AMÉLIORATION DE LA METHODE UTA PAR
INTRODUCTION D'UNE DOUBLE FONCTION D'ERREURS

CAHIER N° 49
octobre 1983

J. SISKOS
D. YANNACOPOULOS
<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>1</td>
</tr>
<tr>
<td>RESUME</td>
<td>II</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. UTA AMELIOREE VERSUS UTA</td>
<td>3</td>
</tr>
<tr>
<td>3. UN EXEMPLE NUMERIQUE</td>
<td>5</td>
</tr>
<tr>
<td>4. UNE EXPERIMENTATION ET CONCLUSION</td>
<td>7</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>12</td>
</tr>
</tbody>
</table>
IMPROVING THE UTA ORDINAL REGRESSION METHOD
BY INTRODUCTION OF A DOUBLE ERROR FUNCTION

ABSTRACT

This paper presents an improved version of the UTA method performing an ordinal regression analysis using more powerful linear programming formulations. The ordinal variable to be analysed is a pre-order relation whereas the independent variables are criteria, i.e. quantitative and/or qualitative monotone variables. The method is illustrated on a simple numerical example. Finally, experimental results are given, showing, upon three distinct indicators, the superiority of the adjustments obtained with the new method.

Key-words: Ordinal regression; Additive utility; Multicriteria analysis.
AMELIORATION DE LA METHODE UTA
PAR INTRODUCTION D'UNE DOUBLE FONCTION D'ERREURS

RESUME

Ce cahier présente une version améliorée de la méthode UTA effectuant une analyse de régression ordinaire à l'aide de formulations de programmation linéaire différentes. La variable ordinaire à analyser est un préordre de préférence et les variables indépendantes des critères multiples, c'est-à-dire des variables quantitatives et/ou qualitatives. La méthode est illustrée sur un exemple numérique simple. Finalement, on montre la supé-
1. INTRODUCTION

Le problème de régression ordinaire, tel qu'il est traité par la méthode UTA (cf. Jacquet-Lagrèze et Siskos [2]), s'explicite de la façon suivante : soit $(\succ, \sim)$ une structure de préférences sous forme de préordre donnée sur un ensemble d'objets ou actions où $\succ$ désigne la préférence et $\sim$ l'indifférence ; on cherche à ajuster des fonctions d'utilité additives, sur critères multiples, de façon que la structure préférentielle résultant de ces utilités soit aussi compatible que possible avec la structure initiale.

Appelons $A = \{a, b, c, \ldots\}$ l'ensemble des actions sur lequel est donnée la structure $(\succ, \sim)$. Soit $g_1, g_2, \ldots, g_n$ une famille de $n$ critères d'évaluation ; chacun est défini ici sous la forme d'une application réelle monotone $g_i : A \times \{g_i^*, g_i^\#\} \rightarrow \mathbb{R}$ de façon à ce que $g_i(a)$ représente l'évaluation de l'action $a$ sur le critère $g_i$. Les valeurs $g_i^* = \min_{a \in A} g_i(a)$, $g_i^\# = \max_{a \in A} g_i(a)$ représentent respectivement le niveau le moins et le plus préféré du critère.

Au niveau d'un seul critère, les préférences s'explicite de la façon suivante :

\[ a \succ b \iff g_i(a) > g_i(b) \quad (1) \]
\[ a \sim b \iff g_i(a) = g_i(b) \quad (2) \]

ce qui signifie que chaque critère définit sur l'ensemble $A$ une relation de préordre $(\succ, \sim)$.

Une fonction d'utilité en présence de certitude est une application réelle $u : \times_{i=1}^n [g_i^*, g_i^\#] \rightarrow \mathbb{R}$ définie par les relations :

\[ a \succ b \iff u[g_i(a)] > u[g_i(b)] \quad (3) \]
\[ a \sim b \iff u[g_i(a)] = u[g_i(b)] \quad (4) \]
Dans la version d'UTA développée dans ce cahier, nous proposons une fonction d'erreurs composée de deux erreurs potentielles par action. Cela permettrait de mieux stabiliser la position des points autour de la courbe (figure 1). Ainsi, l'utilité d'une action $a \in A$ sera remplacée dans les relations (3)-(4) par: $u[a] + \sigma^+(a) - \sigma^-(a)$, $\sigma^+(a) \geq 0$, $\sigma^-(a) \geq 0$.

Dans la section suivante, nous présentons succinctement le modèle UTA ainsi que les améliorations introduites. La section 3 permet une comparaison des deux modèles à l'aide d'un exemple numérique à cinq actions et trois critères. Dans la dernière section, nous montrons expérimentalement la supériorité du nouveau modèle sur l'ancien en utilisant trois indicateurs différents.

2. **UTA AMELIOREE VERSUS UTA**

UTA utilise une formulation de programmation linéaire pour estimer les utilités marginales $u_1$ sous les conditions (5)-(7). La technique d'estimation présume la discrétisation de chaque intervalle de variation des critères

$$[g_{i1}^*, g_{i1}^+] = [g_{i1} = g_{i1}^1, g_{i1}^2, ..., g_{i1}^{a_i} = g_{i1}^*]$$

et l'introduction de contraintes $u_1(g_{i1}^{j+1}) \geq u_1(g_{i1}^j)$, $\forall j$ afin de préserver la monotonie des critères. Le nombre de points équidistants $a_i$ peut se calculer par algorithme dans la mesure de l'information disponible; pour les critères quantitatifs, nous employons la technique de l'interpolation linéaire.

Suivant les prescriptions (3)-(7), l'ancien modèle UTA se schématisse en quatre étapes:

1) Exprimer, dans l'ordre imposé par le préordre initial $(\succ, \succeq)$, les utilités des actions $u[a]$, $a \in A$ en fonction des utilités marginales $u_1(g_{i1}^j)$. 
2) Définir une chaîne sur le préordre allant de la tête à la queue en écrivant, pour chaque paire \((a, b)\) d'actions consécutives, les expressions analytiques :

\[
\Delta(a, b) \equiv u[g(a)] - u[g(b)] + \sigma(a) - \sigma(b).
\]  \hspace{1cm} (9)

Le nombre de ces expressions est égal au nombre d'actions moins 1.

3) Résoudre le dual du programme linéaire :

Minimiser \( F = \sum_{a \in A} \sigma(a) \)

sous les contraintes (d'après l'étape 2)

\[
\begin{align*}
\Delta(a, b) & \geq 0 \text{ si } a \succ b \\
\Delta(a, b) & = 0 \text{ si } a \sim b \\
u_i(g_i^{j+1}) - u_i(g_i^j) & \geq 0 \forall i \text{ et } j \\
\Sigma_i u_i(g_i^x) & = 1 \\
u_i(g_i^x) & = 0, u_i(g_i^j) \geq 0, \sigma(a) \geq 0 \forall a \in A, \forall i \text{ et } j
\end{align*}
\]

\( \delta \) une petite valeur positive.

4) Tester l'existence de solutions optimales multiples. En cas de dégénérescence, trouver celles des solutions qui maximisent et/ou minimisent \( u_i(g_i^x) \) pour chaque \( i \).

Les modifications intégrées dans le nouveau modèle sont, par étape, les suivantes :

1) Les contraintes de monotonicité des critères sont prises en compte dans les transformations de variables

\[
w_{ij} = u_i(g_i^{j+1}) - u_i(g_i^j) \geq 0 \forall i \text{ et } j. \]  \hspace{1cm} (10)

Les utilités \( u[g(a)] \) deviennent donc fonctions de \( w_{ij} \), c'est-à-dire, comme \( u_i(g_i^1) = 0 \), on a, pour \( j > 1 \) :

\[
u_i(g_i^j) = \sum_{k=1}^{j-1} w_{ik}. \]  \hspace{1cm} (11)
2) Introduction d'une double fonction d'erreurs : poser pour chaque paire d'actions consécutives \((a, b)\) dans la chaîne

\[ \Delta(a, b) = u[g(a)] - u[g(b)] + \sigma^+(a) - \sigma^-(a) - \sigma^+(b) + \sigma^-(b). \] (12)

3) Résoudre le primal du programme linéaire :

Minimiser \( F = \sum_{a \in A} [\sigma^+(a) + \sigma^-(a)] \)

sous les contraintes

\( \Delta(a, b) \geq \delta \) si \( a \succ b \)

\( \Delta(a, b) = 0 \) si \( a \sim b \)

\( \sum_j w_{ij} = 1 \)

\( w_{ij} \geq 0, \sigma^+(a) \geq 0, \sigma^-(a) \geq 0 \) \( \forall a \in A, \forall i \) et \( j \)

\( \delta \) une petite valeur positive.

4) Sans changement.

Remarque : Cette formulation s'apparente à celles développées en "goal programming" ([1]). Il est donc facile de prouver qu'à l'optimum on aura \( \sigma^+(a) \cdot \sigma^-(a) = 0, \forall a \in A \), c'est-à-dire au moins l'une des erreurs nulle.

3. UN EXEMPLE NUMERIQUE

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 (1) prix (en Francs), (2) temps du trajet (en minutes) et (3) confort (possibilités de s'asseoir), attribue aux cinq modes possibles le classement (pré-ordre) suivant : RER \( \succ \) (METRO 1ère \( \sim \) METRO 2e) \( \succ \) BUS \( \sim \) TAXI (cf. Tableau 1). Pour le critère confort, une échelle qualitative 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.
Tableau 1 : Problème d'analyse du classement de modes de transport

<table>
<thead>
<tr>
<th>Modes de transport</th>
<th>Rangs</th>
<th>Prix (Francs)</th>
<th>Temps (minutes)</th>
<th>Confort (qualitatif)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RER</td>
<td>1</td>
<td>3</td>
<td>10</td>
<td>+</td>
</tr>
<tr>
<td>METRO 1ère</td>
<td>2</td>
<td>4</td>
<td>20</td>
<td>++</td>
</tr>
<tr>
<td>METRO 2e</td>
<td>2</td>
<td>2</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>BUS</td>
<td>3</td>
<td>6</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>TAXI</td>
<td>4</td>
<td>30</td>
<td>30</td>
<td>+++</td>
</tr>
</tbody>
</table>

La première étape d'UTA consiste à expliciter les utilités des cinq actions. Nous avons retenu, pour le calcul analytique des utilités, les échelles discrètes suivantes :

\[
[g_{1*}, g_1^+] = [30, 16, 2]
\]
\[
[g_{2*}, g_2^+] = [40, 30, 20, 10]
\]
\[
[g_{3*}, g_3^+] = [0, +, ++, +++]
\]

d'où, par interpolation linéaire pour le critère \( g_1 \), on trouve :

\[
u[g(\text{RER})] = .07 u_1(16) + .93 u_1(2) + u_2(10) + u_3(+)
\]
\[
u[g(\text{METRO 1})] = .14 u_1(16) + .86 u_1(2) + u_2(20) + u_3(++)
\]
\[
u[g(\text{METRO 2})] = u_1(2) + u_2(20) + u_3(0) = u_1(2) + u_2(20)
\]
\[
u[g(\text{BUS})] \quad .29 u_1(16) + .71 u_1(2) + u_2(40) + u_3(+)
\]
\[
\quad = .29 u_1(16) + .71 u_1(2)
\]
\[
u[g(\text{TAXI})] = u_1(30) + u_2(30) + u_3(++) = u_2(30) + u_3(++).
\]

Ensuite, en remplaçant les \( u_i(g_j^{}) \) par les \( w_{ij} \) (cf. relations 10-11) :

\[
u[g(\text{RER})] = w_{11} + .93 w_{12} + w_{21} + w_{22} + w_{23} + w_{31}
\]
\[
u[g(\text{METRO 1})] = w_{11} + .86 w_{12} + w_{21} + w_{22} + w_{31} + w_{32}
\]
\[
u[g(\text{METRO 2})] = w_{11} + w_{12} + w_{21} + w_{22}
\]
\[
u[g(\text{BUS})] = w_{11} + .71 w_{12}
\]
\[
u[g(\text{TAXI})] = w_{21} + w_{31} + w_{32} + w_{33}.
\]
Finalement, en passant par l'étape 2 de comparaison des actions par paires (RER ⊃ METRO 1 ∧ METRO 2 ⊃ BUS ⊃ TAXI), on arrive bien au programme linéaire du nouveau modèle UTA (δ est pris égal à .1) :

<table>
<thead>
<tr>
<th>$w_{12}$</th>
<th>$w_{21}$</th>
<th>$w_{23}$</th>
<th>$w_{31}$</th>
<th>$w_{32}$</th>
<th>$w_{33}$</th>
<th>variables $σ^+$ et $σ^-$</th>
<th>Signe</th>
<th>Second membre</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.07</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0 -1 -1 -1 1</td>
<td>≥</td>
<td>.1</td>
</tr>
<tr>
<td>0</td>
<td>-.14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1 -1 -1 1</td>
<td>=</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>.29</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0 -1 -1 1</td>
<td>≥</td>
<td>.1</td>
</tr>
<tr>
<td>1</td>
<td>.71</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1 -1 -1</td>
<td>≥</td>
<td>.1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>=</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>F</td>
<td></td>
</tr>
</tbody>
</table>

A partir de la solution optimale du programme linéaire, on trouve : $u_1(2) = .5$, $u_1(16) = 0$, $u_1(30) = 0$, $u_2(10) = .2$, $u_2(20) = .1$, $u_2(30) = .1$, $u_2(40) = 0$, $u_3(0) = 0$, $u_3(+) = 0$, $u_3(++) = 0$, $u_3(++++) = .3$ et $F^* = \min F = 0$ mais cette solution n'est pas unique. L'ancienne version donne $u_1(2) = .56$, $u_1(16) = 0$, $u_1(30) = 0$, $u_2(10) = .14$, $u_2(20) = 0$, $u_2(30) = 0$, $u_2(40) = 0$, $u_3(0) = 0$, $u_3(+) = 0$, $u_3(++) = .08$, $u_3(++++) = .3$ et $F^* = 0$.

La restitution numérique du préordre est, pour les deux modèles, comme suit : Ancienne version : RER + .6, METRO 1ère + .56, METRO 2e + .56, BUS + .4, TAXI + .3 ; UTA amélioré : RER + .7, METRO 1ère + .6, METRO 2e + .6, BUS + .5, TAXI + .4.

Une analyse post-optimale, au sens où elle a été définie par l'étape 4, donnerait un ensemble représentatif de solutions optimales dont on pourrait éventuellement retenir, comme solution unique, le centre de gravité (voir [3] pour une discussion sur ce point).

4. UNE EXPERIMENTATION ET CONCLUSION

Dans le but d'évaluer quantitativement les avantages offerts par la nouvelle version d'UTA vis-à-vis de l'ancienne, nous avons procédé à une
expérimentation des deux modèles sur des données simulées. En effet, nous avons considéré un ensemble de dix actions évaluées sur six critères (tableau 2) sur lequel nous avons généré des préordres aléatoires (de l'ordre d'une vingtaine). Les évaluations des actions ont été gardées inchangées et nous avons pris pour chaque essai \( \delta = .05 \). Les échelles des critères sont définies dans le sens : \( 4 \succ 3 \succ 2 \succ 1 \).

**Tableau 2 : Données de l'expérimentation**

<table>
<thead>
<tr>
<th>Actions</th>
<th>Critères</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 4 2 3 4 4</td>
</tr>
<tr>
<td>2</td>
<td>3 3 3 2 3 2</td>
</tr>
<tr>
<td>3</td>
<td>1 4 2 2 4 4</td>
</tr>
<tr>
<td>4</td>
<td>2 3 1 3 ? 3</td>
</tr>
</tbody>
</table>
### Tableau 3 : Résultats sommaires de l'expérimentation

<table>
<thead>
<tr>
<th>Indicateur</th>
<th>UTA amélioré</th>
<th>UTA classique</th>
<th>Bilan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nombre moyen d'itér. simplexe</td>
<td>11.1</td>
<td>21.5</td>
<td>baisse de 48 %</td>
</tr>
<tr>
<td>τ de Kendall moyen</td>
<td>.69</td>
<td>.63</td>
<td>croissance de 10 %</td>
</tr>
<tr>
<td>Indice de conversion moyenne ( L^* )</td>
<td>24</td>
<td>53</td>
<td>baisse de 25 %</td>
</tr>
</tbody>
</table>
 continu pour l'ancienne version et une courbe en pointillés pour la nouvelle). Le deuxième jeu de données provient de l'31 n. 127 et comporte
Figure 3: Nouvelle version d'UTA versus l'ancienne sur le problème du choix de points de vente [3].
REFERENCES


