

REVUE DE STATISTIQUE APPLIQUÉE

J. SISKOS

Analyses de régression et programmation linéaire

Revue de statistique appliquée, tome 33, n° 1 (1985), p. 41-55

http://www.numdam.org/item?id=RSA_1985__33_1_41_0

© Société française de statistique, 1985, tous droits réservés.

L'accès aux archives de la revue « Revue de statistique appliquée » (<http://www.sfds.asso.fr/publicat/rsa.htm>) 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/>

ANALYSES DE REGRESSION ET PROGRAMMATION LINEAIRE(*)

J. SISKOS

*Professeur à l'Ecole Polytechnique de Crète
Département "Science des Décisions"
Iroon Polytechniou 37, Chania, Crète, Grèce*

RESUME

Cet article entreprend un tour d'horizon des modèles de régression multiple qui utilisent des formulations de programmation linéaire continue. Il développe successivement le modèle de WAGNER en régression linéaire qui minimise la somme des écarts absolus puis une nouvelle version de la méthode UTA en régression ordinale qui explique un préordre par des variables qualitatives et/ou quantitatives monotones. Les deux modèles sont illustrés sur un exemple numérique simple. Il a été constaté que les modèles de programmation linéaire sont très puissants, tant au niveau de leur structure (adjonction de nouvelles contraintes) qu'au niveau de leur flexibilité (possibilité d'effectuer des analyses de stabilité).

ABSTRACT

This paper presents a survey of multiple regression models which use linear programming formulations. It develops successively the model of WAGNER in linear regression, which minimizes the sum of absolute deviations, and a new version of the UTA method in ordinal regression, which explains a weak-order by qualitative and/or quantitative monotone variables. The models are illustrated by a simple numerical example. It has been emphasized that linear programming models are very powerful, on structural level (possibility to add new constraints) as much as on the level of flexibility (possibility to perform stability analyses).

1. INTRODUCTION ET PERSPECTIVE HISTORIQUE

L'usage de la programmation linéaire dans les analyses de régression a été introduit, pour la première fois, par CHARNES *et al.* [2], [3] en programmation linéaire à objectifs multiples ou "goal programming". Cette technique consiste à analyser les incompatibilités inhérentes à un système de contraintes linéaires. En effet, dans [2], CHARNES, COOPER et FERGUSON proposent, en 1955, une formulation de programmation linéaire pour ajuster un modèle linéaire consistant à prédire les salaires de nouveaux embauchés dans une entreprise ; le modèle explique à la fois une variable numérique (salaires-types) et une variable ordinale (hiérarchie d'entreprise donnée) par des variables numériques comme l'initiative, l'expérience professionnelle, le niveau intellectuel, . . . Les auteurs cherchent à minimiser la somme des écarts absolus entre les salaires-types et les salaires prédits.

En 1958, KARST [9] propose un modèle de régression linéaire avec une seule variable explicative dans lequel il minimise la somme des écarts absolus des résidus.

(*) Cet article a été écrit lorsque l'auteur travaillait au LAMSADE, Université de Paris-Dauphine. L'auteur tient à remercier Catherine DUPONT, Pascal DRAPIER et deux referees anonymes dont les remarques et suggestions ont contribué à améliorer une version précédente de cet article.

Mots-clés : Régression Linéaire, Régression Ordinale, Codage Additif, Programmation Linéaire.

Key-words : Linear Regression, Ordinal Regression, Additive Scaling, Linear Programming.

Un an plus tard, WAGNER [22] généralise le problème pour le cas de la régression linéaire multiple. KELLEY [10] s'intéresse au modèle particulier de régression linéaire qui minimise le critère de CHEBYSHEV, c'est-à-dire le maximum des déviations absolues.

En-dehors de la souplesse classique des modèles de programmation linéaire (possibilité d'ajouter des contraintes supplémentaires, positivité des coefficients. . .), certaines études comparatives ont dégagé, par simulation, quelques avantages des estimateurs des moindres écarts absolus sur ceux des moindres carrés. En effet, KIOUNTOUZIS [11] a démontré expérimentalement que les deux estimateurs possèdent pratiquement les mêmes propriétés statistiques, ceci dans quatre hypothèses différentes sur la distribution de probabilité des écarts (lois: uniforme, normale, Laplace, Cauchy). WILSON [23], en outre, prouve, par le même type d'expérience, que, pour le cas de la distribution normale, les estimateurs de programmation linéaire présentent une meilleure "efficacité relative" vis-à-vis des estimateurs de moindres carrés.

Très récemment, FREED et GLOVER [6] ont développé des modèles linéaires d'analyse discriminante utilisant également des techniques de "goal programming".

En régression ordinale, SRINIVASAN et SHOCKER [19] présentent le modèle ORDREG, application directe des idées de CHARNES *et al.* [2], dans lequel la variable à expliquer est un ensemble de paires de stimuli comparés. SISKOS [16] présente, plus tard, UTA II dans le cas où la variable à expliquer est un pré ordre total. Il s'agit, dans les deux cas, de modèles linéaires restreints à des variables numériques.

En régression qualitative, problème plus récent et plus difficile à résoudre, KRUSKAL [13] propose le modèle MONANOVA, une méthode d'analyse de la variance où l'on explique une variable préordinaire totale par des variables nominales. L'auteur introduit un algorithme de codage additif visant à minimiser, pas à pas, une fonction d'erreurs définie par lui-même et qu'il nomme "le stress"; il s'agit d'une heuristique qui ne peut garantir des codages optimaux globaux car la fonction d'erreurs n'est pas convexe. En 1976, YOUNG, DE LEEUW et TAKANE [24] présentent un algorithme de régression pas à pas intitulé MORALS qui cherche à expliquer un préordre par n'importe quel type de variables en minimisant (le plus souvent localement) un critère du type des moindres carrés. Cette méthode de régression présente les mêmes inconvénients que MONANOVA car, d'une part, elle aboutit à des solutions parfois arbitraires (non optimales) et, d'autre part, elle ne permet pas d'analyses de stabilité sur le modèle estimé (cf. MACQUIN [14]).

Le problème de codage en statistique a fait l'objet de recherches dans des contextes variés (cf. CAZES *et al.* [1], DUPONT-GATELMAND [4], MACQUIN [14], SAPORTA [15], TENENHAUS [20], [21], . . .).

JACQUET-LAGREZE et SISKOS [7] proposent, en 1978, la première procédure rigoureuse (UTA) qui permet, grâce à la programmation linéaire, de dégager un ou plusieurs codages additifs optimaux en régression ordinale. La méthode cherche à expliquer un préordre total par des variables qualitatives et/ou quantitatives également monotones (cf. critères). Le problème de la stabilité du codage estimé a été résolu, dans UTA, par une heuristique proposée dans le contexte de l'analyse post-optimale multiparamétrique en programmation linéaire. Cette première version d'UTA a été récemment améliorée par SISKOS et YANNACOPOULOS [18].

Dans la section 2, nous présentons le modèle de WAGNER [22] en régression linéaire multiple, premier modèle utilisant la programmation linéaire. Certains développements concernant la dualité en programmation linéaire y sont utilisés (voir par exemple FAURE [5]). La section 3 est consacrée à la dernière version de la méthode UTA ; un exemple numérique est donné à titre d'illustration (section 4). En conclusion, nous signalons les voies actuelles de mise en application de la méthodologie UTA.

2. LE MODELE DE WAGNER EN REGRESSION LINEAIRE

On considère un ensemble I d'individus sur lequel on mesure n variables "indépendantes" numériques. Soit x_{ij} , $i = 1, 2, \dots, |I|$ et $j = 1, 2, \dots, n$ l'ensemble de ces mesures et y_i , $i = 1, 2, \dots, |I|$ les mesures associées à la variable numérique que l'on cherche à expliquer.

WAGNER [22] cherche à estimer les coefficients $b_1, b_2, \dots, b_n$ du modèle linéaire général (le terme constant b_0 y est intégré en prenant x_n comme vecteur-unité)

$$Y = b_1 X_1 + b_2 X_2 + \dots + b_n X_n \quad (1)$$

en minimisant la somme des écarts absolus

$$[\text{MIN}] \sum_{i=1}^{|I|} |y_i - \sum_{j=1}^n b_j x_{ij}|. \quad (2)$$

Les coefficients b_j peuvent être de signe quelconque dans ce modèle. Pour simplifier ici, nous supposons que $b_j \geq 0$ pour tout j .

Pour résoudre le problème (2), il suffit de résoudre le programme linéaire (3)-(5) ci-dessous :

$$[\text{MIN}] \sum_{i=1}^{|I|} (\sigma_i^+ + \sigma_i^-) \quad (3)$$

sous les contraintes

$$\begin{aligned} \sum_{j=1}^n b_j x_{1j} + \sigma_1^+ - \sigma_1^- &= y_1 \\ &\dots\dots\dots \\ \sum_{j=1}^n b_j x_{ij} + \sigma_i^+ - \sigma_i^- &= y_i \end{aligned} \quad (4)$$

$$\begin{aligned} \sum_{j=1}^n b_j x_{|I|j} + \sigma_{|I|}^+ - \sigma_{|I|}^- &= y_{|I|} \\ b_j &\geq 0, j = 1, 2, \dots, n ; \sigma_i^+ \geq 0, \sigma_i^- \geq 0, i = 1, 2, \dots, |I| \end{aligned} \quad (5)$$

Pour démontrer l'équivalence entre les formulations (2) et (3)-(5), il suffit de poser, pour tout i ,

$$\sigma_i^- = \frac{1}{2} \left[\left| \sum_j b_j x_{ij} - y_i \right| + \left(\sum_j b_j x_{ij} - y_i \right) \right] \quad (6)$$

$$\sigma_i^+ = \frac{1}{2} \left[\left| \sum_j b_j x_{ij} - y_i \right| - \left(\sum_j b_j x_{ij} - y_i \right) \right] \quad (7)$$

et de remplacer σ_i^+ et σ_i^- dans les relations (3)-(5).

A l'optimum, on aura bien sûr $\sigma_i^+ \cdot \sigma_i^- = 0$ pour tout i (cf. [3]), à savoir que l'une au moins des erreurs par observation sera nulle.

Le système (4) comprend autant de contraintes linéaires d'égalité que le nombre d'individus observés, ce qui paraît gênant car, en programmation linéaire, ce nombre détermine la taille-mémoire d'ordinateur à réserver pour traiter un problème de ce type.

Pour pallier à cet inconvénient, WAGNER propose de résoudre le programme linéaire dual en introduisant autant de variables que le nombre des contraintes d'égalité du primal et autant de contraintes linéaires que le nombre des variables du primal. Il vient donc :

$$[\text{MAX}] \sum_i y_i d_i \quad (8)$$

sous les contraintes :

$$\sum_i x_{ij} d_i \leq 0 \quad j = 1, 2, \dots, n \quad (9)$$

$d_i \leq 1, -d_i \leq 1, d_i$ sans restriction de signe

$$\text{pour tout } i = 1, 2, \dots, |I| \quad (10)$$

d'où, en procédant à la transformation des variables

$$f_i = d_i + 1, i = 1, 2, \dots, |I|, \quad (11)$$

on obtient le programme linéaire

$$[\text{MAX}] \sum_i y_i f_i - \sum_i y_i \quad (12)$$

sous les contraintes :

$$\sum_i x_{ij} f_i \leq \sum_i x_{ij}, j = 1, 2, \dots, n \quad (13)$$

$$0 \leq f_i \leq 2, i = 1, 2, \dots, |I| \quad (14)$$

La complexité du programme (12)-(14) est définie par n contraintes seulement, du type (13), les autres, du type (14), n'étant que des contraintes de bornage que nous pouvons gérer par des codes spéciaux en programmation linéaire. Pour les coefficients b_j dont la positivité n'est pas nécessaire, nous aurions posé l'égalité en (13). Enfin, pour passer de la solution optimale du programme linéaire (12)-(14) à celle du programme (3)-(5) concernant $(b_1, b_2, \dots, b_n), (\sigma_i^+, \sigma_i^-, i = 1, 2, \dots,$

II) optima, il suffit d'attribuer (à un signe près) à ces variables les valeurs des coûts marginaux des variables hors base-simplexe correspondantes (voir [5]).

Pour ajuster le modèle $Y = b_1 X_1 + b_2 X_2 + b_3 X_3 + b_0$, $b_j \geq 0$, $j = 0, 1, 2, 3$ sur les données ci-dessous provenant de l'exemple de la section 4 (la quatrième colonne unitaire se reporte à b_0)

y	x_1	x_2	x_3	x_4
1	3	1	3	1
2	4	2	2	1
2	2	2	4	1
3	6	4	4	1
4	30	3	1	1
$\Sigma : 12$	45	12	14	5

nous devons résoudre le programme linéaire suivant :

$$[\text{MAX}] f_1 + 2 f_2 + 2 f_3 + 3 f_4 + 4 f_5 - 12$$

sous les contraintes :

$$3 f_1 + 4 f_2 + 2 f_3 + 6 f_4 + 30 f_5 \leq 45$$

$$f_1 + 2 f_2 + 2 f_3 + 4 f_4 + 3 f_5 \leq 12$$

$$3 f_1 + 2 f_2 + 4 f_3 + 4 f_4 + f_5 \leq 14$$

$$f_1 + f_2 + f_3 + f_4 + f_5 \leq 5$$

$$0 \leq f_i \leq 2, i = 1, 2, 3, 4, 5$$

Ce programme fournit une solution optimale unique ($f_1 = 0$, $f_2 = 1.46$, $f_3 = 2$, $f_4 = .46$, $f_5 = 1.08$, variable d'écart de la 3^e contrainte = .16), d'où, par lecture des coûts marginaux des variables hors base, on en déduit :

$$b_1 = .06$$

$$b_2 = .44$$

$$b_3 = 0$$

$$b_0 = .88$$

$$\sigma_1^- = .5, \sigma_3^+ = .12 \quad \text{et} \quad \sum_i (\sigma_i^+ + \sigma_i^-) = .62.$$

L'estimateur de moindres écarts absolus est donc

$$Y = .06 X_1 + .44 X_2 + .88$$

On en déduit les estimations suivantes :

$$\bar{y}_1 = 1.5 \quad (\text{d'où } \sigma_1^- = .5)$$

$$\bar{y}_2 = 2$$

$$\bar{y}_3 = 1.88 \quad (\text{d'où } \sigma_3^+ = .12)$$

$$\bar{y}_4 = 3$$

$$\bar{y}_5 = 4$$

3. LA REGRESSION ORDINALE

En régression ordinale, on s'intéresse au problème suivant :

Problème du codage additif : Etant donné Y une variable préordinaire à expliquer et $X_1, X_2, \dots, X_n$ n variables explicatives monotones (cf. critères), toutes ces variables étant définies sur un ensemble I d'individus, estimer les codages $X_1^*, X_2^*, \dots, X_n^*$ de $X_1, X_2, \dots, X_n$, et un jeu de paramètres $b_1, b_2, \dots, b_n$ (à des contraintes de normalisation près) de façon à ce que le modèle additif

$$Y^* = b_1 X_1^* + b_2 X_2^* + \dots + b_n X_n^* \quad (15)$$

soit aussi compatible que possible avec le préordre Y .

YOUNG, DE LEEUW et TAKANE [24] ont proposé la méthode MORALS, dite de moindres carrés alternés, qui est fondée sur les recherches de KRUSKAL [12], [13]. A l'exception des variables quantitatives pour lesquelles les auteurs de MORALS ne cherchent pas de codage (ils posent $X_j^* = X_j$ et cherchent à estimer uniquement b_j), les contraintes de normalisation choisies dans MORALS exigent que les variables codées X_j^* soient centrées et réduites.

Cette approche, qui vise à obtenir successivement les codages de Y , soit $\tilde{Y}$, puis $X_1^*, X_2^*, \dots, X_n^*$ en cherchant à minimiser le carré de la norme usuelle

$$\|\tilde{Y} - Y^*\|^2 = \|\tilde{Y} - \sum_{j=1}^n b_j X_j^*\|^2 \quad (16)$$

ne garantit pas l'optimum de la norme (16) et ne fournit pas d'analyse de stabilité du codage estimé (voir MACQUIN [14]).

La méthode UTA de JACQUET-LAGREZE et SISKOS [7] utilise la programmation linéaire et propose un système de contraintes de normalisation différentes. Elle code en outre les variables numériques, ce qui lui permet de généraliser la régression ordinale linéaire.

Soit $[x_j^1, x_j^2, \dots, x_j^{\alpha_j}]$ l'ensemble des modalités ordonnées d'une variable ordinale X_j . Lorsque X_j est numérique, x_j^k , $k = 1, 2, \dots, \alpha_j$ représentent des points équidistants; α_j est donc le nombre de points sur lesquels on cherche à estimer le codage de cette variable et est fixé par algorithme aussi grand que possible (jusqu'à une certaine limite) dans la mesure de l'information disponible, ceci compte-tenu de l'ensemble I . On suppose par la suite que x_j^1 est le niveau le moins préféré de X_j ⁽¹⁾.

Appelons x_j^{*k} le codage du niveau x_j^k de la variable X_j . Lorsque X_j est numérique, le codage de la variable X_j pour une valeur comprise entre x_j^k et x_j^{k+1} se fait par interpolation linéaire (cf. Fig. 1). Nous désignerons par x_{ij}^* le codage de l'appréciation x_{ij} de l'individu $i \in I$ sur la variable X_j .

Les contraintes de normalisation d'UTA sont les suivantes :

$$x_j^{*1} = 0, j = 1, 2, \dots, n \quad (17)$$

$$x_j^{*\alpha_j} = 1, j = 1, 2, \dots, n \quad (18)$$

$$b_1 + b_2 + \dots + b_n = 1 \quad (19)$$

(1) On pourrait envisager un modèle dans lequel le sens de variation des préférences sur une variable ne serait pas prédéterminé mais à estimer par ce modèle.

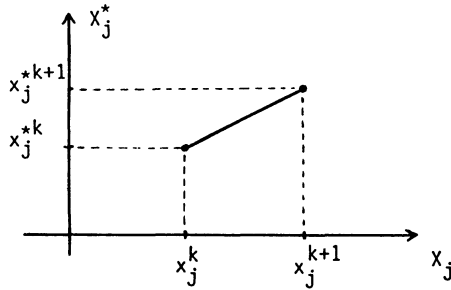


Figure 1. – Codage de X_j entre x_j^k et x_j^{k+1}

En fait, si dans (15) on remplace $b_j X_j^*$ par X_j^* (ce qui revient à faire rentrer b_j dans le codage de la variable X_j , et à abandonner la contrainte de normalisation (18), $x_j^{\alpha_j}$ étant maintenant égal à b_j) les relations (18) et (19) sont alors remplacées par la relation :

$$x_1^{\alpha_1} + x_2^{\alpha_2} + \dots + x_n^{\alpha_n} = 1 \quad (20)$$

et avec ce nouveau codage, (15) est remplacé par :

$$Y^* = X_1^* + X_2^* + \dots + X_n^* \quad (15 \text{ bis})$$

Pour revenir au codage initial (15), avec des codages x_j^* normalisés par (18), il suffit, après estimation des x_j^{*k} , $j = 1, 2, \dots, \alpha_j$ sous les contraintes (17) et (20), de poser $b_j = x_j^{\alpha_j} \neq 0$ et de remplacer x_j^{*k} par $x_j^{*k}/x_j^{\alpha_j}$.

Dans la suite, on estimera les x_j^{*k} sous les contraintes (17) et (20), et comme on vient de le préciser, on pourra a posteriori revenir au codage initial avec le calcul des b_j .

Posons maintenant :

$$y_i^* = \sum_{j=1}^n x_{ij}^* \quad (21)$$

avec y_i^* le codage de y_i . Comme Y représente un préordre sur I , y_i représente le rang de l'individu i dans ce préordre. On supposera, par la suite, que les individus dans I sont rangés selon les valeurs décroissantes de y_i , $i \in I$ et que $1 \leq y_i \leq Q$ pour tout i ($Q \leq |I|$ et entier).

Dans la version originale d'UTA (cf. [7]), il existe une seule fonction d'erreurs $\sigma : I \rightarrow [0, 1]$ où $\sigma(i) \geq 0$ représente l'erreur éventuelle qu'il faut ajouter au codage y_i^* de l'individu i pour que celui-ci puisse récupérer son rang dans le préordre. En effet, les codages y_i^* , pour tout $i \in I$, seraient sans erreurs si tous les points (y_i, y_i^*) de la figure 2 appartenaient à une courbe strictement croissante, c'est-à-dire que plus les rangs y_i des individus diminuent plus les codages y_i^* augmentent (cas des points noirs de la figure 2). Si, maintenant, pour un individu donné i occupant le rang y_i , les appréciations x_{ij} , $j = 1, 2, \dots, n$ sont telles que le codage résultant

$y_i^* = \sum_j x_{ij}^*$ est trop faible par rapport à son rang (cas des points blancs situés à gauche de la courbe de la figure 2 : individus "sur-estimés"), il y a un écart positif $\sigma(i)$ entre le codage des ex-aequos de i , qui eux sont situés sur la courbe, et le codage y_i^* .

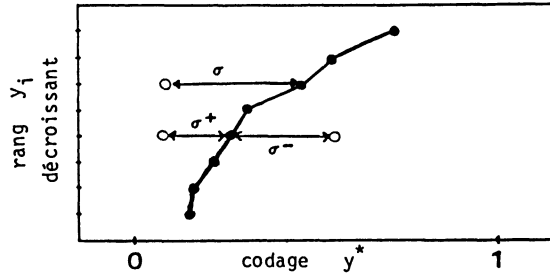


Figure 2. — Codage vs préordre en régression ordinale.

L'idée de la première version du modèle UTA était de remplacer tous les y_i^* par $y_i^* + \sigma(i)$ et de chercher à minimiser la somme des $\sigma(i)$. Néanmoins, cette fonction d'erreurs n'est pas suffisante pour minimiser complètement la dispersion des points blancs autour de la courbe de régression. Le problème est posé notamment par des points qui risquent d'être situés à droite de la courbe (individus "sous-estimés") pour lesquels il faudrait soustraire des erreurs positives σ^- plutôt que d'ajouter aux autres des erreurs du type σ^+ (cf. Fig. 2).

Dans la version actuelle d'UTA (cf. [18]), appelons-la UTA*, nous proposons une fonction d'erreurs composée de deux erreurs potentielles par individu (Fig. 2). Cela permet de mieux stabiliser la position des points autour de la courbe de régression (voir [18]). Ainsi, le codage d'un individu $i \in I$ sera remplacé, dans la relation (21), par $y_i^* + \sigma^+(i) - \sigma^-(i)$ avec $\sigma^+(i) \geq 0, \sigma^-(i) \geq 0$.

Afin de tenir compte du caractère monotone de chacune des variables, nous imposons le système de contraintes suivant :

$$x_j^{*k+1} - x_j^{*k} \geq 0, k = 1, 2, \dots, \alpha_j - 1 \text{ et} \quad (22)$$

$$j = 1, 2, \dots, n$$

Il est possible également de remplacer certains ou tous les zéros du second membre par des petits ϵ .

Introduisons la transformation de variables

$$w_{jk} = x_j^{*k+1} - x_j^{*k}, k = 1, 2, \dots, \alpha_j - 1 \text{ et} \quad (23)$$

$$j = 1, 2, \dots, n$$

Comme, d'après (17), $x_j^{*1} = 0$ pour tout j , il vient

$$x^{*k} = \sum_{l=1}^{k-1} w_{jl}, k = 2, \dots, \alpha_j \quad (24)$$

La relation de normalisation (20) devient donc

$$\sum_{j=1}^n \sum_{k=1}^{\alpha_j-1} w_{jk} = 1 \quad (25)$$

La méthode UTA* cherche à estimer w_{jk} pour tout j et k puis x^{*k} à l'aide de (24). Elle peut être schématisée en quatre étapes :

1) Exprimer, dans l'ordre imposé par le préordre à expliquer ($1 = y_1 \leq y_2 \leq \dots \leq y_{|I|} = Q$), les codages y_i^* , $i \in I$ des individus en fonction des variables w_{jk} suivant d'abord (21) avec interpolation linéaire au cas où certains x_{ij}^* ne coïncident pas avec des modalités x_j^{*k} , puis (24).

2) Générer le système de $|I| - 1$ contraintes linéaires, noté (C), par application de l'algorithme :

a) $i = 1$.

b) Poser (introduction des doubles erreurs) :

$$\left. \begin{aligned} y_i^* - y_{i+1}^* + \sigma^+(i) - \sigma^-(i) - \sigma^+(i+1) + \sigma^-(i+1) &\geq \delta \\ \text{si } y_{i+1} > y_i & ; \text{ sinon } (y_{i+1} = y_i) : \\ y_i^* - y_{i+1}^* + \sigma^+(i) - \sigma^-(i) - \sigma^+(i+1) + \sigma^-(i+1) &= 0 \end{aligned} \right\} \quad \text{C)}$$

c) $i := i + 1$

si $i = |I|$: Arrêt, Sinon : Aller en b)

δ est une valeur positive suffisamment petite pour différencier les Q classes du préordre.

3) Résoudre le programme linéaire ;

$$[MIN] F = \sum_{i \in I} [\sigma^+(i) + \sigma^-(i)]$$

sous les contraintes :

Système (C)

$$\sum_{j=1}^n \sum_{k=1}^{\alpha_j-1} w_{jk} = 1$$

$$w_{jk} \geq 0, \sigma^+(i) \geq 0, \sigma^-(i) \geq 0 \quad \forall i \in I, \forall j \text{ et } k.$$

4) Tester l'existence de solutions optimales multiples (problème de dégénérescence) et/ou de solutions "peu écartées" de l'optimum (analyse de stabilité, cf. [7], [16], [17]). Rechercher éventuellement (en cas d'instabilité) n solutions-codages caractéristiques, celles qui maximisent respectivement $\sum_k w_{jk}$, $j = 1, 2, \dots, n$, c'est-à-dire les "poids" des variables explicatives. Par "peu écartée", nous entendons une solution-codage dont la somme des erreurs F est ϵ -supérieure à l'optimum.

4. UN EXEMPLE NUMERIQUE

Nous reprenons l'exemple dans [18]. Considérons le cas d'un individu dont on veut analyser les choix de modes de transport domicile-lieu de travail pendant les heures d'affluence. Cet individu s'intéressant uniquement aux trois critères

TABLEAU 1
Problème d'analyse du classement de modes de transport

I	Modes de transport	Y (Rangs)	X ₁ Prix (Francs)	X ₂ Temps (minutes)	X ₃ Confort (qualitatif)
1	RER	1	3	10	+
2	METRO 1ère	2	4	20	++
3	METRO 2e	2	2	20	0
4	BUS	3	6	40	0
5	TAXI	4	30	30	+++

(cf. variables monotones) (1) prix du titre de transport (en Francs), (2) temps de trajet (en minutes) et (3) confort (possibilité de s'asseoir) attribue, aux cinq modes, les rangs de préférence du tableau 1.

Les variables X₁, X₂ sont numériques et monotones. Pour X₃, une échelle qualitative ordinale a été utilisée :

0 : aucune chance de s'asseoir,
+ : peu de chances de s'asseoir,
++ : grandes chances de trouver une place assise,
+++ : place assise certaine.

La structure de l'information nous permet de retenir les ensembles de modalités ordonnés ci-dessous :

$$[x_1^1, x_1^2, x_1^3] = [30, 16, 2], \alpha_1 = 3$$

$$[x_2^1, x_2^2, x_2^3, x_2^4] = [40, 30, 20, 10], \alpha_2 = 4$$

$$[x_3^1, x_3^2, x_3^3, x_3^4] = [0, +, ++, +++], \alpha_3 = 4.$$

$\alpha_1 = 3$, car si l'on avait pris $\alpha_1 > 3$, soit $\alpha_1 = 4$, l'ensemble des modalités pour $\tilde{X}_1$ serait $[30, 20\frac{2}{3}, 11\frac{1}{3}, 2]$ et il n'y aurait pas du tout d'information sur la modalité $20\frac{2}{3}$.

Commençant par l'étape 1 de la méthode et sachant que $x_j^{*1} = 0$, $j = 1, 2, 3$, nous obtenons les expressions des codages y_i^* des cinq modes (pour x_{1i} la technique de l'interpolation linéaire a été utilisée ; cf. Fig. 1).

$$y_1^* = .07 x_1^{*2} + .93 x_1^{*3} + x_2^{*4} + x_3^{*2}$$

$$y_2^* = .14 x_1^{*2} + .86 x_1^{*3} + x_2^{*3} + x_3^{*3}$$

$$y_3^* = x_1^{*3} + x_2^{*3}$$

$$y_4^* = .29 x_1^{*2} + .71 x_1^{*3}$$

$$y_5^* = x_2^{*2} + x_3^{*4}$$

et, en remplaçant les x_j^{*k} par les w_{ij} (cf. relations (23)-(24)), nous avons :

$$\begin{aligned}
y_1^* &= w_{11} + .93 w_{12} + w_{21} + w_{22} + w_{23} + w_{31} \\
y_2^* &= w_{11} + .86 w_{12} + w_{21} + w_{22} + w_{31} + w_{32} \\
y_3^* &= w_{11} + w_{12} + w_{21} + w_{22} \\
y_4^* &= w_{11} + .71 w_{12} \\
y_5^* &= w_{21} + w_{31} + w_{32} + w_{33}
\end{aligned}$$

Ensuite, après l'étape 2 de comparaison des individus par paires successives, on arrive au programme linéaire suivant (plusieurs zéros sont omis dans la matrice, δ est pris égal à .05)

w_{11}	w_{12}	w_{21}	w_{22}	w_{23}	w_{31}	w_{32}	w_{33}	variables	σ^+ et σ^-	Signe	Second membre
0	.07	0	0	1	0	-1	0	1 -1 -1 1		$\geq$	.05
0	-.14	0	0	0	1	1	0	1 -1 -1 1		=	0
0	.29	1	1	0	0	0	0	1 -1 -1 1		$\geq$	.05
1	.71	-1	0	0	-1	-1	-1	1 -1 -1 1		$\geq$	.05
1	1	1	1	1	1	1	1			=	1
								1 1 1 1 1 1 1 1		F	

La solution optimale donne (variables non nulles) :

$$\begin{aligned}
w_{11} &= .5 \\
w_{21} &= .05 \\
w_{23} &= .05 \\
w_{33} &= .4 \\
F^* &= \text{MIN } F = 0
\end{aligned}$$

ce qui correspond au codage des variables (d'après (23)-(24))

$$\begin{aligned}
x_1^{*1} &= 0 & x_2^{*1} &= 0 & x_3^{*1} &= 0 \\
x_1^{*2} &= .5 & x_2^{*2} &= .05 & x_3^{*2} &= 0 \\
x_1^{*3} &= .5 & x_2^{*3} &= .05 & x_3^{*3} &= 0 \\
x_2^{*4} &= .1 & x_3^{*4} &= .4
\end{aligned}$$

et à une restitution numérique parfaite du préordre.

En étape 4 (analyse de stabilité), nous recherchons un ensemble caractéristique de tels codages qui garantit également que la somme des erreurs est nulle. En général, on transforme l'objectif F en une contrainte du type

$$\sum_{i \in I} [\sigma^+(i) + \sigma^-(i)] \leq F^* + \epsilon \quad (26)$$

Comme ici $F^* = 0$, on remplace partout dans la matrice σ^+ et σ^- par des zéros et on cherche trois solutions qui maximisent respectivement : $w_{11} + w_{12}$, $w_{21} + w_{22} + w_{23}$, $w_{31} + w_{32} + w_{33}$.

Nous avons donc à résoudre les trois programmes linéaires suivants :

	w_{11}	w_{12}	w_{21}	w_{22}	w_{23}	w_{31}	w_{32}	w_{33}	Signe	Second membre
	0	.07	0	0	1	0	-1	0	$\geq$	.05
	0	-.14	0	0	0	1	1	0	=	0
	0	.29	1	1	0	0	0	0	$\geq$	.05
	1	.71	-1	0	0	-1	-1	-1	$\geq$	.05
	1	1	1	1	1	1	1	1	=	1
MAX	1	1							+	premier objectif
MAX			1	1	1				+	second objectif
MAX						1	1	1	+	troisième objectif

La solution de ces programmes se fait à partir de l'optimum du précédent. Nous trouvons trois solutions complètement contrastées indiquant l'instabilité du codage estimé :

1^{ère} solution : $w_{11} = .7625$, $w_{12} = .175$, $w_{23} = .0375$, $w_{31} = .025$
2^e solution : $w_{11} = .05$, $w_{22} = .05$, $w_{23} = .9$
3^e solution : $w_{11} = .35625$, $w_{12} = .175$, $w_{23} = .0375$, $w_{31} = .025$, $w_{33} = .40625$.

Prenons comme solution unique le codage correspondant au centre de gravité de ces trois codages :

$$\begin{aligned}w_{11} &= .390 , w_{12} = .116 \\w_{21} &= 0 , w_{22} = .017 , w_{23} = .325 \\w_{31} &= .017 , w_{32} = 0 , w_{33} = .135\end{aligned}$$

Il vient alors (cf. (24)) :

$X_1 \rightarrow X_1^*$	$X_2 \rightarrow X_2^*$	$X_3 \rightarrow X_3^*$	$I \rightarrow Y^*$
30 $\rightarrow$ 0	40 $\rightarrow$ 0	0 $\rightarrow$ 0	RER $\rightarrow$.856
16 $\rightarrow$.390	30 $\rightarrow$ 0	+ $\rightarrow$.017	M1 $\rightarrow$.523
2 $\rightarrow$.506	20 $\rightarrow$.017	++ $\rightarrow$.017	M2 $\rightarrow$.523
	10 $\rightarrow$.342	+++ $\rightarrow$.152	BUS $\rightarrow$.473
			TAXI $\rightarrow$.152

Pour passer maintenant de cette forme de codage (vérifiant (15bis), (17) et (20)) à la forme normalisée (17) – (19), il suffit, comme on l'a indiqué au § 3, après la formule (15 bis), de diviser les codages de chaque variable par $x^{*aj} \neq 0$. Nous en déduisons le modèle additif

$$Y^* = .506 X_1^* + .342 X_2^* + .152 X_3^*$$

avec X_1^* , X_2^* données en figure 3 et X_3^* le codage suivant :

$x_3 \rightarrow x_3^*$
 $0 \rightarrow 0$
 $+$ $\rightarrow .112$
 $++ \rightarrow .112$
 $+++ \rightarrow 1$

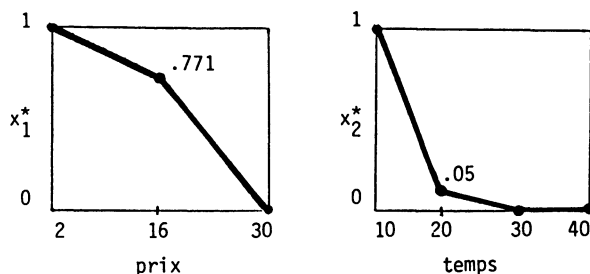


Figure 3. — Codage normalisé des variables numériques prix et temps.

5. CONCLUSION

Nous avons présenté un bref aperçu des méthodes en analyse de régression intégrant des formulations de programmation linéaire continue. Nous avons pu constater la puissance de ces modèles, tant au niveau de leur structure (des contraintes peuvent être ajoutées ou supprimées) qu'au niveau de leur flexibilité (possibilité d'effectuer des analyses de stabilité).

La méthode UTA, en régression ordinale, en particulier, est programmée dans son ensemble (les quatre étapes) en Basic et fonctionne en mode conversationnel sur micro-ordinateur ZENITH 100 (ou compatible IBM).

Il existe actuellement trois différentes facettes de mise en application de la méthodologie UTA comme :

- outil de description ou procédure cognitive ;
- modèle prédictif des choix de l'individu fondé sur la cohérence de situations décisionnelles passées ;
- méthode d'apprentissage itératif des préférences de l'individu, mise au point sous forme d'une procédure du type essai-erreur.

Plusieurs applications réelles ont été effectuées ou sont actuellement en cours (cf. JACQUET-LAGREZE et SISKOS [8], SISKOS [17]).

REFERENCES

- [1] P. CAZES, A. BAUMERDER, S. BONNEFOUS et J.P. PAGES. — Codage et analyse des tableaux logiques: Introduction à la pratique des variables qualitatives, *BURO*, Cahier n° 27, 1977, p. 3-47.
- [2] A. CHARNES, W.W. COOPER et R.O. FERGUSON. — Optimal estimation of executive compensation by linear programming, *Management Science*, 1 (2), 1955, p. 138-151.
- [3] A. CHARNES et W.W. COOPER. — *Management models and industrial applications of linear programming: Vol. 1*, Wiley, New York, 1961.
- [4] C. DUPONT-GATELMAND. — Une méthode de classification automatique sur variables hétérogènes, *Revue de Statistique Appliquée*, Vol XXVII, 1979, p. 23-37.
- [5] R. FAURE. — La programmation linéaire appliquée, *Que sais-je ?* n° 1776, PUF, Paris, 1979.
- [6] N. FREED et F. GLOVER. — Simple but powerful goal programming models for discriminant problems, *European Journal of Operational Research*, 7, 1981, p. 44-60.
- [7] E. JACQUET-LAGREZE et J. SISKOS. — Assessing a set of additive utility functions for multicriteria decision-making: The UTA method, *European Journal of Operational Research*, 10, 1982, p. 151-164 ; version anglaise révisée du : Cahier LAMSADE n° 16, Université de Paris-Dauphine, 1978.
- [8] E. JACQUET-LAGREZE et J. SISKOS. — *Méthode de décision multicritère*, Hommes et Techniques, Paris, 1983.
- [9] O.J. KARST. — Linear curve fitting using least deviations, *Journal of the American Statistical Association*, 53, 1958, p. 118-132.
- [10] J. E. KELLEY. — An application of linear programming to curve fitting, *Journal of Industrial and Applied Mathematics*, 6, 1958, p. 15-22.
- [11] E.A. KIOUNTOUZIS. — Linear programming techniques in regression analysis, *Applied Statistics*, 1, 1973, p. 69-73.
- [12] J.B. KRUSKAL. — Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, *Psychometrika*, 29, 1964, p. 1-27.
- [13] J.B. KRUSKAL. — Analysis of factorial experiments by estimating monotone transformations of the data, *Journal of the Royal Statistical Society, Series B*, 27, 1965, p. 251-263.
- [14] A. MACQUIN. — *La régression qualitative avec variables ordinales: Problèmes méthodologiques et applications*, Thèse de 3^e cycle, Université de Paris-Dauphine, 1980.
- [15] G. SAPORTA. — *Liaison entre plusieurs ensembles de variables et codages de données qualitatives*, Thèse de 3^e cycle, Université de Paris VI, 1975.
- [16] J. SISKOS. — Comment modéliser les préférences au moyen de fonctions d'utilité additives, *RAIRO Recherche Opérationnelle*, 14, 1980, p. 53-82.
- [17] J. SISKOS. — Application de la méthode UTA I à un problème de sélection de points de vente mettant en jeu des critères multiples, *RAIRO Recherche Opérationnelle*, 17, 1983, 121-136.

- [18] J. SISKOS et D. YANNAKOPOULOS. — *Amélioration de la méthode UTA par introduction d'une double fonction d'erreurs*, Cahier LAMSADE n° 49, Université de Paris-Dauphine, 1983.
- [19] V. SRINIVASAN et A.D. SHOCKER. — Estimating the weights for multiple attributes in a composite criterion using pairwise judgments, *Psychometrika*, 38, 1973, p. 473-493.
- [20] M. TENENHAUS. — Analyse en composantes principales d'un ensemble de variables nominales et numériques, *Revue de Statistique Appliquée*, Vol. XXV, n° 2, 1977, p. 39-56.
- [21] M. TENENHAUS, La régression qualitative. — *Revue de Statistique Appliquée*, Vol. XXVII, n° 2, 1979, p. 5-21.
- [22] H.M. WAGNER, Linear programming techniques for regression analysis, *Journal of the American Statistical Association*, 54, 1959, p. 206-212.
- [23] H.G. WILSON, Least-squares versus minimum absolute deviations estimation in linear models, *Decision Sciences*, 9, 1978, p. 322-335.
- [24] F.W. YOUNG, J. DE LEEUW et Y. TAKANE. — Regression with qualitative and quantitative variables: An alternating least squares method with optimal scaling features, *Psychometrika*, 41, 1976, p. 505-529.