

M. DESPLAS

Matrice pseudo-inverse de Moore-Penrose et variables duales généralisées en programmation mathématique

RAIRO. Recherche opérationnelle, tome 26, n° 4 (1992), p. 313-360

http://www.numdam.org/item?id=RO_1992__26_4_313_0

© AFCET, 1992, 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/>

MATRICE PSEUDO-INVERSE DE MOORE-PENROSE ET VARIABLES DUALES GÉNÉRALISÉES EN PROGRAMMATION MATHÉMATIQUE (*)

par M. DESPLAS ⁽¹⁾

Communiqué par J.-Y. JAFFRAY

Résumé. — *Cet article étend la définition et l'utilisation des variables duales ainsi que l'analyse de sensibilité classique des paramètres d'un programme mathématique, linéaire ou non linéaire, en n'importe quel point du domaine des points réalisables. Cette extension revêt une certaine importance car elle donne les mêmes informations que celles dont on dispose habituellement soit aux points extrêmes du domaine réalisable soit au point solution optimale. On donne les preuves des fondements théoriques de l'approche proposée ainsi que des exemples numériques qui confirment son applicabilité comme outil d'analyse.*

Mots clés : Matrice pseudo-inverse de Moore-Penrose; programmation mathématique; dualité; analyse de sensibilité; théorème de l'enveloppe; tableaux du simplexe.

Abstract. — *The Moore-Penrose inverse and generalized dual variables in mathematical programming. This paper extends the definition and usage of dual variables as well as the sensitivity of the parameters of a linear or a non linear mathematical program at any point within the feasible set. This extension is of major importance as it provides information beyond that available from either the extreme points of the feasible set or the optimal solution point. The theoretical basis of the proposed approach is proven and numerical examples are provided illustrating its applicability as an analytical tool.*

Keywords : Moore-Penrose pseudo-inverse matrix; mathematical programming; duality; sensitivity analysis; envelop theorem; simplex tableaux.

(*) Reçu juin 1991, accepté en mars 1992.

⁽¹⁾ Professeur à l'Université de Paris-2, Directeur du Département d'Informatique, et Directeur de la Recherche à l'Institut de Management Public (IDMP), Paris.

INTRODUCTION

Les valeurs des variables duales et l'analyse de sensibilité sont très importantes dans les modélisations de problèmes sous forme de programmes linéaires ou non linéaires. Cependant leurs définitions ne sont actuellement précises qu'aux sommets du domaine des points réalisables. Cet article présente une approche débarrassée de cette restriction par une extension de la définition des variables duales et de l'analyse de sensibilité en n'importe quel point sur la frontière ou à l'intérieur du domaine réalisable. Outre son intérêt théorique, l'approche proposée procure d'incontestables avantages si l'on se réfère au concept de « rationalité limitée » de H. Simon où les firmes n'optimisent pas nécessairement leur comportement. Pouvant alors cependant calculer les valeurs des variables duales et effectuer une analyse de sensibilité en une solution non optimale (c'est-à-dire dans ou sur la frontière du domaine réalisable) elles peuvent cependant disposer d'informations utiles pour la prise de décision.

L'organisation de cet article est la suivante. Après avoir rappelé dans la section 1 des notions classiques dans la littérature, on présentera les définitions nécessaires et on comparera l'approche proposée à l'approche classique de la programmation linéaire ou non linéaire. Les variables duales λ_i et λ'_j rappelées dans la section 1 sont à nouveau définies de façon compatible dans la section 2 sans avoir à restreindre leur définition aux points extrêmes du domaine réalisable et leur signification économique est explicitée en termes de l'approche proposée. La section 3 présente une définition alternative des mêmes variables duales. De plus, on généralisera le théorème de l'enveloppe dans la section 4, où l'on démontrera aussi quelques théorèmes soutenant la validité théorique de l'approche proposée. La section 5 généralise les résultats à l'analyse de sensibilité sur les paramètres de formulations linéaires et non linéaires tandis que la section 6 donne plusieurs exemples numériques illustrant l'applicabilité et l'utilité pratique de l'approche proposée. Enfin, une dernière section figure en guise de conclusion.

L'objet de cet article est de définir les valeurs des variables duales en programmation mathématique (*valeurs* des « ressources » et *coûts d'opportunité* des « activités », qui sont, à l'optimum d'un programme linéaire par exemple, les variables duales et les variables d'écart du programme dual) en *n'importe quel point réalisable* et pas nécessairement au point solution optimale, ce qui est un résultat classique en programmation linéaire ou non linéaire. L'interprétation économique des valeurs des variables duales à l'optimum d'un programme linéaire ou non linéaire (coefficients de Kuhn-Tucker) sera élargie de façon à inclure les cas où la solution « concrète » d'un problème

réel (que nous appellerons par la suite «solution de *statu quo*») ne peut être prouvée comme étant optimale.

Quel est l'effet sur la fonction objectif d'une variation de ses coefficients ou des coefficients des contraintes et comment peut-on réagir à une telle variation quand la situation actuelle n'est pas optimale? Soit

$$\text{Max} \{f(x) \mid x_j \geq 0, j=1, 2, \dots, n, g_i(x) \leq b_i, i=1, 2, \dots, m\}$$

la formulation d'un programme non linéaire. Que peut-on faire si b_i devient $b_i + \delta b_i$, $i=1, 2, \dots, m$? Si x est la solution de *statu quo*, quel sera l'ajustement δx et quel effet cet ajustement aura-t-il sur la fonction objectif f ? Si l'on dénote cette variation de la fonction f par

$$\delta f = \text{grad } f \cdot \delta x$$

nous montrerons que

$$\delta f = \lambda \cdot \delta b$$

par une définition appropriée du vecteur λ de composantes λ_i . Cette définition de λ nous permet donc de prévoir l'effet sur la fonction objectif d'une variation dans les seconds membres des contraintes, *sans avoir à évaluer le nouveau point $x + \delta x$ d'ajustement* aux variations des contraintes, mais en n'envisageant que l'effet de δb sur f .

Ceci peut permettre d'évaluer l'impact d'une politique économique par exemple consistant à modifier des contraintes (restrictions du crédit, incitations à des économies d'énergie, température portée de 20°C à 19°C dans les lieux publics, bureaux ou locaux privés, etc.) sur une fonction d'objectif collectif comme une fonction de bien-être social, sans se préoccuper de façon numérique de l'ajustement δx que la collectivité apportera à son comportement initial x , *qui n'a pas beaucoup de raisons d'être optimal*.

La détermination des valeurs de λ n'est *unique* que si l'on fait une hypothèse sur le comportement de la collectivité, à savoir sur le passage de x à $x + \delta x$. En fait, à chaque déplacement δx est associée une définition précise de λ mais on préconisera la détermination de λ associée au déplacement δx de *norme minimale*: on peut penser en effet que la situation de *statu quo* x n'étant généralement pas optimale, le décideur (en l'occurrence la collectivité) considère que la fonction objectif $f(x)$ est en fait secondaire et que sa réaction à une modification des contraintes s'effectuera en modifiant le moins possible son comportement antérieur, donc en minimisant $\|\delta x\|$. Cela nous conduira à une définition *unique* des composantes du vecteur λ et de celles d'un autre

vecteur noté λ' : nous montrerons que les variables λ_i sont de véritables *variables duales* et que les variables λ'_j sont de véritables *coûts d'opportunité* dans le sens où les premières ont l'interprétation économique classique des multiplicateurs de Lagrange, comme valeurs des ressources, et dans le sens où les secondes sont une mesure du coût de l'introduction ou de la variation des activités. De plus les unes et les autres satisfont les relations d'exclusion classiques vérifiées par les coefficients de Kuhn-Tucker définis à l'optimum ou en un point extrême du domaine des réalisables.

L'application au cas de la firme peut s'effectuer exactement dans les mêmes termes si l'on admet l'hypothèse de *rationalité limitée* de H. Simon selon laquelle par exemple les entreprises sont moins attachées à rechercher, à chaque période, l'optimum d'une fonction objectif unique qu'à s'écarter le moins possible de certains objectifs fixés, généralement proches de ceux de la période qui précède, de sorte que l'ajustement consistera à effectuer des changements δx de leurs activités les plus faibles possible. Suivant que la firme aura alors pour but de varier le moins possible ses activités pour s'adapter aux nouvelles contraintes qui lui sont imposées (ce que nous supposons dans ce qui suit) ou qu'elle cherchera à minimiser l'écart entre ses objectifs et ce qu'elle peut réaliser (ce que l'on fait en Goal Programming) elle sera confrontée aux fonctions objectifs suivantes : minimiser $\|\delta x\|$ dans le premier cas, minimiser $\|Ax - b\|$ dans le dernier cas, lorsque les contraintes sont linéaires du type $g_i(x) = A_i x \leq b_i$.

Dans les deux cas interviennent les matrices inverses généralisées de Moore-Penrose (cf. Campbell & Meyer [4] et Ijiri [6]) et il est donc assez naturel que la définition, donnée dans cet article, de la dualité en dehors de l'optimum soit associée aux inverses généralisées de certaines matrices que l'on précisera, et en particulier à l'inverse généralisée de Moore-Penrose. Ce choix implique que la définition des « variables duales généralisées » (λ_N, λ'_N) par les formules suivantes sera *unique* (l'indice N sera défini plus loin) :

$$(\lambda_N, \lambda'_N) = \delta \cdot M^+ = \delta \cdot M^T (MM^T)^{-1}$$

où δ est le gradient de la fonction objectif $f(x)$, M la matrice des gradients des contraintes saturées au point considéré de statu quo et M^T sa transposée.

Les développements qui suivent ont été largement inspirés par les publications de Rosen et de Strang (réf. [10] et [11] respectivement). Dans la méthode du gradient projeté, Rosen évoque le caractère de variables duales de certaines expressions qu'il obtient dans l'application de sa méthode. Dans son livre, Strang identifie des variables duales dans une étape intermédiaire de la méthode de Karmarkar de résolution d'un programme linéaire. Ces deux

références ne concernent que des contraintes linéaires. Ce papier a pour objectif d'étendre ces deux remarques au cas de la programmation non linéaire, et de façon plus générale de les appliquer à des situations qui ne correspondent pas automatiquement à des points extrêmes du domaine réalisable ou à des points où les coefficients de Kuhn-Tucker sont bien définis.

1. NOTATIONS ET RÉSULTATS CLASSIQUES

Considérons un programme non linéaire sous forme standard, comportant n variables $x_j \geq 0$ et m contraintes inégalités $g_i(x) \leq b_i$. Ce programme peut s'écrire

$$\text{Max} \{ f(x) \mid g(x) \leq b, x \geq 0 \}$$

On a séparé contraintes de structure $g_i(x) \leq b_i$ et contraintes de signe $x_j \geq 0$ pour interpréter séparément les variables duales λ_i associées aux premières et λ'_j associées aux dernières. Après introduction des variables d'écart x'_i ($i = 1, 2, \dots, m$), regroupées en un vecteur x' , le programme s'écrit :

$$\text{Max} \{ f(x) \mid g(x) + x' = b, x \geq 0, x' \geq 0 \}$$

Les contraintes de structure sont alors représentées par $x'_i \geq 0$ exactement sous la même forme que les contraintes de signe $x_j \geq 0$: les variables duales associées aux premières sont les multiplicateurs λ_i et les variables duales associées aux dernières sont les multiplicateurs λ'_j de sorte que les conditions d'optimalité de Kuhn-Tucker au point optimal $(x^*, x'^*, \lambda^*, \lambda'^*)$ s'écrivent :

$\lambda_i^* \geq 0, x'_i{}^* \geq 0, \lambda_i^* \cdot x'_i{}^* = 0, \lambda'_j{}^* \geq 0, x_j^* \geq 0, \lambda'_j{}^* \cdot x_j^* = 0$	(1)
--------------------------------------------------------------------------------------------------------------------------------------------	-----

Expression géométrique des conditions de Kuhn et Tucker :

Si l'on intègre les contraintes de signe $x \geq 0$ sous la forme des contraintes $g(x) \leq b$ en les écrivant $h_j(x) \leq 0$ [en posant $h_j(x) = -x_j, j = 1, 2, \dots, n$] alors les conditions d'optimalité de Kuhn et Tucker expriment que le gradient de l'objectif appartient au cône engendré par les gradients des contraintes saturées :

$$\frac{\partial f^*}{\partial x} = \sum_{i=1}^m \lambda_i^* \frac{\partial g_i}{\partial x} + \sum_{j=1}^n \lambda'_j{}^* \frac{\partial h_j}{\partial x} = \lambda^* \cdot \frac{\partial g(x^*)}{\partial x} - \lambda'^* \tag{2}$$

avec $\lambda_i^* > 0$ et $\lambda'_j{}^* > 0$ pour les contraintes saturées et $\lambda_i^* = 0$ et $\lambda'_j{}^* = 0$ pour les contraintes non saturées. La relation (2) ne fait que traduire la définition

des multiplicateurs λ' intervenant dans les conditions (1) de Kuhn-Tucker par :

$$\lambda' = - \frac{\partial L}{\partial x}, \quad \text{où } L(x, \lambda) = f(x) - \lambda \cdot [g(x) - b]$$

La définition donnée ci-dessous des variables (λ_i, λ'_j) en un point quelconque permettra encore d'écrire une formule comparable, mais certains des (λ_i, λ'_j) pourront être négatifs car on n'est pas à l'optimum du programme :

$$\frac{\partial f(x)}{\partial x} = \sum_{i=1}^m \lambda_i \frac{\partial g_i}{\partial x} + \sum_{j=1}^n \lambda'_j \frac{\partial h_j}{\partial x} = \lambda \cdot \frac{\partial g(x)}{\partial x} - \lambda' \quad (3)$$

Théorème de l'enveloppe :

Nous rappelons le « théorème de l'enveloppe » sous sa forme classique [12] pour pouvoir le comparer à la généralisation que nous en donnons ci-dessous. Considérons le programme paramétrique suivant :

$$\text{Max } \{ f(x, \alpha) \mid g_i(x, \alpha) \geq 0, i = 1, 2, \dots, m, x \geq 0 \}$$

Soit $x^*(\alpha)$ la solution optimale et $\lambda^*(\alpha)$ le vecteur des coefficients de Kuhn-Tucker des contraintes associées aux valeurs α des paramètres. On notera :

$F(\alpha) = f[x^*(\alpha), \alpha]$ la valeur optimale de la fonction objectif f

$\Phi(x, \lambda, \alpha) = f(x, \alpha) - \lambda \cdot g(x, \alpha)$ la fonction de Lagrange de ce problème

$\Psi(\alpha) = f[x^*(\alpha), \alpha] - \lambda^*(\alpha) \cdot g[x^*(\alpha), \alpha] = \Phi[x^*(\alpha), \lambda^*(\alpha), \alpha]$.

THÉORÈME DE L'ENVELOPPE : *Supposons que les fonctions F et Ψ sont continûment différentiables. Nous avons alors :*

$$\frac{\partial F(\alpha)}{\partial \alpha_k} = \frac{\partial \Psi(\alpha)}{\partial \alpha_k} = \frac{\partial \Phi(x, \lambda, \alpha)}{\partial \alpha_k} \quad (k = 1, 2, \dots, l)$$

ou

$$F_\alpha = \Psi_\alpha = \Phi_\alpha$$

avec

$$\frac{\partial \Phi}{\partial \alpha_k} = \frac{\partial f(x, \alpha)}{\partial \alpha_k} - \lambda \cdot \frac{\partial g(x, \alpha)}{\partial \alpha_k}$$

Le théorème de l'enveloppe étudie la sensibilité de la fonction objectif aux variations des paramètres α , en supposant que l'on calcule l'optimum $x^*(\alpha)$ du programme à chaque étape. Quand α devient $\alpha + \delta\alpha$ l'optimum $x^*(\alpha)$

devient $x^*(\alpha + \delta\alpha)$ et le théorème prend en compte de tels ajustements. Si on suppose que nous ne sommes pas au point optimum pour α , pas plus que pour $\alpha + \delta\alpha$, il reste à préciser le passage δx d'un point de statu quo $x(\alpha)$ à un autre point de statu quo $x(\alpha + \delta\alpha) = x(\alpha) + \delta x$. Nous verrons qu'il est possible de généraliser le théorème de l'enveloppe ailleurs qu'au point optimum quand le changement δx est *orthogonal à l'intersection des contraintes saturées*.

Avant de définir les variables duales λ_i et λ'_j en n'importe quel point réalisable, nous rappelons les définitions des coefficients de Kuhn-Tucker (et les conditions d'optimalité) sous une forme qui nous permettra de définir de façon simple la dualité ailleurs qu'à l'optimum. Pour simplifier, nous admettrons la forme la plus forte de condition de qualification des contraintes, c'est-à-dire la condition que les gradients des contraintes saturées à l'optimum soient linéairement indépendants (ce qui implique qu'il existe une matrice Jacobienne de plein rang en lignes) — cette hypothèse est souvent admise dans les publications de recherche sur l'analyse de sensibilité en programmation non linéaire (cf. [5]).

2. DÉFINITION DES VARIABLES DUALES λ_i ET λ'_j EN N'IMPORTE QUEL POINT RÉALISABLE ET CONDITIONS D'OPTIMALITÉ DE KUHN-TUCKER

Supposons qu'au point de statu quo (x, x') soient saturées M contraintes de structure ($M \leq m$) et N contraintes de signe ($N \leq n$). Nous avons alors :

$$\left. \begin{array}{ll} x'_i = 0 & \text{pour } i = 1, 2, \dots, M \\ x'_i > 0 & \text{pour } i = M + 1, M + 2, \dots, m \\ x_j = 0 & \text{pour } j = 1, 2, \dots, N \\ x_j > 0 & \text{pour } j = N + 1, N + 2, \dots, n \end{array} \right\} \quad (4)$$

On regroupe les variables nulles x'_i en un vecteur x'_N de format $(M, 1)$, les variables non nulles x'_i en un vecteur x'_B de format $(m - M, 1)$, les variables nulles x_j en un vecteur x_N de format $(N, 1)$ et les variables non nulles x_j en un vecteur x_B de format $(n - N, 1)$.

En séparant les contraintes saturées et non saturées au point de statu quo (x, x') considéré, les contraintes du programme s'expriment alors ainsi en ce

point :

$$\left. \begin{aligned} g_S(x_B, x_N) + x'_N &= b_S \\ g_{NS}(x_B, x_N) + x'_B &= b_{NS} \\ x_B > 0, \quad x_N &= 0, \quad x'_B > 0, \quad x'_N = 0 \end{aligned} \right\} \quad (5)$$

(Les indices S et NS traduisent respectivement la saturation et la non-saturation des contraintes au point de statu quo).

Les contraintes non-saturées ne jouent aucun rôle et ne seront pas violées par un petit déplacement $(\delta x, \delta x')$ au voisinage du point (x, x') . Comme $x_N = 0$ et $x'_N = 0$ seules seront admissibles des perturbations δx_N et $\delta x'_N$ *non-négatives* de sorte que les N variables x_N et les M variables x'_N doivent être des *variables libres* (ou *variables hors base*) sinon elles risqueraient d'être rendues négatives par le jeu des M contraintes saturées pour certains choix des autres variables (x_B, x'_B) . Comme seules les $(M + N)$ contraintes saturées sont restrictives il devrait y avoir exactement M variables x_B and $m - M$ variables x'_B *basiques* pour pouvoir appliquer le théorème des fonctions implicites. Ce n'est le cas que si le point de statu quo est un sommet, mais ce n'est généralement pas le cas si ce point est sur une face ou sur une arête (c'est-à-dire ne sature pas un nombre suffisant de contraintes), de sorte que nous considérerons les $m + n - (M + N)$ variables non nulles (x_B, x'_B) comme *basiques* et les $M + N$ variables nulles (x_N, x'_N) comme *hors base*.

(a) Cas classique

Si $m = m + n - (M + N)$, i. e. $n = M + N$, le théorème des fonctions implicites permet alors de résoudre le système :

$$g_S(x_B, x_N) + x'_N = b_S$$

car on pourrait alors trouver une matrice Jacobienne J inversible (lorsque le point considéré n'est pas un point dégénéré). Ce sera le cas si le point de statu quo est un sommet du domaine réalisable et d'une façon générale lorsque les conditions d'application du théorème des fonctions implicites sont remplies, avec les définitions classiques des variables basiques ou hors base. On peut alors exprimer implicitement les variables basiques x_B comme fonctions des variables hors base de la façon suivante :

$$x_B = h(x_N, x'_N)$$

Ceci revient en effet à résoudre le système $\delta g_S + \delta x'_N = 0$ ou, de façon équivalente, le système linéaire :

$$J \cdot \delta x_B + C \cdot \delta x_N + \delta x'_N = 0 \tag{6}$$

en posant :

$$J = \frac{\partial g_S}{\partial x_B} \quad \text{et} \quad C = \frac{\partial g_S}{\partial x_N} \tag{7}$$

L'objectif s'écrit alors :

$$f(x) = f(x_B, x_N) = f[h(x_N, x'_N), x_N] = F(x_N, x'_N)$$

Dans tout ce qui suit, on simplifiera les notations en posant :

$$\delta_B = \frac{\partial f}{\partial x_B} \quad \delta_N = \frac{\partial f}{\partial x_N}$$

La solution unique de l'équation (6) est :

$$\delta x_B = -J^{-1} C \delta x_N - J^{-1} \delta x'_N \tag{8}$$

On posera :

$$\frac{\delta f}{\delta x_N} = \frac{\partial F}{\partial x_N} = \frac{\partial f}{\partial x_N} + \frac{\partial f}{\partial x_B} \cdot \frac{\partial h}{\partial x_N} = \delta_N + \delta_B \cdot \frac{\delta x_B}{\delta x_N} \tag{9}$$

$$\frac{\delta f}{\delta x'_N} = \frac{\partial F}{\partial x'_N} = \frac{\partial f}{\partial x'_N} + \frac{\partial f}{\partial x_B} \cdot \frac{\partial h}{\partial x'_N} = \frac{\partial f}{\partial x'_N} + \delta_B \cdot \frac{\delta x_B}{\delta x'_N} \tag{10}$$

Les symboles $\delta f/\delta x_N$, $\delta f/\delta x'_N$, $\delta x_B/\delta x_N$, $\delta x_B/\delta x'_N$ représentent les ajustements de f et des variables x_B aux variations des variables hors base x_N et x'_N impliqués par les relations $g_S(x_B, x_N) + x'_N = b_S$. Les ajustements $\delta x_B/\delta x_N$, $\delta x_B/\delta x'_N$ proviennent de l'équation (6) ou aussi bien de l'équation (8) d'où leurs valeurs :

$$\frac{\delta x_B}{\delta x_N} = -J^{-1} C \quad \frac{\delta x_B}{\delta x'_N} = -J^{-1} \tag{11}$$

On déduit les valeurs de $\delta f/\delta x_N$ et de $\delta f/\delta x'_N$ qui s'écrivent alors :

$$\frac{\delta f}{\delta x_N} = \delta_N - \delta_B J^{-1} C \quad \frac{\delta f}{\delta x'_N} = -\delta_B J^{-1} \tag{12}$$

On reconnaît là respectivement les multiplicateurs λ'_N et λ_N de Kuhn-Tucker qui s'écrivent :

$$\lambda'_N = -\frac{\delta f}{\delta x_N} = \delta_B J^{-1} C - \delta_N, \quad \text{multiplicateurs des contraintes } x_N \geq 0 \quad (13)$$

$$\lambda_N = -\frac{\delta f}{\delta x'_N} = \delta_B J^{-1}, \quad \text{multiplicateurs des contraintes } x'_N \geq 0 \quad (14)$$

auxquels on ajoute les multiplicateurs $\lambda'_B = -\delta f/\delta x_B$ des contraintes $x_B \geq 0$ et $\lambda_B = -\delta f/\delta x'_B$ des contraintes $x'_B \geq 0$. Les conditions d'optimalité expriment que l'on ne peut améliorer (ici augmenter) l'objectif f par action sur les variables et donc que :

$$\lambda'_N \geq 0, \quad \lambda_N \geq 0, \quad \lambda'_B \geq 0, \quad \lambda_B \geq 0.$$

Les relations d'exclusion à l'optimum expriment qu'en outre on doit avoir $\lambda_{B_i} = 0$ et $\lambda'_{B_j} = 0$ car $x'_{B_i} > 0$ et $x_{B_j} > 0$ par définition. Les conditions d'optimalité s'expriment donc par les relations (1).

(b) Généralisation

Mais généralement on a trop de variables *basiques* (x_B, x'_B) au point de statu quo considéré de sorte que $n > M + N$: nous avons *a priori* n degrés de liberté dans le choix de x , et $M + N$ contraintes sur les variables *hors base* ; il y a donc $n - (M + N) > 0$ degrés de liberté et le système (6) possède plusieurs solutions car la matrice Jacobienne J n'est plus inversible. On supposera seulement que la matrice J est de plein rang en lignes, *i.e.* que les contraintes saturées sont linéairement indépendantes (condition de qualification des contraintes au point de statu quo). J possède une inverse à droite de Moore-Penrose J^+ car JJ^T est inversible :

$$J^+ = J^T (JJ^T)^{-1}$$

La solution δx_B de l'équation (6) est alors :

$$\delta x_B = -J^+ C \delta x_N - J^+ \delta x'_N \quad (15)$$

δx_B est donc l'unique *solution des moindres carrés de longueur minimale* et on en déduit les valeurs de $\delta x'$ par

$$\delta x'_B = -U \delta x_B - V \delta x_N \quad (16)$$

où

$$U = \frac{\partial g_{NS}}{\partial x_B} \quad \text{et} \quad V = \frac{\partial g_{NS}}{\partial x_N} \tag{17}$$

Pour ce choix de J^+ remplaçant J^{-1} dans l'application du théorème des fonctions implicites, x_B s'exprime alors *théoriquement* en fonction des variables *hors base* sous la forme suivante :

$$x_B = h^+ (x_N, x'_N)$$

dont le report dans l'objectif donne l'expression de la fonction objectif en fonction des seules variables *hors base* sous la forme :

$$f(x) = f(x_B, x_N) = f[h^+ (x_N, x'_N), x_N] = F(x_N, x'_N)$$

On posera :

$$\frac{\delta f}{\delta x_N} = \frac{\partial F}{\partial x_N} = \frac{\partial f}{\partial x_N} + \frac{\partial f}{\partial x_B} \cdot \frac{\partial h^+}{\partial x_N} = \frac{\partial f}{\partial x_N} + \frac{\partial f}{\partial x_B} \cdot \frac{\delta x_B}{\delta x_N} \tag{18}$$

$$\frac{\delta f}{\delta x'_N} = \frac{\partial F}{\partial x'_N} = \frac{\partial f}{\partial x'_N} + \frac{\partial f}{\partial x_B} \cdot \frac{\partial h^+}{\partial x'_N} = \frac{\partial f}{\partial x'_N} + \frac{\partial f}{\partial x_B} \cdot \frac{\delta x_B}{\delta x'_N} \tag{19}$$

Ici encore les symboles $\delta f/\delta x_N$, $\delta f/\delta x'_N$, $\delta x_B/\delta x_N$, $\delta x_B/\delta x'_N$ représentent les ajustements de f et des variables x_B aux variations des variables *hors base* x_N et x'_N impliqués par les relations $g_S(x_B, x_N) + x'_N = b_S$. (On déduira plus loin, en (51), les ajustements $\delta x'_B/\delta x_N$, $\delta x'_B/\delta x'_N$ concernant les contraintes non saturées $g_{NS}(x_B, x_N) + x'_B = b_{NS}$). Ces ajustements se déduisent de l'identité

$$g_S[h^+ (x_N, x'_N), x_N] + x'_N = b_S$$

qui implique :

$$0 = \frac{\delta g_S}{\delta x_N} = J \cdot \frac{\delta x_B}{\delta x_N} + C \tag{20}$$

$$0 = \frac{\delta g_S}{\delta x'_N} + I = J \cdot \frac{\delta x_B}{\delta x'_N} + I \tag{21}$$

Il s'ensuit les relations suivantes :

$$\frac{\delta x_B}{\delta x_N} = -J^+ C \quad \frac{\delta x_B}{\delta x'_N} = -J^+ \quad (22)$$

$$\frac{\delta f}{\delta x_N} = \frac{\partial f}{\partial x_N} - \frac{\partial f}{\partial x_B} \cdot J^+ C \quad \frac{\partial f}{\delta x'_N} = - \frac{\partial f}{\partial x_B} \cdot J^+ \quad (23)$$

Variations de $f(x)$ associées aux variations des variables x_{Bj} , x'_{Bi} , x_{Nk} , x'_{Ni} :

La condition (6) : $J \cdot \delta x_B + C \cdot \delta x_N + \delta x'_N = 0$ doit être vérifiée et on a :

$$\delta f = \delta_B \delta x_B + \delta_N \delta x_N$$

Dans les exemples qui suivent, les variations δf sont des variations *compensées* de façon à maintenir satisfaites les contraintes saturées lorsque les variables x_{Bj} , x'_{Bi} , x_{Nk} , x'_{Ni} varient *une à une* mais en tenant compte des variations des autres variables impliquées par le maintien des contraintes saturées.

— Quand x_{Bj} varie seule de δx_{Bj} la variation de x_B est alors $\delta x_B = (0, 0, \dots, \delta x_{Bj}, \dots, 0)^T$: x'_B , x'_N et x_N restent fixées et les contraintes de structure doivent être respectées. L'adaptation consiste en un *déplacement sur l'intersection des contraintes saturées*. Il s'ensuit une variation « compensée » δx_B solution de l'équation $J \cdot \delta x_B = 0$:

$$\delta x_B = (I - J^+ J) \delta x_B, \quad \delta f = \delta_B \delta x_B = \delta_B (I - J^+ J) \delta x_B$$

et

$$\frac{\delta f}{\delta x_B} = \delta_B (I - J^+ J)$$

— Quand x'_{Bi} varie seule de $\delta x'_{Bi}$ il n'y a aucune compensation δx_B , δx_N et δf car $\delta x'_B$ n'intervient pas dans (6). Il s'ensuit que :

$$\delta f = 0 \quad \text{et} \quad \frac{\delta f}{\delta x'_B} = 0$$

— Quand x'_{Ni} varie seule de $\delta x'_{Ni}$ la variation de x'_N est alors $\delta x'_N = (0, 0, \dots, \delta x'_{Ni}, \dots, 0)^T$: tout se passe comme si la i -ème contrainte de structure se déplaçait, son second membre passant de b_i à $b_i - \delta x'_{Ni}$. L'adaptation consiste en un *déplacement orthogonal à la nouvelle intersection des contraintes saturées*. Ce déplacement est la solution de longueur minimale de l'équation $J \delta x_B = -\delta x'_N$, i. e. $\delta x_B = -J^+ \delta x'_N$.

Il s'ensuit que :

$$\delta f = \delta_B \cdot \delta x_B = -\delta_B J^+ \delta x'_N \quad \text{et} \quad \frac{\delta f}{\delta x'_N} = -\delta_B J^+$$

— Quand x_{Nk} varie seule de δx_{Nk} la variation de x_N est alors $\delta x_N = (0, 0, \dots, \delta x_{Nk}, \dots, 0)^T$: tout se passe comme si la k -ième contrainte de signe se déplaçait, son second membre passant de 0 à $0 - \delta x_{Nk}$. L'adaptation consiste en un *déplacement orthogonal à la nouvelle intersection des contraintes saturées*. Ce déplacement est la solution de longueur minimale de l'équation $J \delta x_B = -C \delta x_N$ ou $\delta x_B = -J^+ C \delta x_N$.

Il s'ensuit que :

$$\delta f = \delta_B \cdot \delta x_B + \delta_N \cdot \delta x_N = (-\delta_B J^+ C + \delta_N) \delta x_N \quad \text{et} \quad \frac{\delta f}{\delta x_N} = \delta_N - \delta_B J^+ C$$

On pose :

$$\lambda = (\lambda_N, \lambda_B) \quad \lambda' = (\lambda'_B, \lambda'_N)$$

et on définit les variables duales λ et λ' par :

$$\lambda'_B = -\frac{\delta f}{\delta x_B} = \delta_B \cdot (J^+ J - I) \quad (24)$$

$$\lambda'_N = -\frac{\delta f}{\delta x_N} = \delta_B \cdot J^+ C - \delta_N \quad (25)$$

$$\lambda_B = -\frac{\delta f}{\delta x'_B} = 0 \quad (26)$$

$$\lambda_N = -\frac{\delta f}{\delta x'_N} = \delta_B J^+ \quad (27)$$

λ'_B est le multiplicateur des $n - N$ variables *basiques* $x_B > 0$

λ'_N est le multiplicateur des N variables *hors base* $x_N = 0$

λ_B est le multiplicateur des $m - M$ variables *basiques* $x'_B > 0$

λ_N est le multiplicateur des M variables *hors base* $x'_N = 0$.

Remarques: 1. A l'optimum du programme on a un extremum libre de la fonction objectif f -compensée, *i. e.* $\delta f=0$ ou encore :

$$\frac{\delta f}{\delta x_B} = 0 \quad \text{et} \quad \frac{\delta f}{\delta x'_B} = 0, \quad \text{i. e.} \quad \lambda'_B = 0 \quad \text{et} \quad \lambda_B = 0.$$

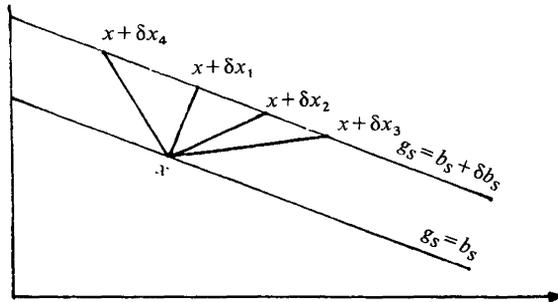
De plus il faut que

$$\frac{\delta f}{\delta x_N} \leq 0 \quad \text{et} \quad \frac{\delta f}{\delta x'_N} \leq 0, \quad \text{i. e.} \quad \lambda'_N \geq 0 \quad \text{et} \quad \lambda_N \geq 0.$$

Les relations d'exclusion des équations (1) sont alors satisfaites et l'on retrouve les conditions d'optimalité de Kuhn-Tucker dans le cas où la condition de qualification des contraintes choisie est la plus restrictive, impliquant l'indépendance linéaire des gradients des contraintes saturées au point optimum.

2. On verra ultérieurement sur les « tableaux généralisés du simplexe » que, lorsque les contraintes saturées sont desserrées de δb_s , et pour un déplacement δx_1 orthogonal à l'intersection des contraintes saturées ainsi déplacées de δb_s , on a cependant la relation globale $\lambda'_B \cdot \delta x_B = 0$ même si l'on n'a pas $\lambda'_{Bj} \cdot \delta x_{Bj} = 0$ pour chaque variable individuelle. Ce déplacement δx_1 est l'ajustement $\delta x_B = J^+ \delta b_s$ et l'on a $\lambda'_B \cdot \delta x_B = 0$ à cause de la relation $J^+ J J^+ = J^+$.

Pour tout autre déplacement $\delta x_2, \delta x_3, \delta x_4, \dots$ sur l'intersection des contraintes saturées déplacées de δb_s on a en général $\lambda'_B \cdot \delta x_B \neq 0$. En effet ces déplacements $\delta x_2, \delta x_3, \delta x_4, \dots$ etc, correspondent aux ajustements $\delta x_B = J^+ \cdot \delta b_s + (I - J^+ J) z_B$ pour diverses valeurs de z_B . On a alors $\lambda'_B \cdot \delta x_B = \delta_B \cdot (J^+ J - I) \cdot z_B \neq 0$ (*cf.* la figure ci-dessous).



3. DÉFINITION ALTERNATIVE DES MÊMES VARIABLES DUALES $(\lambda_B, \lambda_B, \lambda_N, \lambda_N)$.

On peut définir d'une autre façon les mêmes valeurs des variables duales en un point quelconque (de statu quo) en termes de projection du gradient de l'objectif sur l'espace des lignes et sur le noyau de la matrice M des gradients des contraintes saturées au point considéré. Cette matrice M est constituée par la superposition des gradients des contraintes de structure $g_S(x_B, x_N) \leq b_S$ et des contraintes de signe $-x_N \leq 0_S$ saturées au point considéré.

Au point de statu quo (non obligatoirement optimal) les variables satisfont les contraintes suivantes :

$$\begin{aligned} g_S(x_B, x_N) &= b_S && (M \text{ contraintes de structure}) \\ -x_N &= 0_S && (N \text{ contraintes de signe}) \end{aligned}$$

La matrice M des gradients de ces contraintes saturées est alors :

$$M = \begin{array}{cc|cc} & & & & \uparrow \\ & & J & C & M \\ & & \hline & & 0 & -I & \downarrow \\ & & & & \uparrow \\ & & & & N \\ & & & & \downarrow \\ \leftarrow & n-N & \rightarrow & N & \rightarrow \end{array}$$

On en déduit la matrice pseudo-inverse de Moore-Penrose $M^+ = M^T (MM^T)^{-1}$:

$$M^+ = \begin{array}{cc|cc} & & & & \uparrow \\ & & J^+ & J^+ C & n-N \\ & & \hline & & 0 & -I & \downarrow \\ & & & & \uparrow \\ & & & & N \\ & & & & \downarrow \\ \leftarrow & M & \rightarrow & N & \rightarrow \end{array} \tag{29}$$

Dans le cas où x est un sommet non dégénéré, *i.e.* $n = M^+ N$, la condition de qualification des contraintes se ramène à l'existence d'une matrice Jacobienne inversible; alors $J^+ = J^{-1}$ et $M^+ = M^{-1}$. Soit $\text{Im}(M^T)$ l'espace des lignes et $\text{Ker}(M)$ le noyau de la matrice M : le projecteur sur $\text{Im}(M^T)$ est

M^+M et le projecteur sur l'espace orthogonal complémentaire $\text{Ker}(M)$ est $I - M^+M = -R$. On peut vérifier que :

$$M \cdot R = 0, \quad R \cdot M^+ = 0, \quad MM^+ = I \tag{30}$$

Si on suppose que les lignes de J sont linéairement indépendantes, les $M+N$ lignes de M sont aussi linéairement indépendantes. Seules $M+N$ colonnes sur les n colonnes de M sont linéairement indépendantes de sorte que les $n - (M+N)$ autres colonnes sont des combinaisons linéaires des $M+N$ colonnes indépendantes. L'équation $MR=0$ donne n combinaisons linéaires nulles des colonnes de M . La décomposition de R en :

$$R = M^+ M - I = \begin{array}{c} \begin{array}{|cc|} \hline J^+ & J-I \\ \hline 0 & 0 \\ \hline \end{array} \begin{array}{l} \uparrow \\ n-N \\ \downarrow \\ \uparrow \\ N \\ \downarrow \end{array} \end{array} \tag{31}$$

$\leftarrow n-N \rightarrow \leftarrow N \rightarrow$

prend en compte $n - N$ de ces combinaisons linéaires nulles (M d'entre elles sont des répliques ou des combinaisons des $n - (M+N)$ combinaisons linéaires nulles). Si on ajoute à R un certain nombre ($m - M - N$) de colonnes de zéros on obtient la matrice S suivante :

$$S = \begin{array}{c} \begin{array}{|cc|} \hline J^+ & J-I \\ \hline 0 & 0 \\ \hline \end{array} \begin{array}{l} \uparrow \\ n-N \\ \downarrow \\ \uparrow \\ N \\ \downarrow \end{array} \end{array} \tag{32}$$

$\leftarrow n-N \rightarrow \leftarrow m-M \rightarrow$

On pose :

$$(\lambda'_B | \lambda_N | \lambda'_N | \lambda_B) = \left(\begin{array}{c|c} \delta_B & \delta_N \end{array} \right) \cdot \begin{array}{|c|c|c|c|} \hline J^+ & J-I & J^+ & J^+ C \\ \hline 0 & 0 & -I & 0 \\ \hline \end{array} \quad (33)$$

← n-N → ← N →
← n-N → ← M → ← N → ← m-M →

On a :

$$(\lambda'_B | \lambda_B = 0 | \lambda_N | \lambda'_N) = \delta \cdot (S | M^+)$$

ou, en spécifiant les formats :

$$\left(\begin{array}{c|c|c|c} \lambda'_B & \lambda_B = 0 & \lambda_N & \lambda'_N \end{array} \right) = \begin{array}{|c|c|c|c|} \hline \leftarrow n-N \rightarrow & \leftarrow m-M \rightarrow & \leftarrow M \rightarrow & \leftarrow N \rightarrow \\ \hline \end{array}$$

$$= (\delta_B | \delta_N) \cdot \begin{array}{|c|c|c|c|} \hline J^+ & J-I & 0 & J^+ C \\ \hline 0 & 0 & 0 & -I \\ \hline \end{array} \quad (34)$$

← n-N → ← m-M → ← M → ← N →

L'identification avec les définitions précédentes de $(\lambda'_B | \lambda_B)$ et $(\lambda_N | \lambda'_N)$ par (24), (25), (26), (27) est immédiate et l'on a donc la définition générale des variables duales en un point quelconque donnée par :

$$(\lambda'_B | \lambda_B) = \delta \cdot S \quad (\lambda_N | \lambda'_N) = \delta \cdot M^+ \quad (35)$$

Remarque : On peut évaluer la sensibilité des variables (λ, λ') à des variations de la fonction objectif f mesurées par $\delta = \text{grad } f$:

$$\frac{\partial}{\partial \delta} (\lambda'_B | \lambda_B) = S \quad \frac{\partial}{\partial \delta} (\lambda_N | \lambda'_N) = M^+ \quad (36)$$

THÉORÈME : Les définitions (24-25-26-27) ou (35) des variables duales en un point quelconque impliquent que le gradient de l'objectif s'exprime comme combinaison linéaire des gradients des contraintes saturées en ce point, les variables λ_i et λ'_j étant les coefficients (non nécessairement positifs) de cette combinaison linéaire :

$$\frac{\partial f(x)}{\partial x} = \sum_{i=1}^m \lambda_i \frac{\partial g_i}{\partial x} + \sum_{j=1}^n \lambda'_j \frac{\partial h_j}{\partial x} = \lambda \cdot \frac{\partial g(x)}{\partial x} - \lambda' \quad (3)$$

Démonstration: D'après les définitions (24-25-26-27) ou (35) des variables λ_i et λ'_j on a :

$$\lambda = (\lambda_B, \lambda_N) = (0, \delta_B J^+)$$

$$\lambda' = (\lambda'_B, \lambda'_N) = \delta_B J^+ (J, C) - (\delta_B, \delta_N) = \lambda_N (J, C) - (\delta_B, \delta_N)$$

Comme $\lambda_B = 0$, on peut écrire :

$$\lambda' = \lambda_N (J, C) - (\delta_B, \delta_N) + \lambda_B (U, V)$$

d'où l'on tire :

$$\delta = (\delta_B, \delta_N) = \lambda_N (J, C) + \lambda_B (U, V) - \lambda' = \lambda \cdot \frac{\partial g}{\partial x} - \lambda'$$

Cette formule (3) sera vérifiée sur les applications numériques données à la fin.

4. GÉNÉRALISATION DU THÉORÈME DE L'ENVELOPPE

Pour démontrer rapidement le « *théorème de l'enveloppe généralisé* » dans le cas où la solution courante de statu quo n'est pas la solution optimale, et n'est pas non plus un sommet du domaine réalisable, nous donnerons une version paramétrique du problème d'optimisation ; les paramètres sont deux vecteurs α et β . Le programme s'écrit alors :

$$\text{Max} \{ f(x, \alpha) \mid g(x, \alpha) \leq \beta, x \geq 0 \}$$

Après avoir introduit les variables d'écart et séparé les variables en (x_B, x'_B, x'_N, x_N) , la séparation des contraintes saturées des contraintes non saturées donne la forme générale des équations (5) :

$$\left. \begin{array}{lll} g_S(x_B, x_N, \alpha) = \beta_S & g_S(x_B, x_N, \alpha) + x'_N = \beta_S & \text{avec } x'_N = 0 \\ g_{NS}(x_B, x_N, \alpha) < \beta_{NS} & g_{NS}(x_B, x_N, \alpha) + x'_B = \beta_{NS} & \text{avec } x'_B > 0 \\ -x_B < 0_{NS} & -x_B + y'_B = 0_{NS} & \text{avec } y'_B > 0 \\ -x_N = 0_S & -x_N + y'_N = 0_S & \text{avec } y'_N = 0 \end{array} \right\} (37)$$

A partir du point de statu quo considéré nous faisons varier les paramètres et/ou les seconds membres des contraintes des quantités $(\delta\alpha, \delta\beta_S, \delta\beta_{NS}, \delta 0_{NS}, \delta 0_S)$ et, ce faisant, on peut voir les ajustements de la position de statu quo. Nous supposons les variations suffisamment petites pour ne pas altérer la structure basique du point de statu quo : on suppose que $\delta x'_N = 0$ et $\delta y'_N = 0$.

Ce problème de statique comparative implique l'introduction des variables duales $(\lambda'_B, \lambda'_N, \lambda_N, \lambda_B)$ définies par (35) ou par (24-25-26-27) où J est supposée à lignes linéairement indépendantes. Après avoir effectué les variations $(\delta\alpha, \delta\beta_S, \delta\beta_{NS}, \delta\theta_{NS}, \delta\theta_S)$ les équations (37) deviennent :

$$\left. \begin{aligned} g_S + \delta g_S + x'_N + \delta x'_N &= \beta_S + \delta\beta_S \\ g_{NS} + \delta g_{NS} + x'_B + \delta x'_B &= \beta_{NS} + \delta\beta_{NS} \\ -(x_B + \delta x_B) + y'_B + \delta y'_B &= 0_{NS} + \delta 0_{NS} \\ -(x_N + \delta x_N) + y'_N + \delta y'_N &= 0_S + \delta 0_S \end{aligned} \right\} \quad (38)$$

ou, en tenant compte de (7), (17), (37) et en posant

$$\left. \begin{aligned} \frac{\partial g_S}{\partial \alpha} = D \quad \frac{\partial g_{NS}}{\partial \alpha} = W \\ \delta g_S + \delta x'_N = J \delta x_B + C \delta x_N + D \delta \alpha + \delta x'_N = \delta \beta_S \\ \delta g_{NS} + \delta x'_B = U \delta x_B + V \delta x_N + W \delta \alpha + \delta x'_B = \delta \beta_{NS} \\ -\delta x_B + \delta y'_B = \delta 0_{NS} \\ -\delta x_N + \delta y'_N = \delta 0_S \end{aligned} \right\} \quad (40)$$

Dans les équations (40) précédentes on a supposé que toutes les variations étaient infiniment petites de sorte que δg_S and δg_{NS} pouvaient être considérées comme des différentielles. Par permutations des équations (40) on obtient :

$$J \delta x_B + C \delta x_N = \delta \beta_S - D \delta \alpha - \delta x'_N \quad (41.1)$$

$$0 \delta x_B - I \delta x_N = \delta 0_S - \delta y'_N \quad (41.2)$$

$$U \delta x_B + V \delta x_N = \delta \beta_{NS} - W \delta \alpha - \delta x'_B \quad (41.3)$$

$$-I \delta x_B + 0 \delta x_N = \delta 0_{NS} - \delta y'_B \quad (41.4)$$

La résolution du système (41.1 - 41.2) donne la solution $(\delta x_B, \delta x_N)$, dont le report dans (41.3 - 41.4) donnera la solution $(\delta x'_B, \delta y'_B)$: la solution $(\delta x_B, \delta x_N)$ du système (41.1 - 41.2) seule importe et ce système peut s'écrire sous forme matricielle :

$$M \cdot \delta x = \begin{bmatrix} J & C \\ 0 & -I \end{bmatrix} \begin{bmatrix} \delta x_B \\ \delta x_N \end{bmatrix} = \begin{bmatrix} \delta \beta_S - D \delta \alpha - \delta x'_N \\ \delta 0_S - \delta y'_N \end{bmatrix} = u \quad (42)$$

La matrice J est supposée de plein rang en lignes, M est alors aussi de plein rang en lignes et MM^T est inversible: on pose $M^+ = M^T (MM^T)^{-1}$ donnée par (29).

La solution générale de l'équation $M \cdot \delta x = u$ est :

$$\delta x = M^+ u + (I - M^+ M) z, \quad z \text{ arbitraire} \quad (43)$$

Pour $u=0$, le déplacement δx s'effectue sur l'intersection des contraintes saturées. Pour $z=0$, le déplacement δx est orthogonal à cette intersection des contraintes saturées.

Si z est quelconque différent de zéro, alors le déplacement δx est aussi quelconque (en particulier on caractérisera un déplacement sur l'intersection des contraintes saturées au point de statu quo et un autre déplacement orthogonal à cette intersection). Comme nous avons :

$$-Rz = \begin{bmatrix} I - J^+ J & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_B \\ z_N \end{bmatrix} = \begin{bmatrix} (I - J^+ J) z_B \\ 0 \end{bmatrix} \quad (44)$$

La solution (43) s'explique ainsi :

$$\begin{bmatrix} \delta x_B \\ \delta x_N \end{bmatrix} = \begin{bmatrix} J^+ \delta \beta_S + (I - J^+ J) z_B + J^+ C \delta 0_S - (J^+ D \delta \alpha + J^+ \delta x'_N + J^+ C \delta y'_N) \\ -\delta 0_S + \delta y'_N \end{bmatrix} \quad (45)$$

ou, compte tenu de (41.2) :

$$\delta x_B = J^+ \delta \beta_S - J^+ C \delta x_N - J^+ \delta x'_N + (I - J^+ J) z_B - J^+ D \delta \alpha \quad (45')$$

La fonction objectif $f(x_B, x_N, \alpha)$ devient $f(x_B + \delta x_B, x_N + \delta x_N, \alpha + \delta \alpha) = f + \partial f$ où :

$$\delta f = \delta_B \cdot \delta x_B + \delta_N \cdot \delta x_N + f_\alpha \delta \alpha$$

On déduit de (45') :

$$\delta f = \lambda_N \delta \beta_S + \lambda'_N \delta 0_S - \lambda_N \delta x'_N - \lambda'_N \delta y'_N - \lambda'_B z_B + (f_\alpha - \lambda_N D) \delta \alpha \quad (46)$$

ou encore, compte tenu de l'équation (41.2) :

$$\delta f = \lambda_N \delta \beta_S - \lambda'_N \delta x_N - \lambda_N \delta x'_N - \lambda'_B z_B + (f_\alpha - \lambda_N D) \delta \alpha \quad (47)$$

On déduit de (46) et (47) les interprétations suivantes de λ_N et λ'_N :

$$\lambda_N = \frac{\delta f}{\delta \beta_S} \quad \lambda'_N = \frac{\delta f}{\delta 0_S} \quad (48)$$

$$\lambda_N = - \frac{\delta f}{\delta x'_N} \quad \lambda'_N = - \frac{\delta f}{\delta x_N} \quad (49)$$

Remarques : 1. Le système (41.3 - 41.4) donne la solution $(\delta x'_B, \delta y'_B)$:

$$\begin{bmatrix} \delta x'_B \\ \delta y'_B \end{bmatrix} = - \begin{bmatrix} U & V \\ -I & 0 \end{bmatrix} \begin{bmatrix} \delta x_B \\ \delta x_N \end{bmatrix} + \begin{bmatrix} \delta\beta_{NS} - W\delta\alpha \\ \delta 0_{NS} \end{bmatrix}$$

D'après (43) on a :

$$\begin{bmatrix} \delta x_B \\ \delta x_N \end{bmatrix} = M^+ u + (I - M^+ M) z = \begin{bmatrix} J^+ & J^+ C \\ 0 & -I \end{bmatrix} \begin{bmatrix} \delta\beta_S - D\delta\alpha - \delta x'_N \\ \delta 0_S - \delta y'_N \end{bmatrix} + \begin{bmatrix} (I - J^+ J) z_B \\ 0 \end{bmatrix}$$

d'où :

$$\begin{bmatrix} \delta x'_B \\ \delta y'_B \end{bmatrix} = \begin{bmatrix} \delta\beta_{NS} - W\delta\alpha - U(I - J^+ J) z_B \\ \delta 0_{NS} + (I - J^+ J) z_B \end{bmatrix} - \begin{bmatrix} UJ^+ & UJ^+ C - V \\ -J^+ & -J^+ C \end{bmatrix} \begin{bmatrix} \delta\beta_S - D\delta\alpha - \delta x'_N \\ \delta 0_S - \delta y'_N \end{bmatrix} \quad (50)$$

comme $\delta x_N = \delta y'_N - \delta 0_S$ d'après (41.2) on déduit

$$\delta x'_B = \delta\beta_{NS} - W\delta\alpha - UJ^+ \delta\beta_S + UJ^+ D\delta\alpha - UJ^+ (I - J^+ J) z_B + (UJ^+ C - V) \delta x_N + UJ^+ \delta x'_N$$

et les variations compensées

$$\frac{\delta x'_B}{\delta x_N} = UJ^+ C - V \quad \frac{\delta x'_B}{\delta x'_N} = UJ^+ \quad (51)$$

2. La formule (45') permet de confirmer les relations (22) et la formule (47) permet de retrouver les relations (24), (25), (26), (27).

Cas particuliers

(a) Si on prend $z=0$, alors $\delta x = M^+ u$ et on obtient la solution δx de longueur minimale; cette solution appartient à l'espace des lignes de M car sa projection sur le noyau de M est nulle :

$$(I - M^+ M) \delta x = 0, \quad \text{par la relation } M^+ M M^+ = M^+$$

Ceci montre que le déplacement δx est alors orthogonal à l'intersection des contraintes saturées au point de statu quo. La solution (43) de longueur minimale s'obtient en posant $z_B=0$ dans (45') et les formules (46) et (47) donnant les variations de la fonction objectif dues à ce déplacement deviennent :

$$\delta f = \lambda_N \delta\beta_S + \lambda'_N \delta 0_S - \lambda_N \delta x'_N - \lambda'_N \delta y'_N + (f_\alpha - \lambda_N D) \delta\alpha \quad (52)$$

$$\delta f = \lambda_N \delta\beta_S - \lambda'_N \delta x_N - \lambda_N \delta x'_N + (f_\alpha - \lambda_N D) \delta\alpha \quad (53)$$

(b) Si on envisage un déplacement sur l'intersection des contraintes saturées, même en laissant varier ces contraintes ($\delta\beta_S \neq 0$, $\delta\theta_S \neq 0$), (dans ce cas $u \neq 0$), la solution (43) implique plusieurs possibilités différentes de déplacement suivant les valeurs de z : on donnera à z les valeurs successives e_1, e_2, \dots, e_{n-N} , c'est-à-dire les vecteurs unitaires portés par les axes, z se réduisant à z_B , la partie de z concernant les seules composantes de δx_B . Ceci nous donnera les *taux de substitution* des variables δx_B correspondant à ces changements impliqués par ces vecteurs e_j (ceci se voit très facilement sur les tableaux du simplexe généralisés donnés plus loin). Dans ces conditions, on a pour l'ensemble des composantes de δx_B , les variations :

$$(I - J^+ J) \cdot (e_1, e_2, \dots, e_{n-N}) = (I - J^+ J) \cdot I = I - J^+ J$$

de sorte que pour les variations successives $\delta x_B = (e_1, e_2, \dots, e_{n-N})$ on a les ajustements successifs des composantes δx_{Bj} correspondants déduits de (45'); ces ajustements correspondent aux *variations nettes* de δx_B correspondant aux variations unitaires de chacune de ses composantes respectivement. Ceci revient à dire

« que se passe-t-il pour $\delta x_B = z_B = e_1, e_2, \dots, e_{n-N}$? »

Quels sont les *ajustements nets* de δx_B à ces variations, compte tenu du respect des contraintes saturées, c'est-à-dire pour un déplacement sur l'intersection de ces contraintes ?

L'équation (45') s'écrit alors :

$$\delta x_B = (I - J^+ J) \delta x_B - J^+ C \delta x_N - J^+ \delta x'_N + J^+ \delta \beta_S - J^+ D \delta \alpha \quad (45'')$$

ou

$$J^+ J \delta x_B + J^+ C \delta x_N + J^+ \delta x'_N = J^+ \delta \beta_S - J^+ D \delta \alpha \quad (54)$$

La formule (47) donnant la variation de la fonction objectif due à ce déplacement δx_B « compensé » s'écrit alors :

$$\delta f = \lambda_N \delta \beta_S - \lambda'_N \delta x_N - \lambda_N \delta x'_N - \lambda'_B \delta x_B + (f_\alpha - \lambda_N D) \delta \alpha \quad (55)$$

De plus les équations (41.3) et (41.4) donneront la sensibilité des *variables basiques* x'_B et y'_B à ce déplacement. En particulier, compte tenu de (45''), (41.3) devient :

$$\begin{aligned} U(I - J^+ J) \delta x_B - UJ^+ C \delta x_N - UJ^+ \delta x'_N + V \delta x_N + \delta x'_B \\ = -UJ^+ \delta \beta_S + UJ^+ D \delta \alpha + \delta \beta_{NS} - W \delta \alpha \quad (56) \end{aligned}$$

ou encore

$$\begin{aligned}
 & -U(J^+ J - I)\delta x_B + (V - UJ^+ C)\delta x_N - UJ^+ \delta x'_N + \delta x'_B \\
 & = -UJ^+ \delta\beta_S + \delta\beta_{NS} + (UJ^+ D - W)\delta\alpha \quad (57)
 \end{aligned}$$

(c) Si on impose enfin aux paramètres α de rester fixes ($\delta\alpha=0$), alors les équations (54), (55) et (57) peuvent se représenter en tableaux de type «tableaux du simplexe», obtenus à partir des tableaux initiaux par prémultiplication par une matrice appropriée; tous ces tableaux sont dérivés des tableaux équivalents de la programmation linéaire où la matrice classique de prémultiplication

B^{-1}	0
$c_B B^{-1}$	1

est décomposée par blocs :

J^{-1}	0	o
$-UJ^{-1}$	I	o
$c_B J^{-1}$	0	1

pour tenir compte de la séparation entre les contraintes saturées et non saturées au point de statu quo, et peut se généraliser ainsi :

J^+	0	o
$-UJ^+$	I	o
$\delta_B J^+$	0	1

Les équations (54), (55) et (57) s'écrivent alors :

$$J^+ J \delta x_B + J^+ C \delta x_N + J^+ \delta x'_N = J^+ \delta\beta_S \quad (54')$$

$$\lambda'_B \delta x_B + \lambda'_N \delta x_N + \lambda_N \delta x'_N = \lambda_N \delta\beta_S - \delta f \quad (55')$$

$$\begin{aligned}
 & -U(J^+ J - I)\delta x_B + (V - UJ^+ C)\delta x_N - UJ^+ \delta x'_N + \delta x'_B \\
 & = -UJ^+ \delta\beta_S + \delta\beta_{NS} \quad (57')
 \end{aligned}$$

La formule (55') peut être développée sous la forme :

$$\delta_B (J^+ J - I) \delta x_B + (\delta_B J^+ C - \delta_N) \delta x_N + \delta_B J^+ \delta x'_N = \delta_B J^+ \delta \beta_S - \delta f \quad (55'')$$

Les formules (54'), (57') et (55'') se traduisent en «tableaux du simplexe généralisés» donnés ci-après. Ces tableaux se simplifient dans le cas de la programmation linéaire où $\delta_B = c_B$ et où les matrices de gradients J, C, U, V reconstituent la matrice A du programme linéaire.

$\leftarrow M \rightarrow \leftarrow m-M \rightarrow \leftarrow 1 \rightarrow$	$\leftarrow n-N \rightarrow \leftarrow N \rightarrow \leftarrow M \rightarrow \leftarrow m-M \rightarrow$	
		=

J^+	0	o
$-UJ^+$	I	o
$\delta_B J^+$	o	I

δx_B	δx_N	$\delta x'_N$	$\delta x'_B$	
J	C	I	0	δb_S
U	V	0	I	δb_{NS}
$-\delta_B$	$-\delta_N$	o	o	$-\delta f$

δx_B	δx_N	$\delta x'_N$	$\delta x'_B$	
$\leftarrow n-N \rightarrow$	$\leftarrow N \rightarrow$	$\leftarrow M \rightarrow$	$\leftarrow m-M \rightarrow$	

$J^+ J$	$J^+ C$	J^+	0	$J^+ \delta b_S$
$-U (J^+ J - I)$	$-UJ^+ C + V$	$-UJ^+$	I	$-UJ^+ \delta b_S + \delta b_{NS}$
$\delta_B (J^+ J - I)$	$\delta_B J^+ C - \delta_N$	$\delta_B J^+$	o	$\delta_B J^+ \delta b_S - \delta f$

λ'_B	λ'_N	λ_N	λ_B
--------------	--------------	-------------	-------------

Remarque : La première ligne du tableau ci-dessus n'est autre que l'équation (54') dans le cas d'un déplacement sur l'intersection des contraintes saturées. Dans le cas d'un déplacement *orthogonal* à cette intersection c'est l'équation (45') qui régit ce déplacement avec $z_B = 0$ (et $\alpha = 0$): la matrice unité I remplace la matrice $J^+ J$ dans la première ligne du tableau et alors $\delta x_B = J^+ \delta b_S$, ce qui implique la relation d'exclusion $\lambda'_B \cdot \delta x_B = 0$.

Ce produit de matrices est à comparer avec le produit similaire permettant d'obtenir en programmation linéaire le tableau du simplexe d'une base réalisable à partir du tableau initial :

J^{-1}	0	o
$-UJ^{-1}$	I	o
$c_B J^{-1}$	o	1

x_B	x_N	x'_N	x'_B	
J	C	I	0	b_S
U	V	0	I	b_{NS}
$-c_B$	$-c_N$	o	o	$-f$

=

x_B	x_N	x'_N	x'_B	
I	$J^{-1} C$	J^{-1}	0	$J^{-1} b_S$
0	$-UJ^{-1} C + V$	$-UJ^{-1}$	I	$-UJ^{-1} b_S + b_{NS}$
o	$c_B J^{-1} C - c_N$	$c_B J^{-1}$	o	$c_B J^{-1} b_S - f$
λ'_B	λ'_N	λ_N	λ_B	

Relations d'exclusion

La dernière ligne du «tableau du simplexe généralisé» donne les valeurs des variables duales telles qu'elles sont définies par les équations (24), (25), (26) et (27) et on trouve dans la dernière colonne les valeurs des variations des variables primales quand on effectue un déplacement *orthogonal* à l'intersection des contraintes saturées ($J^+ J$ est remplacé par I dans la première ligne du tableau) en supposant que la structure basique reste inchangée ($\delta x_N = 0, \delta x'_N = 0$):

$$\lambda'_B = \delta_B (J^+ J - 1) \quad \lambda'_N = \delta_B J^+ C - \delta_N \quad \lambda_N = \delta_B J^+ \quad \lambda_B = 0$$

$$\delta x_B = J^+ \delta b_S \quad \delta x_N = 0 \quad \delta x'_N = 0 \quad \delta x'_B = -UJ^+ \delta b_S + \delta b_{NS}$$

On observe que les relations d'exclusion sont *presque* toutes satisfaites car on a :

$$\lambda'_{Nj} \cdot \delta x_{Nj} = 0 \quad \lambda_{Ni} \cdot \delta x'_{Ni} = 0 \quad \lambda_{Bi} \cdot \delta x'_{Bi} = 0$$

sans avoir obligatoirement : $\lambda'_{Bj} \cdot \delta x_{Bj} = 0$

Cependant, en sommant par rapport aux indices, on obtient :

$$\lambda'_N \cdot \delta x_N = 0 \quad \lambda_N \cdot \delta x'_N = 0 \quad \lambda_B \cdot \delta x'_B = 0$$

et on a aussi la dernière condition $\lambda'_B \cdot \delta x_B = 0$ due à la relation $J^+ J J^+ = J^+$.

Remarque : Dans le cas général où δx_B est donné par l'équation (45') avec $z_B \neq 0$ cette dernière relation d'exclusion $\lambda'_B \cdot \delta x_B = 0$ n'est pas satisfaite :

$$\lambda'_B \cdot \delta x_B = \delta_B \cdot (J^+ J - I) \cdot z_B$$

Dans ce cas, on ne peut constater les relations de dualité qui suivent.

Dualité (dans le cas d'un déplacement orthogonal à l'intersection des contraintes saturées)

La linéarisation des contraintes permet d'exhiber une *dualité* comparable à la dualité en programmation linéaire en termes de l'égalité à l'optimum des valeurs des objectifs des programmes primal et dual :

$$c \cdot x^* = \lambda^* \cdot b$$

Ici cette égalité reste vraie mais maintenant en termes de *variations* des fonctions objectifs au voisinage du point de statu quo :

$$\delta \cdot \delta x = \lambda \cdot \delta \beta$$

(qui, dans le cas d'un programme linéaire, se réduit à : $c \cdot \delta x = \lambda \cdot \delta b$).

$$\delta \cdot \delta x = \delta_B \cdot \delta x_B + \delta_N \cdot \delta x_N = \delta_B J^+ \delta \beta_S$$

$$\lambda \cdot \delta \beta = \lambda_B \cdot \delta \beta_{NS} + \lambda_N \cdot \delta \beta_S = \delta_B J^+ \delta \beta_S$$

Cette égalité se voit aussi sur les «tableaux du simplexe généralisés» correspondant à ce cas (*voir* la remarque qui suit ce tableau du simplexe généralisé). Elle confirme la structure *duale* des variables λ et λ' par rapport aux variables *primales* x et x' .

Théorème de l'enveloppe généralisé en n'importe quel point de statu quo

Considérons le programme paramétrique suivant :

$$\text{Max } \{ f(x, \alpha) \mid g(x, \alpha) \leq \beta, x \geq 0 \}$$

Supposons que l'on se place en un point de statu quo $x(\alpha, \beta)$, non nécessairement optimal, pour des valeurs données des paramètres α et β . On déduit les valeurs des variables duales $\lambda(\alpha, \beta)$ des équations (26), (27) ou (35). Il est alors possible de généraliser le théorème de l'enveloppe sous la forme suivante :

THÉORÈME : Soit $x(\alpha, \beta)$ un point fixé de statu quo, $\lambda(\alpha, \beta)$ et $\lambda'(\alpha, \beta)$ les variables duales définies par (24), (25), (26) et (27) pour les valeurs fixées des paramètres α et β . On notera :

$F(\alpha, \beta) = f[x(\alpha, \beta), \alpha]$ la valeur de l'objectif en ce point de statu quo

$\Phi(x, \lambda, \alpha, \beta) = f(x, \alpha) - \lambda \cdot [g(x, \alpha) - \beta]$ la fonction de Lagrange du problème

$\Psi(\alpha, \beta) = f[x(\alpha, \beta), \alpha] - \lambda(\alpha, \beta) \cdot \{g[x(\alpha, \beta), \alpha] - \beta\} = \Phi[x(\alpha, \beta), \lambda(\alpha, \beta), \alpha, \beta]$

Supposons que ces fonctions soient continûment différentiables. Nous avons alors :

$$\frac{\partial F(\alpha, \beta)}{\partial \alpha_k} = \frac{\partial \Psi(\alpha, \beta)}{\partial \alpha_k} = \frac{\partial \Phi(x, \lambda, \alpha, \beta)}{\partial \alpha_k} \tag{58}$$

ou :

$$F_\alpha = \Psi_\alpha = \Phi_\alpha, \quad \text{avec} \quad \frac{\partial \Phi}{\partial \alpha_k} = \frac{\partial f(x, \alpha)}{\partial \alpha_k} - \lambda \cdot \frac{\partial g(x, \alpha)}{\partial \alpha_k}$$

De la même façon, on a :

$$\frac{\partial F(\alpha, \beta)}{\partial \beta_i} = \frac{\partial \Psi(\alpha, \beta)}{\partial \beta_i} = \frac{\partial \Phi(x, \lambda, \alpha, \beta)}{\partial \beta_i} \tag{59}$$

c'est-à-dire en particulier :

$$\frac{\partial f[x(\alpha, \beta), \alpha]}{\partial \beta_i} = \lambda_i$$

On a de plus :

$$\frac{\partial f[x(\alpha, \beta), \alpha]}{\partial x_j} = \frac{\partial \Phi(x, \lambda, \alpha, \beta)}{\partial x_j} = -\lambda'_j \tag{60}$$

Démonstration: Avec l'hypothèse supplémentaire de structure basique inchangée qui implique $\delta x_N = 0$ et $\delta x'_N = 0$ l'expression (45') s'écrit :

$$\delta x_B = J^+ \delta \beta_S + (I - J^+ J) z_B - J^+ D \delta \alpha$$

d'où il s'ensuit, parce que $J^+ J J^+ = J^+$:

$$\begin{aligned} \lambda'_B \cdot \delta x_B &= \delta_B (J^+ J - I) \cdot [J^+ \delta \beta_S + (I - J^+ J) z_B - J^+ D \delta \alpha] \\ &= \delta_B (J^+ J - I) z_B = \lambda'_B \cdot z_B \end{aligned}$$

La formule (46) donne la signification économique classique des variables duales λ_N et λ'_N comme mesures de la productivité marginale (en terme de fonction objectif) de la variation des seconds membres des contraintes ou comme mesures de la sensibilité de l'objectif à ces variations :

$$\lambda_N = \frac{\delta f}{\delta \beta_S} \quad (61)$$

$$\lambda'_N = \frac{\delta f}{\delta \theta_S} \quad (62)$$

On a vu que $\lambda_B = 0$ par (26). Il s'ensuit que l'équation (47) peut s'écrire sous la forme :

$$\delta f = \lambda_N \cdot \delta \beta_S + \lambda_B \cdot \delta \beta_{NS} - \lambda'_N \cdot \delta x_N - \lambda'_B \cdot \delta x_B + (f_\alpha - \lambda_N D - \lambda_B W) \cdot \delta \alpha \quad (63)$$

ou

$$\delta f = \lambda \cdot \delta \beta - \lambda' \cdot \delta x + (f_\alpha - \lambda \cdot g_\alpha) \cdot \delta \alpha \quad (64)$$

On déduit les relations (58), (59) et (60) de la formule (64), ce qui achève la démonstration.

Remarques: 1. Les équations (60) indiquent quelles variables x_j il est intéressant d'accroître à partir du point de statu quo pour améliorer la fonction objectif.

2. Les équations (58) peuvent aussi d'obtenir directement car, d'après l'équation (45)

$$\boxed{\frac{\delta x}{\delta \alpha}} = \boxed{\frac{\delta x_B / \delta \alpha}{\delta x_N / \delta \alpha}} = \boxed{\begin{matrix} -J^+ D \\ 0 \end{matrix}}$$

et

$$\begin{aligned} \frac{\delta f}{\delta \alpha} &= \frac{\partial f}{\partial x} \cdot \frac{\delta x}{\delta \alpha} + \frac{\partial f}{\partial \alpha} = -\delta_B J^+ D + f_\alpha \\ &= -\lambda_N D + f_\alpha = -\lambda_N D - \lambda_B W + f_\alpha = f_\alpha - \lambda \cdot g_\alpha \end{aligned}$$

Application du théorème de l'enveloppe à la programmation linéaire

On rappelle les résultats classiques d'analyse de sensibilité obtenus à l'optimum x^* d'un programme linéaire, $\text{Max} \{ cx \mid Ax \leq b, x \geq 0 \}$, en notant $z^* = cx^*$:

$$\frac{\partial z^*}{\partial b_i} = \lambda_i^*, \quad \frac{\partial z^*}{\partial c_j} = x_j^*, \quad \frac{\partial z^*}{\partial a_{kl}} = -\lambda_k^* x_l^*$$

où les λ_i^* sont les valeurs optimales des variables duales.

Le théorème de l'enveloppe généralisé peut s'appliquer en posant $\beta = b$ et $\alpha = (A, c)$:

$$\frac{\partial f}{\partial b_i} = \lambda_i, \quad \frac{\partial f}{\partial c_j} = x_j, \quad \frac{\partial f}{\partial a_{kl}} = \frac{\partial \Phi}{\partial a_{kl}} = -\lambda_k x_l \tag{65}$$

5. ANALYSE DE SENSIBILITÉ

Le programme non linéaire considéré maintenant s'obtient à partir de la version paramétrique du problème d'optimisation donnée pour le théorème de l'enveloppe en posant $\alpha = 0$ et $\beta = b$. La forme générale des équations (37) devient :

$$\left. \begin{array}{lll} g_S(x_B, x_N) = b_S & g_S(x_B, x_N) + x'_N = b_S & \text{avec } x'_N = 0 \\ g_{NS}(x_B, x_N) < b_{NS} & g_{NS}(x_B, x_N) + x'_B = b_{NS} & \text{avec } x'_B > 0 \\ -x_B < 0_{NS} & -x_B + y'_B = 0_{NS} & \text{avec } y'_B > 0 \\ -x_N = 0_S & -x_N + y'_N = 0_S & \text{avec } y'_N = 0 \end{array} \right\} \tag{37'}$$

Comment varie la solution $(x \mid x' \mid f) = (x_B, x_N \mid x'_B, x'_N \mid f)$ lorsque l'on applique certaines perturbations aux seconds membres des contraintes, c'est-à-dire lorsque les contraintes varient un peu (problème n° 1), ou lorsque, sans déplacer les contraintes, on effectue un déplacement de la solution compatible avec les contraintes maintenues, c'est-à-dire quand le déplacement s'effectue

sur l'intersection des contraintes saturées (problème n° 2)? L'analyse de sensibilité s'effectue de façon différente dans les problèmes n° 1 et n° 2.

Problème n° 1

A partir du point considéré on effectue des perturbations ($\delta b_S, \delta b_{NS}, \delta \theta_{NS}, \delta \theta_S$) des seconds membres des contraintes et on constate l'ajustement de la solution par le passage du point de statu quo (x_B, x_N, x'_N, x'_B) au point ($x_B + \delta x_B, x_N + \delta x_N, x'_N + \delta x'_N, x'_B + \delta x'_B$); on a vu que cet ajustement s'effectue par un déplacement δx orthogonal à l'intersection des contraintes saturées au point de statu quo. On suppose que les perturbations sont assez faibles pour ne pas altérer la structure basique de la solution de sorte que le nouveau point de statu quo sature les mêmes contraintes de structure ou de signe. On aura donc encore :

$$x'_N + \delta x'_N = 0 \quad y'_N + \delta y'_N = 0 \quad i. e. \quad \delta x'_N = 0 \quad \delta y'_N = 0.$$

1. Quel est l'effet sur l'objectif de ce desserrement des contraintes ?
2. Comment s'effectue, dans ce desserrement des contraintes le déplacement du point de statu quo ?

Deux cas sont à considérer selon que le point de statu quo est un sommet (non dégénéré) du domaine réalisable ($n = M + N$) ou non ($n > M + N$). Dans le premier cas le desserrement des contraintes à structure basique inchangée conduira à un autre sommet qui est en fait le sommet défini par l'intersection des mêmes contraintes mais déplacé par le desserrement de ces mêmes contraintes. Dans le second cas d'un point non extrême le desserrement des contraintes à structure basique inchangée autorise le passage à un autre point saturant les *mêmes* contraintes; mais ici l'intersection des contraintes ne se réduit pas à un seul point et *tous* les points de cette intersection « déplacée » par le déplacement des contraintes saturées sont *a priori* accessibles. On a vu que la définition choisie des variables duales ($\lambda'_B, \lambda'_N, \lambda_N, \lambda_B$) implique un déplacement δx orthogonal à l'intersection des contraintes saturées. Dans les deux cas nous aurons le résultat suivant :

PROPOSITION 1 : *Dans le cas d'un desserrement des contraintes à structure basique inchangée au point de statu quo, la variation de la fonction objectif δf est donnée par :*

$$\delta f = \lambda_N \cdot \delta b_S + \lambda'_N \cdot \delta \theta_S$$

Il s'ensuit l'interprétation économique classique des variables duales (λ_N, λ'_N) :

$$\lambda_N = \frac{\delta f}{\delta b_S} \quad \lambda'_N = \frac{\delta f}{\delta \theta_S} \tag{66}$$

Le déplacement δx d'ajustement s'effectue perpendiculairement à l'intersection des contraintes saturées.

Démonstration: Nous sommes dans le cas particulier **(a)** dans lequel $\delta\alpha = 0$. De plus la structure basique inchangée implique que $\delta x'_N = 0$ et $\delta y'_N = 0$. L'équation (52) s'écrit alors:

$$\delta f = \lambda_N \cdot \delta b_S + \lambda'_N \cdot \delta \theta_S$$

L'équation $M \cdot \delta x = u$ a pour solution $\delta x = M^+ u$ (car $z = 0$) où $u = \begin{bmatrix} \delta b_S \\ \delta \theta_S \end{bmatrix}$.

On a vu que δx est alors orthogonal à l'intersection des contraintes saturées.

Problème n° 2

On maintient les contraintes telles qu'elles sont et on effectue un déplacement du point de statu quo sur l'intersection des contraintes saturées en ce point. Les différentes variables du problème sont amenées à changer et on peut calculer les taux de substitution des variables entre elles. De plus on peut rechercher l'effet de ce déplacement et de ces variations sur la valeur de la fonction objectif. Nous aurons alors le résultat suivant:

PROPOSITION 2: Dans le cas d'un déplacement sur l'intersection des contraintes saturées la variation de la fonction objectif δf est donnée par:

$$\delta f = -\lambda_N \cdot \delta x'_N - \lambda'_N \cdot \delta x_N - \lambda'_B \cdot \delta x_B$$

Il s'ensuit que les variables duales indiquent, à partir du point de statu quo considéré, quelles variables hors base x'_{Ni} ou x_{Nj} devraient avantageusement être portées à un niveau positif et quelles variables basiques x_{Bj} devraient s'accroître pour améliorer la valeur de la fonction objectif, car:

$$\lambda_{Ni} = -\frac{\delta f}{\delta x'_{Ni}} \quad \lambda'_{Nj} = -\frac{\delta f}{\delta x_{Nj}} \quad \lambda'_{Bj} = -\frac{\delta f}{\delta x_{Bj}} \tag{67}$$

Démonstration: Nous sommes dans le cas particulier **(b)** dans lequel $\delta\alpha = 0$ et $\delta\beta_S = 0$. L'équation (55) s'écrit:

$$\delta f = -\lambda_N \cdot \delta x'_N - \lambda'_N \cdot \delta x_N - \lambda'_B \cdot \delta x_B$$

Les résultats de la proposition 2 s'en déduisent immédiatement.

6. APPLICATIONS NUMÉRIQUES

Nous envisagerons trois exemples, le premier emprunté à la programmation linéaire, un second exemple de programmation non linéaire et le dernier exemple portera sur la théorie du consommateur (où le comportement de ce dernier s'avèrera être un comportement à inertie).

1. Programme linéaire

L'exemple simple ci-dessous permettra de comparer les valeurs des variables duales (λ, λ') définies par les équations (24 à 27) ou les équations (35) aux valeurs des variables duales obtenues de façon classique à l'optimum d'une part et en un sommet quelconque d'autre part (*cf.* [2]). Cet exemple permettra aussi de constater que ces valeurs des variables duales ailleurs qu'en un sommet ont la signification économique habituelle $\partial f/\partial b$ de productivité marginale en termes de fonction objectif des desserrements unitaires des contraintes. Enfin, l'on donnera l'un des «tableaux généralisés du simplexe» associé à un point de statu quo non extrême et situé sur une face (ou une arête) du domaine des points réalisables.

Considérons le programme linéaire suivant :

$$\begin{array}{ll}
 \text{Max } x_1 + x_2 & \text{Max } x_1 + x_2 \\
 x_1 + 4x_2 \leq 12 & x_1 + 4x_2 + x'_1 = 12 \\
 2x_1 + 3x_2 \leq 12 & 2x_1 + 3x_2 + x'_2 = 12 \\
 4x_1 + x_2 \leq 20 & 4x_1 + x_2 + x'_3 = 20 \\
 x_1 \geq 0, x_2 \geq 0 & x_1 \geq 0, x_2 \geq 0, x'_1 \geq 0, x'_2 \geq 0, x'_3 \geq 0
 \end{array}$$

La figure 1 visualise les sommets réalisables O, A, B, C, D (le sommet B est optimal) ainsi que les arêtes OA, AB, BC, CD et OD. On vérifiera ci-dessous que dans les cas où elles existent (cas des sommets), les valeurs duales «classiques» (λ, λ') lues dans la dernière ligne du tableau du simplexe de la base correspondant à ce sommet sont égales aux valeurs données par les formules de définition (24 à 27) ou (35) qui s'écrivent alors :

$$\lambda = [c_B J^+, \lambda_B = 0] \quad \lambda' = [c_B (J^+ J - I), c_B J^+ C - c_N]$$

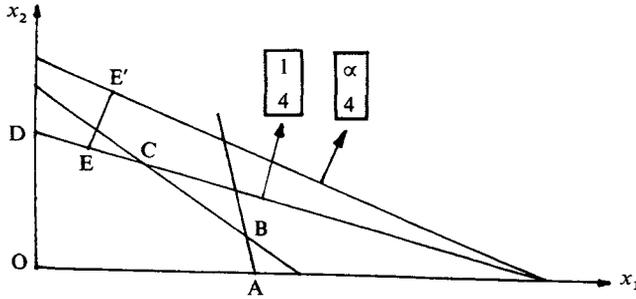


Figure 1

1.1. Calcul des valeurs duales en B (point extrême optimal), en C (point extrême non optimal) et en E (point non extrême)

Dans les cas des points extrêmes la matrice M est inversible et se ramène à la matrice de base (souvent notée B) de sorte que les équations (24 à 27) redonnent les résultats classiques :

$$\lambda_N = c_B B^{-1}, \quad \lambda_B = 0, \quad \lambda'_B = 0, \quad \lambda'_N = c_B B^{-1} N - c_N$$

Par exemple au sommet optimal B on a : $M = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} = J$ et les formules (24 à 27) donnent les valeurs :

$$\lambda_N = (3/10, 1/10), \quad \lambda_B = (0), \quad \lambda'_B = (0, 0), \quad \lambda'_N \text{ vide}$$

Le tableau du simplexe correspondant est alors :

	x_1	x_2	x'_1	x'_2	x'_3	
x'_1	0	0	1	-3/2	1/2	4
x_2	0	1	0	2/5	-1/5	4/5
x_1	1	0	0	-1/10	3/10	24/5
	0	0	0	3/10	1/10	28/5
	λ'_B		λ_B	λ_N		

Au sommet **non optimal** C la matrice M est encore inversible et se ramène à une matrice de base non optimale B :

$$M = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} = J \text{ et les formules (24 à 27) donnent les valeurs:}$$

$$\lambda_N = (-1/5, 3/5), \quad \lambda_B = (0), \quad \lambda'_B = (0, 0), \quad \lambda'_N \text{ vide}$$

Le tableau du simplexe correspondant est alors :

	x_1	x_2	x'_1	x'_2	x'_3	
x_2	0	1	2/5	-1/5	0	12/5
x_1	1	0	-3/5	4/5	0	12/5
x'_3	0	0	2	-3	1	8
	0	0	-1/5	3/5	0	24/5
	λ'_B		λ_N		λ_B	

Sur l'arête CD , on prend comme exemple de point de statu quo le point E de coordonnées $(2, 5/2)$; en ce point la matrice des gradients des contraintes saturées est

$$M = (1 \ 4) = J; \quad M^+ = 1/17 \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

et les formules (24 à 27) donnent les valeurs :

$$\lambda_N = (5/17), \quad \lambda_B = (0, 0), \quad \lambda'_B = (-12/17, 3/17), \quad \lambda'_N \text{ vide.}$$

Le tableau du simplexe généralisé s'obtient ainsi :

1/17	0	0	0
4/17	0	0	0
-14/17	1	0	0
-8/17	0	1	0
5/17	0	0	1

1	4	1	0	0	δb_1
2	3	0	1	0	δb_2
4	1	0	0	1	δb_3
-1	-1	0	0	0	$-\delta f$

	δx_B		$\delta x'_N$	$\delta x'_B$		
δx_1	1/17	4/17	1/17	0	0	$\delta b_1/17$
δx_2	4/17	16/17	4/17	0	0	$4 \delta b_1/17$
$\delta x'_2$	20/17	-5/17	-14/17	1	0	$\delta b_2 - 14 \delta b_1/17$
$\delta x'_3$	60/17	-15/17	-8/17	0	1	$\delta b_1 - 8 \delta b_1/17$
	-12/17	3/17	5/17	0	0	$5 \delta b_1/17 - \delta f$
	λ'_B		λ_N	λ_B		

On remarque sur le tableau précédent que le déplacement δx associé aux desserrements δb des contraintes est bien perpendiculaire à la contrainte saturée (la première contrainte DC) car :

$$\delta x = \delta b_1/17 \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

Dans ce cas d'un déplacement orthogonal à l'intersection des contraintes saturées on peut vérifier les relations d'exclusion et les résultats de dualité énoncés précédemment.

1.2. Analyse de sensibilité

1.2.1. On déplace les contraintes de δb (ou de $\delta 0$ pour les contraintes de signe).

Les ajustements δx s'effectuent perpendiculairement à la contrainte saturée DC et on peut évaluer analytiquement ces déplacements à l'aide de la figure 2. De plus si on envisage la saturation de la seule contrainte CD, le tableau ci-dessus redonne ces ajustements :

$$\delta x_1 = \delta b_1/17 \quad \delta x_2 = 4 \delta b_1/17$$

Il s'ensuit que :

$$\delta f = \delta x_1 + \delta x_2 = 5 \delta b_1/17$$

et

$$\frac{\delta f}{\delta b_1} = \frac{5}{17} = \lambda_1, \quad \frac{\delta f}{\delta b_2} = 0 = \lambda_2, \quad \frac{\delta f}{\delta b_3} = 0 = \lambda_3$$

et d'une façon générale :

$$\frac{\delta f}{\delta b} = \lambda$$

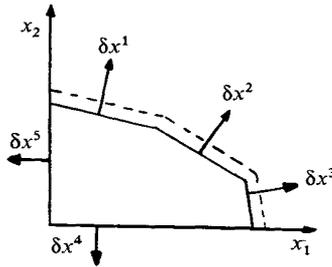


Figure 2

1.2.2. On fait varier un coefficient de la matrice A du programme linéaire.

On fait varier par exemple le coefficient $a_{11}=1$ de la matrice A en le laissant paramétré ($a_{11}=\alpha$); le point de statu quo se déplace de E en E' sur la figure 1 quand la pente de la première contrainte varie, de façon à ce que $EE'=\delta x$ soit orthogonal à la nouvelle position de la contrainte saturée (ce déplacement est le déplacement de longueur minimale). Les coordonnées du point E' sont :

$$2 + \frac{2\alpha(1-\alpha)}{\alpha^2+16}, \quad 2.5 + \frac{8(1-\alpha)}{\alpha^2+16}$$

et la valeur de la fonction objectif est alors :

$$f = 4.5 + \frac{2(1-\alpha)(4+\alpha)}{\alpha^2+16}$$

On vérifie que la valeur de la dérivée $\partial f/\partial \alpha$ évaluée au point E (c'est-à-dire pour $\alpha=1$) est :

$$\left. \frac{\delta f}{\delta \alpha} \right|_{\alpha=1} = -\frac{10}{17} = -\frac{5}{17} \cdot 2 = -\lambda_1 \cdot x_1$$

La vérification de la formule (3) est immédiate si on prend les valeurs des variables (λ, λ') dans leurs définitions (24-25-26-27) ou dans la dernière ligne du « tableau généralisé du simplexe » :

$$(1 \ 1) = \frac{5}{17}(1 \ 4) - \left(-\frac{12}{17} \ \frac{3}{17} \right)$$

2. Programme non linéaire

Considérons l'exemple non linéaire suivant, paramétré par deux coefficients $c_1 = c > 0$ et $b_3 = b > 0$:

$$\left. \begin{array}{l} \text{Max } cx_1 + x_2 \\ (x_1 + 1)^2 + x_2^2 \leq 10 \quad (1) \\ (x_1 - 2)^2 + x_2^2 \leq 5 \quad (2) \\ x_1^2 + x_2^2 \leq b \quad (3) \\ x_1 \geq 0, \quad x_2 \geq 0 \end{array} \right\}$$

Ce programme permet d'illustrer les résultats précédents lorsque varient les paramètres c et b . La figure 3 donne la représentation graphique de ce programme :

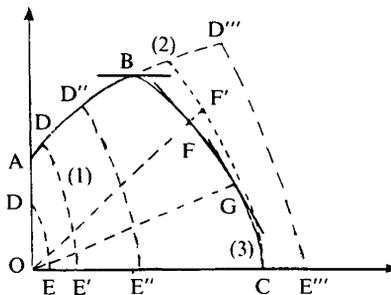


Figure 3

2.1. *Calcul des variables duales à l'optimum et en un point quelconque de statu quo*

Pour $c \geq 1/2$ la solution optimale est un en point G situé sur l'arc BC. (La solution est F pour $c = 1$ par exemple). Les coordonnées du point optimal G sont :

$$x_1^* = c \sqrt{b/(1+c^2)} \quad x_2^* = \sqrt{b/(1+c^2)}$$

et les coordonnées du point F sont $x_1 = x_2 = \sqrt{b/2}$.

Supposons que $c \neq 1$ et que le point choisi de statu quo soit précisément le point F. La contrainte (3) est saturée au point F et :

$$J = \sqrt{2b} \begin{bmatrix} 1 & 1 \end{bmatrix} \quad J^+ = \frac{1}{2\sqrt{2b}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Les formules (24 à 27) donnent les évaluations des variables duales :

$$\lambda'_B = \delta_B (J^+ J - I) = [(1-c)/2, (c-1)/2] = [\lambda'_1, \lambda'_2]$$

$\lambda'_N = \delta_B J^+ C - \delta_N$ est vide car le vecteur δ_N et la matrice C sont vides

$$\lambda_B = [0, 0] = [\lambda_1, \lambda_2]$$

$$\lambda_N = \delta_B J^+ = \frac{(1+c)}{2\sqrt{2b}} = [\lambda_3]$$

Il est clair que, pour $c \neq 1$, le point F ne peut être optimal car l'un ou l'autre de λ'_1 ou λ'_2 sera négatif. Par contre pour $c = 1$ le point F est optimal et on a $\lambda'_B = 0$. Cette condition d'optimalité s'écrit :

$$\lambda'^*_B = \delta_B J^+ J - \delta_B = 0 \quad \text{d'où} \quad \delta_B = \lambda^*_N J$$

ou

$$\text{grad } f = \lambda^*_N \cdot \text{grad } g_S$$

Pour $c \neq 1$, au point optimal G on a $J = 2 \sqrt{b/(1+c^2)} \cdot \begin{bmatrix} c & 1 \end{bmatrix}$ et

$$\lambda'^*_B = \delta_B (J^+ J - I) = [0, 0] = [\lambda'^*_1, \lambda'^*_2]$$

$$\lambda'^*_N = \delta_B J^+ C - \delta_N \text{ est vide}$$

$$\lambda^*_B = (0, 0) = [\lambda^*_1, \lambda^*_2]$$

$$\lambda^*_N = \delta_B J^+ = \sqrt{1+c^2}/\sqrt{4b} = [\lambda^*_3]$$

On vérifie que les conditions d'optimalité, $[x^*_B > 0, \lambda^*_B = 0]$, sont satisfaites; la condition $[\lambda'^*_B = 0]$ signifie qu'au point G le gradient de la fonction objectif

appartient au cône engendré par les gradients des contraintes saturées :

$$\text{grad } f = \lambda_3^* \cdot \text{grad } g_3$$

$$\begin{bmatrix} c & 1 \end{bmatrix} = \sqrt{1+c^2}/\sqrt{4b} \begin{bmatrix} 2x_1^* & 2x_2^* \end{bmatrix}$$

2.2. Tableaux du simplexe généralisés

$1/\sqrt{8b}$	0	0	0
$1/\sqrt{8b}$	0	0	0
$-1-1/\sqrt{2b}$	1	0	0
$-1+2/\sqrt{2b}$	0	1	0
$(1+c)/\sqrt{8b}$	0	0	1

δx_1	δx_2	$\delta x'_3$	$\delta x'_1$	$\delta x'_2$
$\sqrt{2b}$	$\sqrt{2b}$	1	0	0
$\sqrt{2b}+2$	$\sqrt{2b}$	0	1	0
$\sqrt{2b}-4$	$\sqrt{2b}$	0	0	1
$-c$	-1	0	0	0
				$-\delta f$

	δx_1	δx_2	$\delta x'_3$	$\delta x'_1$	$\delta x'_2$
δx_1	1/2	1/2	$1/\sqrt{8b}$	0	0
δx_2	1/2	1/2	$1/\sqrt{8b}$	0	0
$\delta x'_1$	1	-1	$-1-1/\sqrt{2b}$	1	0
x'_2	-2	2	$-1+2/\sqrt{2b}$	0	1
	$\frac{1-c}{2}$	$\frac{c-1}{2}$	$\frac{(1+c)}{\sqrt{8b}}$	0	0
					$\frac{(1+c)}{\sqrt{8b}} \delta b_3 - \delta f$
	λ'_1	λ'_2	λ_3	λ_1	λ_2

Ce tableau donne les ajustements $\delta x = \delta b_3 / \sqrt{8b} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ aux variations

$\delta b = \begin{bmatrix} \delta b_1 \\ \delta b_2 \\ \delta b_3 \end{bmatrix}$ des contraintes : le déplacement δx est orthogonal à la troisième contrainte déplacée de δb_3 .

La dernière ligne du tableau entraîne la relation

$$\delta f = c \delta x_1 + \delta x_2 = (1+c)/\sqrt{8b} \delta b_3 = \lambda_3 \delta b_3$$

qui se déduit des relations d'exclusion : cette dernière ligne représente en effet la relation

$$\frac{1-c}{2} \delta x_1 + \frac{c-1}{2} \delta x_2 + \frac{(1+c)}{\sqrt{8b}} \delta x'_2 + 0 \delta x'_1 + 0 \delta x'_2 = \frac{(1+c)}{\sqrt{8b}} \delta b_3 - \delta f$$

dont le membre de gauche est nul du fait des valeurs des variations δx et $\delta x'$ des variables x et x' .

Les valeurs des variables duales (λ, λ') se lisent dans la dernière ligne du « tableau du simplexe généralisé » et on peut aussi vérifier la formule (3) :

$$(c - 1) = \frac{1+c}{2\sqrt{2b}}(2x_1 - 2x_2) - \left(\frac{1-c}{2} \quad \frac{c-1}{2} \right)$$

Il suffit de remplacer x_1 et x_2 par leurs valeurs $(\sqrt{b/2}, \sqrt{b/2})$ au point F pour constater la validité de la formule (3).

2.3. Analyse de sensibilité

2.3.1. On déplace les contraintes de δb

On a vu que l'ajustement $\delta x = \delta b_3 / \sqrt{8b} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ s'effectue perpendiculairement à la contrainte saturée, la troisième. Le tableau du simplexe généralisé donne ces ajustements δx ainsi que la variation de la fonction objectif $\delta f = [(1+c)/\sqrt{8b}] \delta b_3$. On peut vérifier que :

$$\frac{\delta f}{\delta b_1} = \lambda_1, \quad \frac{\delta f}{\delta b_2} = \lambda_2, \quad \frac{\delta f}{\delta b_3} = \lambda_3$$

2.3.2. On fait varier un coefficient d'une contrainte

Supposons que la troisième contrainte s'écrive, par exemple :

$$ax_1^2 + x_2^2 \leq b$$

et que cette contrainte soit saturée au point de statu quo F'. Les coordonnées de ce point F' sont alors :

$$x_1 = x_2 = \sqrt{b/(1+a)}$$

et

$$J = 2\sqrt{b/(1+a)} \begin{bmatrix} a & 1 \end{bmatrix} \quad J^+ = \frac{\sqrt{1+a}}{2(1+a^2)\sqrt{b}} \begin{bmatrix} a \\ 1 \end{bmatrix}$$

Le théorème « généralisé » de l'enveloppe énonce qu'au point F (où $a = 1$) on a :

$$\left. \frac{\delta f}{\delta a} \right|_{a=1} = \left. \frac{\partial f}{\partial a} \right|_{a=1} - \sum_i \lambda_i \cdot \left. \frac{\partial g_i}{\partial a} \right|_{a=1}$$

Ceci se vérifie numériquement car $\lambda_1 = \lambda_2 = 0$, et nous avons alors :

$$\left. \frac{\delta f}{\delta a} \right|_{a=1} = \left. \frac{\partial f}{\partial a} \right|_{a=1} - \lambda_3 \cdot \left. \frac{\partial g_3}{\partial a} \right|_{a=1}$$

$$f = cx_1 + x_2 = (1+c)[(1+a)/b]^{-1/2}$$

$$\frac{\delta f}{\delta a} = -\frac{1}{2}(1+c)b^{1/2}(1+a)^{-3/2}$$

$$\left. \frac{\delta f}{\delta a} \right|_{a=1} = -\frac{1}{4}(1+c)(b/2)^{1/2}$$

$$J = \begin{bmatrix} 2ax_1 & 2x_2 \end{bmatrix} = 2[b/(1+a)]^{-1/2} \begin{bmatrix} a & 1 \end{bmatrix}$$

$$J^+ = \frac{1}{2(1+a^2)} [(1+a)/b]^{1/2} \begin{bmatrix} a \\ 1 \end{bmatrix}$$

$$\lambda_3 = \delta_B J^+ = \frac{1+ac}{2(1+a^2)} [(1+a)/b]^{1/2}$$

$$\frac{\partial g_3}{\partial a} = x_1^2 = \frac{b}{1+a}, \quad \frac{\partial f}{\partial a} = 0,$$

$$\frac{\delta f}{\delta a} - \lambda_3 \frac{\partial g_3}{\partial a} = -\frac{1+ac}{2(1+a^2)} [(1+a)/b]^{-1/2}$$

Il s'ensuit que :

$$\left. \frac{\partial f}{\partial a} \right|_{a=1} - \lambda_3 \cdot \left. \frac{\partial g_3}{\partial a} \right|_{a=1} = -\frac{1+c}{4} [b/2]^{1/2}$$

Cette égalité vérifie le théorème de l'enveloppe dans cet exemple de programmation non linéaire.

optimale, *i. e.* $x = x^*$. On sait en effet que x^* est tel que

$$u'_1 = \lambda^* p_1 \quad \text{et} \quad u'_2 = \lambda^* p_2$$

Le report dans les formules de définition (24-25-26-27) de λ donne immédiatement :

$$\lambda = \frac{\lambda^* p_1^2 + \lambda^* p_2^2}{p_1^2 + p_2^2} = \lambda^*$$

2. On peut aussi vérifier la validité de la formule (3) sur l'exemple à deux biens de consommation. Nous allons la vérifier dans le cas de n biens.

$$\begin{aligned} J^+ J - I &= \frac{p^T p}{\|p\|^2} - I \\ \lambda' &= (\lambda'_B, \lambda'_N) = \left(u' \frac{p^T p}{\|p\|^2} - u', \text{ vide} \right) \\ \lambda &= (\lambda_B, \lambda_N) = \left(0, \frac{u' p^T}{\|p\|^2} \right) \\ u' &= \left[\frac{u' p^T}{\|p\|^2} p - \left(u' \frac{p^T p}{\|p\|^2} - u' \right) \right] \end{aligned} \quad (3)$$

Interprétation : On peut supposer que le consommateur ait, à une certaine époque $t=0$, où il disposait d'un revenu r , pour des prix p_1^0 et p_2^0 des deux biens, calculé sa consommation optimale et obtenu un panier de consommation x . Puis les prix changent, devenant p_1 et p_2 , et son revenu ne change pas. Pour une raison ou une autre (manque de temps, inefficacité du marché, indisponibilité de moyens de calcul numérique, ...), le consommateur n'ajuste pas sa consommation au nouvel optimum x^* et garde (si c'est possible) son ancienne consommation optimale x comme point de statu quo. Sa consommation n'est donc plus optimale pour le nouveau système de prix. Si, à ce moment, son revenu change, passant de r à r_1 , les prix demeurant inchangés, que fera-t-il ? Dans la théorie classique de la consommation, le consommateur recalcule son nouveau panier de consommation optimal x^{1*} . Ce n'est probablement pas ce qui se passe dans la vie réelle car le consommateur ne peut pas recalculer perpétuellement son optimum x^{1*} de consommation. Nous envisageons le cas où il ne recalcule pas son optimum x^{1*} mais agit de façon conservatrice (comportement à inertie). Le consommateur doit adapter sa consommation selon ses disponibilités et en respectant la nouvelle contrainte de budget mais *il change sa consommation aussi peu que possible*

par rapport à sa consommation antérieure x . Dans ce cas le consommateur minimise $\|\delta x\|$ et sa consommation devient x^1 . Ce comportement justifie cependant le calcul des variables duales λ qui conservent leur signification économique, à savoir l'utilité marginale du revenu, *quelle que soit la fonction d'utilité u du consommateur*. On remarque que le même comportement, si le consommateur avait calculé son optimum de consommation préalable x^* , l'aurait conduit à passer au point x^2 sur la figure 4, et non pas à son nouvel optimum de consommation x^{1*} .

Remarque : On peut vérifier analytiquement la sensibilité des multiplicateurs λ définis par les équations (24-25-26-27) aux variations $u' = \text{grad } u$ de la fonction d'utilité :

$$\frac{\partial \lambda}{\partial u'} = p^+ = \frac{p}{\|p\|^2} \quad \text{ou} \quad \frac{\partial \lambda}{\partial u_i} = \frac{p_i}{\|p\|^2}$$

Ce résultat est à comparer au résultat équivalent de la théorie classique, déduit de la condition d'optimalité $u_i = \lambda^* p_i$:

$$\frac{\partial \lambda^*}{\partial u_i} = \frac{1}{p_i}$$

CONCLUSION

Les développements précédents conduisent à définir la dualité en tout point réalisable en programmation mathématique – linéaire et non linéaire – à l'aide de la matrice M des gradients des contraintes saturées en ce point sous la forme :

$$(\lambda, \lambda') = \delta \cdot M^+$$

où $\delta = \text{grad } f$ et où M^+ est une inverse à droite quelconque de M . Cependant, pour des raisons dont l'interprétation économique est évidente, parmi toutes les inverses à droite M^+ de la matrice M il est préférable de choisir l'inverse à droite de Moore-Penrose; celle-ci d'ailleurs, par son unicité, conduit à l'unicité de la définition de la dualité en tout point réalisable par

$$(\lambda, \lambda') = \delta \cdot M^+ = \delta \cdot M^T (MM^T)^{-1}$$

La contrepartie de ce choix est que le déplacement à partir de la solution de statu quo est, en analyse de sensibilité, *orthogonal à l'intersection des contraintes saturées*. Ce changement se justifie par le fait que le nouveau

point de statu quo n'a pas plus de raisons que l'ancien d'être optimal : ce comportement exprime le conservatisme bien connu des comportements humains.

Cette dualité généralise au cas non linéaire – en retrouvant à l'optimum les multiplicateurs de Kuhn-Tucker – la définition des valeurs duales pour un point sommet, c'est-à-dire pour une base non optimale, définies en programmation linéaire (*cf.* Baumol-Fabian). Elle permet en outre de définir les valeurs duales ailleurs qu'en un sommet : sur une arête, sur une face du domaine réalisable, tant en programmation non linéaire qu'en programmation linéaire.

Ces évaluations duales ont la même interprétation économique que les variables duales classiques, multiplicateurs de Lagrange ou multiplicateurs de Kuhn-Tucker. Elles vérifient des relations d'exclusion tout comme les variables duales classiques d'un programme linéaire (et on les retrouve d'ailleurs comme cas particulier) et comme les multiplicateurs de Kuhn-Tucker (que l'on retrouve aussi comme cas particulier en programmation non linéaire).

On peut étendre les tableaux du simplexe ailleurs qu'en un point extrême. Il est enfin possible d'étendre le *théorème de l'enveloppe* de telle sorte que son utilisation en analyse de sensibilité reste valable.

L'intérêt qu'il y a à connaître les valeurs des variables duales à l'optimum est universellement reconnu par les utilisateurs de techniques d'optimisation en planification et gestion économiques. Cet intérêt n'est pas ici remis en cause, mais plutôt l'intérêt réel de l'optimum lui-même, qu'il faut connaître et atteindre pour en déduire une évaluation des « ressources » et des « activités » par les variables duales.

(i) Cet optimum est-il correctement défini ? N'a-t-on pas oublié une ou plusieurs contraintes ? A-t-on bien recensé l'ensemble des activités mises en œuvre ? Certaines activités peuvent être en apparence secondaires et pourtant souvent indispensables aux activités retenues ; elles peuvent aussi être un sous-produit de la mise en œuvre des activités retenues ; l'intégration dans l'objectif du programme des rendements de ces activités mènerait probablement à un autre optimum, ce qui changerait toutes les évaluations duales attachées à cet optimum.

(ii) Cet optimum est-il accessible ? La grande quantité d'algorithmes proposés en programmation non linéaire ainsi que les mesures de leurs performances respectives montrent que l'on n'a pas vraiment de procédure systématique efficiente de recherche de l'optimum sauf dans certains cas plus

simples. D'ailleurs cet optimum n'est-il pas seulement un optimum local? Tout ceci lui ôte beaucoup d'intérêt par rapport à une solution courante, considérée comme solution de statu quo et qui donne probablement une valeur de l'objectif assez correcte par rapport à celle que donnerait l'optimum.

(iii) Les incertitudes enfin sur les paramètres du modèle rendent illusoire la nécessaire définition précise et rigoureuse de cet optimum. Bien sûr, toutes les analyses de sensibilité essayent de venir à l'aide de l'optimum en essayant de sécuriser dans l'incertitude des paramètres, mais on ne peut être incertain que sur l'un d'eux en même temps; l'on est pourtant souvent incertain sur plusieurs d'entre eux à la fois.

(iv) Dans la pratique cherche-t-on vraiment à optimiser un objectif unique? Dans le cas d'une firme par exemple il semble que les praticiens se fixent plusieurs objectifs qu'ils cherchent à atteindre au plus près; des techniques du type programmation multi-objectif, goal programming ou analyse multicritère semblent mieux adaptées à leurs problèmes de décision que la recherche d'un optimum d'un objectif unique. Cet optimum perd par là-même beaucoup de sa crédibilité.

(v) Les ressources disponibles en quantités supérieures à ce qui est nécessaire à l'optimum ont une valeur duale nulle et elles ne prennent brutalement de valeur que lorsque la contrainte correspondante, en se resserrant, se sature dans la diminution de ces ressources. Cette brutale discontinuité est assez contestable dans la pratique où une ressource utile n'est jamais vraiment excédentaire et on devrait l'évaluer — non pas à sa valeur de marché — mais à une valeur dépendant de l'usage qui en est fait en liaison avec les autres ressources, sans pour autant lui donner une valeur nulle sous le prétexte que les autres ressources sont plus difficiles à acquérir qu'elle-même. Il est important de reconnaître qu'en pratique on peut souvent desserrer TOUTES les contraintes et qu'une situation figée comme l'optimum est trop manichéenne: l'acier ne vaut rien s'il reste un gramme d'acier à l'optimum parce que l'énergie et la main-d'œuvre étaient insuffisantes pour utiliser à bon escient ce gramme d'acier!! A-t-on vraiment mesuré la quantité consommée d'acier avec une précision indiscutable? Ne pouvait-on pas obtenir quelques minutes de travail supplémentaires et trouver quelques kilowatts de plus?... La réponse est généralement oui car on peut toujours acquérir des ressources même à des taux plus élevés.

En réalité ne peut-on reconnaître que la planification ou la gestion ne sont pas optimales? Parce que l'optimum est incertain, parce qu'il est coûteux à obtenir et parce que souvent il n'apporte pas énormément par rapport à une

solution heuristique, il semble qu'il faut accepter de reconnaître que la solution mise en œuvre est non seulement réalisable – et pour cause puisque c'est finalement celle qui est mise en pratique – mais aussi non optimale.

Mais il serait alors fâcheux de ne pas disposer de l'apport économique des évaluations duales. Comment les définir et les déterminer si l'on n'est pas à l'optimum? Comment faire en sorte que les ressources indispensables aient tout de même une valeur même si d'autres ressources sont plus contraignantes? En dehors d'une disponibilité vraiment excédentaire (comme l'air et l'eau dans les problèmes de production de nature économique, car ils peuvent être des facteurs limitatifs dans les problèmes techniques de production), toute ressource utile doit être valorisée en tenant compte de ce que l'on peut, en fait, acquérir pratiquement toutes les ressources.

Dans ces conditions il apparaît plus correct de :

(i) reconnaître que la solution de *statu quo* n'est pas optimale ;

(ii) mettre en seconds membres des contraintes du modèle, non pas les quantités « disponibles » des ressources, puisque l'on peut souvent en acquérir davantage, mais les quantités *effectivement utilisées* dans la mise en œuvre du *statu quo* x . Si vraiment certaines ressources ne s'y prêtent pas parce que leurs disponibilités sont sans limites, alors elles auront des évaluations duales nulles. En dehors de ces dernières contraintes on considèrera que *toutes les contraintes sont saturées* et leurs gradients, regroupés en une matrice M , permettront de définir les évaluations duales des ressources par λ_N et les évaluations duales des activités par λ'_N définies par

$$(\lambda_N, \lambda'_N) = \delta M^+ = M^T (MM^T)^{-1}.$$

Dès lors, pour évaluer les ressources par λ , il suffit de recenser les contraintes où ces ressources interviennent et, pour évaluer les activités par λ' il suffit de constater si elles sont ou non mises en œuvre en x . Les évaluations (λ, λ') ont les mêmes propriétés que les variables duales définies en programmation linéaire ou non linéaire.

BIBLIOGRAPHIE

1. M. AOKI, Introduction to Optimization Techniques, *Mac Millan*, 1971.
2. W. J. BAUMOL et T. FABIAN, Decomposition Pricing for Decentralization and External Economies, *Management Sci.*, 11, (1), September 1964.
3. J. BOUZITAT, Emploi des matrices pseudo-inverses en programmation quadratique convexe (contraintes quadratiques), *Cah. Groupe Math. Econ.*, n° 3, juin 1980.
4. S. L. CAMPBELL et C. D. MEYER, Generalized Inverses of Linear Transformations, *Pittman*, 1979.

5. A. V. Fiacco, Introduction to Sensitivity and Stability Analysis in Nonlinear Programming, *Academic Press*, 1983.
6. Y. Ijiri, Objectifs et contrôle de gestion, *Dunod*, 1970.
7. D. Lacaze, Prix duaux et dégénérescence en programmation mathématique, *Cah. Groupe Math. Econ.*, n° 8, avril 1972.
8. L. D. Pyle, The Generalised Inverse in Linear Programming. Basic Structure, *SIAM*, 22, n° 3, May 1972.
9. L. D. Pyle et R. E. Cline, The Generalised Inverse in Linear Programming. Interior Gradient Projection Method, *SIAM*, 24, n° 4, June 1973.
10. J. B. Rosen, The Gradient Projection Method for Non Linear Programming. *SIAM*, 1960.
11. G. Strang, Introduction to Applied Mathematics, *Wellesley Cambridge Press*, 1986.
12. A. Takayama, Mathematical Economics, 2nd ed, *Cambridge University Press*. 1988.
13. D. J. Wilde et C. S. Beightler, Foundations of Optimization, *Prentice Hall*, 1967.