Revue française d'automatique informatique recherche opérationnelle. Mathématique

DANG TRAN DAC

Brève communication. Décomposition en programmation convexe

Revue française d'automatique informatique recherche opérationnelle. Mathématique, tome 7, nº R1 (1973), p. 68-75

<http://www.numdam.org/item?id=M2AN_1973__7_1_68_0>

R.A.I.R.O.
(7e année, avril 1973, R-1, p. 68-75)

DECOMPOSITION EN PROGRAMMATION CONVEXE

par DANG TRAN DAC [1]

Résumé. — *On propose des méthodes nouvelles de décomposition en programmation convexe.*

INTRODUCTION

Soit Y un sous-ensemble convexe fermé de $\mathbf{R}^m$, v une fonction concave définie sur Y, continue *non différentiable*, la littérature montre que si l'on cherche à jouer sur les structures des problèmes de programmation convexe, on peut décomposer un grand nombre des problèmes en le ramenant au problème suivant :

$$(P) : \max\ (v(y)/y \in Y)$$

Le PHD de Hogan [3] est consacré essentiellement à l'étude numérique de ce problème. Plus exactement, si l'on considère la méthode de Frank-Wolfe modifiée (on remplace la dérivée par la dérivée directionnelle) alors cette méthode célèbre en général ne converge plus. Hogan montre que moyennant une approximation de la dérivée directionnelle on peut faire converger le procédé.

La méthode de Frank-Wolfe n'est pas la seule méthode de déplacements admissibles candidate à la résolution du problème (P). En particulier, dans le cas ou Y *n'est pas un polyèdre convexe* ou même lorsque Y est simplement non borné, cette méthode n'est plus valable.

Hogan dit alors « les autres méthodes de déplacements admissibles sont des candidates logiques au traitement numérique de (P) ».

(1) Département de Mathématiques Appliquées, Université de Clermont-Ferrand.

Le premier objet de ce travail est de montrer que l'on peut effectivement modifier les autres méthodes de déplacements admissibles habituelles et que celles-ci convergent. Le formalisme est alors un formalisme général qui permet, en particulier de retrouver la méthode de Hogan.

D'autre part, dans la méthode de Hogan ainsi que dans les méthodes exposées ici, on suppose pouvoir calculer le maximum de v sur un segment. Vu la complexité de v, ceci est pratiquement impossible. Hogan le signale d'ailleurs et laisse ouvert ce problème. Nous proposons ici un procédé de dichotomie qui suppose simplement le calcul de v en différents points d'un segment et nous montrons la convergence du procédé.

Dans tout ce qui va suivre, plusieurs résultats de la théorie des multi-applications ou encore des fonctions dites multivoques seront utilisées, en particulier, ceux qui sont adaptés à des fonctions à valeur extrémale et à des multiapplications définies par des inégalités [voir [3], [5]].

Nous rappelons ici les définitions suivantes :

Soient X un sous-ensemble de $\mathbf{R}^m$, $\mathcal{T}(X)$ l'ensemble des parties de X, Y un sous-ensemble de $\mathbf{R}^n$ et $\Gamma : Y \to \mathcal{T}(X)$ une multiapplication.

Définition 1 : Γ est dite fermée en $\bar{y} \in Y$ si pour toute suite $\{ y_k \}$ dans Y telle que :

$$\bar{y} = \lim_{k\to\infty} y_k, \quad \bar{x} = \lim_{k\to\infty} x_k, \quad x_k \in \Gamma(y_k)$$

Alors $\bar{x} \in \Gamma(\bar{y})$.

Définition 2 : Γ est dite ouverte en $\bar{y} \in Y$ si pour toute suite $\{ y_k \}$ dans Y telle que $\bar{y} = \lim_{k\to\infty} y_k$ alors pour $\bar{x} \in M(\bar{y})$, il existe un entier m et une suite $\{ x_k \}$ dans X tels que :

$$k \geqslant m, x_k \in \Gamma(y_k), \bar{x} = \lim_{k\to\infty} x_k$$

Définition 3 : Γ est dite continue en $\bar{y} \in Y$ si Γ est ouverte et fermée en $\bar{y}$.

Définition 4 : Γ est dite uniformément bornée près de $\bar{y} \in Y$ s'il existe $N\bar{y}$ un voisinage de $\bar{y}$ tel que $U\,\Omega(y)$ est borné.

$$y \in N\bar{y}$$

I. THEOREME GENERAL DE CONVERGENCE D'UNE CLASSE DE METHODES DE DIRECTIONS ADMISSIBLES

1) Hypothèses et notations

Soient Y un sous-ensemble de $\mathbf{R}^n$ et f une fonction concave définie sur Y à valeurs réelles.

Considérons le problème d'optimisation suivant :

$$(P) : \text{Max}\,[f(y)/y \in Y]$$

On définit (comme dans [4]) les méthodes de directions admissibles comme la composition de deux multiapplications $A = M.D.$ où D est la multiapplication de la recherche des directions admissibles et M la multiapplication du choix des points successifs $[D : Y \to \mathcal{P}(\mathbf{R}^n), M : Y \times D(Y) \to \mathcal{P}(Y)$ où $\mathcal{P}(\mathbf{R}^n)$ et $\mathcal{P}(Y)$ sont respectivement l'ensemble des parties de $\mathbf{R}^n$ et Y.]

Posons :

$V = \{ y \in Y : f(y) \geqslant f(y_1) \}$ où y_1 est le point de départ de l'algorithme

$F = \{ y^* \in Y \; : f(y^*) \geqslant f(y) \; \forall \, y \in Y \}$; $V \setminus F = \{ y \in V \; : \; y \notin F \}$

$\alpha(y, d) = \text{Sup}\,(\alpha/y + \alpha d \in Y), \; y \in V, \; d \in D(y)$

$f'(y/d)$: la dérivée de f au point y suivant la direction d.

On supposera, dans ce paragraphe, les hypothèses suivantes :

i) V et $\bigcup_{y \in V} D(y)$ sont bornés, f est concave et continue sur Y.

ii) Pour toute suite $\{ (y_k, d_k) \}_{k \in \mathbf{N}}$ telle que :

$$y_k \in V, d_k \in D(y^k), (\bar{y}, \bar{d}) = \lim_{k \to \infty} (y^k, d^k) \text{ avec } \bar{y} \notin F$$

alors il existe un nombre $n > 0$ et un entier k_0 tels que :

$$k \geqslant k_0, \alpha(y^k, d^k) \geqslant n$$

iii) Pour $y \notin F$, $D(y)$ est non vide et la multiapplication D est fermée en y.

iv) Pour $y \notin F$, $d \in D(y)$ on a $f'(y/d) > 0$.

REMARQUES :

a) l'hypothèse i) est en général supposée dans la pratique, elle dépend du problème posé;

b) les hypothèses ii) et iii) dépendent de l'algorithme que l'on construit;

c) l'hypothèse iv) est, en général, la conséquence immédiate du fait que y n'est pas une solution optimale.

2) Deux algorithmes de directions admissibles

Maintenant, pour toute multiapplication D, définissons alors les deux algorithmes suivants qui représentent une grande classe de méthodes de directions admissibles :

Algorithme 1

A partir d'un point arbitraire $y_1 \in Y$, nous allons construire une suite $\{y_n\}^{n \in N}$ de la manière suivante :

i) $y_n \to d_n \in D(y_n)$

ii) $(y_n, d_n) \to y_{n+1} \in M(y_n, d_n)$ défini par la relation :

$$y_{n+1} = y_n + \alpha_n^* d_n$$

où α_n^* maximise $f(y_n + d_n)$ sur $[0, \alpha(y_n, d_n)]$

Algorithme 2

A partir d'un point arbitraire $y_1 \in Y$, nous allons construire une suite $\{y_n\}_{n \in N}$ de la manière suivante :

i) $y_n \to d_n \in D(y_n)$

ii) $(y_n, d_n) \to y_{n+1} \in M(y_n, d_n)$ défini par :

$$y_{n+1} = y_n + \alpha_n^* d_n$$

$$\alpha_n^* = \min\left(\alpha(y_n, d_n), \rho_n^c\right)$$

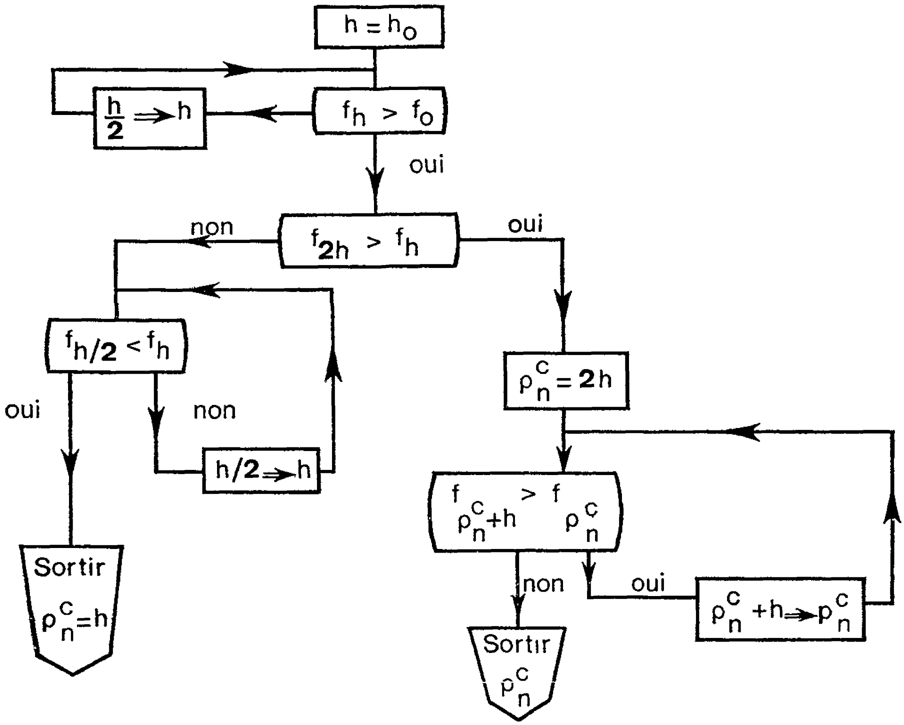
où le choix de ρ_n^c est déterminé par l'algorithme appelé Ro. 5 [voir [1]].

Soient y_n et d_n fixés, posons :

$$f\alpha \begin{cases} = f(y_n + \alpha d_n) & \text{si } y_n + d_n \notin Y, \alpha \geqslant 0 \\ = -\infty & \text{sinon} \end{cases}$$

REMARQUE : Au point de vue de calcul algorithmique, ce qui est important comme on le voit ci-dessous, c'est de comparer les différentes valeurs de f; ceci est immédiat dans Y où f prend des valeurs finies, d'autre part, on sait que $f(y) = -\infty$ si $y \notin Y$ ce qui peut se faire par un test.

On a alors l'organigramme [pour Ro. 5] suivant :



3) Théorème général de convergence pour les algorithmes 1 et 2

Supposons que les hypothèses i), ii), iii), iv) soient vérifiées. Alors, pour chaque algorithme 1 et 2, la limite de toute sous-suite convergente de $\{ y_n \{_{n \in N}$ et il en existe au moins une est solution optimale.

II. EXEMPLES DU CHOIX DES DIRECTIONS ADMISSIBLES VÉRIFIANT LES CONDITIONS DU THEOREME

Lorsque la fonction à optimiser f n'est pas dérivable et si l'on remplace $< f'(y), d >$ par la dérivée directionnelle $f'(y/d)$ alors les méthodes des directions admissibles [comme Frank-Wolfe ...] ne convergent pas en général. Néanmoins, si on fait une certaine modification, par exemple, en utilisant une approximation de la dévirée directionnelle, on peut alors donner des théorèmes de convergence [une telle approximation peut être interprétée comme une procédure antizigzag permettant d'assurer la fermeture de la multiapplication D].

Dans ce paragraphe, nous allons utiliser le théorème précédent pour démontrer la convergence des nouvelles méthodes que nous proposons ici et qui sont comme on peut le constater des méthodes de décomposition.

1° Établissement du problème

Considérons le problème d'optimisation concave suivant :

$$(*) \qquad \text{Max}\,(v(y)/y \in Y)$$

avec :

$$(**) \quad v(y) = \text{Sup}\,[f(x, y)/x \text{ tel que } g(x, y) \leqslant 0],\ Y = \{ y \in \mathbf{R}^n : 1(y) \leqslant 0 \}$$

où f est une fonction définie sur $\mathbf{R}^1 \times \mathbf{R}^n$ à valeur réelle, g une fonction définie sur $\mathbf{R}^1 \times \mathbf{R}^n$ à valeur dans $\mathbf{R}^s$ et 1 une fonction définie sur $\mathbf{R}^n$ à valeur dans $\mathbf{R}^q$.

Soient 1_i, $i = 1 \ldots q$ et g_i, $i = 1 \ldots s$ les composantes de 1 et g.

Posons :

$V = \{ y \in Y : v(y) \geqslant v(y_1) \}$ où y_1 est le point de départ de l'algorithme.

$$F = \{ y^* \in Y : v(y^*) \geqslant v(y)\ \forall\ y \in Y \},\ M(y) = \{ x \in \mathbf{R}^1 : g(x, y) \leqslant 0 \text{ et } v(y) \leqslant f(x, y) \}$$

$$\Omega(y) = \{ x \in \mathbf{R}^1 : g(x, y) \leqslant 0 \}$$

$$V_0 = \{ y \in \mathbf{R}^n : g(x, y) < 0 \text{ pour au moins un } x \in \mathbf{R}^1 \},\ Y_0 = \{ y : 1_i(y) < 0,\ i = 1 \ldots q \}$$

F est donc l'ensemble des solutions optimales de (*) et $M(y)$ l'ensemble des solutions de (**) pour chaque y fixé de Y.

REMARQUE 1 : On se servira du résultat suivant [voir [3]] qui permet d'expliciter la dérivée directionnelle de v :

Si $-f$, g sont convexes, différentiables sur $\mathbf{R}^1 \times N\bar{y}$ où $N\bar{y}$ est un voisinage de $\bar{y}$ et si $v(\bar{y})$ fini, $\bar{x} \in M(\bar{y})$ alors pour $\bar{y} \in \text{Int}\,(V_0)$ on a :

$$\mathring{v}(\bar{y}/d) = \nabla_y f(\bar{x}, \bar{y}) \cdot d + \text{Sup}\,[\nabla_x f(\bar{x}, \bar{y}) \cdot w / \nabla_x g_k(\bar{x}, \bar{y}) \cdot w \leqslant - \nabla_y g_k(\bar{x}, \bar{y}) \cdot d]$$

avec k tel que :

$$g_k(\bar{x}, \bar{y}) = 0.$$

REMARQUE 2 : On s'intéresse plutôt par la suite à une approximation de $\mathring{v}(y/d)$ notée $\mathring{v}_1(y/d)$ définie par :

$$\mathring{v}_1(y/d) = \nabla_y f(x, y) \cdot d + \text{Sup}\,[\nabla_x f(x, y) \cdot w / g(xy) + \nabla_x g(x, y) \cdot w \leqslant - \nabla_y f(x, y]) \cdot d$$

REMARQUE 3 : On remarquera, immédiatement, que :

$$\mathring{v} \geqslant \mathring{v}_1$$

REMARQUE 4 : Dans [3] Hogan montre aussi que :

$$y^* \in F \Leftrightarrow \operatorname{Sup} [\mathring{v}_1(y^*/z - y^*)/z \in Y] = 0$$

2° **Algorithme A** (méthode du type de linéarisation des contraintes [1]).
Tout d'abord, considérons le problème auxiliaire suivant, pour

$$y \in Y, \quad x \in M(y).$$

Problème *Ky*. Déterminer $(\bar\sigma, \bar d, \bar w) \in P(x, y)$ tel que

$$\bar\sigma \geqslant \sigma \,\forall\, (\sigma, d, w) \in P(x, y)$$

avec la multiapplication P définie par :

$$P(x, y) = \begin{cases} (\sigma, d, w) : l(y) + l'(y) \cdot d + \sigma \leqslant 0, \ \|d\| \leqslant M \\ \qquad \nabla_y f(x, y) \cdot d + \nabla_x f(x, y) \cdot w \geqslant \sigma \\ \qquad g(x, y) + \nabla_x g(x, y) \cdot w + \nabla_y g(x, y) \cdot d \leqslant 0, \ \|w\| \leqslant N \end{cases}$$

où M et N sont deux constantes positives.

REMARQUE : $\bar\sigma$ est, en fait, défini comme :

$$\bar\sigma = \max (\sigma(d)/\|d\| \leqslant M)$$

avec :

$$\bar\sigma(d) = \min (-\alpha(d), \mathring{v}_1(y/d)), \alpha(d) = \max [l_i(y) + l_i'(y).\ d/i = 1 \dots q]$$

Les algorithmes que nous proposons sont alors les algorithmes 1 et 2 dans lesquels la multiapplication D de la recherche des directions admissibles est définie par la relation :

$$\bar d \in D(y) \Leftrightarrow \exists\, (\bar\sigma, \bar d) : (\bar\sigma, \bar d, \bar w) \text{ est une solution du problème } Ky.$$

Théorème 1 : Supposons que V soit borné, la multiapplication Ω soit uniformément borné en tout $y \in V$, $-f$, g, 1 soient convexes, différentiables, Y_0 non vide et $Y \subset \operatorname{Int}(V_0)$. Alors, pour chaque algorithme, la limite de toute sous-suite convergente de $\{ y_n \}_{n \in \mathbf{N}}$ et il en existe au moins une est solution optimale.

3° **Algorithme B** (méthode du type Frank-Wolfe généralisée [1]).

Cette fois-ci, la multiapplication D est définie par la relation :

$$\bar{d} \in D(y) \Leftrightarrow \bar{d} \text{ maximise } \mathring{v}_1(y/d) \text{ sur } \Omega(y)$$

$$\text{avec } \Omega(y) = \{ d \in \mathbf{R}^n : y + d \in Y, \|d\| \leqslant L \}$$

où L est une constante positive fixée et on obtient un théorème de convergence identique.

REMARQUE : On peut modifier les contraintes du problème en introduisant des contraintes d'égalité affines.

BIBLIOGRAPHIE

[1] J. CEA, *Optimisation, théorie et algorithmes* [1971, Dunod].

[2] W. HOGAN, *Directionnal derivatives for extremal value functions with applications to the completely convex case* [Graduate School of Management, Université of California].

[3] W. HOGAN, *Points to set maps in mathematical programming* [Feb. 1971, WMSI, Université of California].

[4] W. I. ZANGWILL, *Non linear programming (a unified approach)* [Prentice Hall, 1969].

[5] HUARD : *Cours D.E.A., programmation mathématique* [*à paraître*].