

THESE DE DOCTORAT D'ETAT ES SCIENCES MATHÉMATIQUES

présentée

A LA FACULTE DES SCIENCES

de PARIS

par

M. Jean SALENÇON

pour obtenir

le grade de Docteur Es-Sciences

2ème Thèse : Propositions données par la Faculté

Programmation linéaire et non linéaire.

Soutenue le 22 septembre 1969 devant la Commission d'Examen

MM. GERMAIN	Président
SIESTRUNCK	Examineurs
TEMAM	
MANDEL	

RESULTATS FONDAMENTAUX EN PROGRAMMATION LINEAIRE

Les problèmes de programmation linéaire ont vu leur importance s'affirmer avec le développement des calculs économétriques. La méthode fondamentale pour trouver l'optimum de la fonction économique (ou fonction objective) d'un programme linéaire, est la méthode du simplexe ou de Dantzig. C'est en 1948 que George Dantzig a présenté son célèbre mémoire ⁽¹⁾ et depuis lors de nombreuses variantes ont été publiées permettant de résoudre plus facilement des problèmes ayant des structures particulières.

G. Th. Guilbaud, Kuhn, Tucker, Charnes, Orden, Cooper, Henderson sont parmi ceux qui ont puissamment contribué au développement de la théorie et de ses applications.

1. - ENONCE MATHEMATIQUE GENERAL DES PROGRAMMES LINEAIRES.

La programmation linéaire est l'étude et la résolution d'un problème du type suivant : optimiser une forme linéaire lorsque les variables sont soumises à un certain nombre de contraintes linéaires.

Par exemple le problème de maximisation :
maximiser la forme linéaire (appelée aussi fonction objective),

$$B = \sum_{i=1}^n a_{0i} x_i \quad i = 1, 2, \dots, n$$

sous les contraintes

$$\sum_{i=1}^n a_{ji} x_i \leq a_{j0} \quad j = 1, 2, \dots, m \quad (1)$$

et les relations de non-négativité :

$$x_i \geq 0 \quad i = 1, 2, \dots, n$$

a_{0i} , a_{j0} , a_{ji} sont connues et les x_i sont les inconnues.

On peut donner une autre formulation du problème :
 En introduisant des variables supplémentaires dites variables d'écart (Slack variables), les contraintes peuvent être écrites sous la forme

$$\left. \begin{array}{l} \sum_{i=1}^n a_{ji} x_i + x_{n+j} = a_{j0} \\ x_{n+j} \geq 0 \end{array} \right\} \quad j = 1, 2, \dots, m$$

et avec la représentation matricielle (1) devient :

$$\left. \begin{array}{l} B = a_{0\cdot}^T x \\ \text{contraintes :} \\ Ax = a_{\cdot 0} \\ x \geq 0 \end{array} \right\} \quad (2)$$

où

$$\begin{aligned} x^T &= (x_1, x_2, \dots, x_n, x_{n+1}, \dots, x_{n+m}) \\ a_{0\cdot}^T &= (a_{01}, a_{02}, a_{0n}, a_{0,n+1}, \dots, a_{0,n+m}) \\ \text{avec } a_{0,n+1} &= \dots = a_{0,n+m} = 0 \end{aligned}$$

$$A, \text{matrice } m \times (m+n), = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & 1 & \dots & 0 \\ & & & & & 0 & 1 & 0 \\ a_{m1} & a_{m2} & \dots & a_{mn} & 0 & \dots & 1 \end{pmatrix}$$

Tout vecteur x dont les composantes satisfont les contraintes et les conditions de non négativité est appelé une solution admissible (feasible vector). Si ce vecteur réalise l'optimum, c'est une solution admissible optimale.

Une solution admissible est dite "de base", si elle comporte n composantes nulles et si les vecteurs colonnes de A correspondant aux composantes non nulles de x sont indépendants.

Un problème de programmation linéaire peut aussi se formuler sous la forme (1') dite à contraintes mixtes : maximiser

$$B = \sum_{i=1}^n a_{0i} x_i$$

sous les contraintes :

$$\sum_{i=1}^n a_{ji} x_i \leq a_{j0} \quad j = 1, 2, \dots, m_1 \quad \{M_1\}$$

$$\sum_{i=1}^n a_{ji} x_i = a_{j0} \quad j = m_1 + 1, m_1 + 2, \dots, m \quad \{M_2\}$$

$$x_i \geq 0 \quad i = 1, 2, \dots, n_1 \quad \{N_1\}$$

$$x_i \text{ non limité} \quad i = n_1 + 1, n_1 + 2, \dots, n \quad \{N_2\}$$

(1')

Un tel problème peut toujours se ramener à la formulation (1) en remarquant qu'une contrainte avec égalité correspond à deux contraintes avec inégalités larges de sens opposés, et qu'une variable non limitée est la différence de deux variables limitées. Il peut aussi être avantageux de considérer directement ces problèmes dans leur formulation (1'). Dans cet exposé sommaire, nous ne traiterons, sauf mention expresse du contraire que les problèmes sous la forme (1).

2. - LE THEOREME FONDAMENTAL ET LE CRITERIUM DU SIMPLEXE.

2.1. - Théorème fondamental :

Enoncé : S'il existe une solution admissible optimale x^0 , alors il existe aussi une solution admissible de base optimale, pour laquelle la fonction objective a la même valeur optimale.

Ce résultat dont on peut trouver la démonstration par exemple dans (2), est valable dans tous les cas. Sa signification est la suivante : dans l'espace à n dimension (x_1, x_2, \dots, x_n) , le système des contraintes et des conditions de non-négativité est représenté par l'intérieur d'un polyèdre convexe (borné ou non) ; l'optimum de B sera toujours atteint pour

le sommet du polyèdre le plus extrême par rapport à la famille de plans $B(x_1, x_2, \dots, x_n) = Cte$ (celui ci peut ne pas être unique).

Les différentes méthodes de la programmation linéaire ont pour but ~~trouver~~ de trouver un algorithme efficace pour déterminer le plus rapidement possible ce sommet extrême.

2.2. - Le critérium du simplexe.

Nous excluons maintenant les cas de dégénérescences sur lesquels nous reviendrons plus tard : on dit qu'il y a dégénérescence si le vecteur a_i peut être représenté par une combinaison linéaire positive de moins de m colonnes de la matrice A ; cela revient à dire que par un sommet du polyèdre convexe limitant le domaine possible, passent plus de n plans de contraintes.

Le principe de la méthode du simplexe, exposé par exemple dans le cas du problème de maximisation, consiste, partant d'une solution admissible de base à construire une autre solution admissible de base pour laquelle la fonction objective ait une valeur supérieure (c'est-à-dire : on part d'un sommet du polyèdre et on se déplace de sommet en sommet voisin en augmentant la valeur z de la fonction objective à chaque pas ; il est clair que c'est le caractère convexe du polyèdre -intersection de domaines convexes- qui assure qu'on atteindra ainsi le maximum maximorum).

Plus en détails les choses se présentent de la façon suivante :

Nous partons d'une solution admissible de base qui, après nouvelle indication si nécessaire (.) se met sous la forme :

$$\left. \begin{array}{ll} x_i = 0 & i = 1, 2, \dots, n \\ x_i > 0 & i = n + 1, \dots, n + m \end{array} \right\} \quad (3)$$

(.) Aucune différence n'est plus faite entre les inconnues d'origine du problème et les variables d'écart.

vérifiant

$$Ax = a_{\cdot 0} \quad (4)$$

nous décomposons la matrice A en les deux sous matrices (5) :
 ${}_1A$ ($m \times n$), et ${}_2A$ ($m \times m$) correspondant à $i = n + 1, \dots, n + m$
 et qui donc n'est singulière puisque la solution de départ est
de base.

D'une façon générale la fonction objective s'écrit :

$$z = \sum_{i=1}^{m+n} a_{0i} x_i \quad (6)$$

On peut toujours par des combinaisons appropriées des
 lignes de A et des composantes correspondantes de $a_{\cdot 0}$, mettre
 A dans (4) sous la forme :

$$A = \left(\begin{array}{cccccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 \\ a_{m1} & a_{m2} & \dots & a_{mn} & 0 & 1 \end{array} \right) \quad (7)$$

$\underbrace{\hspace{15em}}_{{}_1A} \quad \underbrace{\hspace{5em}}_{{}_2A}$

Définissons également les vecteurs :

$${}_1a_{0\cdot}^T = (a_{01}, \dots, a_{0n}) \quad {}_2a_{0\cdot}^T = (a_{0,n+1}, \dots, a_{0,n+m})$$

$${}_1x = (x_1, \dots, x_n) \quad {}_2x = (x_{n+1}, \dots, x_{n+m}) ;$$

(4) s'écrit alors :

$${}_1A {}_1x + {}_2A {}_2x = a_{\cdot 0} \quad (8),$$

et (6) devient :

$$z = {}_1a_{0\cdot}^T {}_1x + {}_2a_{0\cdot}^T {}_2x \quad (9).$$

Ceci permet d'obtenir l'expression de z en fonction
 des n variables x_i ($i = 1, \dots, n$)

$$z = ({}_1a_{0\cdot}^T - {}_2a_{0\cdot}^T {}_2A^{-1} {}_1A) {}_1x + {}_2a_{0\cdot}^T {}_2A^{-1} a_{\cdot 0} \quad (10) ;$$

le second terme est le terme constant, valeur de z pour la
 solution admissible de base considérée, soit a_{00} . Comme

${}_2A = I$, on a :

$$a_{00} = {}_2a_{0\cdot}^T \cdot a_{\cdot 0} = \sum_{j=1}^{j=m} a_{0,n+j} a_{j0} \quad (11)$$

et
$$z = a_{00} - \alpha_{0i} x_i \quad (12)$$

avec
$$\alpha_{0i} = \sum_{j=1}^m a_{0,n+j} a_{ji} - a_{0i} \quad (13)$$

Sur l'expression (12) on voit que $-\alpha_{0i} x_i$ représente l'accroissement de z lorsque les variables $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ étant maintenues nulles, x_i croit de 0 à x_i , les valeurs de x_{n+1}, \dots, x_{n+m} étant modifiées de façon à ce que les contraintes (4) demeurent satisfaites (ce qui correspond au déplacement du point représentatif sur une arête du polyèdre).

Il est clair que si $\alpha_{0i} > 0 \quad \forall i = 1, 2, \dots, n$, on a atteint le maximum : la solution admissible de base considérée réalise le maximum.

On appelle critérium du simplexe l'inégalité $\alpha_{0i} > 0$, $i = 1, \dots, n$, ou encore :

$${}_1a_{0\cdot}^T - {}_2a_{0\cdot}^T \cdot {}_2A^{-1} \cdot {}_1A \leq 0.$$

Si au moins un $\alpha_{0i} < 0$, la maximum n'est pas atteint. Pour tout x_i tel que $\alpha_{0i} < 0$, il existe au moins un $a_{ji} > 0$ dans A mise sous la forme (7). En effet s'il n'en est pas ainsi x_i peut croître indéfiniment toutes les contraintes demeurant satisfaites, et z croît alors indéfiniment aussi : il n'y a pas de maximum.

Soit x_e une variable telle que $\alpha_{0e} < 0$; il est possible de trouver une variable x_{n+f} non nulle dans la solution admissible de base de départ, que l'on peut annuler et remplacer par $x_e > 0$ dans la nouvelle solution admissible de base, sans qu'aucune des autres variables de base ne devienne négative. On montre qu'il suffit pour cela de déterminer l'indice f , comme l'indice j correspondant au minimum de a_{je}/a_{je} ($a_{je} > 0$), (*), pour $j = 1, \dots, m$, et de prendre pour nouvelle valeur de x_e la valeur de ce minimum.

La nouvelle solution admissible de base est alors constituée de :

(.) Les a_{j0} sont nécessairement > 0 , puisque A a la forme (7) et que x est admissible.

$$\begin{aligned}
 x_i &= 0 & i &= 1 \dots e-1, e+1, \dots n, n+f \\
 x_i &\neq 0 & i &= e, n+1, \dots n+f-1, n+f+1, \dots n+m.
 \end{aligned}$$

les valeurs des variables non nulles sont :

$$(14) \quad \begin{cases} x_i = x_{n+k} = a'_{k0} = a_{k0} - \frac{a_{f0}}{a_{fe}} a_{ke} & i = n+1, \dots, n+f-1, n+f+1, \dots, n+m \\ x_i = x_e = a'_{e0} = \frac{a_{f0}}{a_{fe}} & i = e \end{cases}$$

et les termes de la nouvelle matrice A , (soit A') mise sous la forme (7) mais sans nouvelle indication sont :

$$(15) \quad \begin{cases} a'_{ji} = a_{ji} - a_{fi} \frac{a_{je}}{a_{fe}} & i = 1, 2 \dots n+m \\ a'_{ei} = \frac{a_{fi}}{a_{fe}} & j = 1, \dots, f-1, f+1, \dots, n+m \end{cases}$$

La nouvelle valeur de z est

$$z = a_{00} - \alpha_{0i} \frac{a_{f0}}{a_{fe}} \quad (16)$$

Nous n'avons pas encore précisé le choix de x_e parmi les x_i pour lesquels $\alpha_{0i} < 0$.

Etant donné la signification de α_{0i} , il est assez raisonnable de choisir comme indice e celui qui correspond au minimum de α_{0i} (négatif). C'est le critère de Dantzig. En fait, comme cela est signalé dans (3), il serait plus logique de choisir pour e la valeur qui correspond au maximum de

$-\alpha_{0i} \frac{a_{f0}}{a_{fe}}$; mais cela nécessiterait d'effectuer avant le choix et donc pour tous les e possibles, la recherche du f correspondant : il n'est pas évident que du point de vue de la rentabilité (temps de calcul) cette méthode serait meilleure. (On peut expliquer ainsi pour atteindre l'optimum, la différence entre les deux critères : le critère de Dantzig correspond à choisir le nouveau sommet du polyèdre sur l'arête issue du sommet précédent le long de laquelle l'augmentation unitaire de z est la plus forte, tandis qu'il serait plus logique de choisir comme nouveau sommet, celui des sommets voisins qui correspond à la plus forte augmentation de z , du sommet initial au nouveau).

L'algorithme du simplexe aboutit en un nombre fini d'opérations ; en effet il y a un nombre fini de solutions admissibles de base, et chacune est atteinte au plus une fois puisque z croît strictement à chaque étape. L'intérêt de la méthode tient à ce qu'en fait, le nombre d'itérations est très nettement inférieur au nombre de toutes les admissibles solutions de base.

Nous n'en dirons pas beaucoup plus sur le simplexe, aujourd'hui très classique.

2.3. - Dégénérescences.

Il y a dégénérescence lorsque par un sommet du polyèdre limitant le domaine admissible dans l'espace (x_1, x_2, \dots, x_n) , passent plus de n plans de contraintes.

Cette propriété est équivalente à chacune des deux suivantes :

. le vecteur a_{j_0} s'exprime en fonction de $(m - 1)$ colonnes de la matrice A ;

. le minimum de $\frac{a_{j_0}}{a_{je}}$, $a_{je} > 0$, est atteint simultanément pour plusieurs valeurs de j .

Il apparaît alors que le critérium du simplexe n'est plus une condition nécessaire pour l'obtention du maximum.

Charnes (⁴) a développé une méthode permettant de traiter ces problèmes, dite méthode des ϵ . Elle est fondée sur le fait que par un léger déplacement des plans de contraintes on rend impossible que plus de n plans concourent au même sommet critique.

2.4. - La solution admissible de base, de départ.

Lorsque dans la formulation (1) du problème, les a_{j_0} sont positifs, on peut prendre comme solution admissible de base initiale,

$$x_i = 0 \quad i = 1, 2, \dots, n$$

$$x_i = a_{j_0} \quad i = n + j \quad j = 1, 2, \dots, m.$$

S'il n'en est pas ainsi on utilise diverses méthodes pour la détermination d'une solution admissible de base initiale : la méthode dite des variables d'écart artificielles (ou M - method) (⁵), ou l'algorithme dû à Dantzig.

2.4. - Perfectionnements de la méthode du simplexe.

D'autres méthodes issues et perfectionnées du simplexe sont utilisées. Citons la méthode de la mise en facteur de l'inverse (Revised simplex method), développée par Dantzig (⁶) ; l'algorithme de décomposition (méthode des sous-matrices) dû à Dantzig et Wolfe (⁷). En effet dans certains cas pratiques où les problèmes à résoudre comportent un grand nombre de variables et un grand nombre de contraintes, l'algorithme du simplexe devient inutilisable même si l'on dispose d'un équipement de calcul très puissant ; l'idée de la méthode de décomposition, est de décomposer le problème en un certain nombre de problèmes partiels que l'on résout séparément et de construire la solution du problème global à partir des solutions des problèmes partiels.

L'algorithme du duoplexe. L'idée fondamentale de cette méthode dont le but est d'arriver plus rapidement à l'optimum, vient de ce qu'il est possible dans de nombreux cas de trouver rapidement un plan de contrainte qui contient le point optimal cherché. Signalons simplement à propos de cette méthode que dans la première étape, dans le but de trouver un plan de contraintes qui ait des chances de contenir l'optimum (maximum), on détermine le plan de contraintes dont l'angle de la normale intérieure au polyèdre avec le gradient de la fonction objective est le plus grand.

Il convient de mentionner également l'existence d'une méthode, due à Gomory (⁸), pour résoudre les problèmes de programmation linéaire dans le cas où en plus des contraintes linéaires comme dans (1), on impose aux inconnues d'origine du problème d'être à valeurs entières.

3. - DUALITE EN PROGRAMMATION LINEAIRE.

Nous allons exposer avec quelques détails le concept de dualité en programmation linéaire.

3.1. - Programmes duals.

On dit que deux programmes linéaires sont duals s'il existe entre eux la relation suivante :

$$\left. \begin{array}{l}
 \text{Problème de maximisation :} \\
 \text{maximiser} \\
 B = a_0^T x \\
 \text{sous les contraintes} \\
 Ax \leq a_0 \\
 x \geq 0
 \end{array} \right\} (17)$$

$$\left. \begin{array}{l}
 \text{Problème de minimisation :} \\
 \text{minimiser} \\
 C = w^T a_0 \\
 \text{sous les contraintes} \\
 A^T w \geq a_0 \\
 w \geq 0
 \end{array} \right\}$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{m1} & a_{m2} & & a_{mn} \end{pmatrix}$$

3.2. - Propriétés.

On a les propriétés suivantes :

Lemme 1 : Pour tout couple de solutions admissibles x et w des deux problèmes on a l'inégalité

$$a_0^T x \leq w^T a_0 \quad (18)$$

Lemme 2 : Si x^0 et w^0 sont des solutions admissibles et si on a :

$$a_0^T x^0 = w^{0T} a_0 \quad (19)$$

alors x^0 et w^0 sont optimales.

Lemme 3 : Le système

$$Kz \geq 0 \quad z \geq 0 \quad (K^T = -K)$$

où K est une matrice carrée antisymétrique, a des solutions z_0 telles que :

$$Kz_0 + z_0 > 0. \quad (20)$$

Ce résultat, dont la démonstration est due à Tucker (⁹) a pour conséquences pour les problèmes (17) :

∃ $w^0 \geq 0, x^0 \geq 0, \lambda^0 \geq 0$ vérifiant

$$\begin{aligned} a_{\cdot 0} \lambda^0 &\geq A x^0 \\ A^T w^0 &\geq a_{0 \cdot} \lambda^0 \\ a_{0 \cdot} x^0 &\geq a_{\cdot 0}^T w^0 \end{aligned} \quad (21)$$

et

$$\begin{aligned} w^0 + a_{\cdot 0} \lambda^0 &> A x^0 \\ x^0 + A^T w^0 &> a_{0 \cdot} \lambda^0 \\ \lambda^0 + a_{0 \cdot}^T x^0 &> a_{\cdot 0}^T w^0. \end{aligned} \quad (22)$$

Alors

Lemme 4 : Si $\lambda^0 > 0$, il existe des solutions admissibles optimales x^0 et w^0 pour les programmes en dualité (17), qui vérifient :

$$\begin{aligned} a_{0 \cdot}^T x^0 &= a_{\cdot 0}^T w^0 \\ w^0 + a_{\cdot 0} &> A x^0 \\ A^T w^0 + x^0 &> a_{0 \cdot}. \end{aligned} \quad (23)$$

Lemme 5 : si $\lambda^0 = 0$ alors

- a/ - au moins l'un des deux problèmes linéaires en dualité n'a pas de solutions admissibles.
- b/ - Si le problème de maximisation a des solutions admissibles, l'ensemble de ces solutions admissibles n'est pas borné, et l'ensemble des $a_{0 \cdot}^T x$ n'est pas borné supérieurement.
(résultats analogues si c'est le problème de minimisation qui a des solutions admissibles).
- c/ - Aucun des deux problèmes n'a de solutions admissibles optimales.

De tout cela on déduit que : ou bien les deux problèmes en dualité ont des solutions admissibles optimales, ou bien aucun des deux n'en a. Dans le premier cas, le maximum et le minimum sont égaux et leur valeur commune est la valeur optimale des deux problèmes duals.

Ce qui permet d'énoncer les deux théorèmes fondamentaux de la théorie :

Une solution admissible x^0 est optimale si et seulement si il existe une solution admissible w^0 telle que :

$$a_{\cdot 0}^T w^0 = a_0^T x^0 \quad (24) ;$$

(énoncé identique en permutant les rôles de x^0 et w^0 , (24) demeurant inchangée).

Une condition nécessaire et suffisante pour que l'un (et par conséquent les deux) des problèmes duals ait des solutions admissibles optimales est que les deux problèmes aient des solutions admissibles.

Remarque :

Si le problème primal est posé sous la forme (1') du § 1, la dualisation se fait comme suit :
minimiser

$$C = \sum_{j=1}^m a_{j0} w_j$$

sous les contraintes :

$$\sum_{j=1}^m a_{ji} w_j \geq a_{0i} \quad i = 1, 2, \dots, n_1 \quad \{N_1\}$$

$$\sum_{j=1}^m a_{ji} w_j = a_{0i} \quad i = n_1 + 1, n_1 + 2, \dots, n \quad \{N_2\}$$

$$w_j \geq 0 \quad j = 1, 2, \dots, m_1 \quad \{M_1\}$$

$$w_j \text{ non limité} \quad j = m_1 + 1, \dots, m \quad \{M_2\}$$

(17')

3.3. - La méthode du simplexe duale.

En application de la théorie de la dualité donnée au § 3.2, il apparaît que pour la résolution de certains problèmes, il pourra être utile d'employer la méthode du simplexe et de résoudre le problème dual du problème posé. Compte tenu des relations existant entre les deux problèmes, on obtient de cette manière l'optimum du problème posé et aussi une solution admissible optimale de ce problème.

La description pratique de cette méthode est donnée par exemple dans (2) ; on pourra se reporter à (10) pour des précisions supplémentaires. Lemke (11) a montré qu'il existait entre les bases pour le système primal et pour le dual, une certaine complémentarité qui permet d'interpréter l'algorithme du simplexe appliqué au dual comme une suite de changements de bases pour le système primal. Ainsi l'utilisation du simplexe pour le problème dual s'y fait par des règles automatiques sur les x_i comme dans le programme primal, sans qu'il soit besoin d'introduire les w_i ; on obtient en quelque sorte une autre méthode de résolution du problème primal mais la validité de celle-ci ne se démontre commodément que si l'on remarque qu'elle correspond à l'application du simplexe au problème dual.

PROGRAMMATION NON LINEAIRE

Les recherches sur la programmation non linéaire se sont heurtées à des difficultés mathématiques plus importantes que pour la programmation linéaire et la théorie qui a fait de gros progrès ne peut pas être considérée comme achevée.

Nous présentons ici les résultats fondamentaux et donnons un aperçu, plus ou moins détaillé suivant les cas, sur les méthodes employées.

1. - PROBLEMES DE PROGRAMMATION NON LINEAIRE EN GENERAL.

L'extension du problème d'optimisation linéaire, au cas non linéaire, conduit si on la fait de la façon la plus générale, à poser le problème suivant :

minimiser la fonction objective

$$G(x)$$

soumis aux contraintes

$$g_j(x) \leq 0 \quad j = 1, 2, \dots, m$$

$$x_i \geq 0 \quad i = 1, 2, \dots, n$$

(1)

sans imposer à $G(x)$ et $g_j(x)$ de conditions particulières.

On est loin à l'heure actuelle de disposer d'une méthode de résolution de ce problème, et on doit se contenter dans la plupart des cas d'utiliser des méthodes par essais de valeurs, qui ne permettent bien souvent que d'atteindre le voisinage d'un optimum secondaire.

Par contre de bonnes méthodes existent dans le cas de la programmation convexe que nous allons exposer.

2. - PROGRAMMATION CONVEXE.

2.1. - Définition.

Soient $F(x)$ et $f_j(x)$, $j = 1, 2, \dots, m$, des fonctions convexes. Un programme convexe sera :

$$\left. \begin{array}{l} \text{minimiser } F(x) \\ \text{sous les contraintes} \\ f_j(x) \leq 0 \quad j = 1, 2, \dots, m \\ x_i \geq 0 \quad i = 1, 2, \dots, n \end{array} \right\} \quad (2)$$

(Pour un problème de maximisation $F(x)$ devrait être concave).

En raison de la convexité des f_j , l'ensemble R des solutions admissibles est convexe.

2.2. - Le théorème de Kuhn et Tucker.

Le théorème de Kuhn et Tucker (¹²) est le théorème central de la théorie des programmes convexes. C'est en quelque sorte la généralisation de la théorie des multiplicateurs de Lagrange au cas où les contraintes ne sont pas formées que d'égalités mais également d'inégalités. Le théorème donne des conditions nécessaires et suffisantes pour qu'un vecteur x soit solution du problème (2) (c'est-à-dire solution admissible optimale).

Introduisons m nouvelles variables u_1, \dots, u_m , correspondant aux multiplicateurs de Lagrange et soit Φ la fonction de Lagrange généralisée :

$$\Phi(x, u) = F(x) + \sum_{j=1}^m u_j f_j(x) \quad (3)$$

L'énoncé du théorème de Kuhn-Tucker est le suivant :
Un vecteur \hat{x} est solution de (2) si et seulement si $\exists \hat{u}$ tel que :

$$\left. \begin{array}{l} \hat{x} \geq 0, \quad \hat{u} \geq 0, \\ \Phi(\hat{x}, u) \leq \Phi(\hat{x}, \hat{u}) \leq \Phi(x, \hat{u}) \\ \forall x \geq 0, \quad \forall u \geq 0 \end{array} \right\} \quad (4)$$

On associe au problème de minimisation (2), un problème de minimax, puisque (4) signifie que $\Phi(x, \hat{u})$ a en (\hat{x}, \hat{u}) pour x positif, un minimum par rapport à x , et que $\Phi(\hat{x}, u)$ y a un maximum par rapport à u .

La démonstration est donnée dans (13) sous la forme due à Slater (14) qui s'applique aux fonctions convexes quelconques, alors que la démonstration originale de Kuhn et Tucker s'applique uniquement au cas où F et f_{ij} sont dérivables; une hypothèse de régularité sur les f_{ij} est nécessaire — $\exists \bar{x}$ tel que $f_j(\bar{x}) < 0, \forall j$ — sauf si ces f_j sont linéaires.

Dans le cas où F et f_{ij} sont dérivables, (4) est équivalent aux conditions de Kuhn-Tucker :

$$\left. \begin{aligned} \frac{\partial \Phi(\hat{x}, \hat{u})}{\partial x_i} &\geq 0 & (5) \\ \hat{x}_i \frac{\partial \Phi(\hat{x}, \hat{u})}{\partial x_i} &= 0 & (6) \\ \hat{x}_i &\geq 0 & (7) \end{aligned} \right\} \quad i = 1, 2, \dots, n$$

$$\left. \begin{aligned} \frac{\partial \Phi(\hat{x}, \hat{u})}{\partial u_j} &\leq 0 & (8) \\ \hat{u}_j \frac{\partial \Phi(\hat{x}, \hat{u})}{\partial u_j} &= 0 & (9) \\ \hat{u}_j &\geq 0 & (10) \end{aligned} \right\} \quad j = 1, 2, \dots, m$$

On peut donner une interprétation géométrique des conditions de Kuhn-Tucker (5) à (10) : Remplaçant par Φ par son expression (3) on voit que : (8) $\Leftrightarrow f_j(x) \leq 0$, donc (7) et (8) signifient que \hat{x} est admissible.

$$\text{Posons } \frac{\partial \Phi(\hat{x}, \hat{u})}{\partial x_i} = \hat{v}_i, \quad (11)$$

on voit que (5) donne :

$$\text{grad } F(\hat{x}) + \sum_{j=1}^m \hat{u}_j \text{grad } f_j(\hat{x}) = \sum_{i=1}^n \hat{v}_i e_i \text{ avec } \hat{v}_i \geq 0, \forall i \quad (12)$$

Si \hat{M} est l'ensemble des j tels que $f_j(\hat{x}) = 0$,
et \hat{N} l'ensemble des i tels que $\hat{x}_i = 0$,
(9) s'écrit :

$$\hat{u}_j = 0 \quad j \in \hat{M} \quad (13)$$

et (6) :

$$\hat{v}_i = 0 \quad i \in \hat{N} \quad (14),$$

donc (5) à (10) s'écrivent aussi :
 \hat{x} est admissible

$$\left. \begin{aligned} -\text{grad } F(\hat{x}) &= \sum_{j \in \hat{M}} \hat{u}_j \text{grad } f_j(\hat{x}) + \sum_{i \in \hat{N}} \hat{v}_i (-e_i) \\ \hat{u}_j &\geq 0, j \in \hat{M}; \hat{v}_i \geq 0, i \in \hat{N}. \end{aligned} \right\} \quad (15)$$

et réciproquement, (15) est équivalent à (5) à (10).

(15) fournit l'interprétation géométrique : au point optimal \hat{x} , le gradient négatif de la fonction objective à minimiser doit être représentable par une combinaison linéaire non négative des normales extérieures des surfaces frontières sur lesquelles se trouve \hat{x} ; si \hat{x} est intérieur au domaine admissible, le gradient doit s'annuler.

Il est intéressant d'indiquer comment sont modifiées les conditions de Kuhn-Tucker dans le cas où le problème n'est pas posé sous la forme (2) :

pour les variables x_i qui ne subissent pas de contraintes de signe, (5) devient

$$\frac{\partial \Phi(\hat{x}, \hat{u})}{\partial x_i} = 0 \quad (5')$$

et évidemment dans (7) \hat{x}_i n'est pas limitée ;

pour les contraintes d'indice j qui sont des égalités, (8) devient :

$$\frac{\partial \Phi(\hat{x}, \hat{u})}{\partial u_j} = 0 \quad (8'),$$

et dans (10) \hat{u}_j n'est pas limité.

2.3. - Dualité en programmation convexe.

Le concept de dualité est moins développé en programmation non linéaire qu'en programmation linéaire. L'intérêt en est moindre car il n'y a plus symétrie parfaite entre les problèmes primal et dual. Des travaux de Dorn, Dennis, Wolfe, Hanson, Mangasarian ont été consacrés à ce sujet.

2.3.1. - Les principes de la dualité.

La dualité consiste à associer deux problèmes de programmation dont l'un, le primal, est un problème de minimisation (maximisation) sous contraintes et l'autre, le dual, est un problème de maximisation (minimisation) sous contraintes, de telle sorte que l'existence de la solution d'un des deux problèmes assure l'existence de la solution de l'autre, et que les extremums des deux problèmes soient égaux (¹⁵).

2.3.2. - Théorème de dualité de Wolfe (¹⁶).

Enoncé : Si x^0 est une solution du problème primal :

$$\left. \begin{array}{l} \text{minimiser } F(x) \\ \text{sous les contraintes} \\ f_j(x) \leq 0 \quad j = 1, 2, \dots, m \end{array} \right\} \quad (16)$$

où F et f_j sont des fonctions convexes différentiables de x , alors il existe un m -vecteur $u^0 \geq 0$ tel que (x^0, u^0) est une solution du problème dual; maximiser

$$\psi(x, u) = F(x) + \sum_{j=1}^m u_j f_j(x)$$

sous les contraintes

$$\frac{\partial \psi(x, u)}{\partial x_i} = 0 \quad i = 1, 2, \dots, n \quad (17)$$

$$u_j \geq 0 \quad j = 1, 2, \dots, m \quad (18),$$

et on a $F(x^0) = \psi(x^0, u^0)$.

Le théorème est une conséquence directe du théorème de Kuhn et Tucker dont il n'est presque qu'une transposition.

2.3.3. - Réciproque du théorème de dualité.

Cette réciproque est due à Hanson ⁽¹⁷⁾ et Mangasarian⁽¹⁵⁾
 Enoncé :

Soit (x^0, u^0) une solution du problème dual où $F(x)$ et $f_j(x)$ sont des fonctions convexes de classe C^2 de x . Le vecteur x^0 est une solution du problème primal si au moins une des conditions suivantes est vérifiée :

Soit $F(x)$ est strictement convexe au voisinage de x^0 soit au moins une des $f_j(x)$ correspondant à $u_j^0 \neq 0$ est strictement convexe au voisinage de x^0 .

2.3.4. - Commentaires.

Ce théorème et sa réciproque ne sont pas totalement satisfaisants, non pas tant par la présence de x dans le programme dual et l'absence de symétrie, mais surtout par la nécessité des hypothèses supplémentaires énoncées dans la réciproque par Mangasarian, qui paraissent trop restrictives : en particulier la réciproque sous cette forme ne s'applique pas au cas de la programmation linéaire car ces hypothèses sont alors violées.

Dans le cas où le problème primal n'est pas posé sous la forme (16), mais sous la forme (2) par exemple, (17) et (18) sont à transformer comme on l'a fait au § 2.2 pour les conditions de Kuhn-Tucker, mais la fonction à minimiser doit elle-même être modifiée. Nous verrons ce point plus en détails dans le cas de la programmation quadratique.

3. - PROGRAMMATION QUADRATIQUE.

3.1. - Généralités.

On considère maintenant le cas de la programmation quadratique qui revêt une importance particulière du point de vue des algorithmes de résolution dont on dispose.

On suppose que la fonction objective $F(x)$ est une fonction convexe quadratique $Q(x)$:

$$Q(x) = a_0^T \cdot x + x^T Cx \quad (19)$$

(C = matrice définie positive ou semi-définie positive) ;
de plus les contraintes sont linéaires, de la forme :

$$Ax \leq a_{.0} \quad (20)$$

et $x \geq 0 \quad (21).$

Plusieurs formulations des contraintes sont d'ailleurs possibles, comme dans le cas de la programmation linéaire : certaines des contraintes (20) pouvant être des égalités et certaines des variables x_i pouvant être sans contraintes de signe (21). Sans introduire de forme normale à contraintes mixtes comme dans le cas de la programmation linéaire, on considère outre le problème de type I correspondant à (20) et (21), le problème de type II :

minimiser

$$Q(x) = a_0^T \cdot x + x^T Cx$$

sous les contraintes

$$Ax = a_{.0} \quad (22)$$

$$x \geq 0 \quad (23) ;$$

et le problème de type III :

minimiser

$$Q(x) = a_0^T \cdot x + x^T Cx$$

sous les contraintes

$$Ax \leq a_{.0} \quad (24).$$

Le domaine R délimité par les contraintes est un polyèdre convexe comme pour les programmes linéaires.

La différence qui apparaît avec les programmes linéaires est la suivante : dans le cas de fonctions objectives linéaires, le minimum est toujours atteint en un sommet du polyèdre ; pour une fonction objective quadratique au contraire, le minimum sous contraintes imposées peut être atteint sur une arête ou à l'intérieur du polyèdre même s'il est unique. Autrement dit dans les programmes linéaires non dégénérés il y a toujours une solution admissible optimale où n contraintes sont vérifiées avec l'égalité, dans le cas de programmes quadratiques n contraintes au plus sont vérifiées avec égalité par la solution optimale.

3.2. - Expression des conditions de Kuhn et Tucker.

La fonction de Lagrange est

$$\Phi(x, u) = Q(x) + \sum_{j=1}^m u_j (a_{ji} x_i - a_{j0})$$

$$\Phi(x, u) = a_{0.}^T x + x^T Cx + u^T (Ax - a_{.0}) \quad (25).$$

En posant :

$$\frac{\partial \Phi}{\partial x} = v \quad \text{et} \quad - \frac{\partial \Phi}{\partial u} = y \quad (26)$$

ou

$$v = a_{0.} + 2 Cx + A^T u \quad (27)$$

et

$$y = - Ax + a_{.0} \quad (28),$$

les conditions de Kuhn-Tucker deviennent :
pour le problème I :

un vecteur \hat{x} est solution du problème de minimisation I si et seulement si il existe un n -vecteur \hat{v} et deux m -vecteurs \hat{y} et \hat{u} tels que :

$$Ax + y = a_{.0} \quad (29)$$

$$2 Cx - v + A^T u = - a_{.0}. \quad (30)$$

$$x \geq 0, y \geq 0, v \geq 0, u \geq 0 \quad (31)$$

$$x^T v + y^T u = 0 \quad (32)$$

I

de même pour le type II où les vecteurs \hat{v} , \hat{u} doivent vérifier :

$$A_x = a_{.0} \quad (33)$$

$$2 Cx - v + A^T u = - a_{.0}. \quad (34)$$

$$x \geq 0, v \geq 0 \quad (35)$$

$$x^T v = 0 \quad (36)$$

II

et pour le type III.

Cette formulation est due à Barankin et Dorfman (¹⁸). On voit sur (32) que pour chaque couple de variables x_i, v_i et u_i, y_i l'une au moins doit être nulle. Donc seules les solutions du système linéaire de $(n + m)$ équations formé par (29), (30) et (31) pour lesquelles au plus $(n + m)$ variables sont non nulles sont à considérer ; ce sont les solutions de base du système parmi lesquelles on doit faire un choix pour que (32) soit également vérifiée. On pourra donc pour cela utiliser une version modifiée du simplexe.

3.3. - Dualité en programmation quadratique.

Les résultats que nous avons énoncés au § 2.3 sont applicables ici.

Le dual du problème III, en appliquant les formules (17) et (18) s'écrit :

maximiser

$$\psi(x, u) = a_{.0}^T x + x^T Cx + u^T (Ax - a_{.0}) \quad (37)$$

sous les contraintes

$$\frac{\partial \psi}{\partial x} = a_{.0} + 2 Cx + A^T u = 0 \quad (38)$$

$$u \geq 0 \quad (39),$$

soit en tenant compte de (38) dans (37) :
maximiser

$$\phi(x, u) = -x^T Cx - a_{\cdot 0}^T u \quad (40)$$

sous les contraintes
(38) et (39) :

c'est la formulation donnée par Dorn (19) ;
pour le problème de type I le dual est :
maximiser

$$\phi(x, u) = -x^T Cx - a_{\cdot 0}^T u \quad (40)$$

sous les contraintes

$$2 Cx + A^T u + a_{\cdot 0} \geq 0 \quad (41)$$

$$u \geq 0 \quad (42) ;$$

pour le dual du problème II , il suffit de supprimer la condition (42).

Les propriétés sont alors les suivantes :

Lorsque le domaine admissible d'un deux problèmes est vide, celui de l'autre est vide aussi ou bien la fonction objective n'est pas bornée sur ce domaine ;

si la fonction objective d'un des problèmes n'est pas bornée dans le domaine admissible, le domaine de l'autre problème est vide ;

lorsqu'un problème possède une solution, l'autre en possède une également et les valeurs extrémales sont égales ;

si \hat{x} est solution du problème primal, \hat{x} est aussi solution de la partie en \hat{x} du problème dual, mais la réciproque n'est vraie que si C est strictement définie positive.

Pour les mêmes raisons qu'au § 2.3.4, la réciproque énoncée ci-dessus n'est pas satisfaisante.

3.4.-Quelques méthodes de résolution des programmes quadratiques.

3.4.1. - La méthode d'Hildreth et d'Esopo.

Nous parlerons d'abord de la méthode d'Hildreth et d'Esopo (²⁰, ²¹, ²²), qui utilise les résultats concernant la dualité dont nous avons traité au § précédent. C'est une méthode asymptotique où les itérations à réaliser sont très simples, ce qui rend commode son emploi sur les calculateurs, mais qui ne s'applique qu'au cas où la matrice C est strictement définie positive.

Le problème est mis sous la forme III
minimiser

$$Q(x) = a_0^T x + x^T C x$$

sous les contraintes

$$Ax \leq a_0 \quad (24),$$

où C est définie positive, et on fait l'hypothèse supplémentaire sur les contraintes qu'il existe $\epsilon < 0$, tel que si l'on remplace les contraintes (24) par les contraintes (43) il existe encore des points admissibles :

$$Ax \leq a_0 - \epsilon \quad (43)$$

Les conditions de Kuhn-Tucker s'écrivent :

$$Ax + y = a_0 \quad (44)$$

$$2 Cx + A^T u = -a_0 \quad (45)$$

$$u \geq 0, \quad y \geq 0 \quad (46)$$

$$u^T y = 0 \quad (47)$$

III

Puisque C^{-1} existe, (45) donne :

$$x(u) = -\frac{1}{2} C^{-1} (A^T u + a_0) \quad (48)$$

qui combiné avec (44), (46), (47), transforme celles-ci en les conditions de Kuhn-Tucker pour le problème

$$\min \{ \varphi(u) = h^T u + u^T G u \mid u \geq 0 \} \quad (49),$$

$$\text{où } h = \frac{1}{2} AC^{-1} a_{\cdot 0} + a_{\cdot 0} \quad (50)$$

$$\text{et } G = \frac{1}{4} AC^{-1} A^T \quad (51),$$

qui est le programme dual (40) ^{(38) et (39)} de (24) formulé sous forme de minimisation, à une constante près.

On a le théorème :

\hat{x} est solution du problème initial posé, si et seulement si $\hat{x} = x(\hat{u})$ où \hat{u} est solution du problème (49).

Le problème est donc ramené à la résolution du problème (49), pour lequel les contraintes sont simples.

La méthode de Hildreth et d'Esopo est simple et proche de l'algorithme de Gauss Seidel utilisé dans la résolution des systèmes linéaires.

Partant d'une solution admissible u^0 , par exemple $u^0 = 0$, on construit u^1 de la façon suivante : on minimise (u) successivement par rapport à chaque composante u_i de u sous la condition $u_i \geq 0$, les autres composantes restant fixes égales à leur dernière valeur calculée. De même pour la détermination de u^{p+1} à partir de u^p .

On obtient ainsi une suite de points admissibles u^p et on démontre que la suite décroissante $\varphi(u^p)$ converge vers le minimum $\varphi(\hat{u})$, et que la suite $x(u^p)$ converge vers $\hat{x} = x(\hat{u})$.

3.4.2. - La méthode de Beale.

Nous allons donner quelques détails sur la méthode due à Beale (23, 24), qui est une extension de la méthode du simplexe.

Le problème est mis sous la forme II

$\min \{Q(x)\}$
sous les contraintes

$$Ax = a_{\cdot 0} \quad (22)$$

$$x \geq 0 \quad (23)$$

n variables, m contraintes, $m < n$.

On part d'une solution admissible de base du système (22) (23) : $(n - m)$ variables sont nulles dans cette solution et les m autres sont positives ; les premières sont les variables indépendantes, soit $z_h = x_{m+h}$, $h = 1, 2, \dots, n - m$, les autres les variables de base ou dépendantes soit x_1, x_2, \dots, x_m .
 Exprimant x_g en fonction de z_h par (22) on peut éliminer les x_g dans Q qui devient :

$$Q(x_1, x_2, \dots, x_n) = Q^1(z_1, \dots, z_{n-m}) \\ = C_{00}^1 + 2 \sum_{i=1}^{n-m} C_{i0}^1 z_i + \sum_{h=1}^{n-m} \sum_{i=1}^{n-m} C_{hi}^1 z_i z_h \quad (52)$$

(où $C_{ih}^1 = C_{hi}^1$).

Le coefficient total de z_h est

$$C_{h0}^1 + C_{h1}^1 z_1 + \dots + C_{h,n-m}^1 z_{n-m}, \quad (53)$$

il représente la dérivée $\frac{1}{2} \frac{\partial Q^1}{\partial z_h}$ et en particulier au point d'essai on a :

$$\frac{1}{2} \frac{\partial Q^1}{\partial z_h} = C_{h0}^1 \quad \text{et} \quad Q = C_{00}^1 \quad (54)$$

Alors si $\forall h = 1, \dots, n-m, \frac{\partial Q^1}{\partial z_h} \geq 0$, le point d'essai est la solution optimale ;

par contre si certaines variables z_h correspondent à :

$$\frac{1}{2} \frac{\partial Q^1}{\partial z_h} = C_{h0}^1 < 0,$$

cela prouve que l'on peut abaisser la valeur de Q en rendant positif un de ces z_h , les autres variables z étant maintenues nulles, et les variables dépendantes, variant comme dans le simplexe pour continuer à satisfaire (22).

Supposons par exemple que $C_{10}^1 < 0$ et faisons croître z_1 . Comme dans le cas du simplexe la question qui se pose est : de combien peut-on faire croître z_1 ? Deux circonstances peuvent se produire qui conduisent à arrêter la croissance de z_1 :

a/ - une variable de base s'annule (comme dans le cas linéaire) ;

b/ - la dérivée $\frac{\partial Q^1}{\partial z_1}$ s'annule avant qu'une variable de base ne s'annule.

Dans le cas a/, le deuxième point d'essai sera :

$$z_1 = \lambda > 0, x_2, \dots, x_m \neq 0, x_1 = 0, z_2 = z_3 = \dots = z_{n-m} = 0,$$

et la transformation s'effectuera comme dans la méthode du simplexe pour se ramener à une situation analogue à celle dont on est parti.

Dans le cas b/ qui est spécifique :

$\frac{\partial Q^1}{\partial z_1}$ s'annule à l'intérieur du domaine admissible ; on introduit alors une nouvelle variable u_1 , sans contrainte de signe,

$$u_1 = \frac{1}{2} \frac{\partial Q^1}{\partial z_1}, \quad (55)$$

variable libre ;

celle-ci est liée aux variables indépendantes par la relation linéaire

$$u_1 = C_{10}^1 + \sum_{h=1}^{n-m} C_{1h}^1 z_h \quad (56).$$

Le deuxième point d'essai correspond alors à :

$u_1, z_2, \dots, z_{n-m} = 0$ ce sont les variables indépendantes en ce point d'essai,

et $z_1, x_1, \dots, x_m \neq 0$, s'exprimant en fonction des variables indépendantes par (56) et (22) : ce sont les variables dépendantes au nombre de $(m+1)$. Le nouveau point d'essai se traitera comme le précédent avec une variable de base supplémentaire et u_1 comme variable indépendante, Q étant exprimé en fonction des variables indépendantes sous la forme Q^2 . La condition d'optimum (de Kuhn-Tucker) pour la variable libre u_1 est

$\frac{\partial Q^2}{\partial u_1} = 0$. On applique de plus la règle additionnelle :

Lorsque cela est possible on commence toujours par faire varier les variables libres pour passer d'un point d'essai à un autre. Lorsque la dérivée de Q par rapport à toutes les variables libres s'annule, on peut prendre dans la base une variable réelle indépendante.

L'algorithme est très bien décrit, sur un exemple dû à Beale, dans (5).

On démontre que le point optimal est atteint après un nombre fini d'itérations (importance de la règle additionnelle).

3.4.3. - Les méthodes de gradient

Nous parlerons maintenant des méthodes de gradient. L'application de cette catégorie importante de méthode d'optimisation n'est pas limitée aux programmes quadratiques et est valable pour les programmes convexes en général. Il n'est pas question d'exposer ici toutes les méthodes de ce type et nous nous limiterons à traiter, sans entrer dans les détails, de quelques unes qui nous paraissent représentatives.

3.4.3.1. - Une méthode différentielle : la méthode du gradient de Lagrange.

Soit à minimiser la fonction convexe $F(x)$
sous les contraintes convexes

$$\begin{aligned} f_j(x) &\leq 0 \\ x_i &\geq 0 \end{aligned} \quad (2)$$

L'idée consiste, partant d'un premier point d'approximation x^0 à aboutir au minimum du problème (2), en suivant une trajectoire $x(t)$ projection d'une ligne de plus grande pente de la surface $y = F(x)$.

La trajectoire est définie par :

$$\frac{dx(t)}{dt} = - \text{grad } F(x) - \sum_{j=1}^m u_j \text{ grad } f_j(x) \quad (57)$$

$$\frac{du_j(t)}{dt} = \begin{cases} f_j(x) & \text{si } f_j(x) > 0 \text{ ou } u_j(x) > 0 \\ 0 & \text{si } f_j(x) \leq 0 \text{ et } u_j(x) \leq 0 \end{cases} \quad (58)$$

On démontre que sous certaines conditions de régularité, ⁽²⁵⁾ les équations différentielles (57) et (58) ont des solutions $x(t)$ et $u(t)$ à partir des données initiales x^0

point d'approximation admissible de départ, et $u^0 \geq 0$ paramètre de Lagrange donné en ce point. Ces solutions ont des limites pour $t \rightarrow \infty$, et on voit que celles-ci ne sont autres que \hat{x} et \hat{u} du théorème de Kuhn Tucker § 2.2, formules (4) à (10) : \hat{x} solution du problème et \hat{u} valeur du multiplicateur de Lagrange.

Il s'agit là d'une méthode différentielle et l'intégration du système (57,58) peut se faire par exemple par la méthode polygonale d'Euler.

On peut remarquer que la trajectoire $x(t)$ n'est pas nécessairement entièrement contenue dans le domaine admissible.

3.4.3.2. - Méthodes de gradient à pas fini

Les méthodes de gradient utilisées en pratique sont surtout les méthodes "à pas long". Par celles-ci on cherche à s'approcher de l'optimum en suivant, non plus une trajectoire plus ou moins bien approchée par de petits segments rectilignes dans l'intégration numérique, mais au contraire une courbe linéarisée formée, par nature, de segments rectilignes.

Nous parlerons ici, pour plus de simplicité, uniquement du cas où les contraintes sont linéaires, mais il faut signaler que Zoutendijk (²⁶) a inclus dans sa méthode qui semble la plus élaborée de ce type, le cas des contraintes convexes non linéaires.

Nous nous bornerons à énoncer les principes généraux sans entrer dans les détails des diverses méthodes (Rosen (²⁷), Frish (²⁸), Zoutendijk).

$$\left. \begin{array}{l} \text{Soit le problème :} \\ \text{minimiser } F(x) \text{ convexe} \\ \text{sous les contraintes} \\ \\ Ax \leq a_0 \end{array} \right\} \quad (59)$$

On admet que $F(x)$ possède un gradient continu sur le domaine admissible.

On part d'un point d'essai admissible : point d'itération initial x^0 .

Pour passer d'un point d'itération x^k au point d'itération suivant x^{k+1} qui est à définir, on chemine suivant un segment rectiligne dont la direction est à déterminer.

Pour cela on détermine en x^k les rayons vecteurs s^k issus de x^k tels que :
 au voisinage de x^k les points situés sur ces rayons vecteurs ($x^k + \lambda s^k$, $\lambda > 0$ suffisamment petit) appartiennent au domaine admissible ;

la condition nécessaire suffisante est

$$\left. \begin{aligned} a_j^T \cdot s^k &\leq 0 && \text{pour tous les } j \text{ tels que} \\ a_j^T \cdot x^k &= a_{j_0} \end{aligned} \right\} \quad (60)$$

ce sont les directions admissibles.

Et parmi celles-ci, les directions le long desquelles en x^k , F est une fonction décroissante de λ

condition : $s^k T \text{ grad } F < 0$;

ce sont les directions utilisables.

On passe de x^k à x^{k+1} en suivant une direction utilisable et les diverses méthodes ne diffèrent que par la façon de choisir cette direction dans l'ensemble des directions utilisables en x^k .

Un fois cette direction choisie, la détermination du point d'itération x^{k+1} suivant se fait toujours de la même manière :

on chemine sur la direction $x^k + \lambda s^k$, tant que F décroît et que l'on reste à l'intérieur du domaine admissible ; autrement dit

si λ' est la valeur pour laquelle les points du rayon vecteur quittent le domaine admissible,

et λ'' la valeur pour laquelle F atteint son minimum sur ce rayon vecteur, déterminée par

$$s^k T \text{ grad } F(x^k + \lambda'' s^k) = 0 ,$$

on prend

$$\lambda^k = \min \{ \lambda', \lambda'' \}$$

et $x^{k+1} = x^k + \lambda^k s^k$.

Si λ^k est infini, cela veut dire que F n'est pas borné.

Il est clair que si au point x^k on a $\text{grad } F = 0$, il n'y a pas de directions utilisables et cela veut dire que x^k est le minimum.

Ces méthodes nécessitent en principe un nombre infini d'itérations, mais dans le cas quadratique, en imposant une condition supplémentaire de conjugaison, on obtient un algorithme fini.

BIBLIOGRAPHIE

- (¹) G.B. Dantzig Programming in a linear structure
Washington D.C. 1948.
- (²) H.P. Künzi, H.G. Tzschach Numerical methods of Mathematical
C.A. Zehnder optimization. Academic Press,
New York 1968
- (³) A. Kaufmann Méthodes et modèles de la recherche
opérationnelle. Dunod. Paris, 1962.
- (⁴) A. Charnes Optimality and degeneracy in linear
programming. *Econometrica* 20 (1952)
pp. 160-170.
- (⁵) H.P. Künzi et Lineare Programmierungen. Zürich 1959.
W. Krelle.
- (⁶) G.B. Dantzig, A. Orden, The generalized simplex method for
et Ph. Wolfe minimizing a linear form under linear
inequalities restraints. The RAND
Corporation R.M. 1264 (1954), *Pacific
J. Math* 5 (1955).
- (⁷) G.B. Dantzig et Decomposition principle for linear
Ph. Wolfe programs. *Operations Res.* 8 (1960)
pp. 101-111.
- (⁸) R.E. Gomory Essentials of an algorithm for inte -
ger solutions to linear programs.
Bul. Amer. Math Soc. 64 (1958)
- (⁹) A.W. Tucker Dual systems of homogeneous linear
relations in Kuhn et Tucker, *Linear
inequalities and related systems*,
Princeton (1956).
- (¹⁰) G.B. Dantzig Linear Programming and extensions,
Princeton University Press, N.J., 1963.

- (¹¹) C.E. Lemke The dual method of solving the linear programming problem - Naval res. logist. quart. 1, 1954 pp. 36-47.
- (¹²) H.W. Kuhn et
A.W. Tucker Non linear programming in Proc. 2nd Berkeley Symp. on Math. Stat and Prob. Berkeley Cal. University Press, 1950.
- (¹³) H.P. Künzi et
W. Krelle La programmation non linéaire,, Gauthier Villars, Paris 1969.
- (¹⁴) M. Slater Lagrange multiplier revisited : A contribution to non-linear programming. Santa Monica, Cal. 1951.
- (¹⁵) O.L. Mangasarian Duality in non linear programming. Quart. Appl. Math. 20, 3, 1962, pp. 300-302. et 21, 3, 1963, p. 252.
- (¹⁶) P. Wolfe A duality theorem for non linear programming. Quart. Appl. Math. 19 1961. pp. 239-244.
- (¹⁷) M.A. Hanson A duality theorem in non linear programming with non linear constraints, Austral. J. Statist, 3, 1961, pp. 64-72.
- (¹⁸) E. Barankin et
R. Dorfman Toward quadratic programming. Los Angeles : Univ. of Calif., Management Sc. Res. Project Res. Rep. 42, 1955.
- (¹⁹) W.S. Dorn Duality in quadratic programming, Quart. Appl. Math. 18, 2, 155-162.
- (²⁰) D.A. D'Esopo A convex programming procedure. Nav. Res. Log. Qu. 6, 1959.
- (²¹) C.G. Hildreth Point estimates of ordinates of concave functions, JASA 49, 1954.
- (²²) C.G. Hildreth A quadratic programming procedure; Nav. Res. Log. Qu. 4, 1957.
- (²³) E.M.L. Beale On minimizing a convex function subject to linear inequalities, J. Roy. Stat. Soc. 17B, 1955.
- (²⁴) E.M.L. Beale On quadratic programming, Nav. Res. Log. Qu. 6, 1959.

- (²³) E.M.L. Beale On minimizing a convex function
subject to linear inequalities,
J. Roy. Stat. Soc. 17 B , 1955
- (²⁴) E.M.L. Beale On quadratic programming, Nav;
Res. Log. Qu. 6 1959.
- (²⁵) K.J. Arrow Studies in linear and non linear
L. Hurwits programming Stanford University
H. Uzawa Press, 1958.
- (²⁶) G. Zoutendijk Methods of feasible directions.
A study in linear and non linear
programming. Elsevier publishing
company, 1960.
- (²⁷) J. B. Rosen The gradient projection method
for non linear programming -
SIAM , 8, 1960 et 3, 1961.
- (²⁸) R.A.K. Frisch The multiplex method for
linear and quadratic programming
Mem. Univers. Social. Fustitute
of Oslo 1957.