REVUE FRANÇAISE D'INFORMATIQUE ET DE RECHERCHE OPÉRATIONNELLE

P. HUARD

Programmation mathématique convexe

Revue française d'informatique et de recherche opérationnelle, tome 2, nº R1 (1968), p. 43-59

<http://www.numdam.org/item?id=M2AN_1968__2_1_43_0>

© AFCET, 1968, tous droits réservés.

L'accès aux archives de la revue « Revue française d'informatique et de recherche opérationnelle » implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

R.I.R.O.
(2e année, N° 7, 1968, p. 43-59)

PROGRAMMATION MATHEMATIQUE CONVEXE

par P. HUARD (*)

Résumé. — *La résolution d'un Programme Mathématique non linéaire par linéarisations successives du problème a déjà été envisagée par plusieurs auteurs. Ce procédé présente des difficultés de convergence pratique, ou conduit à des dimensions excessives du problème linéarisé, lorsqu'on linéarise directement le problème, c'est-à-dire lorsqu'on maximise la fonction critère linéarisée sur le domaine linéarisé. Cet article décrit un autre procédé, qui utilise la méthode dite « des Centres » et qui n'utilise la linéarisation qu'aux niveaux des calculs intermédiaires de cette méthode.*

L'avantage pratique de ce procédé est de conduire à la résolution d'une séquence de Programmes linéaires, de dimension constante et égale, (à une variable et à une contrainte près), à celle du Programme non linéaire d'origine. La convergence de l'algorithme vers l'optimum cherché est établie dans le cas convexe.

1. INTRODUCTION

Nous établissons dans ce qui suit, avec deux variantes, un algorithme de résolution pour les programmes mathématiques convexes, du type :

$$\text{Maximiser } f(x) \text{ sous les conditions}$$

$$g_i(x) \geqslant 0, \qquad i = 1, 2, \dots m$$

où les fonctions numériques f et g_i, définies dans R^n, sont concaves et continûment différentiables. Nous supposons de plus qu'il existe un point $\overset{0}{x} \in R^n$ tel que $g_i(\overset{0}{x}) > 0$, pour l'ensemble des fonctions g_i non affines.

L'intérêt de cet algorithme est de conduire à la résolution d'une suite, théoriquement infinie mais finie pratiquement, de programmes linéaires analogues, c'est-à-dire dont le nombre de contraintes est constant. Plus précisément, ces contraintes linéaires sont obtenues en linéarisant les diverses fonctions f et g_i, en des points tendant vers la solution optimale du problème donné.

(*) Conseiller scientifique au Service des Études de Réseaux et de Calcul Automatique, Direction des Études et Recherches d'Électricité de France. Faculté des Sciences de Lille.

Si certaines fonctions g_i sont de plus affines, les contraintes correspondantes des programmes linéaires ont leurs expressions inchangées, ce qui fait que cet algorithme ne détruit pas la partie linéaire du problème : à la limite, si le problème donné est un programme linéaire, les calculs se ramènent à la méthode simpliciale, avec paramétrisation d'un second membre.

En fait, cet algorithme, sous sa première variante, n'est qu'une application de la méthode des centres telle qu'elle est définie dans [1].

On désigne par cette expression une famille d'algorithmes, permettant de résoudre les programmes mathématiques non linéaires sous des hypothèses assez larges (par exemple : la variable x peut être un élément d'un espace topologique quelconque et la convexité n'est pas nécessaire), chaque algorithme étant défini par le choix d'une fonction devant satisfaire à quelques conditions simples, le choix de cette fonction demeurant par ailleurs assez arbitraire. C'est en utilisant les propriétés de convexité et de différentiabilité du programme envisagé ici que nous avons pu choisir une telle fonction s'adaptant bien aux calculs par linéarisation.

Nous donnons en 2 des rappels concernant la méthode des centres, en nous limitant au minimum nécessaire, et nous établissons en 3 l'adaptation au problème posé ici. On remarquera qu'une partie des calculs de l'algorithme (partie que nous désignons par « algorithme partiel ») consiste à maximiser une fonction concave non continûment différentiable sur un polyèdre, et peut être considérée comme une généralisation de la méthode de Franck et Wolfe [2]. Enfin, la partie 4 est consacrée à quelques considérations pratiques.

Cet article était rédigé lorsque nous avons eu connaissance de celui de Topkis et Veinott [3], dont le passage consacré aux problèmes de maximisation sous contraintes contient un des algorithmes présentés ici. En fait, la variante dite du premier ordre décrite dans [3] utilise — implicitement — la méthode des centres. C'est pourquoi nous avons ajouté le paragraphe 3.5.5, afin de préciser cette liaison.

Notations :

Si $A \subset R_n$, on désigne par $\overset{\circ}{A}$ l'intérieur de A et par $Fr(A)$ la frontière de A.

Si $f : R^n \to R$ est une fonction numérique continûment différentiable, on désigne par $\nabla f(x)$ la valeur du gradient de f calculé au point x.

Si $a, b \in R^n$, leur produit scalaire est noté simplement $a \cdot b$.

2. RAPPELS SUR LA METHODE DES CENTRES

2.1. Problème envisagé

On considère le programme mathématique P suivant :

$$P : \boxed{\begin{array}{c} \text{Maximiser } f(x) \text{ sous les conditions} \\ x \in A \subset R^n \\ x \in B \subset R^n \end{array}}$$

où $f : R^n \to R$ est une fonction numérique continue, bornée supérieurement sur $A \cap B$, telle que :

$$\{ x \mid f(x) = \lambda \} = \mathrm{Fr}\, \{ x \mid f(x) \geqslant \lambda \}, \qquad \forall \lambda \in R \tag{2.1}$$

et où A est un ensemble de R^n vérifiant l'hypothèse suivante :

$$\overset{0}{A} \neq \varnothing \qquad \text{et} \qquad \mathrm{Fr}\,(A) = \mathrm{Fr}\,(\overset{0}{A}) \tag{2.2}$$

Aucune hypothèse n'est faite à présent sur B.

On désignera par $E(\lambda)$ un « tronçon » de A, c'est-à-dire :

$$E(\lambda) = \{ x \mid x \in A, f(x) \geqslant \lambda \} \qquad (\lambda \in R)$$

2.2. Fonction F-distance

On appelle F-distance compatible avec f toute fonction numérique $d : R^n \times R \to R$ satisfaisant aux conditions :

(i) $\quad d(x, \lambda) > 0, \quad \forall \lambda \in R \quad$ et $\quad \forall x \in \overset{0}{E}(\lambda)$

(ii) $\quad d(x, \lambda) = 0, \quad \forall \lambda \in R \quad$ et $\quad \forall x \in \mathrm{Fr}\,[E(\lambda)]$

(iii) $\forall \lambda, \lambda' \in R$, on a :

$$x \in E(\lambda) \quad \text{et} \quad \lambda > \lambda' \quad \Rightarrow \quad d(x, \lambda) \leqslant d(x, \lambda')$$

(iiii) (compatibilité avec f)

Quelle que soit la suite $\lambda_k \in R$, $k = 1, 2, 3, \ldots$ monotone non décroissante, et quelle que soit la suite $\overset{k}{x} \in R^n$, $k = 1, 2, 3\ldots$ telle que :

$$\overset{k}{x} \in \mathrm{Fr}\,[E(\lambda_k)], \forall k,$$

on a

$$\text{“ } f(\overset{k+1}{x}) - f(\overset{k}{x}) \to 0 \text{ quand } k \to \infty \text{ ”} \Rightarrow \text{“ } d(\overset{k+1}{x}, \lambda_k) \to 0 \text{ quand } k \to \infty \text{ ”}$$

2.3. ε-centre d'un tronçon

Étant donné un tronçon $E(\lambda)$ et $\varepsilon \in R$ tel que :

$$0 \leqslant \varepsilon < \sup \{ d(x, \lambda) \mid x \in E(\lambda) \}$$

on appelle ε-centre de $E(\lambda)$ tout point $\bar{x}$ de $E(\lambda)$ tel que :

$$d(\bar{x}, \lambda) \geqslant \sup \{ d(x, \lambda) \mid x \in E(\lambda) \} - \varepsilon$$

Un 0-centre, s'il existe, est appelé un centre.

2.4. Méthode des centres

Cette méthode consiste, après avoir choisi une F-distance d compatible avec f, et une valeur initiale λ_0 convenable, c'est-à-dire telle que :

$$\lambda_0 < \sup \{ f(x) \mid x \in A \cap B \},$$

à résoudre la suite de programmes mathématiques suivante, pour :

$$k = 0, 1, 2, 3, \ldots :$$

$$Q_k : \boxed{\begin{array}{c} \text{Maximiser } d(x, \lambda_k), \text{ à } \varepsilon_k \text{ près sous les conditions} \\ x \in E(\lambda_k) \\ x \in B \end{array}}$$

En d'autres termes, la résolution de Q_k consiste à déterminer un « ε_k-centre relatif » de $E(\lambda_k)$ dans B. Soit $\overset{k+1}{x}$ ce point. La suite des valeurs λ_k est obtenue par :

$$\lambda_k = f(\overset{k}{x}), \qquad k = 1, 2, 3, \ldots$$

La suite des ε_k doit être monotone décroissante, convergeante vers 0 et satisfaire à :

$$0 \leqslant \varepsilon_k < \sup \{ d(x, \lambda_k) \mid x \in E(\lambda_k) \cap B \}$$

Les ε_k-centres relatifs étant des points intérieurs aux tronçons $E(\lambda_k)$, la contrainte $x \in E(\lambda_k)$ est généralement inutile *sur le plan pratique des calculs.*

On obtient une suite de points $\overset{k}{x}$, $k = 1, 2, 3, \ldots$ généralement infinie, qui converge vers une solution optimale $\hat{x}$ du problème P. Les tronçons $E(\lambda_k)$ diminuent par inclusion et tendent vers une limite $E(\hat{\lambda})$ telle que $\overset{0}{E}(\hat{\lambda}) \cap B = \varnothing$. Les valeurs de $\sup \{ d(x, \lambda_k) \mid x \in E(\lambda_k) \cap B$ tendent vers 0 quand $k \to \infty$, et cette remarque permet de définir les ε_k de la façon suivante :

$$\varepsilon_k = (1 - \rho) \sup \{ d(x, \lambda_k) \mid x \in E(\lambda_k) \cap B \}$$

avec $\rho \in]0, 1]$, constante indépendante de k. Dans ces conditions, les points $\overset{k+1}{x}$ vérifient la relation ci-dessous, qui nous sera utile par la suite :

$$(2.3) \quad \overset{k+1}{x} : d\left(\overset{k+1}{x}, \lambda_k\right) \geqslant \rho \sup \{ d(x, \lambda_k) \mid x \in E(\lambda_k) \cap B \}, \qquad \rho \in]0, 1]$$

3. RESOLUTION DU PROBLEME P

3.1. Adaptation de la méthode des centres

Nous nous proposons, dans ce qui suit, d'appliquer la méthode des centres, telle qu'elle est décrite en 2, dans les conditions particulières suivantes :

— I est un ensemble fini d'indices. Par exemple $I = \{ 1, 2, \dots m \}$.

(3.1) f, $g_i : R^n \to R$, $i \in I$, sont des fonctions concaves continûment différentiables.

f satisfait à (2.1) et les g_i à une condition semblable.

$$(3.2) \qquad A = \{ x \mid g_i(x) \geqslant 0, \forall i \in I \}$$

A est donc un ensemble convexe fermé de R^n, ainsi que les tronçons $E(\lambda)$.

$$(3.3) \qquad \overset{0}{A} \neq \varnothing$$

Par suite, puisque A est convexe, on a bien $\mathrm{Fr}\,(A) = \mathrm{Fr}\left(\overset{0}{A}\right)$ et l'hypothèse (2.2) est satisfaite ([1] page 61).

(3.4) B est un polyèdre linéaire fermé de R^n, c'est-à-dire qu'il est défini par un système d'inégalités et d'égalités linéaires. Nous supposons de plus qu'il est borné, donc compact. On peut toujours se ramener à ce cas en ajoutant éventuellement la condition $x \in C$, où C est un pavé de R^n, assez grand pour ne pas modifier la solution du problème P, supposée à distance finie.

$$(3.5) \qquad d(x, \lambda) = \min \{ f(x) - \lambda, g_i(x) \mid i \in I \}$$

Il s'agit bien d'une F-distance compatible avec f ([1] page 59), fonction concave en x pour tout λ fixé, prenant des valeurs $\leqslant 0$ pour tout $x \notin E(\lambda)$. Elle atteint son maximum par rapport à x, sur chaque ensemble $E(\lambda) \cap B$, tel que $\overset{0}{E}(\lambda) \cap B \neq \varnothing$, en un point intérieur à $E(\lambda)$.

— Les différents points $\overset{k+1}{x}$ fournis par la méthode des centres seront déterminés d'après le critère (2.3), c'est-à-dire en résolvant le problème suivant :

$$Q_k : \boxed{\begin{array}{c} \text{Trouver } \overset{k+1}{x} : d\left(\overset{k+1}{x}, \lambda_k\right) \geqslant \rho \max \{ d(x, \lambda_k) \mid x \in E(\lambda_k) \cap B \} \\ \text{avec } \rho \in]0, 1] \text{ constante indépendante de } k \\ \lambda_k = f\left(\overset{k}{x}\right) \end{array}}$$

La suite des problèmes Q_k, $k = 1, 2, 3, \ldots$ représente l'essentiel des calculs à effectuer.

3.2. Linéarisation

Soit $g'_i : R^n \times R^n \to R$, $\forall i \in I$, des fonctions numériques définies par :

$$g'_i(x\,;y) = g_i(y) + \nabla g_i(y) \cdot (x - y) \tag{3.6}$$

et soit $f' : R^n \times R^n \to R$ la fonction numérique définie par :

$$f'(x\,;y) = f(y) + \nabla f(y) \cdot (x - y) \tag{3.7}$$

Les g'_i et f' sont des fonctions affines de x. Posons par ailleurs :

$$d'(x, \lambda\,; y) = \min \{ f'(x\,;y) - \lambda, g'_i(x\,;y) \mid i \in I \} \tag{3.8}$$

d' est une fonction concave et affine par morceaux de x.

Nous avons la relation classique, liée à la concavité des fonctions envisagées :

$$d'(x, \lambda\,; y) \geqslant d(x, \lambda), \qquad \forall x, y \in R^n, \qquad \forall \lambda \in R \tag{3.9}$$

Considérons le programme mathématique $Q'(\lambda; y)$ suivant :

$$Q'(\lambda; y) : \boxed{\begin{array}{l} \text{Maximiser } d'(x, \lambda; y) \text{ sous la condition} \\ \quad x \\ \qquad\qquad x \in B \end{array}}$$

λ et y étant fixés. Ce problème peut se formuler sous la forme d'un programme linéaire en introduisant une variable supplémentaire μ :

$$Q''(\lambda\,; y) \quad \boxed{\begin{array}{l} \text{Maximiser } \mu \text{ sous les conditions} \\ \qquad g'_i(x\,;y) - \mu \geqslant 0, \qquad \forall i \in I \\ \qquad f'(x\,;y) - \mu \geqslant \lambda \\ \qquad\quad x \in B \end{array}}$$

λ et y étant fixés.

3.3. Algorithme partiel

Notons $z(\lambda;\ y)$ une solution optimale du programme linéaire $Q''(\lambda;\ y)$. Considérons l'algorithme suivant, au cours duquel λ demeure constant.

1. Choisir une valeur de départ $\overset{0}{y} \in B$. Faire $h = 0$.
2. Résoudre $Q''(\lambda\,;\overset{h}{y})$. Soit $\overset{h}{z} = z(\lambda\,;\overset{h}{y})$ une solution.
3. Déterminer $\overset{h+1}{y}$: $d(\overset{h+1}{y}, \lambda) = \max \left\{ d(x, \lambda) \mid x \in \left[\overset{h}{y}, \overset{h}{z}\right] \right\}$
4. Aller en (2) avec $h + 1$ au lieu de h.

On obtient ainsi une suite infinie de points $\overset{h}{y}$, avec éventuellement à partir d'un certain rang, des points identiques : c'est le cas si l'on trouve $\overset{h+1}{y} = \overset{h}{y}$ pour un certain h, ce qui entraîne alors que tous les problèmes $Q''(\lambda\,;\overset{h}{y})$ ultérieurs sont identiques.

La suite des valeurs $d(\overset{h}{y}, \lambda)$, $h = 0, 1, 2, \ldots$ est monotone non décroissante d'après la partie (3) de l'algorithme.

3.4. Convergence de l'algorithme partiel

Soit S la suite infinie des indices h. On a, $\forall h \in S$, $\overset{h}{z} \in B$ par définition, et $\overset{h+1}{y} \in B$ également, car B est convexe, $\overset{h+1}{y} \in \left[\overset{h}{y}, \overset{h}{z}\right] \subset B$ et $\overset{0}{y} \in B$. Puisque B est compact, $B \times B$ l'est aussi, et il existe une sous-suite $S' \subset S$ d'indices h tels que :

$$\left(\overset{h}{y}, \overset{h}{z}\right) \to (\bar{y}, \bar{z}) \qquad \text{quand} \qquad h \to \infty,\ h \in S'$$

avec $\bar{y} \in B$, $\bar{z} \in B$

D'une part, on peut écrire : $\forall h \in S$:

$$\begin{aligned} d'(\overset{h}{z}, \lambda\,;\overset{h}{y}) &\geqslant d'(\overset{h}{y}, \lambda\,;\overset{h}{y}) \text{ car } \overset{h}{z} \text{ maximise } d' \text{ sur } B \\ &= d\left(\overset{h}{y}, \lambda\right) \text{ d'après les définitions (3.6) et (3.7).} \end{aligned}$$

Si $h \to \infty$, $h \in S'$, on a à la limite :

$$d'(\bar{z}, \lambda\,;\bar{y}) \geqslant d(\bar{y}, \lambda) \tag{3.10}$$

D'autre part, soit $\theta \in [0, 1]$. On a, $\forall h, h' \in S'$, $h' > h$:

$$\begin{aligned} d\left[\overset{h}{y} + \theta(\overset{h}{z} - \overset{h}{y}), \lambda\right] &\leqslant \max \left\{ d(x, \lambda) \mid x \in \left[\overset{h}{y}, \overset{h}{z}\right] \right\} \\ &= d(\overset{h+1}{y}, \lambda) \text{ par définition} \\ &\leqslant d(\overset{h'}{y}, \lambda) \text{ car } h' \geqslant h + 1, \text{ et la suite des valeurs} \\ &\quad\ d(\overset{h}{y}, \lambda) \text{ est monotone non décroissante} \end{aligned}$$

Par suite, quand $h \to \infty$, $h \in S'$, en passant à la limite, avec θ fixé, on obtient :

$$d[\bar{y} + \theta(\bar{z} - \bar{y}), \lambda] \leqslant d(\bar{y}, \lambda) \qquad \forall \theta \in [0, 1] \tag{3.11}$$

La relation (3.11) entraîne avec (3.10) :

$$d'(\bar{z}, \lambda ; \bar{y}) = d(\bar{y}, \lambda) \tag{3.12}$$

En effet, du fait que f et les g_i sont continûment différentiables, (3.11) entraîne (3.13) ou (3.14) :

$$\exists i_0 \in I \left| \begin{array}{l} g_{i_0}(\bar{y}) = d'(\bar{y}, \lambda ; \bar{y}) = d(\bar{y}, \lambda) \\ \nabla g_{i_0}(\bar{y}) \cdot (\bar{z} - \bar{y}) \leqslant 0 \end{array} \right. \tag{3.13}$$

$$\left| \begin{array}{l} f(\bar{y}) - \lambda = d'(\bar{y}, \lambda ; \bar{y}) = d(\bar{y}, \lambda) \\ \nabla f(\bar{y}) \cdot (\bar{z} - \bar{y}) \leqslant 0 \end{array} \right. \tag{3.14}$$

et puisque $d'(x, \lambda ; y) \leqslant \min \{ f'(x, y) - \lambda, g'_{i_0}(x ; y) \}$, (3.13) ou (3.14), avec (3.10), entraîne (3.12).

Montrons que $\bar{y}$ maximise $d(x, \lambda)$ sur $E(\lambda) \cap B$.

Soit un point quelconque $x \in B$. On a, $\forall h \in S$:

$$d(x, \lambda) \leqslant d'\left(x, \lambda ; \overset{h}{y}\right) \text{ d'après (3.9)}$$

$$\leqslant d'\left(\overset{h}{z}, \lambda ; \overset{h}{y}\right) \quad \text{car} \quad \overset{h}{z} \text{ maximise } d'\left(x, \lambda ; \overset{h}{y}\right) \text{ sur } B \tag{3.16}$$

A la limite, quand $h \to \infty$, $h \in S'$:

$$d(x, \lambda) \leqslant d'(\bar{z}, \lambda ; \bar{y}) = d(\bar{y}, \lambda) \text{ d'après (3.12)}$$

En définitive :

$$d(x, \lambda) \leqslant d(\bar{y}, \lambda), \qquad \forall x \in B \tag{3.17}$$

Puisque la suite des valeurs $d\left(\overset{h}{y}, \lambda\right)$, $h \in S$, est monotone non décroissante, elle converge vers $d(\bar{y}, \lambda)$.

3.5. Remarques sur la résolution de Q_k

3.5.1. Nous venons d'établir en 3.3 et 3.4 un algorithme, dit partiel, qui permet de résoudre le problème Q_k, dans le cas $\rho = 1$, par une séquence infinie de résolutions de programmes linéaires $Q''\left(\lambda_k ; \overset{h}{y}\right)$, sous l'hypothèse de concavité des fonctions f et g_i envisagées.

En fait, il est possible de reprendre la démonstration de la convergence, établie en 3.4, en n'utilisant plus cette hypothèse : le résultat, plus faible, ainsi obtenu, est que l'algorithme converge vers un point stationnaire (cf. 3.5.2).

On peut également, lors de la maximisation de $d(x, \lambda)$ sur le segment $\left[\overset{h}{y}, \overset{h}{z}\right]$, ne pas aller jusqu'au bout de cette maximisation, comme il est indiqué en 3.5.3, sans détruire la convergence de l'algorithme partiel.

Enfin, la méthode des centres n'exige pas pour chaque tronçon de calculer un centre, (c'est-à-dire de prendre $\rho = 1$ dans la définition de Q_k), mais seulement un ε-centre ($\rho \in]0, 1]$). Nous donnons en 5.3.4 un moyen de contrôler si le point $\overset{h}{y}$ trouvé est bien une solution de Q_k, c'est-à-dire vérifiant la condition (2.3).

Nous montrons en 5.3.5 que le premier point $\overset{h}{y}$ trouvé dans l'algorithme partiel peut toujours être considéré comme un ε-centre, c'est-à-dire que l'on peut théoriquement se contenter d'un seul programme linéaire par itération de la méthode des centres.

3.5.2. Abandon de la concavité

Dans l'établissement de la convergence, en 3.4, la concavité des fonctions f et g_i n'est utilisée qu'à la fin, quand on fait appel à la relation (3.9) : on peut établir ainsi que $\bar{y}$ est un maximum global de $d(x, \lambda)$ sur B. Si on abandonne l'hypothèse de concavité, la relation (3.12) demeure par contre valable.

Soit $x \in B$ et $\varepsilon = \|x - \bar{y}\|$ —. Si ε est assez petit, et $\forall h \in S'$ assez grand (c'est-à-dire $\overset{h}{y}$ suffisamment voisin de $\bar{y}$), on a au moins l'un des résultats suivants (3.18) et (3.19) :

$$(3.18) \quad \exists i_0 \in I : \left| \begin{array}{l} d(\overset{h}{y}, \lambda) = g_{i_0}(\overset{h}{y}) \\ d'(x, \lambda ; \overset{h}{y}) = g_{i_0}(\overset{h}{y}) + \nabla g_{i_0}(\overset{h}{y}) \cdot (x - \overset{h}{y}) \end{array} \right.$$

$$(3.19) \quad \left| \begin{array}{l} d(\overset{h}{y}, \lambda) = f(\overset{h}{y}) - \lambda \\ d'(x, \lambda ; \overset{h}{y}) = f(\overset{h}{y}) - \lambda + \nabla f(\overset{h}{y}) \cdot (x - \overset{h}{y}) \end{array} \right.$$

Par ailleurs :

$$(3.20) \qquad d'(x, \lambda ; \overset{h}{y}) \leqslant d'(\overset{h}{z}, \lambda, \overset{h}{y}) \text{ car } \overset{h}{z} \text{ maximise } d' \text{ sur } B$$

En passant à la limite quand $h \to \infty$, $h \in S'$, les relations (3.18) à (3.20) entraînent en tenant compte de (3.12) au moins l'un des deux résultats suivants, correspondant respectivement aux situations (3.18) et (3.19) :

$$(3.21) \qquad \nabla g_{i_0}(\bar{y}) \cdot (x - \bar{y}) \leqslant 0$$

$$(3.22) \qquad \nabla f(\bar{y}) \cdot (x - \bar{y}) \leqslant 0$$

ce qui montre que la limite $\bar{y}$ est un point stationnaire pour la maximisation de $d(x, \lambda)$ dans B.

3.5.3. Maximisation approchée sur un segment

Si on remplace l'ordre (3) de l'algorithme partiel par l'ordre suivant :

(3′) Déterminer $\overset{h+1}{y}$:

$$d\left(\overset{h+1}{y}, \lambda\right) \geqslant d\left(\overset{h}{y}, \lambda\right) + \rho'\left[\max\left\{ d(x, \lambda) \mid x \in \left[\overset{h}{y} ; \overset{h}{z}\right] \right\} - d\left(\overset{h}{y}, \lambda\right)\right]$$

avec $\rho' \in]0, 1]$, ρ' constante indépendante de h, l'algorithme partiel converge encore. En effet, le raisonnement qui aboutit à la relation (3.10) demeure valable, et celui qui aboutit à la relation (3.11) donne dans ces conditions :

$$d\left[\overset{h}{y} + \theta\left(\overset{h}{z} - \overset{h}{y}\right), \lambda\right] \leqslant d\left(\overset{h}{y}, \lambda\right) + \frac{1}{\rho'}\left[d\left(\overset{h'}{y}, \lambda\right) - d\left(\overset{h}{y}, \lambda\right)\right] \quad \forall \theta \in [0, 1]$$

et à la limite, quand $h \to \infty$, $h \in S'$, on retrouve (3.11).

Ce résultat est intéressant sur le plan pratique, car la maximisation de $d(x, \lambda)$ sur un segment est toujours approchée.

3.5.4. Reconnaissance d'un ε-centre

Étant donné $\rho \in]0, 1]$, le fait que pour un rang h^* fini, on ait

$$d\left(\overset{h^*}{y}, \lambda_k\right) \geqslant \rho d(\bar{y}, \lambda_k)$$

est évident. Il reste à pouvoir reconnaître que cette inégalité est satisfaite, autrement dit il faut définir un critère d'arrêt.

Pour cela, calculons une borne supérieure de $d(\bar{y}, \lambda)$ en prenant $x = \bar{y}$ et $\lambda = \lambda_k$ dans la relation (3.16), ce qui donne :

$$d(\bar{y}, \lambda_k) \leqslant d'\left(\overset{h}{z}, \lambda_k ; \overset{h}{y}\right), \qquad \forall h \in S \tag{3.23}$$

Par suite, on peut arrêter l'algorithme partiel quand on a obtenu un point $\overset{h^*}{y}$ tel que :

$$d\left(\overset{h^*}{y}, \lambda_k\right) \geqslant \rho d'\left(\overset{h^*}{z}, \lambda_k ; \overset{h^*}{y}\right) \qquad \rho \in]0, 1] \text{ donné} \tag{3.24}$$

Ce point $\overset{h^*}{y}$ est solution de Q_k et l'on pose :

$$\overset{k+1}{x} = \overset{h^*}{y}$$

3.5.5. La première étape est un ε-centre

Si, dans la résolution de Q_k, on s'arrête systématiquement à la première solution $\overset{1}{y}$ trouvée, que l'on pose $\overset{k+1}{x} = \overset{1}{y}$ pour passer au problème Q_{k+1} suivant, et si $\overset{0}{x} \in B \cap E(\lambda_0)$, on obtient encore une méthode des centres. Il suffit, pour le montrer, d'établir que la première solution de Q_k est un ε_k-centre

relatif pour le tronçon $E(\lambda_k)$, c'est-à-dire que $\overset{k}{x} \in B \cap E(\lambda_k)$, $\forall k$, et que la séquence infinie des ε_k tend vers zéro quand $k \to \infty$.

L'algorithme ainsi défini conduit à une séquence S d'itérations qui se résument comme suit :

Étant donné $\overset{k}{x}$, $\lambda_k = f(\overset{k}{x})$, déterminer $\overset{k}{z}$, pius $\overset{k+1}{x}$, tels que :

$$d'(\overset{k}{z}, \lambda_k ; \overset{k}{x}) = \max \left\{ d'(x, \lambda_k ; \overset{k}{x}) \mid x \in B \right\} \tag{3.25}$$

$$d(\overset{k+1}{x}, \lambda_k) = \max \left\{ d(x, \lambda_k) \mid x \in \left[\overset{k}{x}, \overset{k}{z}\right] \right\} \tag{3.26}$$

Dans ces conditions on vérifie aisément de proche en proche à partir de $\overset{0}{x}$, que $\overset{k}{x} \in B \cap E(\lambda_k)$ et que $d(\overset{k}{x}, \lambda_k) = f(\overset{k}{x}) - \lambda_k = 0$. D'où l'on remarque que :

$$d(\overset{k'}{x}, \lambda_k) \geqslant d(\overset{k+1}{x}, \lambda_k), \qquad \forall k, k' \in S, k' > k \tag{3.27}$$

Par ailleurs, λ_k étant borné supérieurement, il existe une sous-suite $S' \subset S$ telle que :

$$(\overset{k}{x}, \overset{k}{z}, \lambda_k) \to (\bar{x}, \bar{z}, \bar{\lambda}) \qquad \text{quand } k \to \infty, k \in S'$$

On en déduit, puisque $0 = d(\overset{k}{x}, \lambda_k)$, $\forall k$, et en passant à la limite :

$$d(\bar{x}, \bar{\lambda}) = 0 \tag{3.28}$$

En reprenant le raisonnement de 3.4, mais en remplaçant l'indice d'itération h par k, en tenant compte de ce que λ prend une valeur différente λ_k à chaque itération, et en utilisant les remarques (3.27) et (3.28) :

$$d'(\bar{z}, \bar{\lambda} ; \bar{x}) = 0 \tag{3.29}$$

Exprimons une borne supérieure de l'écart ε_k entre la valeur trouvée

$$d(\overset{k+1}{x}, \lambda_k)$$

et la valeur de d en un centre relatif $\overset{k}{c}$ du tronçon $E(\lambda_k)$ dans B, sous l'hypothèse de concavité des fonctions f et g_i :

$$\varepsilon_k = d(\overset{k}{c}, \lambda_k) - d(\overset{k+1}{x}, \lambda_k) \geqslant 0 \tag{3.30}$$

$$d(\overset{k}{c}, \lambda_k) \leqslant d'(\overset{k}{c}, \lambda_k ; \overset{k}{x}) \qquad \text{d'après (3.9), en prenant } x = \overset{k}{c}, \quad \lambda = \lambda_k \quad \text{et} \quad y = \overset{k}{x}$$

$$\leqslant d'(\overset{k}{z}, \lambda_k ; \overset{k}{x}) \text{ car } \overset{k}{z} \qquad \text{maximise } d' \text{ sur } B \tag{3.31}$$

Par ailleurs, d'après (3.27) :

$$d(\overset{k+1}{x}, \lambda_k) \geqslant d(\overset{k}{x}, \lambda_k) = 0, \qquad \forall k \in S \tag{3.32}$$

L'expression (3.30) de l'écart devient, avec (3.31) et (3.32) :

$$\varepsilon_k^{\boxtimes} \leqslant d'(\overset{k}{z}, \lambda_k ; \overset{k}{x}) \qquad \forall k \in S' \tag{3.33}$$

En passant à la limite, et d'après (3.29), la borne supérieure tend vers zéro : il en est de même pour l'écart $\varepsilon_k \geqslant 0$.

3.6. Troncatures par excès et par défaut

Dans la méthode des centres, telle qu'elle est décrite en 2.4, les tronçons sont déterminés par les solutions $\overset{k}{x}$. Plus précisément, une fois trouvé dans B, un ε-centre relatif $\overset{k+1}{x}$ du tronçon $E(\lambda_k)$, le nouveau tronçon est pris égal à :

$$E\left[f(\overset{k+1}{x})\right].$$

Une idée assez naturelle consiste à envisager un procédé de troncature différent, conduisant à l'algorithme suivant :

1. Prendre α et β tels que $f(\hat{x}) \in [\alpha, \beta]$
Choisir $\lambda_0 \in [\alpha, \beta]$. Aller en (2) avec $k = 0$.

2. Déterminer $T_k = E(\lambda_k) \cap B$.
Si $T_k \neq \varnothing$, remplacer α par λ_k.
Si $T_k = \varnothing$, remplacer β par λ_k.
Aller en (3).

3. Prendre $\lambda_{k+1} = (\alpha + \beta)/2$.
Aller en (2) avec $k + 1$ au lieu de k.

Cet algorithme opère avec des valeurs de troncature λ_k ajustées par excès et par défaut et convergeant vers la valeur $\hat{\lambda} = f(\hat{x})$ cherchée.

Pour vérifier si un tronçon $E(\lambda)$ a, avec B, une intersection $T(\lambda)$ vide ou non, il suffit de maximiser $d(x, \lambda)$, définie par (3.5), pour $x \in B$. S'il est possible de trouver un point $\bar{x}$ tel que $d(\bar{x}, \lambda) \geqslant 0$, alors $T(\lambda) \neq \varnothing$. Sinon, c'est-à-dire si $\max \{ d(x, \lambda) \mid x \in B \} < 0$, on en déduit que $T(\lambda) = \varnothing$.

Soit $\bar{x}(\lambda)$ un point qui maximise $d(x, \lambda)$ sur B. Posons $\bar{\mu}(\lambda) = d(\bar{x}(\lambda), \lambda)$. Cette fonction numérique de λ est un critère de vacuité pour $T(\lambda)$:

$$T(\lambda) \neq \varnothing \iff \bar{\mu}(\lambda) \geqslant 0$$

$$T(\lambda) = \varnothing \iff \bar{\mu}(\lambda) < 0$$

Pour chaque valeur λ proposée, on peut utiliser l'algorithme partiel décrit en 3, afin de déterminer $\bar{\mu}(\lambda)$. On calcule par dichotomies successives la valeur de λ qui annule $\bar{\mu}(\lambda)$, en utilisant toute technique d'interpolation utile. En particulier, en remarquant que $\bar{\mu}(\lambda)$ est une fonction concave (résultat classique

sur la paramétrisation du second membre des programmes mathématiques convexes), la valeur de λ donnée par l'interpolation linéaire :

$$\lambda' = \frac{\beta\bar{\mu}(\alpha) - \alpha\bar{\mu}(\beta)}{\bar{\mu}(\alpha) - \bar{\mu}(\beta)}$$

donne une estimation par défaut de $\hat{\lambda}$.

Par ailleurs, la dualité permet d'obtenir une estimation par excès de $\hat{\lambda}$ à chaque résolution de programme linéaire $Q''(\lambda_k\,;\overset{h}{y})$, si cette dernière nous fournit, avec la solution optimale, les multiplicateurs de Kuhn et Tucker correspondants. Les conditions d'optimalité de $Q''(\lambda_k\,;\overset{h}{y})$ donnent, entre autres relations, en posant $B = \{\, x \mid B_l x \geqslant b_l,\ l \in L \,\}$:

$$\exists u^i,\ i \in I\,;\ u^0\,;\ v^l,\ l \in L \text{ tels que :}$$

$$\begin{aligned} u^i &\geqslant 0, \qquad \forall i \in I \\ u^0 &\geqslant 0 \\ v^l &\geqslant 0, \qquad \forall l \in L \\ \sum_{i \in I} u^i\, \nabla g_i(\overset{h}{y}) + u^0\, \nabla f(\overset{h}{y}) &+ \sum_{l \in L} v^l B_l = 0 \\ \sum_{i \in I} u^i + u^0 &= 1 \end{aligned}$$

Si $u^0 \neq 0$ (il est toujours possible de se placer dans ce cas, pour $\overset{h}{y}$ fixé, en prenant λ_k assez grand), les valeurs u^i/u^0, $i \in I$, et v^l/u^0, $l \in L$, fournissent avec $\overset{h}{y}$, une solution réalisable pour le dual D de P. Rappelons que ces deux programmes s'écrivent :

P :

$$\boxed{\begin{array}{ll} \multicolumn{2}{l}{\text{Maximiser } f(x) \text{ sous les conditions}} \\ g_i(x) \geqslant 0, & \forall i \in I \\ B_l x \geqslant b_l, & \forall l \in L \end{array}}$$

D :

$$\boxed{\begin{array}{c} \text{Maximiser } \displaystyle\sum_{i \in I} u^i g_i(x) + f(x) + \sum_{l \in L} v^l (B_l x - b_l) \text{ sous les conditions :} \\ u^i \geqslant 0, \qquad \forall i \in I \\ v^l \geqslant 0, \qquad \forall l \in L \\ \displaystyle\sum_{i \in I} u^i\, \nabla g_i(x) + \nabla f(x) + \sum_{l \in L} v^l B_l = 0 \end{array}}$$

On a la relation classique :

$$\hat{\lambda} \leqslant f(\overset{h}{y}) + \sum_{i \in I} u^i g_i(\overset{h}{y}) + \sum_{l \in L} v^l(B_l \overset{h}{y} - b_l)$$

4. ASPECTS PRATIQUES. PARAMETRISATION

4.1. Aspect fini des calculs

Nous avons vu, au paragraphe 3, que la détermination d'un ε-centre revient à résoudre une suite finale de programmes linéaires, séparés par la recherche du maximum d'une fonction concave sur un segment de droite. Les programmes linéaires peuvent être résolus, par exemple, par la méthode simpliciale. Si les maximisations sur un segment sont réalisées de façon approchée, comme il est indiqué dans la remarque 3.5.3, une méthode classique par dichotomies conduira en définitive à un nombre fini de calculs de valeurs de d.

Enfin, bien que le nombre d'itérations de la méthode des centres, c'est-à-dire le nombre d'ε-centres $\overset{k}{x}$ à calculer, est théoriquement infini en général, il devient fini si on détermine non pas une solution optimale exacte $\hat{x}$ du problème P, mais une valeur approchée $\hat{x}'$ telle que :

$$f(\hat{x}') \geqslant f(\hat{x}) - \varepsilon'$$

avec les conditions :

$$\hat{x}' \in A \cap B \qquad \text{et} \qquad \varepsilon' > 0 \text{ donné}$$

(il en est toujours ainsi en pratique).

Par suite, le calcul d'une solution approchée $\hat{x}'$ peut se faire à l'aide d'une suite finie d'opérations (en particulier de résolutions de programmes linéaires).

4.2. Cas linéaire. Paramétrisation

Si l'on envisage le cas d'un programme P entièrement linéaire, c'est-à-dire où les fonctions f et g_i, $i \in I$, sont affines, le calcul d'un centre se réduit à la résolution d'un simple programme linéaire, car on a :

$$g_i'(x\,;y) \equiv g_i(x), \qquad \forall i \in I \quad \text{et} \quad \forall y$$
$$f'(x\,;y) \equiv f(x), \qquad \forall y$$

et ce programme s'écrit :

$$(Q_k) \quad \boxed{\begin{array}{l} \text{Maximiser } \mu \text{ sous les conditions} \\ \quad g_i(x) - \mu \geqslant 0, \qquad \forall i \in I \\ \quad f(x) - \mu \geqslant \lambda_k \\ \quad x \in B \end{array}}$$

Lorsque l'on a déterminé la solution optimale $(\overset{k+1}{x}, \overset{k+1}{\mu})$ de ce programme linéaire, on peut alors faire varier continûment la valeur de la « troncature » λ, à partir de λ_k, par valeurs croissantes. La valeur de la solution optimale varie avec λ, et en utilisant l'algorithme classique de la méthode simpliciale paramétrée, on peut déterminer une solution optimale extrême $[\bar{x}(\lambda), \bar{\mu}(\lambda)]$, qui varie linéairement par morceaux avec λ. La méthode simpliciale paramétrée fournit directement la suite des valeurs de λ qui déterminent les intervalles de variation linéaire. Il est aisé de voir comme il a été indiqué en 3.6 que $\bar{\mu}(\lambda)$ est une fonction concave décroissante de λ, linéaire par morceaux.

En ajoutant la condition $\mu \geqslant 0$, la paramétrisation s'arrête quand on ne peut plus augmenter λ. On a alors $\hat{x}(\lambda) = \hat{x}$, et $\lambda = f(\hat{x})$.

4.3. Paramétrisation dans le cas non linéaire

Le procédé de paramétrisation décrit en 4.2 n'a qu'un intérêt théorique, puisqu'il n'est défini que pour le cas linéaire et qu'il utilise... la méthode simpliciale. Néanmoins, on peut envisager de l'adapter au cas non linéaire, lorsqu'on détermine des ε-centres définis par la conditions (2.3), ou pratiquement, par la condition (3.24).

En effet, supposons que l'on soit à l'étape k, avec la solution correspondante $\overset{k}{x}$, et la troncature $\lambda_k = f(\overset{k}{x})$. Après résolution d'un premier programme linéaire $Q''(\lambda_k ; \overset{k}{x})$, on obtient les points correspondants $\overset{1}{z}$ et $\overset{1}{y}$. Si la condition (3.24), où l'on remplace h^* par 1, est satisfaite, on peut conserver la linéarisation au point $\overset{k}{x}$, c'est-à-dire ne pas modifier $Q''(\lambda_k ; \overset{k}{x})$, mais paramétrer ce programme linéaire par rapport à λ, à partir de λ_k. On obtient ainsi une suite de points $\overset{h}{z}$, et pour chacun d'entre eux, on détermine le point $\overset{h}{y}$ qui maximise $d(x, \lambda'_h)$ sur $[\overset{k}{x}, \overset{h}{z}]$, en désignant par λ'_h la valeur du paramètre λ correspondant à la solution $\overset{h}{z}$. Quand la condition (3.24), où l'on remplace h^* par h et λ_k par λ'_h, n'est plus satisfaite pour un certain rang h, on choisit pour nouveau point $\overset{k+1}{x}$ de linéarisation le point $\overset{h-1}{y}$, et l'on prend pour nouvelle valeur de troncature $\lambda = f(\overset{k+1}{x})$. On résoud alors $Q''(\lambda ; \overset{k+1}{x})$, et on essaie de nouveau la paramétrisation.

L'intérêt pratique de ce procédé de paramétrisation n'est pas évident, et doit dépendre des caractéristiques du problème P traité : fonctions faiblement non linéaires, beaucoup de contraintes actives à l'optimum, etc... Dans le cas d'un programme linéaire, on retrouve la méthode décrite en 4.2, la linéarisation devenant sans objet, et les points $\overset{h}{y}$ sont de vrais centres. D'un autre côté, dans le cas d'un programme fortement non linéaire (variations importantes des ∇g_i), on retrouvera pratiquement la méthode des centres décrite en 3.1 si pour tout $\overset{k}{x}$, on doit linéariser en tous les points $\overset{h}{y}$.

4.4. Démarrage des programmes linéaires

La résolution d'un problème $Q''(\lambda_k\,;\overset{k}{x})$ par la méthode simpliciale nécessite de satisfaire aux conditions initiales posées par la technique de cette méthode. En particulier, il faut avoir une solution réalisable extrême, dite « de base », comme solution de départ.

En introduisant des variables d'écart y_0, y_i, $i \in I$, la matrice des contraintes du *P.L.* à traiter peut se présenter sous la forme suivante :

x	μ	y	y_0		
A	$\begin{matrix}1\\1\\1\\1\end{matrix}$	$\begin{matrix}+1 & & & \\ & +1 & & \\ & & +1 & \\ & & & +1\end{matrix}$	0	$=$	a
A_0	1	0	$+1$	$=$	a_0
B	0	0	0	$=$	b

où

$$A_i = -\nabla g_i(\overset{k}{x}), \qquad a_i = -\nabla g_i(\overset{k}{x}) \cdot \overset{k}{x} + g_i(\overset{k}{x}), \qquad i \in I$$

$$A_0 = -\nabla f(\overset{k}{x}), \qquad a_0 = -\nabla f(\overset{k}{x}) \cdot \overset{k}{x} + f(\overset{k}{x}) - \lambda_k$$

On suppose ici que le polyèdre B est représenté par $Bx = b$, $x \geqslant 0$.

Si $\overset{0}{x}$ est une solution extrême (solution de base K) du polyèdre B, et si l'on possède le tableau simplicial correspondant, il suffit de rendre nuls les éléments des lignes A_0, A_i situés dans les colonnes d'indices $j \in K$, par les procédés classiques (combinaisons linéaires des lignes entre elles ou prémultiplication matricielle). La base du *P.L.* complet est obtenue en réunissant à K les indices de l'ensemble $I + \{\,_0\,\}$. La solution de départ a pour valeur :

$$\left|\begin{array}{l} x = \overset{0}{x} \\ \mu = \min\left\{ f'(\overset{0}{x}\,;\overset{k}{x}) - \lambda_k, g'_i(\overset{0}{x}\,;\overset{k}{x}) \mid i \in I \right\} \\ y_i = g'_i(\overset{0}{x}\,;\overset{k}{x}) - \mu, \qquad i \in I \\ y_0 = f'(\overset{0}{x}\,;\overset{k}{x}) - \lambda_k - \mu \end{array}\right.$$

On peut vérifier que, à l'exclusion de μ, toutes ces composantes sont bien $\geqslant 0$. Il faut alors faire entrer μ dans la base.

5. REFERENCES

[1] BUI TRONG LIEU et HUARD (P.). « La méthode des Centres dans un espace topologique », *Numerische Mathematik* **8**, 56-57 (1966).

[2] FRANCK (M.) et WOLFE (Ph.). « An algorithm for quadratic programming », *Naval Research Logistics Quaterly*, mars-juin 1956.
Voir également pages 90-91 de :
BERGE (C.) et GHOUILA-HOURI (A.). *Programmes, jeux et réseaux de transport*, Éd. Dunod, Paris, 1962.

[3] TOPKIS (D. M.) et VEINOTT (A. F.). « On the convergence of some feasible direction algorithms for nonlinear programming », *S.I.A.M. Journal on control* **5** (2), 1967, p. 268-279.