

R. ROUSSEAU

K. BEN SALEM

Problème sur l'estimation d'une donnée manquante d'après le critère de la trace minima

Les cahiers de l'analyse des données, tome 18, n° 3 (1993),
p. 359-368

http://www.numdam.org/item?id=CAD_1993__18_3_359_0

© Les cahiers de l'analyse des données, Dunod, 1993, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » 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/>

PROBLÈME SUR L'ESTIMATION D'UNE DONNÉE MANQUANTE D'APRÈS LE CRITÈRE DE LA TRACE MINIMA

[PROB. TRAC. MIN.]

R. ROUSSEAU*
K. Ben SALEM**

1 Énoncé du problème

L'objet du problème est de reconstituer une unique donnée manquante d'un tableau de correspondance, selon une nouvelle méthode, d'après le "critère de la trace minima" (cf. [TRAC. MANQ.] in CAD, Vol.XVII, n°1, 1992).

1.1 Définition de la quantité critère

Soit un tableau k_{IJ} de correspondance quelconque sur $I \times J$, I et J étant des ensembles finis. On note classiquement k le total des $k(i,j)$; et $k(i)$, $k(j)$, respectivement, les totaux de la ligne i et de la colonne j ; (ces totaux étant supposés, dans la suite, être tous non nuls); et on note f_{IJ} la loi de probabilité sur $I \times J$ associée au tableau k_{IJ} : $\forall i \in I, \forall j \in J : f_{ij} = k(i,j) / k$;

On définit la quantité suivante :

$$\text{Crit}(k) = \sum \{ (k(i,j))^2 / (k(i).k(j)) \mid i \in I, j \in J \} - 1 \ .$$

1.1.1 On considère la matrice $f_j^J = f_j^I \circ f_I^J$.

Exprimer $\text{Crit}(k)$ en fonction de $\text{tr}(f_j^J)$, trace de f_j^J , i.e., somme des termes diagonaux de la matrice f_j^J .

Exprimer $\text{Crit}(k)$ en fonction de $\text{Int}(k)$, inertie totale du tableau k_{IJ} , dont on rappelle qu'elle s'écrit notamment sous les deux formes suivantes :

(*) Université Catholique de l'Ouest; Angers.

(**) Faculté des Sciences; Tunis.

$$\text{Int}(k) = \sum \{ f_i \cdot \| f_j^i - f_j \|^2 \mid i \in I \} = \sum \{ f_j \cdot \| f_i^j - f_i \|^2 \mid j \in J \} .$$

1.1.2 On suppose que f_{IJ} est une loi produit sur $I \times J$, i.e.:

$$f_{IJ} = f_I \otimes f_J ;$$

$$\forall i \in I, \forall j \in J : f_{ij} = f_i \cdot f_j ;$$

Quelle est, dans ce cas, la valeur de $\text{Crit}(k)$?

1.1.3 On suppose maintenant que $\text{Crit}(k) = 0$.

Pour i et j quelconques de I et J , exprimer f_{ij} en fonction de f_i et de f_j .

1.2 Reconstitution de l'unique donnée manquante

On considère dorénavant un tableau de correspondance k_{IJ} ayant une unique donnée manquante située dans la case (i_o, j_o) . On suppose de plus que le tableau k_{IJ} n'est pas réduit à la seule ligne i_o ni à la seule colonne j_o ; et n'a ni ligne ni colonne nulle.

On note f_{IJ} la loi de probabilité sur $I \times J$ associée au tableau k_{IJ} dans lequel la valeur manquante reçoit la valeur ZÉRO.

On note a_{IJ} la loi de probabilité ponctuelle sur $I \times J$ concentrée en (i_o, j_o) , i.e. définie par :

$$\forall i \in I, \forall j \in J : \text{ si } (i, j) = (i_o, j_o), a_{ij} = 1 ; \text{ sinon, } a_{ij} = 0 .$$

Pour $x \in [0, 1[$, on note t_{IJ} le tableau de correspondance sur $I \times J$ suivant :

$$t_{IJ} = (1-x) f_{IJ} + x \cdot a_{IJ} .$$

On cherche à déterminer x , de telle sorte que $\text{Crit}(t)$ soit minimum; et, pour être proportionnel à t_{IJ} , le tableau k_{IJ} sera complété par $k(i_o, j_o) = k \cdot x / (1-x)$.

1.2.1 À quoi est égale la somme des éléments du tableau t_{IJ} ?

1.2.2 Pour $i \neq i_o$ et $j \neq j_o$, exprimer $t(i, j)^2 / (t(i)t(j))$ en fonction des f_{ij} , f_i et f_j .

1.2.3 On note: $H = \sum \{ f_{i_o, j} \mid j \in J, j \neq j_o \} = f_{i_o} ;$

$$G = \sum \{ f_{i, j_o} \mid i \in I, i \neq i_o \} = f_{j_o} .$$

- a) Exprimer $t(i_0, j_0)^2 / (t(i_0) t(j_0))$ en fonction de x , H et G .
- b) Pour $j \neq j_0$, exprimer $t(i_0, j)^2 / (t(i_0) t(j))$ en fonction de x , H , $f_{i_0 j}$ et f_j .
- c) Pour $i \neq i_0$, exprimer $t(i, j_0)^2 / (t(i) t(j_0))$ en fonction de x , G , $f_{i j_0}$ et f_i .

$$1.2.4 \text{ On note: } h = \sum \{ f_{i_0 j}^2 / f_j \mid j \in J, j \neq j_0 \} ;$$

$$g = \sum \{ f_{i j_0}^2 / f_i \mid i \in I, i \neq i_0 \} .$$

Comparer h à H ainsi que g à G .

$$1.2.5 \text{ On note } Y = \sum \{ k(i, j)^2 / (k(i)k(j)) \mid i \in I - \{i_0\}; j \in J - \{j_0\} \} .$$

- a) Exprimer $\text{Crit}(t)$ en fonction de x , H , G , h , g et Y .
- b) On note: $U = 1 - H$; et : $V = 1 - G$.

Exprimer $\text{Crit}(t)$, que l'on notera désormais $\text{Cr}(x)$, sous la forme suivante :

$$\text{Crit}(t) = \text{Cr}(x) = (Y - 1) + (P(x) / Q(x)) ;$$

où $P(x)$ et $Q(x)$ sont des trinômes du second degré en x que l'on déterminera sans chercher à les mettre sous la forme explicite $a.x^2 + b.x + c$.

- c) De quel signe sont les deux racines de Q ?

1.2.6 On suppose dans cette question qu'on a les inégalités strictes:

$$h + g < H + G ; \text{ avec : } H > 0 \text{ et } G > 0 .$$

- a) Calculer les valeurs de $\text{Cr}'(0)$ et $\text{Cr}'(1)$, de la dérivée de Cr en $x=0$ et $x=1$.
- b) En déduire le nombre de zéros de Cr' sur l'intervalle $[0, 1[$.
- c) Conclure sur l'existence et l'unicité de la valeur x , sur $[0, 1[$, à attribuer à la case manquante (i_0, j_0) pour donner une valeur minimum au critère Crit .

1.2.7 On se propose, dans cette question, de considérer le cas où $h = H$ et $g = G$; en revenant sur la question 1.2.4.

- a) Quelle est la forme du tableau k_{IJ} si $h=H$?
- b) Quelle est la forme du tableau k_{IJ} si $h=H$ et $g=G$?
- c) Dans ce dernier cas, exprimer $\text{Crit}(t)$ en fonction de x , Y , H , G . Y a-t-il, dans ce cas une valeur de $x \in [0, 1[$ rendant $\text{Cr}(x)$ minimum ?

2 Solution du problème

2.1 Étude de la quantité critère

2.1.1 Exprimer Crit(k) en fonction de la trace $\text{tr}(f_J^J)$

$$\begin{aligned} \text{Crit}(k) &= \sum\{ (k(i,j)/k(i)) \cdot (k(i,j)/k(j)) \mid i \in I, j \in J \} - 1 \\ &= \sum\{ f_j^i \cdot f_i^j \mid i \in I, j \in J \} - 1 \\ &= \sum\{ f_j^j \mid j \in J \} - 1 = \text{tr}(f_J^J) - 1 . \end{aligned}$$

En déduire l'expression de Crit(k) en fonction de Int(k)

On sait que Int(k) est l'inertie du nuage N(J) des variables (colonnes); laquelle est égale à la somme des valeurs propres non nulles et non triviales; i.e.:

$$\text{Int}(k) = \text{tr}(f_J^J) - 1 \quad ; \quad \text{d'où il résulte que : } \text{Crit}(k) = \text{Int}(k) .$$

2.1.2 Si $f_{IJ} = f_I \otimes f_J$, quelle est la valeur de Crit(k)?

On a : $\forall i \in I, \forall j \in J : f_{ij} = f_i \cdot f_j$;

i.e. : $\forall i \in I, \forall j \in J : k(i,j) = k \cdot (k(i)/k) \cdot (k(j)/k) = (k(i) \cdot k(j)) / k$;

$$\begin{aligned} \text{donc : } \text{Crit}(k) &= \sum\{ k(i) \cdot k(j) / k^2 \mid i \in I, j \in J \} - 1 \\ &= \sum\{ f_i \cdot f_j \mid i \in I, j \in J \} - 1 = 1 - 1 = 0 . \end{aligned}$$

2.1.3 On suppose que Crit(k)=0; exprimer f_{ij} en fonction de f_i et f_j

On a : $0 = \text{Int}(k) = \sum \{ f_i \cdot \| f_j^i - f_j \|^2 \mid i \in I \}$;

donc tout individu i , de poids non nul, (ce qu'on a, d'ailleurs supposé être le cas de tout i), est tel que $\| f_j^i - f_j \|^2 = 0$; ou, en d'autres termes, a un profil f_j^i qui coïncide avec le profil f_j , du centre de gravité du nuage N(I). Donc:

$$\begin{aligned} \forall i \in I \quad & : (f_i \neq 0) \Rightarrow f_j^i = f_j \quad ; \\ \forall i \in I, \forall j \in J : (f_i \neq 0) & \Rightarrow f_j^i = f_j \Leftrightarrow (f_{ij}/f_i) = f_j \quad ; \\ \forall i \in I, \forall j \in J : (f_i \neq 0) & \Rightarrow f_{ij} = f_i \cdot f_j \quad ; \end{aligned}$$

cette dernière propriété valant évidemment si $(f_i = 0)$, on a :

$$f_{IJ} = f_I \otimes f_J \quad ;$$

nous avons redémontré l'équivalence générale entre trace=0 et indépendance.

2.2 Reconstitution de l'unique donnée manquante

2.2.1 Valeur de la somme des éléments du tableau t_{IJ}

le tableau t_{IJ} est défini comme barycentre de deux lois de probabilité; t_{IJ} est donc également une loi de probabilité; et la somme demandée est 1.

2.2.2 Pour $i \neq i_o, j \neq j_o$, exprimer $t(i,j)^2 / (t(i).t(j))$ en fonction de f_{IJ}

$$\forall i \in I, \forall j \in J : (i \neq i_o) \Rightarrow t(i,j) = (1-x).f_{ij} + x.a_{ij} = (1-x).f_{ij} ;$$

$$\forall i \in I : (i \neq i_o) \Rightarrow t(i) = (1-x).\sum\{f_{ij} \mid j \in J\} = (1-x).f_i ;$$

de même, $\forall j \in J : (j \neq j_o) \Rightarrow t(j) = (1-x).f_j ;$ donc :

$$(i \neq i_o, j \neq j_o) \Rightarrow t(i,j)^2 / (t(i).t(j)) = ((1-x).f_{ij})^2 / ((1-x).f_i . (1-x).f_j) = f_{ij}^2 / (f_i . f_j).$$

2.2.3.a Exprimer $t(i_o,j_o)^2 / (t(i_o).t(j_o))$ en fonction de x, H et G

$$t(i_o) = \sum\{t(i_o,j) \mid j \in J\} = t(i_o,j_o) + \sum\{t(i_o,j) \mid j \in J, j \neq j_o\} = x + (1-x).H ;$$

on a, de même: $t(j_o) = x + (1-x).G ;$ d'où :

$$t(i_o,j_o)^2 / (t(i_o).t(j_o)) = x^2 / ((x + (1-x).H) . (x + (1-x).G)) .$$

2.2.3.b Pour $j \neq j_o$, exprimer $t(i_o,j)^2 / (t(i_o).t(j))$

$$t(i_o,j) = (1-x).f_{i_o j} ; \quad t(i_o) = x + (1-x).H ;$$

$$t(j) = \sum\{t(i,j) \mid i \in I\} = (1-x).\sum\{f_{ij} \mid i \in I\} = (1-x).f_j ;$$

$$t(i_o,j)^2 / (t(i_o).t(j)) = ((1-x) / (x + (1-x).H)) . (f_{i_o j}^2 / f_j).$$

2.2.3.c Pour $i \neq i_o$, exprimer $t(i,j_o)^2 / (t(i).t(j_o))$

par symétrie, il vient:

$$t(i,j_o)^2 / (t(i).t(j_o)) = ((1-x) / (x + (1-x).G)) . (f_{i j_o}^2 / f_i).$$

2.2.4 Comparer h à H , ainsi que g à G

Puisque $f_{i_o j}$ est l'un des termes de la somme f_j , (qui est non nulle par hypothèse) on a : $f_{i_o j} \leq f_j$; donc:

$$\forall j \in J : (f_{i_o j}^2 / f_j) \leq f_{i_o j} ;$$

$$h = \sum\{(f_{i_o j}^2 / f_j) \mid j \in J\} \leq \sum\{f_{i_o j} \mid j \in J\} = H ;$$

de même, par symétrie, $g \leq G$.

2.2.5 On note $Y = \sum\{k(i,j)^2 / (k(i).k(j)) \mid i \in I, i \neq i_o ; j \in J, j \neq j_o\}$

2.2.5.a Exprimer Crit(t) en fonction de x, H, G, h, g, Y

$$\text{Crit}(t) = \sum \{t(i,j)^2 / (t(i).t(j)) \mid i \in I, j \in J\} - 1$$

Dans la somme double, on distingue:

le terme en (i_0, j_0) , calculé au §2.2.3.a ;

les termes en (i_0, j) , pour $j \neq j_0$, dont on a montré au §2.2.3.b, qu'ils sont proportionnels aux $(f_{i_0, j}^2 / f_j)$, le coefficient étant $((1-x) / (x + (1-x).H))$;

les termes en (i, j_0) , pour $i \neq i_0$, traités semblablement au §2.2.3.b ;

enfin, les termes en (i, j) , pour $i \neq i_0, j \neq j_0$, dont, d'après le §2.2.2, la somme n'est autre que ce qu'on a noté Y.

Il en résulte que l'on a pour Crit(t):

$$\begin{aligned} \text{Crit}(t) &= x^2 / ((x + (1-x).H).(x + (1-x).G)) \\ &+ ((1-x).h / (x + (1-x).H)) + ((1-x).g / (x + (1-x).G)) + Y - 1 . \end{aligned}$$

2.2.5.b Exprimer Crit(t) = Cr(x), comme: (Y-1) + (P(x)/Q(x))

Cr(x) est la somme de (Y-1) et de trois termes dépendant de x, qu'on peut réduire à un même dénominateur; en notant $u=(1-H)$ et $V=(1-G)$, il vient:

$$\begin{aligned} \text{Cr}(x) &= (Y-1) + \\ &(x^2 + (1-x).h.(Vx+G) + (1-x).g.(Ux+H)) / ((Ux+H).(Vx+G)) . \end{aligned}$$

2.2.5.c Étudier le signe des zéros du dénominateur Q(x)

Au §2.2.3.b, dénominateur Q(x) est exprimé comme le produit de deux formes linéaires: les racines de l'équation $Q(x)=0$ sont donc en évidence; il reste à en examiner les valeurs. Considérons le premier facteur; on a :

$$U.x + H = 0 \Leftrightarrow x = -H/U = -H/(1-H) ; \quad \text{où : } 0 \leq H = f_{i_0} \leq 1 ;$$

selon l'hypothèse, faite ci-dessus, que I n'est pas réduit à une seule ligne, i_0 , contenant toute la masse, on a l'inégalité stricte: $(H < 1)$, il y a une racine; celle-ci est strictement négative si H est strictement positif. De même, puisque J n'est pas réduit à une seule colonne, j_0 , il y a une deuxième racine; négative si $G > 0$.

N.B.: Les deux inégalités strictes $(H > 0 ; G > 0)$, résultent de ce qu'on a supposé que f_{IJ} n'a pas de ligne ni de colonne nulle, notamment i_0 et j_0 .

2.2.6 On suppose vérifiées les inégalités strictes: $h+g < H+G$, $0 < H$, $0 < G$

2.2.6.a Calcul de la dérivée, $Cr'(x)$ pour $x=0$ et $x=1$

$$Cr(x) = (Y-1) + (P(x)/Q(x)) \quad ; \quad Cr'(x) = (P'(x).Q(x) - P(x).Q'(x)) / Q(x)^2 ;$$

$$\begin{aligned} P(x) &= x^2 + (1-x).h.(Vx+G) + (1-x).g.(Ux+H) & ; \\ P(0) &= h.G + g.H & ; \quad P(1) = 1 & ; \\ P'(0) &= h.V + g.U - (h.G + g.H) & ; \quad P'(1) = 2 - (h + g) & ; \\ Q(x) &= (x + (1-x).H).(x + (1-x).G) = (U.x + H).(V.x + G) & ; \\ Q(0) &= H.G & ; \quad Q(1) = 1 & ; \\ Q'(0) &= U.G + V.H & ; \quad Q'(1) = 2 - (H + G) & ; \end{aligned}$$

$$\begin{aligned} Cr'(0) &= - (h.G^2 + g.H^2) / (H.G)^2 < 0 & ; \\ Cr'(1) &= (2 - (h + g)) - (2 - (H + G)) = (H + G) - (h + g) > 0 & ; \end{aligned}$$

2.2.6.b En déduire le nombre de zéros de Cr' sur $[0, 1]$

Pour un quotient $(P(x)/Q(x))$ de deux trinômes du second degré, le numérateur de la dérivée est un polynôme du second degré, et a donc, au plus deux zéros; dans le cas présent, sur l'intervalle $[0, 1]$, le numérateur de Cr' change de signe, et le dénominateur, $Q(x)^2$, reste strictement positif; il y donc exactement un zéro de Cr' .

2.2.6.c Existence et unicité du minimum de Cr sur $[0, 1]$

On a les inégalités strictes: $Cr'(0) < 0$ et $Cr'(1) > 0$; avec un seul zéro, x_0 , de Cr' , dans l'intervalle; ce zéro correspond, pour $Cr(x)$ à un unique minimum; on a:

$$t_{IJ} = (1-x_0).f_{IJ} + x_0.a_{IJ} \quad ;$$

et la donnée qui manque dans k_{IJ} doit être suppléée par:

$$k(i_0, j_0) = k \cdot (x_0 / (1-x_0)) \quad ;$$

où k est le total du tableau donné, avec un zéro pour la donnée manquante.

2.2.7 Cas où $h = H$ et $g = G$;

2.2.7.a Forme du tableau k_{IJ} si $h=H$

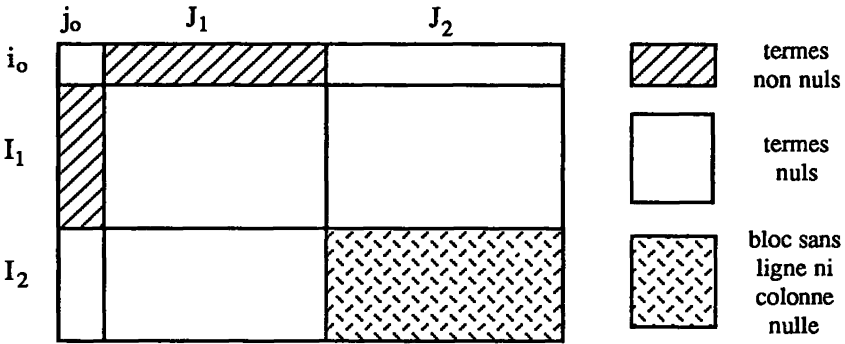
D'après ce qu'on a vu au §2.2.4, $h=H$ implique le système:

$$\forall j \in J - \{j_0\} : (f_{i_0j}^2 / f_j) = f_{i_0j} \quad ;$$

ce système est équivalent au suivant:

$$\forall j \in J - \{j_0\} : (f_{i_0j} = f_j) \vee (f_{i_0j} = 0) ;$$

en d'autres termes, ou bien une colonne j ne comporte qu'un seul terme non nul, qui est f_{i_0j} ; ou bien f_{i_0j} est nul. Cette dernière condition ne peut être réalisée pour tout $j \neq j_0$, car, en ce cas, la ligne i_0 serait nulle.



2.2.7.b Forme du tableau k_{IJ} si $h=H$ et $g=G$

D'après le § précédent, on a les deux systèmes de conditions:

$$\forall j \in J - \{j_0\} : (f_{i_0j} = f_j) \vee (f_{i_0j} = 0) ;$$

$$\forall i \in I - \{i_0\} : (f_{ij_0} = f_i) \vee (f_{ij_0} = 0) ;$$

ce qui correspond au schéma de la figure; où, compte tenu des hypothèses faites, I_1 et J_1 sont nécessairement non vides.

2.2.7.c Calcul de $Cr(x)$ en fonction de x, Y, H, G ; et recherche du minimum

Reprenons la formule du §2.2.5.a, en posant $X = x/(1-x)$; il vient :

$$\begin{aligned} Cr(x) &= x^2 / ((x + (1-x).H).(x + (1-x).G)) \\ &\quad + ((1-x).h / (x + (1-x).H)) + ((1-x).g / (x + (1-x).G)) + Y - 1 \\ &= (Y-1) + X^2/((X+H)(X+G)) + H/(X+H) + G/(X+G) \\ &= (Y-1) + X^2/((X+H)(X+G)) + H/(X+H) + G/(X+G) \\ &= (Y-1) + (X^2 + (H+G) X + 2HG)/(X^2 + (X+G)X + HG) \end{aligned}$$

$$= Y + HG/(X+G)(X+H)$$

quand x croît de 0 à 1, X croît de 0 à l'infini et le critère décroît de $Y+1$ à Y ; toutefois, le minimum correspondant à $x=1$ n'offre pas d'intérêt réel, car, pour $x=1$, la donnée manquante recevrait une valeur infinie.

Le résultat du calcul s'interprète bien d'après la forme du tableau k_{IJ} : celui-ci est décomposé en deux blocs diagonaux:

$$(I_1 \cup \{i_0\}) \times (J_1 \cup \{j_0\}) \quad ; \quad I_2 \times J_2 \quad ;$$

la trace du premier bloc décroît de 1 à 0 quand x croît de 0 à 1 ; la trace du deuxième bloc n'est autre que $(Y-1)$; la trace du tableau composé est la somme des traces des deux sous-tableaux diagonaux, augmentée de 1 ; augmentation qui correspond à l'existence d'une valeur propre non triviale 1, associée à la décomposition en deux blocs.

3 Commentaire pédagogique

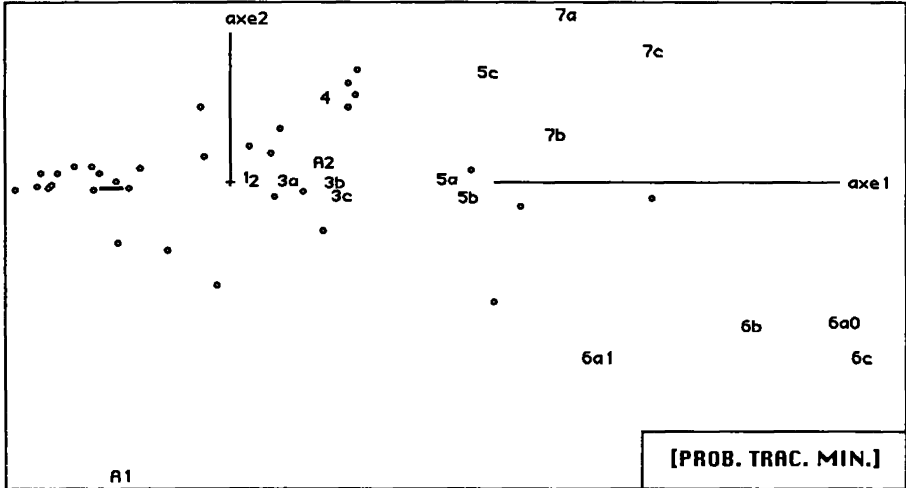
Le problème ci-dessus, (avec un énoncé différent, par quelques détails de celui publié ici,) a été proposé à 35 étudiants, travaillant pendant 3 heures sans documents. Nous croyons pouvoir servir l'enseignement de l'analyse des données en rendant compte brièvement des résultats obtenus. Les questions de la partie 1, définition du critère, seront désignées par les sigles A1, A2, A3; pour la partie 2, reconstitution de la donnée manquante, nous garderons simplement le chiffre, suivi, s'il y a lieu, d'une lettre: 1, 2, 3a, 3b, 3c...

Dans A1, tous les étudiants ont fait correctement les calculs de trace; mais pour l'inertie du nuage des individus, certains ont tenté de reprendre les calculs classiques, mais 3 seulement y sont parvenus; plusieurs ont, toutefois évoqué le résultat classique, sans le redémontrer. Le calcul de la trace dans le cas de l'indépendance a été fait par les 2/3 des étudiants; mais un seul a démontré correctement la réciproque, (trace nulle implique indépendance,) objet de la question A3.

Dans la deuxième partie, {1, 2, 3a} ont été bien réussis par presque tous; {3b, 3c} seulement par les 2/3; {4, 5a, 5b} ont été réussis par la moitié; il est plus surprenant que 1/3 seulement ait réussi 5c: alors que les zéros du dénominateur $Q(x)$ étaient en évidence, beaucoup se sont engagés dans des calculs classiques sur ce trinôme.

La question 6 a mis à rude épreuve l'aptitude des étudiants au calcul algébrique: seulement 1/4 a obtenu des résultats intéressants.

La question 7, (d'ailleurs formulée autrement que ci-dessus,) a pu être abordée, même par des étudiants qui avaient mal calculé dans 6, et une dizaine d'entre eux ont repris utilement les considérations de 4 et de 5.



Ne s'imposait-il pas de soumettre à l'analyse des correspondances l'ensemble de ces résultats?

À cet effet, on a d'abord créé un tableau croisant 35 lignes et 18 colonnes:

{A1, A2, 1, 2, 3a, 3b, 3c, 4, 5a, 5b, 5c, 6a0, 6a1, 6b, 6c, 7a, 7b, 7c},

sur lesquelles il nous suffira de noter qu'est supprimée A3, réussie par un seul étudiant, et que 6a est dédoublée en {6a0 6a1}, calculs de $Cr(0)$ et $Cr(1)$. On a $k(i, j)=1$ si le sujet i a réussi l'item j ; et 0 sinon. Pour A1, n'est compté comme succès qu'un calcul correct de l'inertie, puisque la trace a été vue par tous.

Afin de tenir compte des échecs, on a créé une colonne '-', complément à 18 du total des points obtenus dans les autres colonnes.

L'analyse produit d'abord, comme on l'attendait, un facteur de niveau général (v.p. = 0,31; taux = 48%); auquel la variable '-' est corrélée avec $COR1(-)=938$; tous les items ayant $F1>0$, excepté A1, réussi seulement par 3 sujets dont aucun n'était parmi les plus brillants (points 'o' avec $F1>>0$).

L'axe 2, (v.p. = 0,075; taux = 12%,) est d'une interprétation intéressante: il oppose {6, A1}, $F2<0$, à {4, 5c, 7}, $F2>0$. Un facteur 2 négatif caractérise donc une bonne aptitude au calcul; tandis que ($F2>0$) correspond à des raisonnements généraux fondés sur des calculs très simples.

Référence bibliographique

K. Ben SALEM: «Le critère de la trace minima pour estimer les données manquantes dans un tableau de correspondance», [TRAC. MANQ.], in CAD, Vol.XVII, n°1, pp. 97-112; (1992).