phi(k)})$ avec $\rho_i^{\phi(k)} > o$. <u>Le lemme est donc démontré.</u>

Quitte à extraire de $\{\rho_i^{\phi_q(k)}\}_{k\in\mathbb{N}}$ une nouvelle sous suite définie par $\phi : \mathbb{N} \longrightarrow \mathbb{N}$, on peut supposer d'après le lemme que l'ensemble $\{s_i^{\phi(k+1)} \,|\, i \in I_1\}$ contient au moins un point nouveau par rapport à l'ensemble $\{ s_i^{\phi(k)} \,|\, i=1,\ldots,n\}$. Notons $s_{e(k)}^{\phi(k+1)}$ le dernier élément nouveau introduit et supposons qu'il ait été introduit à l'itération $\psi(k)$ $(\phi(k) \leqq \psi(k) < \phi(k+1))$. On a donc :

$$s_i^{\phi(k+1)} = \begin{cases} s_i^{\psi(k)} & \text{si } i \neq e(k) \\ \bar{s}^{\psi(k)} & \text{si } i = e(k) \end{cases} \qquad i \in I_1$$

De la relation :

$$\alpha^{\phi(k+1)} = \sum_{i\in I_0} \rho_i^{\phi(k+1)} c(s_i^{\phi(k+1)}) + \sum_{i\in I_1} \rho_i^{\phi(k+1)} c(s_i^{\phi(k+1)})$$

on déduit en retranchant la quantité :

$$\alpha^{\psi(k)} = \sum_{i=1}^{n} \rho_i^{\phi(k+1)} (x^{\psi(k)} | \; a(s_i^{\phi(k+1)}))$$

$$\alpha^{\phi(k+1)} - \alpha^{\psi(k)} = \sum_{i\in I_0} \rho_i^{\phi(k+1)} \left[c(s_i^{\phi(k+1)}) - (x^{\psi(k)} | a(s_i^{\phi(k+1)}))\right]$$

$$+ \sum_{\substack{i\in I_1 \\ i\neq e(k)}} \rho_i^{\phi(k+1)} \left[c(s_i^{\psi(k)}) - (x^{\psi(k)} | a(s_i^{\psi(k)}))\right]$$

$$+ \rho_{e(k)}^{\phi(k+1)} \; (\bar{d}^{\psi(k)} - \xi^{\psi(k)})$$

avec $d^k = \bar{d}^k - \xi^k \qquad (o \leqq \xi^k \leqq \eta^k)$.

En tenant compte de la définition de $x^{\psi(k)}$ et en ajoutant $\sum_{i\in I_0} \rho_i^{\phi(k+1)} (x^{\psi(k)}|z)$ aux deux membres il vient :

$$(4\text{-}4) \qquad \alpha^{\phi(k+1)} + (\varepsilon(k)-1)\alpha^{\psi(k)} = B(k) + \rho_{e(k)}^{\phi(k+1)} (d^{\psi(k)} - \xi^{\psi(k)})$$

avec : $\varepsilon(k) = \sum_{i\in I_0} \rho_i^{\phi(k+1)}$

et

$$B(k) = \sum_{i\in I_0} \rho_i^{\phi(k+1)} \left[c(s_i^{\phi(k+1)}) - (x^{\psi(k)} | a(s_i^{\phi(k+1)}) - z)\right]$$

En posant $\omega^* = \operatorname{Max}_{s\in S} |c(s)|$ on a:

$$|B(k)| \leqq \varepsilon(k)(\omega^* + 2 \operatorname{Max}\{|(x^{\psi(k)}|t)| \; | t \in (-a(S)) \cup \{z\}\}$$

D'après le lemme 4-5 il existe $w > o$ telle que : $-A \cup A \subset wA$ avec A enveloppe convexe de $(-a(S)) \cup \{z\}$. On a alors :

$$\operatorname{Max}_{t\in A} |(x^{\psi(k)}|t)| = \operatorname{Max}_{t\in(-A)\cup A} (x^{\psi(k)}| t) \leqq \operatorname{Max}_{t\in wA} (x^{\psi(k)}|t)$$

d'où :

$$\operatorname{Max}\{|(x^{\psi(k)}|t)| \, | t\in -a(S) \cup \{z\}\} \leqq w \operatorname{Max}_{t\in -a(s) \cup \{z\}} (x^{\psi(k)}| t).$$

On a par ailleurs :

$$\underset{t\in -a(S)}{\mathrm{Max}} \; (x^{\psi(k)}|t) \leqq \underset{s\in S}{\mathrm{Max}} \left[c(s)-(x^{\psi(k)}|a(s))\right] + \omega^* = \bar{d}^k + \omega^*$$

et $\quad (x^{\psi(k)}|z) = \alpha^{\psi(k)} \leqq \bar{d}^{\psi(k)} + d^* \qquad$ avec $\alpha^* = (x^*|z) \qquad$ (lemme 4-3)

On a donc : $\quad \underset{t\in A}{\mathrm{Max}} \; |(x^{\psi(k)}|t)| \leqq w\bar{d}^{\psi(k)} + w\,\mathrm{Max}(\omega^*,\alpha^*)$

et par suite :

$$|B(k)| \leqq \varepsilon(k)\left[\omega^* + w\mathrm{Max}(\omega^*,\alpha^*) + w\bar{d}^{\psi(k)}\right]$$

En posant

$$d' = \omega^* + w\mathrm{Max}(\omega^*,\alpha^*) > 0, \text{ (4-4) s'écrit :}$$

$$\alpha^{\phi(k+1)} + (\varepsilon(k)-1)\,\alpha^{\psi(k)} + \varepsilon(k)d' + \rho^{\phi(k+1)}_{e(k)} \;\; \xi^{\psi(k)} \geqq \rho^{\phi(k+1)}_{e(k)} - w\;\varepsilon(k)\bar{d}^{\psi(k)}$$

L'indice e(k) appartenant à I_1, on a :

$$0 < \mu \leqq \rho^{\phi(k+1)}_{e(k)}$$

La suite $\bar{d}^{\psi(k)}$ est donc majorée, pour k suffisamment grand, par la suite $\{\beta_k\}$ avec :

$$\beta_k = \frac{1}{\mu - w\varepsilon(k)} \left(\alpha^{\phi(k+1)} + (\varepsilon(k)-1)\alpha^{\psi(k)} + \varepsilon(k)d' + K\xi^{\psi(k)}\right)$$

On a par construction $\lim\limits_{k\to\infty} \xi^{\psi(k)} = 0$; la convergence de $\{\alpha^k\}$ et celle de $\varepsilon(k)$ vers zéro entraîne $\lim\limits_{k\to\infty} \beta_k = 0$ et par suite $\lim\limits_{k\to\infty} \bar{d}^{\psi(k)} = 0$ ce qui démontre ii) du théorème.

On a toujours $\alpha^k \leqq \alpha$; en faisant un raisonnement analogue à celui du lemme 4-3 on a $\lim\limits_{k\to\infty} \alpha^{\psi(k)} = \alpha$ et $\{\alpha^k\}$ étant croissante on a $\lim\limits_{k\to\infty} \alpha^k = \alpha$ ce qui achève la démonstration du théorème.

Q.E.D.

REMARQUE :

L'hypothèse H3 est, dans le cas de l'algorithme de REMES, beaucoup moins forte que la "condition de HAAR". Le théorème précédent assure la convergence de l'algorithme dès qu'il se "déroule bien"(Hypothèse H3).

REFERENCES :

[1] CARASSO C. : Etude de l'algorithme de REMES en l'absence de conditions de HAAR,Numer.Math., (1972), 165-178.

[2] CARASSO C. : Densité des hypothèses assurant la convergence de l'algorithme de REMES, R.A.I.R.O.,(1972), R.3, 69-84.

[3] CARASSO C. : Thèse, Grenoble (1973).

[4] CARASSO C. : Pré-Publication n°3, Université de Saint-Etienne (1975)

[5] GOLDSTEIN A.A. : Constructive real analysis, Haaper and Row (1967).

[6] LAURENT P.J. : Approximation et Optimisation. Hermann, Paris (1972).

[7] LAURENT P.J. : Conference on Approximation Theory, Austin (Texas)(1973).

[8] ROCKAFELLAR R.T. : Convex analysis, Princ.Univ.Press, (1970).

[9] SINGER I. : Best approximation in normed linear spaces by elements of linear subspaces. Springer-Verlag (1970).