

J. BOUZITAT

Sur la méthode de Wolfe et la méthode de Dantzig en programmation quadratique convexe

RAIRO. Recherche opérationnelle, tome 13, n°2 (1979),
p. 151-184

http://www.numdam.org/item?id=RO_1979__13_2_151_0

© AFCET, 1979, tous droits réservés.

L'accès aux archives de la revue « RAIRO. 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/>

SUR LA MÉTHODE DE WOLFE ET LA MÉTHODE DE DANTZIG EN PROGRAMMATION QUADRATIQUE CONVEXE (*)

par J. BOUZITAT ⁽¹⁾

Résumé. — La méthode de Wolfe et la méthode de Dantzig résolvent les problèmes de programmation quadratique convexe à contraintes linéaires par des algorithmes apparentés à celui du Simplexe. Elles s'appuient sur les conditions de Kuhn et Tucker, qui sont alors nécessaires et suffisantes. Le présent article expose ces méthodes et en donne des justifications théoriques complètes qui, à la connaissance de l'auteur, sont en grande partie originales.

La méthode de Wolfe est ici complétée et se révèle alors plus efficace que ne le laissaient croire des phénomènes de « blocage » apparent, dont l'étude montre qu'ils ne compromettent pas la convergence du processus complété. La « forme courte » de la méthode suffit ainsi pour résoudre tous les problèmes non paramétriques, et la « forme longue » peut être réservée aux problèmes paramétriques.

La justification présentée pour la méthode de Dantzig ne s'appuie pas sur l'étude directe de la suite des tableaux de calcul, mais sur la convexité de la fonction quadratique à minimiser. Elle s'en trouve considérablement allégée.

Enfin la présentation générale des méthodes de Wolfe et de Dantzig est illustrée par un exemple numérique complètement traité, qui permet de les comparer.

Abstract. — Both Wolfe's and Dantzig's methods solve linear-constrained convex quadratic programming problems by simplex-like algorithms. They use the Kuhn-Tucker conditions, which are necessary and sufficient for such problems. This paper presents those two methods, with complete theoretical proofs the greater part of which, to the author's knowledge, is new.

Wolfe's method receives here a complement and is then found to be more efficient than it previously appeared, on account of "blocking" phenomena which are proved not to stop the convergent process of the completed algorithm. The "short form" of the method is consequently applicable to solve any non-parametric problem, and the "long form" may be reserved for parametric problems only.

The present proof of Dantzig's algorithm convergence is not based on the direct study of computation schemata, but uses the convexity of the quadratic function to be minimized, which leads to a quite simple proof.

The general presentation of Wolfe's and Dantzig's methods is illustrated by a numerical problem which is solved by both of them, so as to permit a comparison.

(*) Reçu mars 1978.

(1) Université de Paris-I - Panthéon-Sorbonne.

1. PROBLÈME DE PROGRAMMATION QUADRATIQUE CONVEXE A CONTRAINTES LINÉAIRES

On considère le problème paramétrique suivant :

$$\text{Minimiser } f(\lambda, x) = \lambda \underset{(1, n)}{p}' x + \frac{1}{2} \underset{(n, n)}{x}' C x,$$

sous les contraintes ⁽²⁾ :

$$\underset{(m, n)}{A} x = \underset{(m, 1)}{b} \quad \text{et} \quad x \underset{(n, 1)}{\geq} 0,$$

$x \in \mathbf{R}^n$ étant un vecteur inconnu, λ étant un paramètre réel positif ou nul, _(n, 1)

C étant une matrice symétrique réelle telle que $\underset{(n, n)}{x}' C x \geq 0, \forall x \in \mathbf{R}^n$. _{(1, n) (n, n) (n, 1)}

La forme quadratique $x' C x$ est alors dite semi-définie positive, et elle définit une fonction convexe.

1.1. Propriétés générales

1)

$$[x' C x = 0] \Rightarrow [C x = 0].$$

En effet

$$[x' C x = 0] \Rightarrow [(y + tx)' C (y + tx) = y' C y + 2 ty' C x]$$

et

$$y' C y + 2 ty' C x \geq 0, \forall t \in \mathbf{R} \text{ et } \forall y \in \mathbf{R}^n]$$

$$\Rightarrow [y' C x = 0, \forall y \in \mathbf{R}^n] \Rightarrow [C x = 0].$$

2) Si \bar{x} est solution du problème posé, l'ensemble des solutions est

$$\{ x = \bar{x} + z \mid \lambda p' z = 0, C z = 0, A z = 0, \bar{x} + z \geq 0 \}.$$

(Cet ensemble peut d'ailleurs se réduire à $\{\bar{x}\}$. C'est le cas, en particulier, si $[x' C x = 0] \Rightarrow [x = 0]$.)

En effet tout point x de cet ensemble est solution du problème.

Réciproquement, si \bar{x} et $\bar{x} + z$ sont solutions du problème, il en est de même de $\bar{x} + tz, \forall t \in [0, 1]$ puisque la fonction f est convexe en x , de sorte que

$$a) \quad \lambda p' (\bar{x} + tz) + \frac{1}{2} (\bar{x} + tz)' C (\bar{x} + tz) = \lambda p' \bar{x} + \bar{x}' C \bar{x}, \quad \forall t \in [0, 1],$$

$$(\lambda p' z + \bar{x}' C z) t + \frac{1}{2} (z' C z) t^2 = 0, \quad \forall t \in [0, 1],$$

⁽²⁾ Tous les types de contraintes linéaires peuvent être mis sous cette forme. On pourrait aussi se donner les contraintes sous la forme $Ax \leq b, x \geq 0$ (cf. § 4.1), ou sous d'autres formes plus générales.

d'où

$$z' C z = 0 \quad \text{et} \quad C z = 0 \quad (\text{d'après 1}), \quad \lambda p' z = 0,$$

$$\begin{aligned} \text{b)} \quad A(\bar{x} + tz) &= A\bar{x}, \quad \forall t \in [0, 1] \quad \text{et} \quad A z = 0, \\ \bar{x} + tz &\geq 0, \quad \forall t \in [0, 1] \quad \text{et} \quad \bar{x} + z \geq 0. \end{aligned}$$

$$\text{3)} \quad F(\lambda) = \min_x \{ f(\lambda, x) \mid A x = b \text{ et } x \geq 0 \} = f(\lambda, x_\lambda)$$

est une fonction concave de λ , et $p' x_\lambda$ est une fonction non croissante de λ .

En effet

$$\begin{aligned} F(t\lambda_1 + (1-t)\lambda_2) &= f(t\lambda_1 + (1-t)\lambda_2, x_{t\lambda_1 + (1-t)\lambda_2}) = f(\lambda, x_\lambda) \\ &= t f(\lambda_1, x_\lambda) + (1-t) f(\lambda_2, x_\lambda) \geq t F(\lambda_1) + (1-t) F(\lambda_2), \quad \forall t \in [0, 1]. \end{aligned}$$

D'autre part

$$f(\lambda, x_\lambda) + f(\mu, x_\mu) \leq f(\lambda, x_\mu) + f(\mu, x_\lambda),$$

d'où

$$\lambda p' x_\lambda + \mu p' x_\mu \leq \lambda p' x_\mu + \mu p' x_\lambda,$$

ou

$$(\lambda - \mu)(p' x_\lambda - p' x_\mu) \leq 0.$$

Remarque. — Puisque $f(\lambda, x) \geq f(\lambda, x_\lambda), \forall x \mid A x = b \text{ et } x \geq 0$,

$$[A x = b, x \geq 0, p' x \leq p' x_\lambda] \Rightarrow [x' C x \geq x'_\lambda C x_\lambda],$$

de sorte que

$$\boxed{x'_\lambda C x_\lambda = \min \{ x' C x \mid A x = b, x \geq 0, p' x \leq p' x_\lambda \}}.$$

1.2. Effet de perturbations portant sur la matrice C

Si la forme $x' C x$ est semi-définie positive, et non définie positive, il se peut que la solution du problème posé existe sans être unique. Soit alors \hat{x} la solution de norme $(\hat{x}' \hat{x})^{1/2}$ minimale. Cette solution existe et est unique puisque l'ensemble des solutions est fermé et convexe.

Considérons le problème déduit du problème posé par perturbation de la matrice C , remplacée par $(C + \varepsilon I)$, où I est une matrice unité et où ε est un scalaire positif que l'on fera tendre vers zéro.

La forme $x'(C + \varepsilon I)x$ étant définie positive, ce problème a une solution unique \hat{x}_ε , minimisant la fonction économique

$$f_\varepsilon(\lambda, x) = \lambda p'x + \frac{1}{2}x'(C + \varepsilon I)x = f(\lambda, x) + \frac{\varepsilon}{2}x'x,$$

sous les contraintes $Ax = b$ et $x \geq 0$.

D'après la définition de \hat{x} et \hat{x}_ε ,

$$f(\lambda, \hat{x}) + \frac{\varepsilon}{2}\hat{x}'\hat{x} = f_\varepsilon(\lambda, \hat{x}) \geq f_\varepsilon(\lambda, \hat{x}_\varepsilon) = f(\lambda, \hat{x}_\varepsilon) + \frac{\varepsilon}{2}\hat{x}'_\varepsilon\hat{x}_\varepsilon \geq f(\lambda, \hat{x}) + \frac{\varepsilon}{2}\hat{x}'\hat{x}_\varepsilon.$$

D'où

$$\boxed{\lim_{\varepsilon \rightarrow 0} f_\varepsilon(\lambda, \hat{x}_\varepsilon) = f(\lambda, \hat{x})} \quad \text{et} \quad \boxed{\hat{x}'_\varepsilon\hat{x}_\varepsilon \leq \hat{x}'\hat{x}}$$

(On montre de même que $\hat{x}'_\varepsilon\hat{x}_\varepsilon$ est fonction non croissante de $\varepsilon > 0$.)

Il en résulte que

$$\boxed{\lim_{\varepsilon \rightarrow 0} \hat{x}_\varepsilon = \hat{x}.}$$

En effet, s'il n'en était pas ainsi, l'ensemble des $\hat{x}_\varepsilon | \varepsilon > 0$ étant borné, il existerait une suite convergente $\hat{x}_{\varepsilon_1}, \hat{x}_{\varepsilon_2}, \dots, \hat{x}_{\varepsilon_n}, \dots$ avec $\varepsilon_1 > \varepsilon_2 > \dots > \varepsilon_n > \dots$ et $\varepsilon_n \xrightarrow{n \rightarrow \infty} 0$, telle que $\lim_{n \rightarrow \infty} \hat{x}_{\varepsilon_n} = x_0 \neq \hat{x}$.

Puisque $f_\varepsilon(\lambda, x)$ est fonction continue de ε et de x , on aurait alors

$$f(\lambda, x_0) = \lim_{n \rightarrow \infty} f_{\varepsilon_n}(\lambda, \hat{x}_{\varepsilon_n}) = f(\lambda, \hat{x}), \quad \text{avec} \quad x'_0 x_0 \leq \hat{x}'\hat{x},$$

de sorte que x_0 , vérifiant de plus les contraintes, serait une seconde solution de norme minimum du problème posé. Or il n'en existe pas d'autre que \hat{x} .

Ainsi, le remplacement de la forme $x' C x$ semi-définie positive par $x'(C + \varepsilon I)x$ définie positive ($\varepsilon > 0$) conduit à une solution \hat{x}_ε du problème perturbé qui tend vers la solution \hat{x} de plus petite norme du problème initial lorsque ε tend vers zéro.

Cette méthode de perturbations permet, au moins en théorie, d'approcher d'aussi près qu'on le veut la solution de plus petite norme (supposée finie) d'un problème à forme quadratique semi-définie positive, par la résolution d'un problème à forme quadratique définie positive, et par conséquent à solution unique.

Les calculs liés à l'emploi de cette méthode risquent cependant d'être lourds et difficiles à programmer. Mais certaines démonstrations peuvent s'en trouver simplifiées.

1.3. Conditions de Kuhn et Tucker

Associons au problème posé le *lagrangien* (formé sans tenir compte des conditions de signe $x \geq 0$) :

$$l(x, u) = \lambda \underset{(1, n)}{p'} x + \frac{1}{2} x' \underset{(n, n)}{C} x + u' \left(\underset{(1, m)}{A} x - \underset{(m, 1)}{b} \right),$$

qui est égal à la fonction économique $f(\lambda, x)$ pour tout programme x réalisable.

Les conditions de Kuhn et Tucker s'écrivent :

$$\begin{aligned} \nabla'_x l(x, u) &= \lambda \underset{(n, 1)}{p} + \underset{(n, n)}{C} x + \underset{(n, m)}{A'} u = v \ (\geq 0), \\ &\quad - \underset{(m, 1)}{b} + \underset{(m, n)}{A} x = 0, \\ x &\geq 0, \quad v \geq 0, \quad v_j x_j = 0, \quad \forall j. \end{aligned}$$

Ces conditions sont *nécessaires* parce que les contraintes du problème posé, étant linéaires, sont qualifiées. Elles sont *suffisantes* parce que les contraintes sont linéaires et la fonction économique à minimiser est convexe.

Les conditions de Kuhn et Tucker sont ici linéaires par rapport à l'ensemble des n variables x_j et des $(m+n)$ multiplicateurs u_i et v_j , si l'on excepte les n relations d'exclusion (ou de complémentarité) $v_j x_j = 0$.

Compte tenu des $(n+m)$ égalités linéaires précédentes,

$$2 f(\lambda, x) = \lambda p' x + (\lambda p' + x' C) x = \lambda p' x + (v' - u' A) x$$

$$2 f(\lambda, x) = \lambda p' x - u' b + v' x,$$

le dernier terme (non linéaire) étant nul quand les relations d'exclusion sont vérifiées.

D'après les relations d'exclusion, *dans toute solution des conditions de Kuhn et Tucker il y a au plus $(m+n)$ des $(m+2n)$ variables scalaires u_i (de signes quelconques), $v_j \geq 0$, $x_j \geq 0$ qui soient non nulles.*

Si (x, u, v) et $(x + \delta x, u + \delta u, v + \delta v)$ sont deux solutions du système des $(n+m)$ égalités linéaires précédentes,

$$\begin{cases} C \delta x + A' \delta u = \delta v, \\ A \delta x = 0, \end{cases}$$

d'où

$$\delta v' \cdot \delta x = \delta x' \cdot C \cdot \delta x \geq 0.$$

Si ces deux solutions vérifient *toutes* les conditions de Kuhn et Tucker,

$$v' x = 0, \quad (v' + \delta v')(x + \delta x) = v' \cdot \delta x + \delta v' \cdot x + \delta v' \cdot \delta x = 0,$$

et

$$\forall j \in \{1, 2, \dots, n\}, \quad \left\{ \begin{array}{l} [x_j > 0] \Rightarrow [v_j = 0] \\ \qquad \qquad \Rightarrow [\delta v_j \geq 0] \Rightarrow [v_j \cdot \delta x_j = 0 \text{ et } \delta v_j \cdot x_j \geq 0], \\ [v_j > 0] \Rightarrow [x_j = 0] \\ \qquad \qquad \Rightarrow [\delta x_j \geq 0] \Rightarrow [v_j \cdot \delta x_j \geq 0 \text{ et } \delta v_j \cdot x_j = 0], \end{array} \right.$$

d'où

$$\forall j \in \{1, 2, \dots, n\}, \quad v_j \cdot \delta x_j = 0 \quad \text{et} \quad \delta v_j \cdot x_j = 0,$$

puisque $v' \cdot \delta x$, $\delta v' \cdot x$, $\delta v' \cdot \delta x$ sont positifs ou nuls et de somme nulle.

Ainsi, *s'il existe une solution des conditions de Kuhn et Tucker pour laquelle $x_j > 0$ (ou $v_j > 0$), alors dans toute solution de ces conditions $v_j = 0$ (ou $x_j = 0$).* Cette propriété, classique en programmation linéaire, est ainsi étendue à la programmation quadratique *convexe* à contraintes linéaires.

Il en résulte que l'ensemble des points $(\bar{x}, \bar{u}, \bar{v})$ solutions des conditions de Kuhn et Tucker peut être défini par un système linéaire obtenu à partir de ces conditions par l'annulation d'une variable déterminée (v_j ou x_j) dans chacun des n couples de variables complémentaires.

Cet ensemble est donc un tronçon de l'espace \mathbf{R}^{m+2n} , de sorte qu'il peut être défini à partir de ses points extrêmes, et éventuellement de ses directions d'infinitude extrêmes.

C'est ce qui conduit à chercher des *solutions de base* (au sens de la méthode du Simplexe) du système des conditions linéaires de Kuhn et Tucker ($m+n$ égalités et $2n$ conditions de signe portant sur $m+2n$ variables), en leur imposant de vérifier les n relations d'exclusion $v_j x_j = 0, \forall j$.

Tel est le principe de la résolution des problèmes de programmation quadratique convexe à contraintes linéaires par des *méthodes simpliciales*. Deux de ces méthodes vont être étudiées : la méthode de Wolfe et la méthode de Dantzig.

2. MÉTHODE DE WOLFE COMPLÉTÉE

2.1. Principe

On applique la méthode du Simplexe à un système linéaire déduit des conditions linéaires de Kuhn et Tucker par l'introduction de variables

artificielles positives ou nulles dont on se propose de minimiser la somme, avec une condition supplémentaire excluant la présence simultanée dans la base de deux variables complémentaires x_j et v_j . Les relations d'exclusion $v_j x_j = 0, \forall j$ sont ainsi constamment vérifiées.

L'optimum du problème de programmation quadratique posé, pour lequel les conditions de Kuhn et Tucker sont nécessaires et suffisantes, est atteint quand toutes les variables artificielles sont annulées. (Dès qu'une variable artificielle s'annule en sortant de la base, elle peut être supprimée.)

Mais, en raison de la restriction imposée par le respect des relations d'exclusion, il n'est pas sûr *a priori* que la méthode permette toujours d'atteindre l'optimum quand il existe. Et certains exemples ont conduit à penser que le processus pouvait en effet « se bloquer » pour une valeur strictement positive de la somme des variables artificielles, même quand cette somme a un minimum nul. Nous verrons cependant qu'un tel blocage n'est qu'apparent, et que le processus de calcul peut être aménagé de telle manière qu'il conduise toujours à l'optimum quand cet optimum existe, pourvu que l'éventualité d'un « cyclage » soit écartée (cette réserve, d'une faible importance pratique, accompagnant toujours l'application de la méthode du Simplexe).

La *forme courte* du processus résout le problème de minimisation de $f(\lambda, x)$ pour une valeur déterminée de λ . La *forme longue* du processus résout le même problème pour toutes les valeurs positives ou nulle du paramètre λ .

2.2. Forme courte du processus (pour une valeur donnée de λ).

On considère le système linéaire suivant :

$$\begin{array}{l} C x + A' u - I v \quad + \quad E z = -\lambda p, \\ \begin{array}{cccc} (n, n) & (n, m) & (n, n) & (n, n) \end{array} \quad \begin{array}{c} (n, 1) \\ (n, 1) \end{array} \\ A x \quad + \quad w = b, \\ \begin{array}{ccc} (m, n) & (m, 1) & (m, 1) \end{array} \\ x \geq 0, \quad v \geq 0, w \geq 0, z \geq 0, \end{array}$$

où I est une matrice unité, E est une matrice diagonale dont les éléments diagonaux e_1, e_2, \dots, e_n , égaux à 1 ou à -1 , ne seront fixés qu'à la fin de la première phase du processus (c'est-à-dire après l'annulation de w).

Les coordonnées de w et de z sont des variables artificielles.

On suppose $b \geq 0$, ce qui ne nuit pas à la généralité.

1^{re} phase. — On part du programme de base

$$\begin{cases} Ez = -\lambda p, \\ w = b, \end{cases} \quad \begin{cases} x = 0, \\ u = 0, \\ v = 0, \end{cases}$$

et l'on applique la méthode du Simplexe pour *minimiser la fonction* $\sum_{i=1}^m w_i$, en tenant compte des conditions de signe $w \geq 0$ et $x \geq 0$, mais en laissant les $(m+n)$ variables u_i et v_j hors de la base (donc nulles), et les n variables $e_j z_j$ dans la base sans se préoccuper de leurs signes (puisque les coefficients e_j ne sont pas fixés).

Si les contraintes $\begin{cases} Ax = b \\ x \geq 0 \end{cases}$ du problème initial sont compatibles, ce que l'on suppose, cette première phase du processus conduit, par des échanges entre les w_i et les x_j , à annuler les m variables artificielles w_i , et à supprimer éventuellement des contraintes redondantes parmi les m égalités $Ax = b$, ainsi que les coordonnées correspondantes de u .

A la fin de cette première phase, on dispose d'un programme x réalisable du problème initial, w est nul et peut être supprimé, et la base est constituée par certaines variables x_j (en nombre égal au rang de A) et toutes les variables z_j . On choisit alors les coefficients e_j , égaux à 1 ou à -1 , de telle manière que les valeurs des n variables z_j soient toutes positives ou nulles.

Pour simplifier l'écriture, on supposera que rang de $A = m$.

2^e phase. — A partir du programme de base ainsi obtenu, où $z \geq 0$ et $v = 0$, on applique la méthode du Simplexe pour *minimiser la fonction* $\sum_{j=1}^n z_j$, en tenant compte des conditions de signe $z \geq 0$, $x \geq 0$ et $v \geq 0$, en faisant d'abord entrer dans la base, où elles resteront, les m variables u_i [de signes quelconques ⁽³⁾], et en s'interdisant, à chaque pas du processus, de faire entrer dans la base une variable v_j (ou x_j) si la variable complémentaire x_j (ou v_j) s'y trouve déjà. Ainsi les relations d'exclusion $v_j x_j = 0, \forall j$ sont-elles constamment vérifiées.

Alors, comme on l'établira au paragraphe 2.4.2,

— ou bien le processus conduit à annuler les n variables artificielles z_j , et donne ainsi les solutions extrêmes du problème initial (et de ses conditions de Kuhn et Tucker);

⁽³⁾ On pourrait supprimer, dans les tableaux de calcul ultérieurs, les lignes relatives aux m variables u_i .

– ou bien le processus se termine pour une valeur non nulle de $\sum_{j=1}^n z_j$, et conduit

à montrer que $f(\lambda, x)$ n'a pas de minimum fini dans le problème initial.

Dans le cas où $\left\{ \begin{array}{l} Cr=0 \\ Ar=0 \end{array} \right\} \Rightarrow [r=0]$, et en particulier si $x' C x > 0, \forall x \neq 0$, on établira que $z=0$ à la fin du processus, de sorte que le problème initial admet une solution unique (cf. § 1.1).

2.3. Forme longue du processus (pour λ croissant à partir de 0)

On considère le système linéaire suivant :

$$\begin{array}{r} \boxed{\begin{array}{l} C x + A' u - I v + p \lambda + E z = 0, \\ \begin{array}{cccccc} (n, n) & (n, m) & (n, n) & (n, 1) & (n, n) & (n, 1) \end{array} \\ A x + w = b, \\ \begin{array}{ccc} (m, n) & (m, 1) & (m, 1) \end{array} \\ x \geq 0, w \geq 0, \quad v \geq 0, \quad \lambda \geq 0, \quad z \geq 0, \end{array}} \end{array}$$

où les notations sont celles qui ont été définies au paragraphe 2.2 à propos de la forme courte, et où l'on suppose encore $b \geq 0$.
(m, 1)

1^{re} et 2^e phases. – On part du programme de base (dégénéré) :

$$\left\{ \begin{array}{l} E z = 0, \\ w = b, \end{array} \right. \left\{ \begin{array}{l} x = 0, \\ u = 0, \\ v = 0, \\ \lambda = 0 \end{array} \right.$$

et l'on applique successivement les deux phases de la forme courte du processus pour $\lambda=0$, la variable λ restant hors de la base.

A la fin de ces deux phases, on obtient les solutions extrêmes du problème initial et de ses conditions de Kuhn et Tucker pour la valeur 0 du paramètre λ . (Il est exclu que $f(0, x) = (1/2)x' C x$ puisse décroître indéfiniment puisque $x' C x \geq 0, \forall x$.)

Toutes les variables artificielles w_i et z_j sont alors nulles et peuvent être supprimées.

Pour simplifier l'écriture, on supposera que $\text{rang de } \underset{(m, n)}{A} = m$.

3^e phase. — On part du (ou de l'un des) programme(s) de base ainsi obtenu(s) pour le système linéaire

$$\left\{ \begin{array}{l} Cx + A'u - Iv + p\lambda = 0, \\ Ax \qquad \qquad \qquad = b, \\ x \geq 0, \quad v \geq 0, \lambda \geq 0, \end{array} \right.$$

programme dans lequel $\lambda = 0$ (λ étant hors de la base) et $v_j x_j = 0, \forall j$, et l'on applique la méthode du Simplexe pour *maximiser* λ (ou *minimiser* $-\lambda$), sous ces contraintes linéaires, en cherchant d'abord à faire entrer λ dans la base, et *en s'interdisant, à chaque pas du processus, de faire entrer dans la base une variable v_j (ou x_j) si la variable complémentaire x_j (ou v_j) s'y trouve déjà*, de sorte que $v_j x_j = 0, \forall j$.

Alors, comme on l'établira au paragraphe 2.4.3,

— *ou bien le processus conduit à faire croître indéfiniment λ à partir de 0, en passant par une suite finie de valeurs $0 = \lambda_0 < \lambda_1 < \dots < \lambda_{k-1} < \lambda_k$ (éventuellement réduite à $0 = \lambda_0$), et donne la suite des solutions extrêmes correspondantes du problème initial (et de ses conditions de Kuhn et Tucker), ainsi qu'une direction d'infinitude pour $\lambda > \lambda_k$, à partir de quoi il est facile de déterminer les solutions de ce problème, par interpolation ou par extrapolation linéaire, pour toute valeur positive du paramètre λ ;*

— *ou bien le processus se termine pour $\lambda = 0$, et conduit à montrer que $f(\lambda, x)$ n'a pas de minimum fini dans le problème initial si $\lambda > 0$.*

Dans le cas où $\left\{ \begin{array}{l} Cr = 0 \\ Ar = 0 \end{array} \right\} \Rightarrow [r = 0]$, et en particulier si $x' C x > 0, \forall x \neq 0$, on

établira que le processus ne peut pas se terminer pour $\lambda = 0$, de sorte que le problème initial (dont les contraintes sont supposées compatibles) admet une solution unique (*cf.* § 1.1) pour toute valeur positive ou nulle du paramètre λ . (Pour chaque valeur de λ , cette solution peut être obtenue par la forme courte du processus.)

2.4. Étude de l'efficacité de la méthode de Wolfe complétée

2.4.1. ÉTUDE GÉNÉRALE DES CAS DE « BLOCAGE » DU PROCESSUS

Soit le problème de programmation linéaire suivant :

$ \begin{array}{l} Cx + A'u - Iv + Py = a, \\ \begin{array}{cccc} (n, n) & (n, m) & (n, n) & (n, l) & (n, 1) \end{array} \\ Ax = b, \\ \begin{array}{cc} (m, n) & (m, 1) \end{array} \\ x \geq 0, \quad v \geq 0, \quad y \geq 0, \\ \text{Minimiser } \psi = q'y, \\ \begin{array}{c} (1, l) \end{array} \end{array} $	<p>où rang de $A = m$.</p>
---	---------------------------------------

On cherche à résoudre ce programme linéaire par la méthode du Simplexe, en faisant d'abord entrer dans la base les m variables u_i (de signes quelconques) et en tenant compte de la restriction de Wolfe qui, pour assurer le respect des relations d'exclusion $v_j x_j = 0, \forall j$, interdit de faire entrer dans la base une variable v_j (ou x_j) si la variable complémentaire x_j (ou v_j) s'y trouve déjà.

Soit un tableau du Simplexe conduisant à un « blocage », en ce sens que, compte tenu de la restriction de Wolfe, il est impossible de faire décroître $\psi = q'y$ en faisant croître à partir de 0 l'une des variables hors base qui peuvent entrer dans la base, ce qui se traduit par le fait que les coefficients de ces variables sont positifs ou nuls dans l'expression de ψ en fonction des variables hors base.

Soit $(\bar{x}, \bar{u}, \bar{v}, \bar{y})$ le programme réalisable extrême correspondant à ce tableau, programme où la base ne contient aucun couple de variables complémentaires v_j, x_j , et soit la partition V, W, X de l'ensemble d'indices $\{1, 2, \dots, n\}$ telle que

$$\left\{ \begin{array}{l}
 \bar{v}_j \text{ est dans la base pour } j \in V, \text{ hors de la base pour } j \in W \text{ et pour } j \in X, \\
 \bar{x}_j \text{ est hors de la base pour } j \in V \text{ et pour } j \in W, \text{ dans la base pour } j \in X.
 \end{array} \right.$$

Puisque rang de $A = m$, il y a $(m+n)$ variables de base, dont les m variables \bar{u}_i , et, s'il y a au moins une variable \bar{y}_h dans la base, $\text{card } V + \text{card } X < n$ et $W \neq \emptyset$.

Le programme linéaire considéré et le programme dual s'écrivent :

$ \left\{ \begin{array}{l} Cx + A'u - Iv + Py = a, \\ Ax = b, \\ x \geq 0, \\ u \text{ quelconque,} \\ v \geq 0, \\ y \geq 0, \\ \psi = q'y \text{ minimum,} \end{array} \right. $	$ \left\{ \begin{array}{l} r' \text{ quelconque,} \\ s' \text{ quelconque,} \\ r'C + s'A \leq 0, \\ r'A' = 0, \\ -r'I \leq 0, \\ r'P \leq q', \\ \varphi = r'a + s'b \text{ maximum.} \end{array} \right. $
--	---

Soit (\bar{r}', \bar{s}') le programme de base (en général irréalisable) associé, dans le dual, au programme extrême $(\bar{x}, \bar{u}, \bar{v}, \bar{y})$ et correspondant au même tableau du Simplexe. Ce couple de programmes de base associés vérifie les relations d'exclusion qui impliquent l'égalité des deux fonctions à optimiser : $\bar{\psi} = q' \bar{y} = \bar{r}' a + \bar{s}' b = \bar{\varphi}$.

De plus, pour tout programme (x, u, v, y) réalisable,

$$\psi - \bar{\psi} = \psi - \bar{\varphi} = -\bar{r}' a - \bar{s}' b + q' y = -(\bar{r}' C + \bar{s}' A)x - \bar{r}' A' u + \bar{r}' v + (q' - \bar{r}' P) y,$$

ce qui donne l'expression de ψ en fonction des variables hors base (en vertu des relations d'exclusion entre (\bar{r}', \bar{s}') et $(\bar{x}, \bar{u}, \bar{v}, \bar{y})$).

D'après l'hypothèse de « blocage » qui a été explicitée, et compte tenu de la structure du couple de programmes de base associés $(\bar{x}, \bar{u}, \bar{v}, \bar{y})$ et (\bar{r}', \bar{s}') ,

pour $j \in V$, $\bar{x}_j = 0$,	$\bar{r}' C^j + \bar{s}' A^j q c q$,	$\bar{v}_j \geq 0$ (dans la base),	$\bar{r}_j = 0$,
pour $j \in W$, $\bar{x}_j = 0$,	$\bar{r}' C^j + \bar{s}' A^j \leq 0$,	$\bar{v}_j = 0$,	$\bar{r}_j \geq 0$,
pour $j \in X$, $\bar{x}_j \geq 0$ (dans la base),	$\bar{r}' C^j + \bar{s}' A^j = 0$,	$\bar{v}_j = 0$,	$\bar{r}_j q c q$;
$\bar{u} q c q$ (dans la base),	$\bar{r}' A' = 0$,		
$\bar{y} \geq 0$,	$\bar{r}' P \leq q'$,	$\bar{r}' P \bar{y} = q' \bar{y} = \bar{r}' a + \bar{s}' b$,	

car la restriction de Wolfe interdit de faire entrer dans la base $x_j | j \in V$ et $v_j | j \in X$, mais permet d'y faire entrer $x_j | j \in W$ ou $v_j | j \in W$, ou tout y_h qui ne s'y trouve pas déjà.

Ces relations impliquent (compte tenu du fait que $x' C x \geq 0, \forall x$) :

$$\left\{ \begin{array}{l} (\bar{r}' C + \bar{s}' A) \bar{r} \leq 0 \\ A \bar{r} = 0 \end{array} \right\} \Rightarrow [\bar{r}' C \bar{r} \leq 0] \Rightarrow [\bar{r}' C \bar{r} = 0] \Rightarrow [C \bar{r} = 0],$$

$$\left\{ \begin{array}{l} (\bar{r}' C + \bar{s}' A) \bar{x} = 0 \\ \bar{r}' C = 0 \end{array} \right\} \Rightarrow [\bar{s}' A \bar{x} = 0] \Rightarrow [\bar{s}' b = 0].$$

De plus

$$[\bar{s}' A^x = 0] \Rightarrow [\bar{s}' = 0], \text{ puisque rang de } A^x = \text{rang de } A = m.$$

(1, m)

Ainsi, dans les conditions de « blocage » envisagées, il existe un vecteur \bar{r} tel que (n, 1)

$\left\{ \begin{array}{l} C \bar{r} = 0, \\ (n, n) \end{array} \right. \quad \left\{ \begin{array}{l} \bar{r}_V = 0, \\ \bar{r}_W \geq 0, \end{array} \right.$
$\left\{ \begin{array}{l} A \bar{r} = 0, \\ (m, n) \end{array} \right. \quad \left\{ \begin{array}{l} P' \bar{r} \leq q, \\ (l, n) \end{array} \right. \quad , \quad \left\{ \begin{array}{l} a' \bar{r} = q' \bar{y} \\ (1, n) \end{array} \right. \quad \left\{ \begin{array}{l} \\ (1, l) \end{array} \right.$

et

$$\psi - \bar{\psi} = \bar{r}' v + (q' - \bar{r}' P) y,$$

pour tout programme (x, u, v, y) réalisable.

Si $q' \bar{y} \neq 0$, ou si $q \not\geq 0$, $\bar{r} \neq 0$. Il existe alors, en général, d'autres programmes réalisables (x, u, v, y) donnant la même valeur $\bar{\psi}$ à $\psi = q' y$, soit $(\bar{x} + t\bar{r}, \bar{u}, \bar{v}, \bar{y})$ pour tout scalaire t tel que $t\bar{r}_W \geq 0$ et $\bar{x}_X + t\bar{r}_X \geq 0$. Ces programmes vérifient les relations d'exclusion $\bar{v}_j(\bar{x}_j + t\bar{r}_j) = 0, \forall j$.

S'il y a au moins une variable \bar{y}_h dans la base, $W \neq \emptyset$ et, puisque le programme $(\bar{x}, \bar{u}, \bar{v}, \bar{y})$ est extrême, $\bar{r}_W \neq 0$ de sorte que $t\bar{r}_W \geq 0 \Rightarrow t \geq 0$. En effet, si \bar{r}_W était nul, la condition $\bar{x}_X + t\bar{r}_X \geq 0$ permettrait au scalaire t de varier dans un intervalle encadrant 0, du moins dans le cas général où $\bar{x}_X > 0$ (on peut toujours se ramener à ce cas par l'un des procédés qui permettent d'éviter tout risque de « cyclage » dans l'application de la méthode du Simplexe).

Alors $P'(\bar{x} + t\bar{r}) \leq P'\bar{x} + tq$ et $a'(\bar{x} + t\bar{r}) = a'\bar{x} + tq'y$.

— Si $\bar{r}_X \geq 0$, le scalaire t peut croître indéfiniment à partir de 0.

— Si $\bar{r}_X \not\geq 0$, le scalaire t peut croître de 0 à un maximum

$$\theta = \min \left\{ -\frac{\bar{x}_j}{\bar{r}_j} \mid j \in X \text{ et } \bar{r}_j < 0 \right\}.$$

Dans le second cas, où \bar{r}_j n'est pas positif ou nul pour tout $j \in X$, il existe donc un programme réalisable $(x, u, v, y) = (\bar{x} + \theta\bar{r}, \bar{u}, \bar{v}, \bar{y})$ dans lequel

$$\left\{ \begin{array}{ll} x_V = \bar{x}_V = 0, & v_V = \bar{v}_V \geq 0, \\ x_W = \theta \bar{r}_W \geq 0, & v_W = \bar{v}_W = 0, \\ x_X = \bar{x}_X + \theta \bar{r}_X \geq 0, & v_X = \bar{v}_X = 0, \\ y_Y = \bar{y}_Y \geq 0, & y_{\bar{Y}} = \bar{y}_{\bar{Y}} = 0 \end{array} \right.$$

(si \bar{y}_h est dans la base pour $h \in Y$, hors de la base pour $h \in \bar{Y}$).

Mais le passage de $(\bar{x}, \bar{u}, \bar{v}, \bar{y})$ à $(\bar{x} + \theta\bar{r}, \bar{u}, \bar{v}, \bar{y})$ ne correspond pas à un pas de la méthode du Simplexe si \bar{r}_W a plusieurs coordonnées strictement positives, car plusieurs variables hors base x_j telles que $j \in W$ et $\bar{r}_j > 0$ prennent alors des valeurs strictement positives tandis qu'une variable de base au moins s'annule.

x_j telle que $j \in X$, $\bar{r}_j < 0$ et $\bar{x}_j / |\bar{r}_j|$ minimum parmi les rapports analogues.

Cependant, il existe toujours un programme extrême dans lequel $x_V = 0, v_W = 0, v_X = 0, y_{\bar{Y}} = 0$ et $x_{\bar{Y}} = 0$, et, à partir du tableau du Simplexe correspondant à $(\bar{x}, \bar{u}, \bar{v}, \bar{y})$, ce programme extrême peut être atteint par un ou plusieurs pas de la méthode du Simplexe.

	x'_w	x'_v	v'_w	v'_x	y'_y	
ψ	0	0	$-\bar{r}'_w$	$-\bar{r}'_x$	$-(q' - \bar{r}' P)_y$	$\bar{\psi}$
x_x						\bar{x}_x
v_v						\bar{v}_v
y_y						\bar{y}_y
u						\bar{u}

A chacun de ces pas, x_v, v_w, v_x, y_y restent hors de la base, tandis que des échanges se produisent entre une variable hors base et une variable de base dans l'ensemble des variables appartenant à x_w, x_x, v_v, y_y . Chacun de ces échanges respecte la restriction de Wolfe et ne modifie pas la valeur de la fonction ψ ni son expression en fonction de v et y .

Pour éviter une exploration aléatoire de l'ensemble des programmes extrêmes où $x_v=0, v_w=0, v_x=0, y_y=0$, on peut appliquer à cette exploration la méthode du Simplexe en se proposant de minimiser la variable x_j , qu'il est facile de déterminer à partir du tableau ci-dessus où figurent les \bar{x}_j et les \bar{r}_j pour $j \in X$.

On est assuré de pouvoir faire sortir x_j de la base à la fin de cette exploration, et le tableau du Simplexe ainsi obtenu ne conduit pas à un blocage, puisque x_j et v_j sont alors hors de la base et que $\bar{r}_j < 0$. La restriction de Wolfe permet donc de faire entrer v_j dans la base, et la croissance de v_j à partir de 0 fait décroître $\psi = \bar{\psi} + \bar{r}' v + (q' - \bar{r}' P) y$, puisque $\bar{r}_j < 0$.

(Pour être sûr d'obtenir ainsi une décroissance stricte de la fonction ψ , on peut éventuellement utiliser l'un des procédés qui permettent d'éviter tout risque de « cyclage » dans l'application de la méthode du Simplexe.)

2.4.2. APPLICATION A LA FORME COURTE DU PROCESSUS

Le phénomène de « blocage » qui vient d'être étudié peut a priori se présenter dans la 2^e phase du processus sous forme courte, et les résultats obtenus au paragraphe 2.4.1 peuvent être appliqués à ce cas avec les notations suivantes :

$$P = E, \quad y = z, \quad a = -\lambda p, \quad q' = (1, 1, \dots, 1).$$

$(n, 1) \quad (n, n) \quad (l, 1) \quad (n, 1) \quad (n, 1) \quad (n, 1) \quad (1, l)$

Ainsi, dans les conditions de « blocage » envisagées, il existe un vecteur \bar{r} tel que

$$\left\{ \begin{array}{l} C \bar{r} = 0, \\ (n, n) \\ A \bar{r} = 0, \\ (m, n) \end{array} \right. \quad E \bar{r} \leq \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} \quad -\lambda p' \bar{r} = \sum_{j=1}^n \bar{z}_j.$$

1) Si $\lambda p' \bar{r} = 0$, $\sum_{j=1}^n \bar{z}_j = 0$, ce qui implique $\bar{z}_j = 0, \forall j$. Toutes les variables artificielles sont alors nulles, et le tableau du Simplexe considéré donne une solution extrême du problème initial (et de ses conditions de Kuhn et Tucker). Si cette solution n'est pas unique, le tableau présente une dégénérescence qui permet de déterminer toutes les solutions en poursuivant le processus, compte tenu des propriétés établies au paragraphe 1.3. *Le problème initial est ainsi résolu.*

Telle est, en particulier, la fin du processus si $\lambda = 0$, ou si $\left\{ \begin{matrix} Cr = 0 \\ Ar = 0 \end{matrix} \right\} \Rightarrow [r = 0]$, ce qui est le cas si la forme quadratique $x' C x$ est définie positive (c'est-à-dire strictement positive pour tout $x \neq 0$). Dans ce dernier cas, auquel on peut toujours se ramener par la méthode des perturbations (voir § 1.2), la solution du problème initial est unique (voir § 1.1).

2) Si, au contraire, $\sum_{j=1}^n \bar{z}_j > 0$ au moment du « blocage » (ce qui n'est possible que si $\lambda > 0$ et $x' C x$ semi-définie positive), $\bar{r} \neq 0$ et il existe en général un ensemble infini de programmes réalisables $(\bar{x} + t\bar{r}, \bar{u}, \bar{v}, \bar{z})$, donnant la même valeur à $\sum_{j=1}^n z_j$ et vérifiant les relations d'exclusion $\bar{v}_j(\bar{x}_j + t\bar{r}_j) = 0, \forall j$.

La fonction à minimiser dans le problème initial y prend les valeurs

$$f(\lambda, \bar{x} + t\bar{r}) = \lambda p'(\bar{x} + t\bar{r}) + \frac{1}{2} \bar{x}' C \bar{x} = f(\lambda, \bar{x}) - t \sum_{j=1}^n \bar{z}_j,$$

car

$$C\bar{r} = 0 \quad \text{et} \quad \lambda p' \bar{r} = - \sum_{j=1}^n \bar{z}_j < 0.$$

Deux cas sont alors possibles, d'après l'étude faite au paragraphe 2.4.1 :

– ou bien le scalaire t peut croître indéfiniment à partir de 0 (cas où $\bar{r}_x \geq 0$), et $f(\lambda, \bar{x} + t\bar{r}) \rightarrow -\infty$ quand $t \rightarrow +\infty$, de sorte que la fonction convexe $f(\lambda, x)$ n'a pas de minimum fini dans le problème initial;

– ou bien le scalaire t peut croître de 0 à un maximum

$$\theta = \min \left\{ -\frac{\bar{x}_j}{\bar{r}_j} \mid j \in X \text{ et } \bar{r}_j < 0 \right\} \quad (\text{cas où } \bar{r}_x \not\geq 0),$$

et il est possible de profiter des dégénérescences qui accompagnent le « blocage » pour atteindre, en poursuivant le processus, un programme extrême donnant la même valeur à $\sum_{j=1}^n z_j$, mais tel que le tableau du Simplexe correspondant ne soit

plus « bloqué » (c'est sur ce point que la méthode ici présentée complète la méthode de Wolfe).

Ainsi arrivera-t-on toujours, en un nombre fini de pas, soit à résoudre le problème initial, soit à montrer que $f(\lambda, x)$ n'a pas de minimum fini dans ce problème.

2.4.3. APPLICATION A LA FORME LONGUE DU PROCESSUS

Le phénomène de « blocage » peut *a priori* se présenter, non dans la 2^e phase où $\lambda = 0$ (voir § 2.4.2), mais dans la 3^e phase du processus sous forme longue, et les résultats obtenus au paragraphe 2.4.1 peuvent être appliqués à ce cas avec les notations suivantes :

$$P = \begin{matrix} p \\ (n, 1) \end{matrix}, \quad y = \begin{matrix} \lambda \\ (1, 1) \end{matrix}, \quad a = \begin{matrix} 0 \\ (n, 1) \end{matrix}, \quad q' = \begin{matrix} -1 \\ (1, 1) \end{matrix}.$$

Ainsi, dans les conditions de « blocage » envisagées, il existe un vecteur \bar{r} tel (n, 1)

que

$$\begin{cases} C \bar{r} = 0, \\ (n, n) \\ A \bar{r} = 0, \\ (m, n) \end{cases} \quad p' \bar{r} \leq -1, \quad 0 = -\bar{\lambda}.$$

Donc, ou bien le « blocage » envisagé ne se produit pas, ou bien il se produit pour $\lambda = 0$.

1) *S'il ne se produit pas de « blocage »*, le processus de maximisation de λ , sous les conditions assurant le respect des relations d'exclusion $v_j x_j = 0, \forall j$, conduit à faire croître indéfiniment λ à partir de 0, en passant par une suite finie de valeurs $0 = \lambda_0 < \lambda_1 < \dots < \lambda_{k-1} < \lambda_k$ (éventuellement réduite à $0 = \lambda_0$), et donne la suite des solutions extrêmes correspondantes du problème initial (et de ses conditions de Kuhn et Tucker), soit $(x^0, u^0, v^0), (x^1, u^1, v^1), \dots, (x^k, u^k, v^k)$, complétée par $(\xi^k, \alpha^k, \beta^k)$ tel que, pour tout scalaire $t \geq 0$, $(x^k + t \xi^k, u^k + t \alpha^k, v^k + t \beta^k)$ soit solution du problème initial pour $\lambda = \lambda_k + t$.

En l'absence de dégénérescence, ces suites sont déterminées sans ambiguïté, et c'est ce que l'on suppose ici. Il en est certainement ainsi si $\begin{cases} Cz = 0 \\ Az = 0 \end{cases} \Rightarrow [z = 0]$, et en particulier si la forme quadratique $x' C x$ est définie positive. (S'il en était autrement, le processus permettrait encore de déterminer toutes les solutions extrêmes, nécessairement en nombre fini.)

La restriction qui interdit, à chaque pas du processus, de faire entrer dans la base une variable v_j (ou x_j) si la variable complémentaire x_j (ou v_j) s'y trouve déjà implique :

$$\begin{cases} 0 = v'^h x^h = v'^h x^{h+1} = v'^{h+1} x^h = v'^{h+1} x^{h+1}, & \forall h \in \{0, 1, 2, \dots, k-1\}, \\ 0 = v'^k x^k = v'^k \xi^k = \beta'^k x^k = \beta'^k \xi^k. \end{cases}$$

Les conditions de Kuhn et Tucker (linéaires si l'on excepte les relations d'exclusion) montrent alors que le problème initial, dépendant du paramètre $\lambda \geq 0$, admet comme solutions :

pour $\lambda \in [\lambda_h, \lambda_{h+1}]$,

$$(x, u, v) = \frac{\lambda_{h+1} - \lambda}{\lambda_{h+1} - \lambda_h} (x^h, u^h, v^h) + \frac{\lambda - \lambda_h}{\lambda_{h+1} - \lambda_h} (x^{h+1}, u^{h+1}, v^{h+1}),$$

pour $\lambda \geq \lambda_k$,

$$(x, u, v) = (x^k, u^k, v^k) + (\lambda - \lambda_k) (\xi^k, \alpha^k, \beta^k).$$

Ces solutions s'obtiennent par interpolation linéaire entre deux solutions extrêmes adjacentes, et par extrapolation linéaire à partir de (x^k, u^k, v^k) dans la direction d'infinitude $(\xi^k, \alpha^k, \beta^k)$.

Le problème paramétrique initial est ainsi résolu pour tout $\lambda \geq 0$, lorsqu'il ne se produit pas de « blocage ». Telle est, en particulier, la fin du processus si $\begin{cases} Cr=0 \\ Ar=0 \end{cases} \Rightarrow [r=0]$ (puisque'il est alors impossible que $p' \bar{r} \leq -1$), ce qui est le cas si la forme quadratique $x' C x$ est définie positive.

2) S'il se produit un « blocage » pour $\lambda = 0$ (ce qui n'est possible que si $x' C x$ est semi-définie positive), cette situation se présente après que λ est entré dans la base avec la valeur 0, de sorte que $\text{card } W = 1$. Alors $\bar{r} \neq 0$ et il existe en général un ensemble infini de programmes $(\bar{x} + t\bar{r}, \bar{u}, \bar{v})$ optimaux dans le problème initial pour $\lambda = 0$.

La fonction à minimiser dans le problème initial y prend les valeurs

$$f(\lambda, \bar{x} + t\bar{r}) = \lambda p'(\bar{x} + t\bar{r}) + \frac{1}{2} \bar{x}' C \bar{x} \leq f(\lambda, \bar{x}) - \lambda t,$$

pour $\lambda \geq 0$ et $t \geq 0$, car $C\bar{r} = 0$ et $p'\bar{r} \leq -1$.

Deux cas sont alors possibles, d'après l'étude faite au paragraphe 2.4.1 :

- ou bien le scalaire t peut croître indéfiniment à partir de 0 (cas où $\bar{r}_x \geq 0$), et, si $\lambda > 0$, $f(\lambda, \bar{x} + t\bar{r}) \rightarrow -\infty$ quand $t \rightarrow +\infty$, de sorte que la fonction convexe $f(\lambda, x)$ n'a pas de minimum fini dans le problème initial dès que $\lambda > 0$;
- ou bien le scalaire t peut croître de 0 à un maximum

$$\theta = \min \left\{ -\frac{\bar{x}_j}{\bar{r}_j} \mid j \in X \text{ et } \bar{r}_j < 0 \right\} \quad (\text{cas où } \bar{r}_x \not\geq 0),$$

et le passage de $(\bar{x}, \bar{u}, \bar{v})$ à $(\bar{x} + \theta\bar{r}, \bar{u}, \bar{v})$ correspond ici à un pas de la méthode du Simplexe, avec échange entre l'unique variable $x_j \mid j \in W$ (qui entre dans la base) et une variable $x_j \mid j \in X, \bar{r}_j < 0$ et $\bar{x}_j + \theta\bar{r}_j = 0$ (qui sort de la base). Le tableau du

Simplexe ainsi obtenu, adjacent au précédent, donne encore la valeur 0 au paramètre λ , mais il n'est plus « bloqué ».

Ainsi arrivera-t-on toujours, en un nombre fini de pas, soit à faire croître indéfiniment λ à partir de 0 et à résoudre le problème paramétrique initial pour tout $\lambda \geq 0$, soit à montrer que $f(\lambda, x)$ n'a pas de minimum fini dans ce problème dès que $\lambda > 0$.

2.5. Note sur le calcul de $f(\lambda, x)$ dans la méthode de Wolfe

D'après le calcul fait au paragraphe 1.3, et compte tenu des variables artificielles, coordonnées de w et de z , introduites dans les conditions de Kuhn et Tucker (cf. § 2.2), $(m, 1)$ $(n, 1)$

$$2f(\lambda, x) = \lambda p'x - u'b + v'x - Ez'x + u'w.$$

A la fin de la 1^{re} phase du processus, $w = 0$, et à la fin de la 2^e phase, $z = 0$.
De plus, la restriction de Wolfe impose constamment $v'x = 0$. $(m, 1)$ $(n, 1)$

Il est donc intéressant d'ajouter une ligne dans les tableaux du Simplexe pour calculer, au cours du processus, l'expression $\lambda p'x - u'b$, linéaire en x et u' , qui donnera, à la fin du processus, la valeur optimale de $2f(\lambda, x)$ (quand elle existe). $(n, 1)$ $(1, m)$

Dans la forme longue du processus, $(\lambda p'x - u'b)$ dépend linéairement du paramètre λ dans chaque tableau du Simplexe et, au cours de la 3^e phase, les calculs d'interpolation et d'extrapolation linéaires doivent être effectués à partir de ces expressions linéaires en λ , sans que λ soit remplacé par sa valeur dans chaque tableau. Ainsi, quand le minimum de $f(\lambda, x)$ existe pour tout $\lambda \geq 0$, c'est une fonction concave de λ (cf. § 1.1), quadratique par morceaux.

3. MÉTHODE DE DANTZIG

3.1. Principe ⁽⁴⁾

On considère le problème non paramétrique de programmation convexe à contraintes linéaires, qui coïncide avec celui du paragraphe 1 pour $\lambda = 1$:

$$\text{Minimiser } f(x) = p'x + \frac{1}{2}x' C x,$$

sous les contraintes :

$$Ax = b \quad \text{et} \quad x \geq 0,$$

⁽⁴⁾ Pour faciliter la lecture, on rappelle d'abord ici les résultats obtenus au paragraphe 1.3.

C étant une matrice symétrique réelle telle que $x' C x \geq 0, \forall x \in \mathbf{R}^n$.

La lagrangien (formé sans tenir compte des conditions de signe $x \geq 0$) :

$$l(x, u) = p' x + \frac{1}{2} x' C x + u' (A x - b)$$

est égal à la fonction économique $f(x)$ pour tout programme x réalisable, et les conditions de Kuhn et Tucker, nécessaires et suffisantes (cf. § 1.3), s'écrivent :

$$\begin{aligned} \nabla_x l(x, u) &= p + C x + A' u = v \quad (\geq 0), \\ &\quad \quad \quad (n, 1) \quad (n, n) \quad (n, m) \quad (n, 1) \\ &\quad - b + A x = 0, \\ &\quad \quad \quad (m, 1) \quad (m, n) \\ x &\geq 0, \quad v \geq 0, \quad v_j x_j = 0, \quad \forall j. \end{aligned}$$

Compte tenu des $(n+m)$ égalités précédentes, linéaires en x_j, u_i, v_j ,

$$2 f(x) = p' x - u' b + v' x.$$

S'il existe une solution des conditions de Kuhn et Tucker pour laquelle $x_j > 0$ (ou $v_j > 0$), alors dans toute solution de ces solutions $v_j = 0$ (ou $x_j = 0$) (cf. § 1.3).

L'ensemble des points $(\bar{x}, \bar{u}, \bar{v})$ solutions des conditions de Kuhn et Tucker est donc un tronçon de l'espace \mathbf{R}^{m+2n} , de sorte qu'il peut être défini à partir de ses points extrêmes, et éventuellement de ses directions d'infinitude extrêmes.

C'est ce qui conduit à chercher des solutions de base (au sens de la méthode du Simplexe) du système des conditions linéaires de Kuhn et Tucker ($m+n$ égalités et $2n$ conditions de signe portant sur $m+2n$ variables), en leur imposant de vérifier les n relations d'exclusion $v_j x_j = 0, \forall j$.

Dans la méthode de Dantzig, on y parvient en appliquant à ce système linéaire une méthode analogue à la méthode du Simplexe, de telle manière que les variables x_j soient constamment positives ou nulles tandis que les variables v_j ne sont toutes positives ou nulles que si l'optimum est atteint. Les programmes de base successifs vérifient soit toutes les relations $v_j x_j = 0$ (programmes dits « standard »), soit toutes ces relations sauf une (programmes dits « non standard »). De plus, les passages de programme de base à programme de base adjacent, selon l'algorithme de Gauss, sont organisés de telle manière que le lagrangien $l(x)$, constamment égal à la fonction économique $f(x)$ à minimiser, ne puisse pas croître. Et, pourvu que l'éventualité d'un cyclage soit écartée, le processus conduit, en un nombre fini de pas, soit aux solutions extrêmes du problème posé (et de ses conditions de Kuhn et Tucker), soit à montrer que $f(x)$ n'a pas de minimum fini dans ce problème.

possible compte tenu de la forme du système linéaire considéré. On peut par exemple échanger u_h et v_k aussitôt après avoir échangé w_h et x_k (et, éventuellement, v_k et v_l aussitôt après avoir échangé x_k et x_l). Il est facile de vérifier que, si l'on procède ainsi systématiquement, les deux pivots correspondant à ces deux échanges successifs sont opposés.

Pour simplifier l'écriture, on supposera que $\text{rang de } \underset{(m, n)}{A} = m$.

On dispose alors, à la fin de la première phase, d'un programme standard de départ dans lequel, les m variables artificielles w_i ayant été supprimées, la base est constituée par n variables $x_j \geq 0$ ou v_j de signes quelconques (une et une seule dans chaque couple (v_j, x_j)) et les m variables u_i de signes quelconques (qui resteront ensuite dans la base) ⁽⁵⁾.

2° phase. — On part du programme de base ainsi obtenu pour le système linéaire

$$\begin{cases} -Cx - A'u + v = p, \\ Ax = b, \\ x \geq 0, \end{cases}$$

programme dans lequel $v_j x_j = 0, \forall j$, puisqu'il s'agit d'un programme standard.

Si de plus $v \geq 0$, ce programme est optimal, et le processus se termine aussitôt. Si le problème donné (avec ses conditions de Kuhn et Tucker) admet d'autres solutions, le tableau de calcul présente une dégénérescence qui permet de déterminer toutes les solutions, compte tenu des propriétés établies au paragraphe 1.3.

Si au contraire $v \not\geq 0$, le processus se poursuit par des passages de programme de base à programme de base adjacent, selon l'algorithme de Gauss, comme dans la méthode du Simplexe. Ces passages sont organisés de la façon suivante :

- 1) A PARTIR D'UN PROGRAMME DE BASE STANDARD OU $\exists j \mid v_j < 0$
 - on fait entrer dans la base en croissant une des variables hors base x_j dont la variable complémentaire $v_j = \partial l(x, u) / \partial x_j$ (qui est dans la base) a une valeur négative (par exemple la valeur négative de plus grande valeur absolue), soit $x_h \mid v_h < 0$;
 - on fait sortir de la base celle des variables de base x_j et de la variable v_h , complémentaire de la variable entrante x_h , qui s'annule la première quand x_h croît à partir de 0 (on montrera au paragraphe 3.3 qu'alors v_h ne décroît pas). En cas d'ambiguïté entre v_h et une variable x_j , on choisit v_h comme variable sortante.

⁽⁵⁾ On pourrait supprimer, dans les tableaux de calcul ultérieurs, les lignes relatives aux m variables u_i .

Ainsi est-on assuré que, dans le passage au programme de base adjacent,

- toutes les variables x_j restent positives ou nulles,
- la fonction $f(x)$ à minimiser décroît (sauf dans le cas de dégénérescence où x_h resterait nulle en entrant dans la base), puisque $f(x)$ est constamment égale au lagrangien $l(x, u)$, que x_h croît et $v_h = \partial l(x, u) / \partial x_h$ reste négative ou nulle au cours de ce passage, tandis que les autres variables x_j restent nulles (hors de la base) ou sont associées à des $v_j = \partial l(x, u) / \partial x_j$ qui restent nulles, et qu'enfin $\nabla'_u l(x, u) = Ax - b$ reste nul.

Si la variable sortante est v_h , complémentaire de la variable entrante x_h (le pivot est alors négatif), le nouveau programme de base est standard comme le précédent. (On dira qu'il s'agit d'un passage P_{11} .)

Si la variable sortante est x_k (le pivot est alors positif), le nouveau programme de base est non standard : le couple (v_h, x_h) est dans la base, avec $v_h < 0$, tandis que le couple (v_k, x_k) est hors de la base. (On dira qu'il s'agit d'un passage P_{12} .)

2) A PARTIR D'UN PROGRAMME DE BASE NON STANDARD

- on fait entrer dans la base en croissant la variable v_k figurant dans le couple (v_k, x_k) de variables complémentaires hors base;
- on fait sortir de la base celle des variables de base x_j et de la variable $v_h (< 0)$, figurant dans le couple (v_h, x_h) de variables complémentaires de base, qui s'annule la première quand v_k croît à partir de 0 (on montrera au paragraphe 3.3 qu'alors v_h croît et x_h ne décroît pas). En cas d'ambiguïté entre v_h et une variable x_j , on choisit v_h comme variable sortante.

Ainsi est-on assuré que, dans le passage au programme de base adjacent,

- toutes les variables x_j restent positives ou nulles (sans pouvoir croître indéfiniment puisque v_h ne peut croître au-delà de 0);
- la fonction $f(x)$ à minimiser décroît ou reste constante, puisque $f(x)$ est constamment égale au lagrangien $l(x, u)$, que x_h ne décroît pas (comme on le montrera au paragraphe 3.3) et $v_h = \partial l(x, u) / \partial x_h$ reste négative ou nulle au cours de ce passage, tandis que les autres variables x_j restent nulles ou sont associées à des $v_j = \partial l(x, u) / \partial x_j$ qui restent nulles, et qu'enfin $\nabla'_u l(x, u) = Ax - b$ reste nul.

Si la variable sortante est v_h (le pivot est alors négatif), le nouveau programme de base est standard. (On dira qu'il s'agit d'un passage P_{21} .)

Si la variable sortante est x_l (le pivot est alors positif), $l \neq h$ puisque x_h ne décroît pas au cours du passage, et le nouveau programme de base est non standard comme le précédent : le couple (v_h, x_h) reste dans la base, avec $v_h < 0$, tandis que le couple (v_l, x_l) est hors de la base. (On dira qu'il s'agit d'un passage P_{22} .)

On montrera de plus au paragraphe 3.3 que, si x_h reste constante quand v_k croît à partir de 0, il en est de même pour toutes les variables x_j , de sorte que la variable sortante est v_h , et le nouveau programme de base est standard. Ce cas est le seul où la fonction $f(x)$ à minimiser reste constante (si l'on excepte les cas de dégénérescence), et il s'agit alors d'un passage P_{21} de programme non standard à programme standard.

3) CONVERGENCE DU PROCESSUS

Il résulte de l'étude précédente qu'en l'absence de dégénérescence, les passages de programme de base à programme de base adjacent, selon l'algorithme de Gauss, tels qu'ils sont organisés dans la 2^e phase de la méthode de Dantzig, donnent une suite de programmes réalisables du problème posé, la fonction $f(x)$ à minimiser décroissant strictement à chaque passage, sauf éventuellement dans certains passages P_{21} isolés où elle reste constante.

Il est donc impossible de passer plusieurs fois par un même programme de base, et cela reste vrai même dans les cas de dégénérescence si l'éventualité d'un cyclage est écartée par l'un des procédés connus pour la méthode du Simplexe (cette réserve est d'ailleurs d'une faible importance pratique).

Comme les programmes de base sont en nombre fini, le processus s'achèvera en un nombre fini de passages,

– soit par la détermination des solutions extrêmes du problème posé (et de ses conditions de Kuhn et Tucker), correspondant aux programmes de base standards où $x_j \geq 0$, $v_j \geq 0$, $v_j x_j = 0$, $\forall j$;

– soit par la mise en évidence de programmes réalisables du problème posé, dans lesquels $f(x)$ prend des valeurs arbitrairement petites, de sorte que $f(x)$ n'a pas de minimum fini dans ce problème.

Dans ce second cas, le processus conduit à un programme standard non optimal, à partir duquel il est possible de faire croître indéfiniment une variable hors base x_h , associée à une variable de base v_h qui reste constante et négative, tandis que les autres variables x_j restent positives ou nulles (on ne trouve alors aucun pivot).

Dans le cas où $x' C x > 0$, $\forall x \neq 0$, le problème posé admet une solution unique (cf. § 1.1 et § 3.3).

3.3. Justification théorique de la méthode de Dantzig

Soient δx , δu , δv les vecteurs représentant les variations de x , u , v entre deux programmes qui vérifient les $(n+m)$ égalités linéaires figurant dans les conditions de Kuhn et Tucker du problème posé (cf. § 3.1).

Alors

$$\begin{cases} C \delta x + A' \delta u = \delta v, \\ A \delta x = 0, \end{cases}$$

d'où

$$\delta v' \cdot \delta x = \delta x' \cdot C \cdot \delta x \geq 0,$$

puisque la forme quadratique $x' C x$ est semi-définie positive.

1) Dans chacun des passages de programme de base à programme de base adjacent, tels qu'ils sont organisés dans la 2^e phase de la méthode de Dantzig,

$$\delta v' \cdot \delta x = \delta v_h \cdot \delta x_h \quad (\text{avec les notations du paragraphe 3. 2}),$$

car un facteur est nul dans chacun des autres termes $\delta v_j \cdot \delta x_j (j \neq h)$ du produit scalaire $\delta v' \cdot \delta x$.

Ainsi $\delta v_h \cdot \delta x_h \geq 0$, ce qui montre que, si x_h croît, v_h ne décroît pas.

2) Considérons un passage P_{12} de programme standard à programme non standard, au cours duquel x_h entre dans la base et x_k sort de la base, et le passage (P_{21} ou P_{22}) qui suit P_{12} et au cours duquel v_k entre dans la base en croissant.

Compte tenu du fait que, pour toute base, les valeurs des variables de base sont déterminées sans ambiguïté par celles des variables hors base, si les variables qui se trouvent hors de la base au début du passage P_{12} restent nulles à l'exception de x_h et v_k qui subissent des variations arbitraires δx_h et δv_k , les variables v_h et x_k subissent des variations δv_h et δx_k déterminées en fonction linéaire de δx_h et δv_k :

$$\begin{cases} \delta v_h = \theta \delta v_k + \lambda \delta x_h, \\ \delta x_k = \mu \delta v_k + \nu \delta x_h, \end{cases}$$

$\theta, \lambda, \mu, \nu$ étant des coefficients constants, avec $\nu < 0$ puisque, dans P_{12} , $\delta v_k = 0, \delta x_h > 0, \delta x_k < 0$.

De plus, $\delta v' \cdot \delta x = \delta v_h \cdot \delta x_h + \delta v_k \cdot \delta x_k = \lambda (\delta x_h)^2 + (\theta + \nu) \delta x_h \cdot \delta v_k + \mu (\delta v_k)^2 \geq 0$ quels que soient δx_h et δv_k , ce qui implique $\lambda \geq 0, \mu \geq 0, 4\lambda\mu \geq (\theta + \nu)^2$.

Ainsi, au cours du passage P_{12} , où $\delta v_k = 0$ et $\delta x_h > 0$,

$$\delta v_h = \lambda \delta x_h \geq 0 \quad \text{et} \quad \delta x_k = \nu \delta x_h < 0,$$

et, au cours du passage qui suit P_{12} , où $\delta v_k > 0$ et $\delta x_h = 0$,

$$\delta v_h = \left(\theta - \frac{\lambda\mu}{\nu} \right) \delta v_k \quad \text{et} \quad \delta x_k = -\frac{\mu}{\nu} \delta v_k \geq 0.$$

De plus, d'après les signes de λ , μ , v , et la condition $4\lambda\mu \geq (\theta+v)^2$,

$$\theta - \frac{\lambda\mu}{v} \geq \theta - \frac{(\theta+v)^2}{4v} = -\frac{(\theta-v)^2}{4v} \geq 0.$$

Si $(\theta - (\lambda\mu/v))$ était nul, il en résulterait

$$\theta = v < 0 \quad \text{et} \quad \lambda\mu = v^2 > 0,$$

d'où, au cours du passage qui suit P_{12} , $\delta v_h = 0$, $\delta x_h > 0$, et $\delta v' \cdot \delta x = 0$, ce qui implique $\delta x' \cdot C \cdot \delta x = 0$ et $C \delta x = 0$ (cf. § 1.1), de sorte que $\begin{cases} A' \delta u = \delta v, \\ A \delta x = 0. \end{cases}$

En particulier, les relations $C \delta x = 0$ et $A \delta x = 0$, jointes à la nullité de δx_j au cours de ce passage pour toute variable x_j hors base, déterminent sans ambiguïté $\delta x = 0$, ce qui est contradictoire avec $\delta x_h > 0$.

Donc $\theta - (\lambda\mu/v) > 0$ et $\delta v_h > 0$ au cours du passage qui suit P_{12} .

Il est ainsi établi que (sauf en cas de dégénérescence) :

$$\begin{array}{ll} \delta x_h > 0 & \text{et} \quad \delta v_h \geq 0 \quad \text{au cours du passage } P_{12}, \\ \delta x_h \geq 0 & \text{et} \quad \delta v_h > 0 \quad \text{au cours du passage qui suit } P_{12}. \end{array}$$

Si $\delta v_h = 0$ au cours du passage P_{12} , c'est-à-dire si $\lambda = 0$, $\delta v' \cdot \delta x = 0$, ce qui implique $\delta x' \cdot C \cdot \delta x = 0$ et $C \delta x = 0$, de sorte que $\begin{cases} A' \delta u = \delta v \\ A \delta x = 0 \end{cases}$ au cours de P_{12} .

En particulier, la relation $A' \delta u = \delta v$, jointe à la nullité de δv_j au cours de P_{12} pour toute variable v_j hors base, détermine sans ambiguïté $\delta u = 0$ et $\delta v = 0$. Ainsi, si $\delta v_h = 0$, tous les u_i et tous les v_j restent constants au cours du passage P_{12} .

Ce cas est exclu si $\begin{cases} C \delta x = 0 \\ A \delta x = 0 \end{cases} \Rightarrow [\delta x = 0]$ (ce qui est contradictoire avec

$\delta x_h > 0$), et en particulier si la forme quadratique $x' C x$ est définie positive. Alors $\delta v_h > 0$ au cours de P_{12} , ce qui exclut que la variable x_h puisse croître indéfiniment sans que v_h devienne positif.

Si $\delta x_h = 0$ au cours du passage qui suit P_{12} , c'est-à-dire si $\mu = 0$, $\delta v' \cdot \delta x = 0$, ce qui implique $\delta x' \cdot C \cdot \delta x = 0$ et $C \delta x = 0$, de sorte que $\begin{cases} A' \delta u = \delta v \\ A \delta x = 0 \end{cases}$ au cours de ce passage. En particulier, les relations $C \delta x = 0$ et $A \delta x = 0$, jointes à la nullité de δx_j

au cours de ce passage pour toute variable x_j hors base, déterminent sans ambiguïté $\delta x = 0$. Ainsi, si $\delta x_h = 0$, tous les x_j restent constants au cours du passage qui suit P_{12} , de sorte qu'il s'agit d'un passage P_{21} conduisant à un programme standard : v_h croît à partir d'une valeur négative et sort de la base en s'annulant.

3) Si un passage P_{12} , au cours duquel x_h entre dans la base et x_k sort de la base, est suivi d'un passage P_{22} , au cours duquel v_k entre dans la base et x_l sort de la base, $\delta x_h > 0$, $\delta v_h > 0$ et $\delta x_l < 0$ au cours de P_{22} tandis que $\delta v_k > 0$, et, au cours du passage (P_{21} ou P_{22}) qui suit P_{22} , v_l entre dans la base en croissant.

Si les variables qui se trouvent hors de la base au début du passage P_{22} restent nulles à l'exception de v_k et v_l qui subissent des variations arbitraires δv_k et δv_l , les variables x_h , v_h et x_l subissent des variations δx_h , δv_h et δx_l déterminées en fonction linéaire de δv_k et δv_l . Mais, puisque $\delta x_h > 0$ quand $\delta v_k > 0$ et $\delta v_l = 0$ (au cours de P_{22}), on peut choisir comme variables indépendantes δv_k et δv_l au lieu de δv_k et δv_l , et exprimer δv_h et δx_l en fonction linéaire de δx_h et δv_l :

$$\begin{cases} \delta v_h = \theta' \delta v_l + \lambda' \delta x_h, \\ \delta x_l = \mu' \delta v_l + \nu' \delta x_h, \end{cases}$$

θ' , λ' , μ' , ν' étant des coefficients constants, avec $\begin{cases} \lambda' > 0 \\ \nu' < 0 \end{cases}$ puisque, dans P_{22} , $\delta v_l = 0$, $\delta x_h > 0$, $\begin{cases} \delta v_h > 0, \\ \delta x_l < 0. \end{cases}$

De plus, puisque $\delta x_k = 0$, $\delta v' \cdot \delta x = \delta v_h \cdot \delta x_h + \delta v_l \cdot \delta x_l \geq 0$, $\forall (\delta x_h, \delta v_l)$.

On peut donc reprendre ici, pour P_{22} et le passage qui le suit, l'étude faite en 2) pour P_{12} et le passage qui le suit, en remplaçant θ , λ , μ , ν par θ' , λ' , μ' , ν' et x_k , v_k par x_l , v_l .

Il est ainsi établi que (sauf en cas de dégénérescence) :

$$\begin{cases} \delta x_h > 0 & \text{et} & \delta v_h > 0 & \text{au cours du passage } P_{22}, \\ \delta x_h \geq 0 & \text{et} & \delta v_h > 0 & \text{au cours du passage qui suit } P_{22}, \end{cases}$$

et que δx_h ne peut être nul au cours de ce dernier passage que si c'est un passage P_{21} .

Par récurrence, ces résultats restent valables pour toute suite de passages P_{22} succédant à P_{12} .

3.4. Note sur le calcul de $f(x)$ dans la méthode de Dantzig

D'après le calcul fait au paragraphe 1.3 et compte tenu des variables artificielles, coordonnées de w , éventuellement introduites dans les conditions de Kuhn et Tucker (cf. § 3.2),

$$2f(x) = p'x - u'b + v'x + u'w.$$

A la fin de la 1^{re} phase du processus, $w = 0$, et dans tout programme de base standard, en particulier dans tout programme extrême optimum, $v'x = 0$.

Il est donc intéressant d'ajouter une ligne dans les tableaux de calcul, pour calculer, au cours du processus, l'expression $p'x - u'b$, linéaire en x et

u' , qui donnera, à la fin du processus, la valeur optimale de $2f(x)$ (quand elle existe).

Cette ligne supplémentaire permet de suivre l'évolution de la fonction économique au cours de la 2^e phase du processus, $2f(x)$ étant égal à $(p'x - u'b)$ dans chaque programme standard, et à $(p'x - u'b)$ augmenté du seul terme $v_j x_j$ non nul dans chaque programme non standard.

4. EXEMPLE

4.1. Problème

λ étant un paramètre réel positif ou nul.

$$\text{Minimiser } f(\lambda, x) = -\lambda(5x_1 + 2x_2) + \frac{1}{18}(x_1 - 2x_2)^2,$$

sous les contraintes :

$$\left\{ \begin{array}{l} x_1 + x_2 \leq 12 \\ 4x_1 + x_2 \leq 24 \\ x_1, x_2 \geq 0 \end{array} \right\}, \quad \text{ou} \quad \left\{ \begin{array}{l} x_1 + x_2 + x_3 = 12 \\ 4x_1 + x_2 + x_4 = 24 \\ x_1, x_2, x_3, x_4 \geq 0 \end{array} \right\}$$

(après introduction
des variables d'écart x_3, x_4).

Le lagrangien (formé sans tenir compte des conditions de signe $x_j \geq 0$) s'écrit :

$$l(x, u) = -\lambda(5x_1 + 2x_2) + \frac{1}{18}(x_1 - 2x_2)^2 \\ + u_3(x_1 + x_2 + x_3 - 12) + u_4(4x_1 + x_2 + x_4 - 24),$$

et l'on en déduit les conditions de Kuhn et Tucker (nécessaires et suffisantes) :

$$\left\{ \begin{array}{l} -5\lambda + \frac{x_1}{9} - \frac{2x_2}{9} + u_3 + 4u_4 = v_1, \\ -2\lambda - \frac{2x_1}{9} + \frac{4x_2}{9} + u_3 + u_4 = v_2, \\ \qquad \qquad \qquad \qquad \qquad \qquad u_3 = v_3, \\ \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad u_4 = v_4, \\ \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad x_1 + x_2 + x_3 = 12, \\ \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad 4x_1 + x_2 + x_4 = 24, \\ v_j \geq 0, \quad x_j \geq 0, \quad v_j x_j = 0, \quad \forall j. \end{array} \right.$$

ou

$$\boxed{\begin{array}{l} \frac{x_1}{9} - \frac{2x_2}{9} + v_3 + 4v_4 - v_1 = 5\lambda, \\ -\frac{2x_1}{9} + \frac{4x_2}{9} + v_3 + v_4 - v_2 = 2\lambda, \\ x_1 + x_2 + x_3 = 12, \\ 4x_1 + x_2 + x_4 = 24, \\ v_j \geq 0, \quad x_j \geq 0, \quad v_j x_j = 0, \quad \forall j \in \{1, 2, 3, 4\}, \end{array}}$$

après élimination (immédiate) des deux multiplicateurs u_3 et u_4 , de signes quelconques (*a priori*).

Compte tenu des quatre égalités linéaires précédentes (cf. § 1.3),

$$\boxed{2f(\lambda, x) = -\lambda(5x_1 + 2x_2) - 12v_3 - 24v_4 + \sum_{j=1}^4 v_j x_j,}$$

le dernier terme (non linéaire) étant nul quand les relations d'exclusion, $v_j x_j = 0, \forall j$, sont vérifiées.

4.2. Application de la méthode de Wolfe complétée

Un programme de base réalisable dans le problème initial est en évidence :

$$x_1 = 0, \quad x_2 = 0, \quad x_3 = 12, \quad x_4 = 24,$$

ce qui dispense de la 1^{re} phase du processus.

4.2.1. FORME COURTE DU PROCESSUS POUR $\lambda = 1$

On ajoute les variables artificielles respectives z_1 et z_2 aux premiers membres de la première et de la seconde équation de Kuhn et Tucker, et les termes $-z_1 x_1 - z_2 x_2$ à la dernière expression de $2f(\lambda, x)$ (cf. § 2.5).

Les tableaux de calcul sont les suivants :

	x_1	x_2	v_3	v_4	v_1	v_2	
$z_1 + z_2$	$-\frac{1}{9}$	$\frac{2}{9}$	2	5	-1	-1	7
z_1	$\frac{1}{9}$	$-\frac{2}{9}$	1	4	-1	0	5
z_2	$-\frac{2}{9}$	$\frac{4}{9}$	1	1	0	-1	2
x_3	1	1	0	0	0	0	12
x_4	4	1	0	0	0	0	24
$2f$	5	2	12	24	0	0	0

	x_1	z_2	v_3	v_4	v_1	v_2	
$z_1 + z_2$	0	$-\frac{1}{2}$	$\frac{3}{2}$	$\frac{9}{2}$	-1	$-\frac{1}{2}$	6
z_1	0	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{9}{2}$	-1	$-\frac{1}{2}$	6
x_2	$-\frac{1}{2}$	$\frac{9}{4}$	$\frac{9}{4}$	$\frac{9}{4}$	0	$-\frac{9}{4}$	$\frac{9}{2}$
x_3	$\frac{3}{2}$	$-\frac{9}{4}$	$-\frac{9}{4}$	$-\frac{9}{4}$	0	$\frac{9}{4}$	$\frac{15}{2}$
x_4	$\frac{9}{2}$	$-\frac{9}{4}$	$-\frac{9}{4}$	$-\frac{9}{4}$	0	$\frac{9}{4}$	$\frac{39}{2}$
$2f$	6	$-\frac{9}{2}$	$\frac{15}{2}$	$\frac{39}{2}$	0	$\frac{9}{2}$	-9

	x_4	z_2	v_3	v_4	v_1	v_2	
$z_1 + z_2$	0	$-\frac{1}{2}$	$\frac{3}{2}$	$\frac{9}{2}$	-1	$-\frac{1}{2}$	6
z_1	0	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{9}{2}$	-1	$-\frac{1}{2}$	6
x_2	$\frac{1}{9}$	2	2	2	0	-2	$\frac{20}{3}$
x_3	$-\frac{1}{3}$	$-\frac{3}{2}$	$-\frac{3}{2}$	$-\frac{3}{2}$	0	$\frac{3}{2}$	1
x_1	$\frac{2}{9}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{13}{3}$
$2f$	$-\frac{4}{3}$	$-\frac{3}{2}$	$\frac{21}{2}$	$\frac{45}{2}$	0	$\frac{3}{2}$	$-\frac{35}{2}$ -26

	x_4	z_2	v_3	z_1	v_1	v_2	
$z_1 + z_2$	0	-1	0	-1	0	0	0
v_4	0	$\frac{1}{9}$	1	$\frac{2}{9}$	$-\frac{2}{9}$	$-\frac{1}{9}$	$\frac{4}{3}$
x_2	$\frac{1}{9}$	$\frac{16}{9}$	4	$\frac{4}{9}$	$\frac{4}{9}$	$-\frac{16}{9}$	4
x_3	$-\frac{1}{3}$	$-\frac{4}{3}$	-1	$\frac{1}{3}$	$-\frac{1}{3}$	$\frac{4}{3}$	3
x_1	$\frac{2}{9}$	$-\frac{4}{9}$	-1	$\frac{1}{9}$	$-\frac{1}{9}$	$\frac{4}{9}$	5
$2f$	$-\frac{4}{3}$	$-\frac{3}{2}$	3	-5	5	4	-65

La solution unique du problème initial est atteinte au 4^e tableau (pour $\lambda = 1$) :

$$\bar{x}_1 = 5, \quad \bar{x}_2 = 4, \quad \text{et} \quad f(1, \bar{x}) = -65/2.$$

Remarques. — L'élément encadré dans chaque tableau autre que le dernier est le pivot.

Dès qu'une variable artificielle z_1 ou z_2 sort de la base, elle pourrait être supprimée.

Quand $(-z_1 x_1 - z_2 x_2) \neq 0$ (3^e tableau), sa valeur est inscrite sous la valeur de la partie linéaire de $2f$ calculée par l'algorithme du Simplexe.

Le second tableau de calcul conduit à un blocage apparent, puisque x_1 ou v_1 peuvent seuls entrer dans la base et que leur croissance à partir de 0 ne peut faire décroître $(z_1 + z_2)$. Mais ce blocage n'est qu'apparent, car l'entrée de x_1 dans la base permet de passer à un tableau non bloqué (cf. § 2.4.2).

Si l'on conservait les seules contraintes de signe dans le problème initial, les deux premiers tableaux de calcul resteraient valables si l'on y supprimait les lignes x_3, x_4 et les colonnes v_3, v_4 . Le second tableau permettrait alors une croissance indéfinie de x_1 , au cours de laquelle $f(1, x)$ décroîtrait indéfiniment, avec $2x_2 - x_1 = 9$.

4.2.2. FORME LONGUE DU PROCESSUS

Le programme de base réalisable qui est en évidence dans le problème initial en est une solution extrême pour $\lambda = 0$, ce qui dispense de la 1^{re} et de la 2^e phases du processus. Il ne reste qu'à essayer de faire croître λ à partir de 0.

Les tableaux de calcul se trouvent à la page 181.

Le problème paramétrique initial est résolu par les 4^e et 5^e tableaux (cf. § 2.4.3) :

— pour $\lambda = 0$,

$$\bar{x}_1 = \frac{16}{3}\alpha, \quad \bar{x}_2 = \frac{8}{3}\alpha, \quad \forall \alpha \in [0, 1], \quad \text{et} \quad f(0, \bar{x}) = 0;$$

— pour $\lambda \in [0, 4]$,

$$\left. \begin{array}{l} \bar{x}_1 = \frac{16}{3} \frac{4-\lambda}{4} + 4 \frac{\lambda}{4} = \frac{16-\lambda}{3}, \quad \bar{x}_2 = \frac{8}{3} \frac{4-\lambda}{4} + 8 \frac{\lambda}{4} = \frac{8+4\lambda}{3}, \\ \text{et} \\ f(\lambda, \bar{x}) = -16\lambda \frac{4-\lambda}{4} - (64+18\lambda) \frac{\lambda}{4} = -32\lambda - \frac{\lambda^2}{2}; \end{array} \right\}$$

— pour $\lambda \geq 4$,

$$\bar{x}_1 = 4, \quad \bar{x}_2 = 8, \quad \text{et} \quad f(\lambda, \bar{x}) = -64 - 18\lambda - 18(\lambda - 4) = 8 - 36\lambda.$$

(On pourrait aussi calculer $\bar{x}_3, \bar{x}_4, \bar{v}_1, \bar{v}_2, \bar{v}_3, \bar{v}_4$ en fonction de $\lambda \geq 0$.)

Remarques. — Les trois premiers tableaux correspondent au programme de départ, qui est dégénéré, avec $\lambda = 0$, mais il n'y a pas de cyclage. (On aurait pu commencer par échanger λ et v_2 , au lieu de λ et v_1 .)

	x_1	x_2	v_3	v_4	λ		x_1	x_2	v_3	v_4	v_1		v_2	x_2	v_3	v_4	v_1		
v_1	$\frac{1}{9}$	$\frac{2}{9}$	-1	-4	(5)	0	λ	$-\frac{1}{45}$	$\frac{2}{45}$	$-\frac{1}{5}$	$-\frac{4}{5}$	0	λ	$\frac{1}{12}$	0	$-\frac{1}{4}$	$-\frac{3}{4}$	$\frac{1}{6}$	
v_2	$\frac{2}{9}$	$-\frac{4}{9}$	-1	-1	2	0	v_2	(4)	$-\frac{8}{15}$	$-\frac{3}{5}$	$\frac{3}{5}$	$-\frac{2}{5}$	0	x_1	$\frac{15}{4}$	-2	$-\frac{9}{4}$	$\frac{9}{4}$	
x_3	1	1	0	0	0	12	x_3	1	1	0	0	12	x_3	$-\frac{15}{4}$	3	$\frac{9}{4}$	$-\frac{9}{4}$	$\frac{3}{2}$	
x_4	4	1	0	0	0	24	x_4	4	1	0	0	24	x_4	-15	(9)	9	-9	6	
$2f$	5λ	2λ	12	24	0	0	$2f$	5λ	2λ	12	24	0	0	$-\frac{75\lambda}{4}$	12λ	$(12 + \frac{45\lambda}{4})$	$(24 - \frac{45\lambda}{4})$	$\frac{15\lambda}{2}$	0



	v_2	x_4	v_3	v_4	v_1		v_2	x_4	v_3	x_3	v_1	
λ	$\frac{1}{12}$	0	$-\frac{1}{4}$	$-\frac{3}{4}$	$\frac{1}{6}$	0	λ	$\frac{4}{3}$	$-\frac{1}{3}$	-1	$-\frac{1}{3}$	4
x_1	$\frac{5}{12}$	$\frac{2}{9}$	$-\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{16}{3}$	x_1	0	$\frac{1}{3}$	$-\frac{1}{3}$	0	4
x_3	$\frac{5}{4}$	$-\frac{1}{3}$	$-\frac{1}{4}$	(3)	$-\frac{1}{2}$	4	v_4	$\frac{5}{3}$	$-\frac{1}{9}$	-1	$-\frac{2}{3}$	$\frac{16}{3}$
x_2	$-\frac{5}{3}$	$\frac{1}{9}$	1	-1	$\frac{2}{3}$	$\frac{8}{3}$	x_2	0	$-\frac{1}{3}$	$\frac{4}{3}$	0	8
$2f$	$\frac{5\lambda}{4}$	$-\frac{4\lambda}{3}$	$(12 - \frac{3\lambda}{4})$	$(24 + \frac{3\lambda}{4})$	$-\frac{\lambda}{2}$	-32λ	$2f$	$-\frac{40}{3}$	$(\frac{32}{3} - \lambda)$	$36(-32 - \lambda)$	16	$-128 - 36\lambda$

Le troisième tableau de calcul conduit à un blocage apparent, puisque v_2 ou x_2 peuvent seuls entrer dans la base et que leur croissance à partir de 0 ne peut faire croître λ . Mais ce blocage n'est qu'apparent, car l'entrée de x_2 dans la base permet de passer à un tableau non bloqué (cf. § 2.4.3).

Dans le dernier tableau, la variable hors base v_3 , qui peut croître indéfiniment à partir de 0, est marquée d'une flèche, et la dernière colonne donne les accroissements qui en résultent pour les variables de base et pour $2f$.

Si l'on conservait les seules contraintes de signe dans le problème paramétrique initial, les trois premiers tableaux de calcul resteraient valables si l'on y supprimait les lignes x_3 , x_4 et les colonnes v_3 , v_4 . Le troisième tableau permettrait alors une croissance indéfinie de x_2 , λ restant nul, avec $x_1 - 2x_2 = 0$, et, pour tout $\lambda > 0$, $f(\lambda, x)$ décroîtrait alors indéfiniment.

4.3. Application de la méthode de Dantzig pour $\lambda = 1$

Le fait qu'un programme de base réalisable soit en évidence dans le problème posé dispense de la 1^{re} phase du processus.

Les tableaux de calcul se trouvent à la page 183.

La solution unique du problème posé est atteinte au 4^e tableau :

$$\bar{x}_1 = 5, \quad \bar{x}_2 = 4, \quad \text{et} \quad f(1, \bar{x}) = -65/2.$$

Remarques. — Quand $\sum_{j=1}^4 v_j x_j \neq 0$ (3^e tableau), la valeur du seul terme $v_j x_j$ non nul dans le programme non standard correspondant est inscrite sous la valeur de la partie linéaire de $2f$ calculée par l'algorithme du Simplexe.

Si l'on conservait les seules contraintes de signe dans le problème posé, les deux premiers tableaux de calcul resteraient valables si l'on y supprimait les lignes x_3 , x_4 et les colonnes v_3 , v_4 . Le second tableau permettrait alors une croissance indéfinie de x_1 , au cours de laquelle $f(1, x)$ décroîtrait indéfiniment, la variable v_1 associée à x_1 restant constante et négative, avec $2x_2 - x_1 = 9$.

N.B. : La règle pratique généralement adoptée aurait conduit à commencer par faire entrer dans la base la variable hors base x_1 associée à v_1 , qui est la variable v_j négative de plus grande valeur absolue. Mais cette règle ne s'impose pas: on peut vérifier qu'ici, son application n'abrégierait pas (et n'allongerait pas) le calcul.

En commençant par faire entrer x_2 dans la base, on obtient ici des tableaux de calcul immédiatement comparables à ceux obtenus dans l'application de la forme courte de la méthode de Wolfe (cf. § 4.2). Cette comparaison montre que la

	x_1	x_2	v_3	v_4			x_1	v_2	v_3	v_4	
v_1	$-\frac{1}{9}$	$\frac{2}{9}$	-1	-4	-5	v_1	0	$\frac{1}{2}$	$-\frac{3}{2}$	$-\frac{9}{2}$	-6
v_2	$\frac{2}{9}$	$\frac{4}{9}$	-1	-1	-2	x_2	$-\frac{1}{2}$	$-\frac{9}{4}$	$\frac{9}{4}$	$\frac{9}{4}$	$\frac{9}{2}$
x_3	1	1	0	0	12	x_3	$\frac{3}{2}$	$\frac{9}{4}$	$-\frac{9}{4}$	$-\frac{9}{4}$	$\frac{15}{2}$
x_4	4	1	0	0	24	x_4	$\frac{9}{2}$	$\frac{9}{4}$	$-\frac{9}{4}$	$-\frac{9}{4}$	$\frac{39}{2}$
$2f$	5	2	12	24	0	$2f$	6	$\frac{9}{2}$	$\frac{15}{2}$	$\frac{39}{2}$	-9

	x_4	v_2	v_3	v_4			x_4	v_2	v_3	v_1	
v_1	0	$\frac{1}{2}$	$-\frac{3}{2}$	$\frac{9}{2}$	-6	v_4	0	$-\frac{1}{9}$	$\frac{1}{3}$	$-\frac{2}{9}$	$\frac{4}{3}$
x_2	$\frac{1}{9}$	-2	2	2	$\frac{20}{3}$	x_2	$\frac{1}{9}$	$-\frac{16}{9}$	$\frac{4}{3}$	$\frac{4}{9}$	4
x_3	$-\frac{1}{3}$	$\frac{3}{2}$	$-\frac{3}{2}$	$-\frac{3}{2}$	1	x_3	$-\frac{1}{3}$	$\frac{4}{3}$	-1	$-\frac{1}{3}$	3
x_1	$\frac{2}{9}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$\frac{13}{3}$	x_1	$\frac{2}{9}$	$\frac{4}{9}$	$-\frac{1}{3}$	$-\frac{1}{9}$	5
$2f$	$-\frac{4}{3}$	$\frac{3}{2}$	$\frac{21}{2}$	$\frac{45}{2}$	-35	$2f$	$-\frac{4}{3}$	4	3	5	-65
					-26						

méthode de Dantzig conduit à des calculs plus réduits, en évitant d'introduire des variables artificielles.

(En revanche, la forme longue de la méthode de Wolfe permet de résoudre le problème paramétrique.)

BIBLIOGRAPHIE

1. E. W. BARANKIN et R. DORFMAN, *On Quadratic Programming*, University of California Publications in Statistics, vol. 2, 1958, p. 285-318.
2. J. C. G. BOOT, *Programmation quadratique : algorithmes, anomalies, applications*, North-Holland Publishing Company, 1964 et Dunod, 1968.
3. G. B. DANTZIG, *Quadratic Programming : a Variant of the Wolfe-Markowitz Algorithms*, Research Report 2 of the Operations Research Center of the University of California, Berkeley, 1961.

4. G. B. DANTZIG, *Applications et prolongements de la programmation linéaire*, Princeton University Press, 1963 et Dunod, 1966.
5. W. S. DORN, *Nonlinear Programming: a Survey*, Management Science, vol. 9, 1963, p. 171-208.
6. E. L. KELLER, *The General Quadratic Optimization Problem*, Mathematical Programming, vol. 5, 1973, p. 311-337.
7. H. P. KÜNZI, W. KRELLE et W. OETTLI, *La programmation non linéaire*, Blaisdell Publishing Company, 1966 et Gauthier-Villars, 1969.
8. C. VAN DE PANNE et A. WHINSTON, *The Simplex and the Dual Method for Quadratic Programming*, Report 6314 of the Econometric Institute of the Netherlands School of Economics, Rotterdam, 1963.
9. Ph. WOLFE, *The Simplex Method for Quadratic Programming*, *Econometrica*, vol. 27, 1959, p. 382-398.