

PROBLÈME D'IDENTIFICATION DE MODÈLES DE RÉGRESSION ET SA SOLUTION PAR LA PROGRAMMATION LINÉAIRE¹

Efim A. Galperin et Richard Labonté

Résumé

Etant donné un phénomène physique, économique ou autre, on cherche souvent à construire un modèle représentatif sous forme de système dynamique à coefficients constants ou de système d'équations aux différences finies. On recherche ici, à partir d'une suite d'observations équidistantes dans le temps relatives au phénomène, une équation aux différences à coefficients constants d'ordre minimal ayant la propriété de reproduire les données connues en respectant une certaine précision exigée à l'avance. Le problème est résolu par des méthodes finies de programmation linéaire; il est formulé dans la norme ℓ_∞ comme celui de l'approximation linéaire de Chebyshev qui peut être réduit à un programme linéaire particulier. Dans le cas où les données récentes ont plus d'influence sur le comportement futur du phénomène que les données antérieures, le problème est formulé dans la norme $\ell_{1,\rho}$ et est réduit à un autre programme linéaire. Ces deux programmes étant de structure spéciale, on développe pour chacun un algorithme particulier pour accélérer le processus de résolution sur ordinateur. Sous certaines conditions, ces algorithmes fournissent une méthode efficace de prédiction des données futures dans la classe des modèles de régression associés aux données observées ainsi qu'un instrument d'identification de la nature linéaire stationnaire (si tel est le cas) du phénomène considéré.

¹ Recherche supportée par les subventions A-3492 du Conseil de recherches en sciences naturelles et en génie du Canada et Z-215 de l'Université du Québec à Montréal

1. Introduction

Etant donné une suite d'observations, on recherche habituellement un modèle à coefficients constants de la forme

$$\frac{dx}{dt} = Ax, \quad x(t_0) = x_0, \quad t \geq t_0 \quad (1)$$

$$y(t) = hx, \quad x \in \mathbb{R}^m, \quad y \in \mathbb{R}$$

où $A = A(p)$, $x_0 = x_0(p)$, $h = h(p)$ où $p = (p_1, \dots, p_k)$ est un vecteur de paramètres à identifier en assujettissant la fonction vectorielle $x(t, p) \in \mathbb{R}^m$ aux données observées $y_1, \dots, y_s$ ($y_i = y(t_i)$). Pour ce faire, la méthode des moindres carrés est fréquemment employée et s'avère efficace si, à partir d'une approximation initiale p_0 , les itérations successives convergent vers le vecteur de paramètres unique p^* . Mais dans des cas complexes, il convient de vérifier à prime abord s'il existe un modèle du genre, et si oui, dans quel intervalle de temps il est acceptable.

Supposons les observations *équidistantes dans le temps*. A ce moment, le système discret

$$x_{n+1} = Fx_n, \quad F = \exp(A\Delta t) = \sum_{k=0}^{\infty} \frac{(A\Delta t)^k}{k!} \quad (3)$$

$$y_n = hx_n, \quad x_n \in \mathbb{R}^m, \quad y_n \in \mathbb{R}, \quad n = 0, 1, \dots \quad (4)$$

où $x_n = x(t_n)$, $t = t_0 + n\Delta t$, $y_n = y(t_n)$ (les données) avec A et Δt constants (c.-à-d. F constant), génère les valeurs exactes de $x(t)$ et $y(t)$ du système (1)-(2) aux temps $t_n = t_0 + n\Delta t$. En ne prenant que les deux premiers termes de la série en (3), c.-à-d. $F^* = I + \Delta tA$, alors le système (3) approxime discrètement le système (1), en ce sens que les valeurs de x_n^* tendent vers celles de $x(t_n)$ lorsque Δt tend vers zéro. Donc, si seules des observations discrètes sont disponibles, alors le système (3)-(4) fournit une description complète et exacte du comportement du système (1)-(2) par rapport aux informations contenues dans la suite de données discrètes.

On sait [1,2] qu'un modèle à coefficients constants de la forme (1)-(2) ou (3)-(4) d'ordre m existe *si et seulement si* les observations $y_0, y_1, \dots$ satisfont une équation aux différences à coefficients constants de la forme

$$y_{n+r} = a_1 y_n + a_2 y_{n+1} + \dots + a_r y_{n+r-1} \quad (r \leq m) \quad (5)$$

$$a_i = \text{constante} \quad (i = 1, \dots, r) ; \quad n = 0, 1, 2, \dots$$

En posant

$$\bar{x}_n = (y_n, y_{n+1}, \dots, y_{n+r-1}) , \quad (6)$$

on peut réécrire (5) sous la forme

$$\bar{x}_{n+1} = Q \bar{x}_n , \quad \bar{x}_n \in R^r \quad (r \leq m) \quad (7)$$

$$y_n = h \bar{x}_n , \quad y_n \in R , \quad h = (1, 0, \dots, 0) \quad (8)$$

où

$$Q = \begin{bmatrix} 0 & & & \\ & \ddots & & \\ & & I_{r-1} & \\ a_1 & a_2 & \dots & a_r \end{bmatrix} \quad (9)$$

dont la structure est essentiellement de la forme (3)-(4) mais dans un système de coordonnées différent [2].

Si un modèle existe, il existe toujours une infinité d'autres modèles équivalents par rapport aux "informations" contenues dans la suite de données. Ces modèles forment une classe d'équivalence qui contient des systèmes discrets ou continus, d'ordres et de dynamiques variés, mais tous équivalents au point de vue information. N'importe quel modèle peut être choisi comme représentant de sa classe; par conséquent, nous nous restreindrons dorénavant au modèle le plus simple, celui de la forme (5).

2. Modèles de régression

DÉFINITION 1. Un *modèle dynamique* est une égalité du type (5) qui tient exactement pour quelque valeur de r , dite l'*ordre* du modèle.

Tout modèle dynamique est représentable sous forme (1)-(2) ou (3)-(4). Si (5) ne tient pour aucune valeur de r , alors il n'existe pas de modèle dynamique linéaire à coefficients constants pouvant être associé à la suite de données considérée. Dans ce cas, nous introduisons le modèle fourni par les relations suivantes [3]:

DÉFINITION 2. Un *modèle de régression* est une égalité du type

$$\hat{y}_{n+r} = a_1 y_n + a_2 y_{n+1} + \dots + a_r y_{n+r-1} \quad (n = 0, 1, \dots, N-r) \quad (10)$$

où

$$\eta_n(r, a) = y_{n+r} - \hat{y}_{n+r} ; \quad a = [a_1, \dots, a_r], \quad (11)$$

$$|\eta_n(r, a)| \leq \mu, \quad n = 0, 1, \dots, N-r \quad (12)$$

et $y_0, y_1, \dots, y_N$ sont les *observations* (connues); r est l'*ordre* du modèle; μ est la *précision* du modèle (c.-à-d. l'écart maximal permis entre les valeurs réelles des données et celles reproduites par le modèle de régression); $a = [a_1, \dots, a_r]$ est la *dynamique* du modèle.

Si $\mu = 0$, alors le modèle de régression devient un modèle dynamique, car en ce cas il est régi par un système dynamique ou un système d'équations aux différences. On suppose ici, comme c'est le cas en pratique, qu'un nombre fini d'observations (disons $N+1$) est disponible.

On peut toujours trouver une valeur de r , soit r_0 , pour laquelle il existe un modèle (c.-à-d. un vecteur " a ") avec $\mu = 0$ et ce, quelles que soient les données $y_0, y_1, \dots, y_N$ [3]. Par contre, pour certaines suites d'observations, étant donné $\mu > 0$ et $r < r_0$, aucun modèle de régression n'existe quelles que soient les composantes $a_1, \dots, a_r$ du vecteur " a ". Le problème qui se pose est

donc le suivant: considérant une suite d'observations $y_0, y_1, \dots, y_N$, trouver $r_{\min}$ et les composantes du vecteur $a = [a_1, \dots, a_r]$ tels que le modèle de régression respecte la précision exigée $\mu > 0$. Avec un modèle adéquat, on pourrait ainsi prédire avec confiance les valeurs futures de la suite de données.

2.1. Modèles dans la norme ℓ_∞

A partir de (12), on pourrait tout naturellement exiger que l'erreur sur la reproduction des données connues soit uniformément aussi faible que possible et respecte la précision $\mu > 0$. En introduisant la fonction

$$\xi(r) = \inf_a \|\eta_n(r, a)\|_\infty = \inf_a \max_{0 \leq n \leq N-r} |\eta_n(r, a)| \leq \mu \quad (13)$$

deux problèmes se posent:

- (a) Etant donné r , évaluer $\xi(r)$ et identifier $a_0 = \arg \inf \|\eta_n(r, a)\|_\infty$ pour obtenir les coefficients d'un modèle de régression d'ordre r , si $\xi(r) \leq \mu$.
- (b) Si $\xi(r) > \mu$, trouver r^* (s'il existe) tel que $\xi(r^*) \leq \mu$ pour connaître l'ordre d'un modèle de régression acceptable.

Comme la fonction $\xi(r)$ de (13) est non croissante [3], alors le problème (b) est facilement résoluble si on a une méthode efficace pour résoudre le problème (a).

A remarquer que l'on aurait pu formuler le problème selon la méthode classique des moindres carrés (en norme euclidienne). Cette approche permet également d'identifier un modèle de régression obtenu à l'aide d'une matrice pseudo-inverse. Le désavantage de cette méthode réside dans le fait qu'elle minimise la *moyenne quadratique* des écarts, et non pas l'*écart maximal* qui peut être prononcé même si d'autres sont nuls. Par la méthode présentée ici, on obtient un modèle pour lequel les déviations sont uniformément "petites".

2.2. Programme linéaire correspondant au modèle en ℓ_∞

Le problème (a) de la section 2.1 est celui de l'approximation de Chebyshev résoluble par des méthodes de programmation linéaire finies. Considérant la fonction $\xi(r)$ de (13) comme borne inférieure d'une variable supplémentaire a_{r+1} inconnue, on a ainsi:

$$\xi(r) = \inf_a a_{r+1}(a) , \quad a = [a_1, \dots, a_r] . \quad (14)$$

De (14) et (13), on aboutit à l'inégalité

$$a_{r+1} = \max_{0 \leq n \leq N-r} |\eta_n(r, a)| \geq |\eta_n(r, a)| , \quad n = 0, \dots, N-r . \quad (15)$$

De (10)-(11) et (14)-(15), on obtient le programme linéaire suivant à résoudre pour trouver la dynamique $[a_1, \dots, a_r]$ du modèle de régression:

$$\min \xi = a_{r+1} , \quad (16)$$

$$\text{sous les contraintes} \quad \left\{ \begin{array}{l} a_{r+1} - \eta_n \equiv a_{r+1} + \sum_{i=1}^r y_{n+i-1} a_i - y_{n+r} \geq 0 , \\ a_{r+1} + \eta_n \equiv a_{r+1} - \sum_{i=1}^r y_{n+i-1} a_i + y_{n+r} \geq 0 , \end{array} \right. \quad (17)$$

$$\left\{ \begin{array}{l} a_{r+1} \geq 0 \quad (n = 0, 1, \dots, N-r) . \end{array} \right. \quad (18)$$

On ne doit pas exiger ici la positivité des variables $a_1, \dots, a_r$. Le problème est toujours résoluble et sa solution fournit un modèle de régression optimal dans le cas où toutes les données ont la même importance.

2.3. Modèles dans la norme $\ell_{1,\rho}$

Dans certains contextes (en économie, par exemple), il convient de laisser tomber l'uniformité relative à l'usage de la norme ℓ_∞ de (13), car les valeurs récentes de la suite peuvent avoir une influence plus prononcée sur les valeurs futures que les données antérieures éloignées dans le temps.

On tient compte de ceci en utilisant une suite pondérée positive $\{\rho_n\}_{n=0}^{N-r}$, $\rho_1 > 0$, ce qui conduit naturellement à l'utilisation de la norme $\ell_{1,\rho}$. La relation (13) est donc remplacée maintenant par

$$\xi_1(r) = \inf_a \|\eta_n(r,a)\|_{1,\rho} = \inf_a \sum_{n=0}^{N-r} \rho_n |\eta_n(r,a)| \leq \mu_1. \quad (20)$$

En introduisant $N-r+1$ variables auxiliaires $a_{r+1}, \dots, a_{N+1}$ comme bornes supérieures des écarts

$$|\eta_n(r,a)| \leq a_{n+r+1} \quad (n = 0, \dots, N-r), \quad (21)$$

on peut réduire (20) à un simple problème de programmation linéaire. Ici encore, il s'avère que la fonction $\xi_1(r)$ est non croissante, si on utilise des suites pondérées compatibles [3]. Donc, le problème est d'évaluer $\xi_1(r)$ et d'identifier le vecteur $a = [a_1, \dots, a_r]$.

2.4. Programme linéaire correspondant au modèle en $\ell_{1,\rho}$

De (21), on tire:

$$a_{n+r+1} - \eta_n \geq 0 \quad (n = 0, \dots, N-r), \quad (22)$$

$$a_{n+r+1} + \eta_n \geq 0 \quad (n = 0, \dots, N-r). \quad (23)$$

De (20), (22) et (23), on a le programme suivant:

$$\min \xi_1(r) = \sum_{n=0}^{N-r} \rho_n a_{n+r+1}, \quad (24)$$

$$\text{sous les contraintes} \quad \left\{ \begin{array}{l} a_{n+r+1} - \eta_n \equiv a_{n+r+1} + \sum_{i=1}^r y_{n+i-1} a_i - y_{n+r} \geq 0, \end{array} \right. \quad (25)$$

$$\left\{ \begin{array}{l} a_{n+r+1} + \eta_n \equiv a_{n+r+1} - \sum_{i=1}^r y_{n+i-1} a_i + y_{n+r} \geq 0, \end{array} \right. \quad (26)$$

$$a_{n+r+1} \geq 0, \quad n = 0, 1, \dots, N-r. \quad (27)$$

Ce problème est également toujours résoluble.

2.5. Remarques sur l'application de la méthode du simplexe

On peut grossièrement estimer la complexité de la résolution d'un programme linéaire en calculant le nombre de solutions de base qu'il présente. Examinons simplement le programme obtenu à partir de la norme ℓ_∞ .

On doit d'abord introduire à la place de chacune des variables a_i ($i = 1, \dots, r$) sans restriction de signe une différence de deux variables $a'_i - a''_i$ ($i = 1, \dots, r$) desquelles on devra exiger $a'_i \geq 0$, $a''_i \geq 0$, $i = 1, \dots, r$. Avec la variable originale $a_{r+1} \geq 0$, on obtient ainsi $2r+1$ variables principales. De plus, à partir de (17)-(18), on doit ajouter $2(N-r+1)$ variables d'écart ainsi que $N-r+1$ variables artificielles pour obtenir une solution réalisable initiale; le total des variables est ainsi porté à $3N-r+4$. Comme on a $2(N-r+1)$ contraintes, le nombre total de solutions de base $K(N, r)$ est donné par

$$K(N, r) = \binom{3N-r+4}{2N-2r+2} = \frac{(3N-r+4)!}{(2N-2r+2)!(N-r+2)!}.$$

Numériquement, par exemple, avec $N = 20$ observations et $r = 1, \dots, 5$ pour identifier les modèles d'ordre 1 à 5, on aura $K(20, 1) \cong 9,4 \times 10^{16}$, $K(20, 2) \cong 9,7 \times 10^{16}$, $\dots$, $K(20, 5) \cong 4,8 \times 10^{16}$. Ainsi, l'approche directe aux deux problèmes originaux est peu efficace et on aura recours à des algorithmes spéciaux pour accélérer le processus de résolution sur ordinateur.

3. Programme dual (norme ℓ_∞)

En réécrivant (17)-(18) sous la forme

$$\sum_{i=1}^r y_{n+i-1} a_i + a_{r+1} \geq y_{n+r}, \quad (28)$$

$$-\sum_{i=1}^r y_{n+i-1} a_i + a_{r+1} \geq -y_{n+r}, \quad (29)$$

$$(n = 0, 1, \dots, N-r),$$

on associe à chacune des contraintes (28) une variable duale u_n et à chacune des

contraintes (29) une variable duale v_n ($n = 0, 1, \dots, N-r$). Comme $a_{r+1} \geq 0$ et $a_1, \dots, a_r$ sont sans restriction de signe, on aboutit au problème dual suivant [4]:

3.1. Formulation du dual

$$\max Z = y_r(u_0 - v_0) + y_{r+1}(u_1 - v_1) + \dots + y_N(u_{N-r} - v_{N-r}), \quad (30)$$

$$\text{sous les contraintes} \left\{ \begin{array}{l} \sum_{n=0}^{N-r} y_{n+i-1}(u_n - v_n) = 0, \quad i = 1, \dots, r, \quad (31) \\ \sum_{n=0}^{N-r} (u_n + v_n) \leq 1, \quad (32) \\ u_n \geq 0, \quad v_n \geq 0 \quad (n = 0, 1, \dots, N-r). \quad (33) \end{array} \right.$$

Le programme implique $2(N-r+1)$ variables principales et $r+1$ contraintes; on doit de plus ajouter r variables artificielles et une variable d'écart pour obtenir une solution initiale, d'où un nombre total de $2N-r+3$ variables. On a ainsi le nombre total de solution de base:

$$K^*(N, r) = \frac{(2N-r+3)!}{(r+1)!(2N-2r+2)!} \ll K(N, r)$$

qui est beaucoup moindre que celui du programme primal.

Par exemple, pour $N = 20$, $r = 1, \dots, 5$, on a:

$K^*(20, r)$	861	10 660	91 390	575 757	2 760 681
r	1	2	3	4	5

La complexité du problème est réduite de beaucoup; cependant, certains critères spéciaux de sélection des nouvelles variables de base devront être appliqués plus loin.

3.2. Interprétation géométrique

Les contraintes d'inégalité (32)-(33) restreignent le domaine du problème à la "pyramide" régulière de dimension $2(N-r+1)$ de côté "1" dans la région où toutes les variables sont positives. Les sommets de cette pyramide sont l'origine et tous les points sur les axes de coordonnées à distance unité de l'origine. La

pyramide est évidemment indépendante des données $y_0, y_1, \dots, y_N$ mais les paramètres N , le nombre de données, et r , l'ordre du modèle, déterminent ses dimensions. Par exemple, pour $N = 3$, $r = 1$, la pyramide est de dimension $2(N-r+1) = 6$, comme il est illustré à la Figure 1.

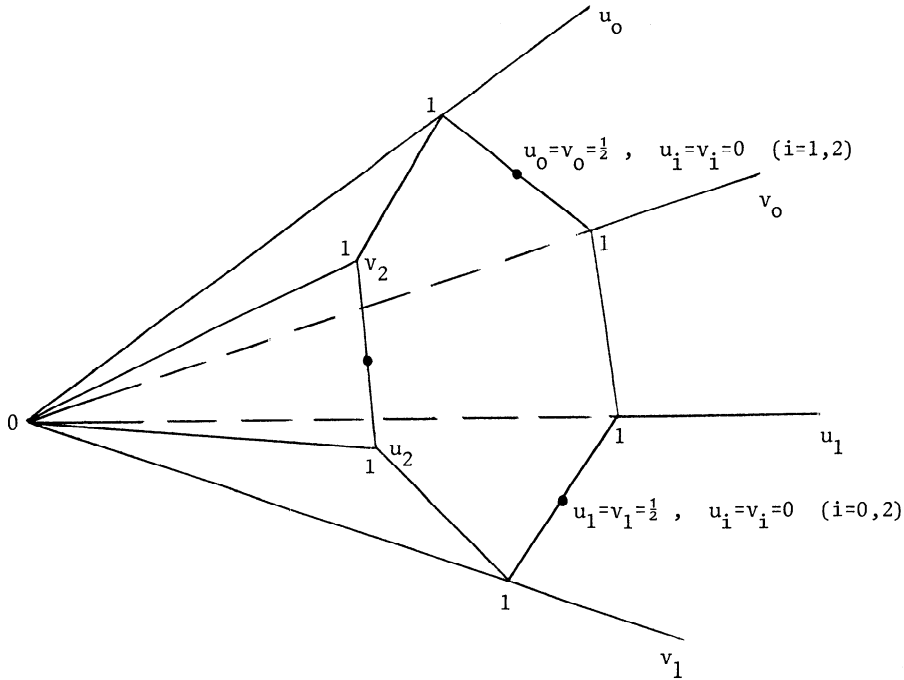


Figure 1

Les contraintes d'égalité (31) font intersection avec les faces de la pyramide et en distinguent un simplexe fermé de dimension $2N-3r+2$ dont le sommet principal est à l'origine. Le plan $Z = 0$ de (30) passe par l'origine et coupe chacune des faces de la pyramide à 45° . Le dual est donc toujours résoluble (l'ensemble des solutions réalisables est borné et évidemment non vide) et on peut toujours trouver un modèle de régression qui respectera ou non la précision exigée. Remarquons aussi que les constatations géométriques précédentes sont indépendantes des données $y_0, y_1, \dots, y_N$ elles-mêmes. Ce qui dépend des données, ce sont les directions des plans de (30)-(31). Par rapport à ces directions, la solution (le maximum) peut se trouver soit à l'origine avec $\max Z = 0$, ce qui fournit un modèle dynamique, soit sur la base de la pyramide avec $\max Z > 0$. Les points symétriques sur la base

(par exemple, $u_0 = v_0 = \frac{1}{2}$, $u_i = v_i = 0$, $i = 1, \dots, N-r$) pour lesquels $Z = 0$ fournissent une solution réalisable indépendante des données, mais pas nécessairement la solution optimale pour maximiser la valeur de la fonction économique.

3.3. Algorithmes de résolution du dual

On suppose ici, du lecteur, une bonne connaissance de la méthode du simplexe. Notons w_{ij} les éléments de la matrice de (31) augmentée de la rangée de "1" de (32) et de la matrice identité d'ordre $(r+1)$ dont les r premières colonnes correspondent aux variables artificielles et la dernière à la variable d'écart de (32):

$$Z \Rightarrow y_r, \dots, y_N, -y_r, \dots, -y_N \quad (34)$$

$$W = [w_{ij}] = \begin{bmatrix} y_0 & \dots & y_{N-r} & -y_0 & \dots & -y_{N-r} & \vdots \\ \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ y_{r-1} & \dots & y_{N-1} & -y_{r-1} & \dots & -y_{N-1} & \vdots \\ 1 & \dots & 1 & 1 & \dots & 1 & \vdots \end{bmatrix} I_{r+1} \quad (35)$$

La ligne (34), les coefficients de la fonction économique, est complétée par $(r+1)$ zéros pour accéder à la dimension $2N-r+3$ de (35). On notera par C le vecteur de dimension $(r+1)$ qui comporte les coefficients de la fonction économique reliés aux variables de base (évidemment, C varie à chaque fois que la base change, c.-à-d. que les positions des colonnes de la matrice identité varient quand on pivote sur un élément de W). On notera les éléments de (34) par Z_j ($j = 0, 1, \dots, 2N-r+3$) et les colonnes de W par W_j ($j = 1, \dots, m+r+1$ avec $m = 2(N-r+1)$).

Comme il est courant dans l'application de la méthode du simplexe, on introduit les scalaires:

$$p_j = C W_j \quad (j = 1, \dots, m+r+1), \quad (36)$$

$$g_j = Z_j - p_j \quad (j = 1, \dots, m).$$

Remarquons que le vecteur ressource coïncide avec W_{m+r+1} , de sorte qu'il est

inutile de l'introduire dans le tableau. Dans le tableau initial, on aura ainsi: $C = 0$, $p_j = 0$ et $g_j = Z_j$. Par (34), il existe assurément des valeurs Z_j positives de sorte que le tableau initial n'est certainement pas optimal; des changements de base sont ainsi nécessaires pour aboutir au critère d'optimalité $g_j \leq 0$ ($j = 1, \dots, m$). On donne plus loin des méthodes itératives accélérant le processus d'optimisation du tableau. Supposons que le tableau optimal ait été obtenu; on retrouve alors les valeurs des variables primales par

$$a_i = p_{m+i} \quad (i = 1, \dots, r), \quad (38)$$

$$a_{r+1} = \max Z = p_{m+r+1} \geq 0. \quad (39)$$

Ceci fournit la dynamique $[a_1, \dots, a_r]$ et la précision a_{r+1} d'un modèle de régression d'ordre r (dans la norme ℓ_∞). Si $a_{r+1} \leq \mu$, la solution est complète; si de plus, $a_{r+1} = 0$, alors on a abouti à un modèle dynamique.

3.3.1. Procédure 1

Par (35), la solution de base initiale se trouve à l'origine: $u_n = v_n = 0$ ($n = 0, 1, \dots, N-r$). Comme le vecteur ressource de (31)-(32), $b = (0, \dots, 0, 1)$, est fortement dégénéré, on constate aisément que le critère usuel de la méthode du simplexe ne désigne aucune variable à exclure de la base et ne fournit aucune possibilité de déplacer la solution de base ailleurs qu'à l'origine, sauf dans le cas où

$$\exists j \text{ tel que } w_{ij} \leq 0 \text{ pour } i = 1, \dots, r \text{ et } w_{r+1,j} > 0. \quad (40)$$

Si (40) se produit, on pivote alors sur l'élément $w_{r+1,j}$ pour déplacer la solution de l'origine et obtenir une solution de base (non dégénérée si $w_{ij} < 0$, $\forall i$) à partir de laquelle on peut appliquer la méthode usuelle du simplexe. Mais il faut auparavant retirer toutes les variables artificielles de la base sinon les équations de (31) pour lesquelles $w_{m+i} = e_i$ et $w_{ij} \neq 0$ dans la colonne j ($j \neq m+i$) qui entre dans la base ne seront plus vérifiées. Considérant cela ainsi que le fait que les valeurs des variables de base apparaissent dans la colonne w_{m+r+1} , on peut

avancer l'algorithme suivant:

ÉTAPE 1. Introduire r colonnes W_j pour lesquelles $w_{ij} \neq 0$ et $1 \leq j \leq m$, $i = 1, \dots, r$, dans la base. (Ainsi, toutes les variables artificielles sont exclues de la base et la colonne $W_{m+r+1} = e_{r+1}$ demeure intacte.) Vérifier la condition d'optimalité

$$g_j \leq 0, \quad j = 1, \dots, m \quad (41)$$

à chaque itération (si elle est satisfaite, le problème est résolu et on obtient un modèle dynamique avec les coefficients (38) et $a_{r+1} = 0$). Si (41) n'est satisfaite en aucun cas, alors après $k \leq r$ changements de base, toutes les variables artificielles seront hors-base, sauf s'il existe une ligne de W où tout élément est nul sauf le "1" de la colonne qui correspond à la variable artificielle. Ces lignes (s'il en existe) sont exemptes de tout calcul ultérieur;

ÉTAPE 2. Vérifier la condition (40). Si elle n'est satisfaite pour aucun $j \in \{1, \dots, m\}$, alors continuer les opérations de pivotage comme dans l'étape 1 en vérifiant (40) et (41) simultanément. Après un nombre fini d'opérations, (40) ou (41) sera satisfaite. Si c'est (41): on arrête, la solution est complète; si c'est (40), alors:

ÉTAPE 3. Exclure la variable d'écart de la base en pivotant sur l'élément $w_{r+1,j} > 0$ de (40). Vérifier (41); si elle n'est pas satisfaite, appliquer alors l'algorithme du simplexe aux colonnes W_j , $j \in \{1, 2, \dots, m, m+r+1\}$, conservant les colonnes W_j ($j = m+1, \dots, m+r$) disponibles pour obtenir par produits scalaires la dynamique (38) du modèle.

La procédure se terminera par (41) en un nombre fini d'itérations. Elle n'est évidemment pas optimale: le nombre d'itérations peut être réduit en choisissant efficacement les éléments sur lesquels effectuer les pivots pour exclure les variables artificielles de la base.

3.3.2. Procédure 2

Une méthode pour choisir l'élément sur lequel pivoter dans l'étape 1 est de prendre w_{ik} tel que:

$$k(i) = \arg \max_{1 \leq j \leq m} |w_{ij}| \quad (i = 1, \dots, r) . \quad (42)$$

On s'assure ainsi que le coût relatif g_j n'augmentera pas considérablement. En fait, comme le pivot transforme g_j en g_j^* suivant

$$g_j^* = g_j - \frac{w_{ij}}{w_{ik}} g_k , \quad (43)$$

on aura

$$|g_j^*| \leq 2 \max_j |g_j| .$$

Cette procédure très simple a fourni de bons résultats quand appliquée au cas de l'équation du pendule (section 3.4).

3.3.3. Procédure 3

Une autre méthode qui améliore le processus est de choisir l'élément pivot w_{ik} de l'étape 1 tel que, pour $i = 1, \dots, r$,

$$k(i) = \arg \min_{s \in \Omega_i} \max_{1 \leq j \leq m} (g_j - \frac{w_{ij}}{w_{is}} g_s) , \quad (44)$$

où

$$\Omega_i = \{s \mid s \notin \text{base}, w_{is} \neq 0\} . \quad (45)$$

S'il arrive, pour certaines valeurs de i et s , que

$$\hat{g} = \max_{1 \leq j \leq m} (g_j - \frac{w_{ij}}{w_{is}} g_s) \leq 0 , \quad (46)$$

alors on a terminé: la solution est complète et on trouve un modèle dynamique par

(38). Si (46) ne se produit pas, alors l'étape 1 sera complétée avec $\hat{g} = \max g_j^* > 0$ et le tableau résultant ne sera pas optimal.

On continue la vérification par (44)-(46) et on pivote s'il arrive en (44) que $\min \max < \hat{g}$ du tableau précédent; on peut économiser du temps de calcul en exigeant ici $g_s \neq 0$ dans la définition (45) de Ω_1 .

Une autre modification est de pivoter si dans (44) on a simplement $\max < \hat{g}$ du tableau précédent. Ce processus assure la décroissance monotone de $\hat{g}$ vers une certaine valeur g_0 sans affecter la colonne associée à la variable d'écart ni la valeur $Z = 0$. Si $\hat{g}_0 > 0$, on enchaîne avec les étapes 2 et 3.

3.4. Application à un système non linéaire: l'équation du pendule

Pour vérifier l'efficacité de la méthode et son habileté à distinguer les phénomènes linéaires des phénomènes non linéaires à partir des observations, on considère le système classique du pendule non linéaire (Fig. 2) et son approximation linéaire, soient les deux équations différentielles du second ordre suivantes:

$$\frac{d^2x}{dt^2} + \omega^2 \sin x = 0, \quad (47)$$

$$\frac{d^2x}{dt^2} + \omega^2 x = 0, \quad (48)$$

avec les conditions initiales $x(0) = x_0 > 0$, $\frac{dx(0)}{dt} = 0$.

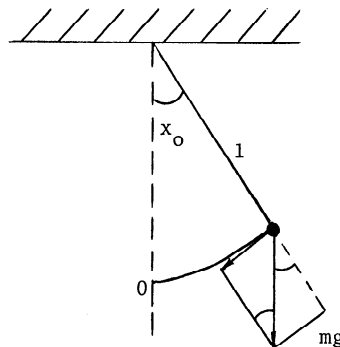


Figure 2

Prenons $\omega = 2\pi$, ce qui correspond à une période $T = 1$ seconde pour un pendule physique de longueur 25 cm avec x_0 assez petit. Les deux questions qui se posent sont les suivantes:

i) Peut-on reconnaître le système linéaire du second ordre à coefficients constants (48) simplement à partir d'une suite d'observations discrètes sur sa trajectoire?

ii) A partir de quelle valeur de $x_0 > 0$ peut-on déceler la nature non linéaire de (47) (qui se révèle par un ordre élevé du modèle de régression correspondant)?

Pour ce faire, on intègre (47) et (48) avec les mêmes conditions initiales par la méthode de Runge-Kutta avec une précision de $(\Delta t)^5$. L'intervalle d'intégration est $[0, 2]$ de longueur $2T = 2$ secondes (2 oscillations). Les accroissements de temps sont (successivement) $\Delta t_1 = 0,025$ et $\Delta t_2 = 0,050$. Donc, on applique les algorithmes précédents avec les nombres d'observations N et les précisions suivantes:

Δt	N	Précision d'intégration
0,025	80	10^{-8}
0,050	40	3×10^{-7}

Tableau 1

Les données discrètes obtenues par intégration sont prises comme observations $y_0, y_1, \dots, y_N$ et soumises à un programme PASCAL utilisant la procédure 2 décrite ci-avant pour évaluer les ordres minimaux de modèles de régression pouvant reproduire les données en fonction d'une précision exigée à l'avance. Les expériences numériques sont établies avec différentes conditions initiales $x_0 > 0$ et différentes valeurs de μ , la précision exigée. Les tableaux suivants donnent l'ordre minimal r d'un modèle de régression (10)-(11)-(12) en fonction de x_0 et de $\mu = 10^{-k}$ pour $k = 2, 3, 4, 5, 6$.

x_0 (rad)	k = 2	3	4	5	6
1,7	2	4	4	6	11
1,5	2	4	4	6	8
1,3	2	3	4	6	8
1,1	2	3	4	5	6
0,9	2	2	4	4	6
0,7	2	2	4	4	6
0,5	2	2	3	4	4
0,3	2	2	2	4	4
0,1	2	2	2	2	2

Tableau 2. Système non linéaire (47); $\Delta t = 0,025$, $N = 80$.

x_0 (rad)	k = 2	3	4	5	6
1,7	4	6	8	11	16
1,5	3	4	6	10	12
1,3	2	4	6	8	12
1,1	2	4	6	6	11
0,9	2	4	4	6	9
0,7	2	3	4	6	6
0,5	2	2	4	4	6
0,3	2	2	2	4	4
0,1	2	2	2	2	4

Tableau 3. Système non linéaire (47); $\Delta t = 0,050$, $N = 40$.

En ce qui concerne le système (48), ces tableaux sont tous composés de "2" de sorte que pour (48), on a $r = r(x_0, \mu) = 2 = \text{constant}$. On a en fait un modèle dynamique car $\mu = 10^{-6}$ est comparable à la précision de la méthode de Runge-Kutta (Tableau 1).

Ces résultats montrent l'efficacité des modèles de régression ainsi que des algorithmes proposés. Dans le cas (47), les ordres minimaux croissants indiquent la non-existence de modèles dynamiques linéaires pour le système (47). En vertu des Tableaux 2 et 3, il est possible d'approximer linéairement (47) par (48) pour de petites oscillations dont l'amplitude x_0 est fonction de la précision exigée μ et de l'accroissement Δt sur les trajectoires discrètes.

4. Programme dual (norme $\ell_{1,\rho}$)

Nous revenons maintenant au programme primal (24)-(27) pour le modèle dans la norme $\ell_{1,\rho}$. A chacune des contraintes de (25)-(26), on associe les variables duales u_n, v_n ($n = 0, 1, \dots, N-r$). Le programme dual contiendra $N-r$ contraintes d'inégalité de plus que le programme dual dans le cas ℓ_∞ et se présentera sous la forme suivante [5]:

4.1. Formulation du dual

$$\text{Max } Z = y_r(u_0 - v_0) + y_{r+1}(u_1 - v_1) + \dots + y_N(u_{N-r} - v_{N-r}), \quad (49)$$

$$\left\{ \begin{array}{l} \sum_{n=0}^{N-r} y_{n+i-1}(u_n - v_n) = 0, \quad i = 1, \dots, r, \\ u_n + v_n \leq \rho_n, \quad n = 0, 1, \dots, N-r, \\ u_n \geq 0, \quad v_n \geq 0, \quad n = 0, 1, \dots, N-r. \end{array} \right. \quad (50)$$

sous les
contraintes

$$u_n + v_n \leq \rho_n, \quad n = 0, 1, \dots, N-r, \quad (51)$$

$$u_n \geq 0, \quad v_n \geq 0, \quad n = 0, 1, \dots, N-r. \quad (52)$$

Le programme implique $2(N-r+1)$ variables principales et $N+1$ contraintes; après avoir ajouté r variables artificielles et $N-r+1$ variables d'écart pour obtenir une solution de base initiale, on aura $3N-2r+3$ variables au total. Comme dans le cas ℓ_∞ , le nombre total de solutions de base du programme dual est beaucoup moindre que celui du programme primal. Le programme est toujours résoluble, et étant donné sa structure spéciale, on peut y appliquer le principe de décomposition de Dantzig [6] pour accélérer le processus de résolution (section 4.3).

4.2. Interprétation géométrique

De (51) on tire:

$$\sum_{n=0}^{N-r} (u_n + v_n) \leq \sum_{n=0}^{N-r} \rho_n = \rho, \quad \rho > 0, \quad (53)$$

ce qui signifie, avec (52), que le domaine du programme dual est compris dans la

"pyramide" régulière de dimension $2(N-r+1)$ dont les arêtes sont de longueur ρ , dans la région où toutes les variables sont positives. Les sommets de cette pyramide sont l'origine et tous les points sur les axes de coordonnées à distance ρ de l'origine. En fait, le domaine déterminé par (51)-(52) est une région convexe bornée dont les sommets incluent l'origine et les points de la forme $(\rho_0, 0, 0, \dots, 0)$, $(0, \rho_0, 0, \dots, 0)$, $(0, 0, \rho_1, 0, \dots, 0)$, $(0, 0, 0, \rho_1, 0, \dots, 0)$, ..., $(0, \dots, 0, \rho_{N-r}, 0)$, $(0, \dots, 0, 0, \rho_{N-r})$. Cette région est évidemment indépendante des données $y_0, y_1, \dots, y_N$, mais les paramètres N et r en déterminent la dimension. Les contraintes d'égalité (50) font intersection avec les faces de la région convexe bornée et en distinguent un simplexe fermé de dimension $2N-3r+2$ dont le sommet principal est à l'origine. Le plan $Z = 0$ de (49) passe par l'origine et coupe chacune des faces de la région à 45° . Le dual est donc toujours résoluble et on peut trouver un modèle qui respectera ou non la précision exigée. Ce qui dépend des données, ce sont les directions des plans de (49)-(50). Par rapport à ces directions, la solution peut se trouver soit à l'origine avec $\max Z = 0$ (modèle dynamique), soit sur la "base" de la région convexe, ce qui correspond en général à un modèle de régression avec $\max Z > 0$. Les points symétriques sur la base (par exemple, $u_0 = v_0 = \rho_0/2$, $u_i = v_i = 0$, $i = 1, \dots, N-r$) pour lesquels $Z = 0$ fournissent une solution réalisable indépendante des données, mais pas nécessairement la solution optimale qui maximise la valeur de Z . On voit que l'interprétation géométrique de ce problème diffère de celle de (30)-(33) par le simple fait que la région convexe bornée constituant l'ensemble des solutions réalisables est dépendante de la suite pondérée $\rho_0, \rho_1, \dots, \rho_{N-r}$ introduite lors de l'utilisation de la norme $\ell_{1,\rho}$.

4.3. Algorithme de décomposition

Posons

$$x = (u_0, u_1, \dots, u_{N-r}, v_0, v_1, \dots, v_{N-r}) , \quad (54)$$

$$q = (y_r, y_{r+1}, \dots, y_N, -y_r, -y_{r+1}, \dots, -y_N) . \quad (55)$$

Le programme dual (49)-(52) peut ainsi être réécrit sous la forme:

$$\max Z = qx , \quad (56)$$

$$\text{sous les contraintes} \left\{ \begin{array}{l} Ax = 0 , \\ Ex \leq p , \\ x \geq 0 , \end{array} \right. \quad (57)$$

$$\begin{array}{l} E = (I_{N-r+1}; I_{N-r+1}) \\ p = (\rho_0, \rho_1, \dots, \rho_{N-r}) \end{array} , \quad (58)$$

$$x \geq 0 , \quad (59)$$

où A est la matrice fournie par (50). En vertu des considérations géométriques précédentes (section 4.2), les relations (58)-(59) déterminent une région convexe bornée dont on notera les sommets $X_1, X_2, \dots, X_M$. Par conséquent, tout point x du simplexe (57)-(59) peut s'exprimer sous la forme:

$$x = \sum_{i=1}^M X_i z_i , \quad \sum_{i=1}^M z_i = 1 , \quad z_i \geq 0 . \quad (60)$$

En substituant (60) dans (56)-(57), on obtient le *z-programme* suivant:

$$\max Z = Sz , \quad S = q(X_1, \dots, X_M) , \quad (61)$$

$$\text{sous les contraintes} \left\{ \begin{array}{l} Pz = 0 , \quad P = (P_1, \dots, P_M) , \quad P_i = AX_i , \\ \sum_{i=1}^M z_i = 1 , \\ z = (z_1, \dots, z_M) \geq 0 . \end{array} \right. \quad (62)$$

$$\sum_{i=1}^M z_i = 1 , \quad (63)$$

$$z = (z_1, \dots, z_M) \geq 0 . \quad (64)$$

Etant donné les r contraintes d'égalité de (62), on introduit r variables artificielles; de plus, comme $X_M = 0$ est un sommet de la région convexe bornée, on aura un vecteur unitaire supplémentaire associé à (63).

La base initiale du *z-programme* est donc sous la forme de la matrice

$$F = \begin{bmatrix} 0 & I_r \\ 1 & 0 \end{bmatrix} .$$

Esquissons ici un tableau du simplexe pour ce programme:

$$Z \Rightarrow S_1 \ S_2 \ \dots \ S_{M-1} \ 0, \quad (65)$$

$$W = [w_{ij}] = \begin{bmatrix} & & & & 0 \\ & & & & \vdots \\ P_1 & P_2 & \dots & P_{M-1} & 0 & I_r \\ \hline 1 & 1 & \dots & 1 & 1 & 0 & \dots & 0 \end{bmatrix}, \quad (66)$$

où:

$$S_i = qX_i \quad (i = 1, \dots, M \text{ et } X_M = 0),$$

$$P_i = AX_i \quad (i = 1, \dots, M \text{ et } X_M = 0).$$

La ligne (65), les coefficients de la fonction économique, est complétée par r "zéros" pour accéder à la dimension $M+r$ de la matrice W de (66).

Notons les colonnes du tableau par W_j ($j = 1, \dots, M+r$) et par C le sous-vecteur du vecteur S qui contient les coefficients de la fonction économique correspondant aux variables de base.

Utilisant le critère d'optimalité habituel de la méthode du simplexe, une solution z_0 sera optimale pour le z -programme si tous les coûts relatifs sont nuls ou négatifs, c.-à-d.

$$\begin{aligned} g_j &= S_j - C W_j \\ &= qX_j - C \begin{matrix} P_j \\ 1 \end{matrix} \\ &= qX_j - C AX_j - S_M \leq 0 \\ &(\forall j, \ j = 1, \dots, M+r). \end{aligned}$$

On obtient ainsi le critère d'optimalité

$$(q-CA)X_j \leq S_M, \quad \forall j, \ j = 1, \dots, M+r. \quad (67)$$

Comme les X_j ($j = 1, \dots, M$) constituent les solutions de base de (58)-(59), alors pour satisfaire à (67), on peut considérer le sous-programme

$$\max Z^* = q^*x, \quad q^* = q - CA, \quad (68)$$

$$\text{sous les contraintes} \quad \begin{cases} Ex \leq p, \\ x \geq 0. \end{cases} \quad (69)$$

$$(70)$$

Soit $x^* = X_k$ la solution optimale à (68)-(70). Si pour ce k on a $g_k = q^*X_k - S_M \leq 0$, ou $Z^* \max \leq S_M$, alors $g_j \leq g_k \leq 0$, $\forall j$, $j = 1, \dots, M+r$, et la base actuelle du z -programme est optimale.

Si $Z^* \max > S_M$, c.-à-d. $g_k = \max_j g_j > 0$, alors par le critère habituel de la méthode du simplexe, le vecteur $P_k = AX_k$ entre la base et le nouveau $S_k = qX_k$ remplace la coordonnée à changer dans le sous-vecteur $C = (C_1, C_2, \dots, C_{r+1})$ de $S = (S_1, S_2, \dots, S_{M+r})$.

La solution au sous-programme (68)-(70) est immédiate. Effectuons la première itération avec la base initiale du z -programme obtenue ci-haut. Ici: $C = 0$, $S_M = 0$, $q^* = q$ et (68)-(70) coïncide avec (56)-(58)-(59). On introduit $N-r+1$ variables d'écart dans (69) pour obtenir une solution initiale. En ce qui concerne le choix des $N-r+1$ variables de base initiales, on procède comme suit: pour chaque ligne, on prend la variable dont l'indice correspond au coefficient de la fonction économique le plus élevé parmi les trois associés à chaque élément non nul de la ligne (le tableau initial est composé de trois matrices identité d'ordre $N-r+1$ côte-à-côte). On obtient ainsi:

$$Z^* \max = \sum_{i=0}^{N-r} \rho_i |y_{i+r}| > S_M = 0.$$

Si $y_{i+r} \geq 0$, $\forall i$, $i = 0, \dots, N-r$, alors $x^* = X_k = (1, \dots, 1, 0, \dots, 0)$ avec $N-r+1$ unités dans X_k .

Concernant le critère de sortie de la base dans le z -tableau, on choisit une colonne correspondant à un élément pivot $w_{ik} \neq 0$, $1 \leq i \leq r$, associé à une variable artificielle dans la base. On demeure ainsi avec $S_M = 0$ et la colonne correspondant à la variable d'écart w_M inchangée, et le procédé retire toutes les

variables artificielles de la base. Ensuite, si nécessaire, on continue les itérations jusqu'à ce que $w_{ik} \leq 0$ ($i = 1, \dots, r$), $w_{r+1,k} > 0$. On pivote alors sur $w_{r+1,k}$ et la solution réalisable écartée de l'origine; on obtient ainsi un modèle de régression avec $\max Z > 0$ en général. De ce point, on peut utiliser si nécessaire le critère ordinaire de sortie de la méthode du simplexe. Pour retracer la précision et la dynamique du modèle, on doit prendre en mémoire toutes les opérations de pivotage sur la base initiale incluant les colonnes associées aux variables artificielles et d'écart. Ainsi, dans le z -tableau optimal, les produits scalaires $\xi_1 = CW_M$ et $a_j = CW_{j+M}$ ($j = 1, \dots, r$) fournissent la précision et la dynamique du modèle d'ordre r à l'aide de la norme $\ell_{1,p}$.

5. Conclusion

Les méthodes présentées traitent une suite quelconque d'observations équidistantes dans le temps indépendamment de la nature (déterministe) du phénomène considéré. Elles peuvent donc être utilisées dans divers domaines en sciences naturelles, en génie ou même dans la recherche socio-économique.

Leur propriété première est d'établir si le phénomène est de nature linéaire ou non linéaire. Si le phénomène est linéaire, les procédures numériques proposées dans cet article identifient systématiquement un modèle de régression qui peut être employé pour évaluer le comportement futur du phénomène (c.-à-d. les données futures de la suite).

Ces méthodes impliquent des algorithmes finis faisant appel à la programmation linéaire. Les modèles de régression obtenus sont uniques (à équivalence près) et les procédures itératives se terminent toujours soit en identifiant un modèle linéaire stationnaire, soit en concluant qu'il n'en existe aucun acceptable.

On peut intuitivement estimer la validité d'un modèle selon le critère suivant: s'il est suffisamment précis pour reproduire les données connues et qu'aucun facteur notable n'influence le phénomène à partir duquel les données sont relevées, alors on peut se servir du modèle pour estimer les données futures avec une

confiance raisonnable.

Soulignons en terminant que les procédures numériques (et leurs variantes) présentées dans cet article réduisent considérablement le temps de calcul nécessaire sur ordinateur pour résoudre les problèmes abordés.

Il reste toutefois beaucoup de problèmes connexes à étudier pour approfondir les résultats obtenus. Un des plus importants est celui qui traite de l'information supplémentaire suffisante pour assurer l'unicité des paramètres du modèle de régression, c'est-à-dire pour relever un seul modèle de la classe d'équivalence associée à la suite d'observations générées par le processus considéré. En effet, on sait qu'un programme linéaire peut avoir plusieurs solutions équioptimales et que la méthode du simplexe aboutit à un des sommets équioptimaux dépendamment de la première solution (réalisable) de base choisie ou des "branchements" possibles au fil des itérations. Plusieurs modèles différents (numériquement) peuvent ainsi possiblement être identifiés en traitant la même suite. Cette situation de non-unicité de la solution est tout-à-fait normale et reflète le fait que l'information contenue dans la suite d'observations traitée est essentiellement incomplète.

Liste des ouvrages cités

- [1] LEE, R.C.K. (1964), *Optimal Estimation, Identification and Control*, Research Monograph No. 28, MIT Press, Cambridge, Mass.
- [2] GALPERIN, E.A. et ZINGER, A. (1978), *Stationarity Conditions for Linear Models*, Proc. of the 2nd Int. Symp. on Large Engineering Systems, Waterloo, Ontario, 207-214.
- [3] GALPERIN, E.A. (1979), *Linear Programming Solution of the Discrete Stationary Linear Predictor Problem*, Proc. of the Int. Conference on Cybernetics and Society, Denver, Col., 707-711.
- [4] GALPERIN, E.A. et LABONTÉ, R. (1980), *Two Algorithms of Solution of the Discrete Stationary Linear Predictor Problem I*, Proc. of the 3rd Int. Symp.

on Large Engineering Systems, St-Johns, Newfoundland, 113-138.

- [5] GALPERIN, E.A. et LABONTE, R. (1980), *Two Algorithms for Solution of the Discrete Stationary Linear Predictor Problem II*, Proc. of the 1980 Joint Automatic Control Conference, San Francisco, Cal., Vol. 2, pp. TP6-G.
- [6] DANTZIG, G.B. (1963), *Linear Programming and Extensions*, Princeton University Press, Princeton, N.J.

Département de mathématiques et
d'informatique
Université du Québec à Montréal
C.P. 8888, Succ. A
Montréal, Québec H3C 3P8

Manuscrit reçu le 20 juillet 1980.
Révisé le 23 juin 1983.