

REVUE DE STATISTIQUE APPLIQUÉE

P. THIONET

Note technique sur l'article de M. Marion : colinéarité et interaction dans l'analyse de modèles linéaires

Revue de statistique appliquée, tome 20, n° 3 (1972), p. 67-70

http://www.numdam.org/item?id=RSA_1972__20_3_67_0

© Société française de statistique, 1972, 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/>

NOTE TECHNIQUE
 SUR L'ARTICLE DE M. MARION :
 COLINÉARITÉ ET INTÉRACTION
 DANS L'ANALYSE DE MODÈLES LINÉAIRES

L'article de M. Marion est très clair et se suffit à lui-même. Un lecteur habitué aux mathématiques s'étonnera toutefois de n'y pas trouver de preuve que l'analyse des données en composantes simples qu'on y décrit est bien possible et unique. On nous donne seulement un algorithme qui, par une suite d'itérations, doit fournir, en fin de calcul, les dites composantes ; aucune indication n'est donnée sur la convergence (ni la vitesse de convergence) de cette suite, même à propos de l'exemple numérique traité.

L'exposé théorique offre cette particularité de s'arrêter au moment où l'on se trouve en présence d'un système de $I + J + K$ équations du 1° degré à $I + J + K$ inconnues, qui est visiblement de rang $I + J + K - 3$, c'est-à-dire impossible ou indéterminé ; pourquoi ce système se trouverait-il doté d'une solution unique dès lors qu'on le résout par itération ?

Telles sont les questions que nous nous sommes posées et auxquelles on va répondre. Nous conservons les notations de M. Marion.

Il faut tout d'abord observer que, dans le modèle $y = m + \alpha + \beta + \gamma + \varepsilon$ on a l'habitude de spécifier que les composantes α, β, γ sont de moyenne nulle ce qui impose 3 conditions de plus et détermine les estimations a, b, c par les moindres carrés. Dans la théorie des plans d'expérience, les n_{ijk} sont des nombres égaux entre eux (d'expériences ou d'observations par case) ; dans la théorie ici décrite, les n_{ijk} sont distincts et il s'en suit que des moyennes pondérées remplacent les moyennes arithmétiques simples des a, b, c , soit :

$$\sum_i n_{i..} a_i = 0, \quad \sum_j n_{.j.} b_j = 0, \quad \sum_k n_{..k} c_k = 0 \quad (T)$$

Aussi bien, l'auteur s'écarte-t-il de ce schéma théorique ; dans l'application numérique, on cherche en vain à vérifier (T) ; ainsi pour les a_i :

$$\begin{aligned} & (- 101, 32 \times 331 \ 734) + (770, 07 \times 98 \ 927) + \\ & (- 137, 06 \times 112 \ 851) = 27 \ 102 \ 067, 95 \end{aligned}$$

ce qui n'est pas du tout nul. Et de même pour les b_j ou les c_k .

Pour reprendre tous les calculs, il nous manque les $n_{ijk}, n_{ij.}, n_{i.k}, n_{.j.}$ Pour refaire la théorie, les notations matricielles seront plus commodes. Nous nous limiterons au cas de deux facteurs.

Formulation matricielle :

Nous poserons :

$N = (n_{ij})$, $N' = (n_{ji})$, $u' = (11 \dots 1)$, $u' N = h'$ (marge horizontale de N)

$N u = v$ (marge verticale de N), $u' N u = n_{..}$ (total général de N),

$Y = (y_{ij})$, $z_1 = (\bar{y}_{i.} - \bar{y}_{..})$ et $z_2 = (\bar{y}_{.j} - \bar{y}_{..})$ sont 2 vecteurs-colonnes.

On désignera par \hat{c} la matrice diagonale d'éléments c_i (en diagonale) et 0.

On aura besoin de la matrice dont tous les éléments sont des 1, soit U; avec :

$$U = (u \ u \ u \ u \ \dots) = \begin{pmatrix} u' \\ \vdots \\ u' \end{pmatrix}$$

On a alors :

$$\begin{bmatrix} n_{1.} & n_{2.} & \dots \\ \dots & \dots & \dots \\ n_{1.} & n_{2.} & \dots \end{bmatrix} = U \hat{v} \quad , \quad \begin{bmatrix} n_{.1} & n_{.2} & \dots \\ \dots & \dots & \dots \\ n_{.1} & n_{.2} & \dots \end{bmatrix} = U \hat{h}$$

Le problème proposé s'énonce : Quels sont les vecteurs inconnus a, b,

$$a = (a_1 \ a_2 \ \dots) \quad , \quad b = (b_1 \ b_2 \ \dots)$$

définis par le système d'équations (des moindres carrés) obtenu par M Marion :

$$(1) \begin{cases} z_1 = a - (U \hat{v} a + U \hat{h} b) / n_{..} + \hat{v}^{-1} N b \\ z_2 = b - (U \hat{v} a + U \hat{h} b) / n_{..} + \hat{h}^{-1} N' a \end{cases}$$

Nous y adjoignons le système (T) défini plus haut et qui s'écrit matriciellement :

$$(T) \quad v' a = 0, \quad h' b = 0$$

impliquant $U \hat{v} a = 0, \quad U \hat{h} b = 0.$

Nous poserons en outre :

$$\hat{v}^{-1} N = V, \quad \hat{h}^{-1} N' = H,$$

de sorte que (1) s'écrive :

$$(2) \begin{cases} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} I & V \\ H & \hat{I} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{ou} \quad \begin{cases} z_1 = a + V b \\ z_2 = H a + b \end{cases} \end{cases}$$

Proposition 1 :

Ce système peut à son tour se décomposer en : une équation en a et une en b :

$$(3) \left\{ \begin{array}{l} (I - V H) a = z_1 - V z_2 \\ (4) \left\{ \begin{array}{l} (I - H V) b = z_2 - H z_1 \end{array} \right. \end{array} \right.$$

Preuve :

Combinaison linéaire entre $(a + b)$ et $(H a + b)$.

Proposition 2 :

Les matrices $I - V H$ et $I - H V$ ne sont pas inversibles

En effet, V est une matrice stochastique (sommées de lignes égales à 1), H aussi, de même pour $V H$, $H V$, $(V H)^p$, $(H V)^p$ quel que soit l'entier positif p .

Ainsi :

$$I + (V H) + \dots + (V H)^p \text{ et } I + (H V) + \dots + (H V)^p$$

ne peuvent tendre vers des limites (resp. $(I - V H)^{-1}$ et $(I - H V)^{-1}$) puisque leurs sommes de lignes tendent vers l'infini avec p .

Proposition 3 :

$$(I - V H)^{-1} (z_1 - V z_2)$$

représente symboliquement l'algorithme de Marion de a , et

$$(I - H V)^{-1} (z_2 - H z_1) \text{ celui de } b.$$

Effectivement, le calcul est celui de :

$$\begin{aligned} (a)_0 &= z_1, \quad (b)_0 = z_2 - H z_1, \\ (a)_1 &= (I + V H) z_1 - V z_2, \quad (b)_1 = (I + H V) z_2 - H (I + V H) z_1, \\ (a)_2 &= (I + V H + V H + V H) z_1 - V (I + H V) z_2 \end{aligned}$$

etc...

Proposition 4 :

$V H$ a pour vecteurs propres associés à la valeur propre unité u (à droite) et v' à gauche ; de même pour $H V$: u (à droite) et h' à gauche. De même pour $(V H)^p$, $(H V)^p$.

En effet :

$$v' (V H) = u' \hat{v} \hat{v}^{-1} N \hat{h}^{-1} N' = u' N \hat{h}^{-1} N' = h' \hat{h}^{-1} N' = u' N' = v'$$

De même $h' (H V) = \dots = h'$. Une fois normés, ces vecteurs s'écrivent : $\frac{v'}{n..}$, $\frac{h'}{n..}$.

Conséquence :

Si $(V H)^p$ admet une limite pour p infini il est bien connu que c'est : $U \hat{v}/n..$; de même pour $(H V)^p$, c'est $U \hat{h}/n..$.

Proposition 5 :

On a :

$$v' a = v' z_1 = v' V b = 0 ; h' b = h' z_2 = h' H a = 0$$

Vérification immédiate.

Proposition 6 :

Si la matrice (stochastique) VH est indécomposable et que (par conséquent) l'unité en soit valeur propre simple, l'algorithme de a converge. De même pour HV et b. La convergence se fait comme ρ^p (module de la plus grande valeur propre après la valeur propre unité).

Considérons la décomposition de Jordan de VH et de $(VH)^p$

$$A + \sum_s B_s \rho_s^p$$

où la matrice A est celle de la Proposition 3 ci-dessus.

Posons $Z = z_1 - V z_2$. Il résulte des propositions 4 et 5 que $A Z = 0$.

Les séries $B_s Z \rho_s^p$ sont convergentes ; leur somme a la vitesse de convergence du terme renfermant le plus grand des ρ_s .

Raisonnement analogue pour HV et $(HV)^p$.

Nota :

1/ On a pu noter quelques réserves concernant le cas où l'unité serait racine multiple de l'équation caractéristique de VH ou HV. Dans la décomposition de Jordan, A n'a plus la forme uv' . Il est vraisemblable que cette restriction n'a aucune portée pratique. Plus inquiétant serait le cas de valeurs propres imaginaires de module égal à 1 ; $(VH)^p$ ne converge pas. Il s'agit de cas très spéciaux. En pratique, il s'agit de dispositions très particulières des termes nuls de VH (ou HV) c'est-à-dire de $N N'$ (ou $N' N$), donc de N (qui n'a pas de termes négatifs).

2/ Il ne semble pas que le passage de 2 à 3 dimensions (s'il est gênant pour les notations matricielles) puisse perturber beaucoup le résultat. Les équations des moindres carrés se mettant sous la forme : $Z = T a$, où T^{-1} n'existe pas, il conviendra de s'assurer que $T^{-1} Z$ existe néanmoins, en général.

On doit toujours ajouter les contraintes du type $v'a = 0$ au système des M.C. Il ne saurait être question de dire que l'existence d'un programme pour ordinateur ait résolu les difficultés théoriques soulevées par M. Marion, n'étant pas habitués à des matrices à 3 dimensions ! On n'a non plus pas abordé ici le problème des interactions évoqué à la fin de l'article.

3/ Signalons que la condition d'orthogonalité paraît être déjà connue. Citons par exemple :

Seber (G. A. F.) Orthogonality in analysis of variance, Annals of Math. Statistics, 35 (June 1964), p. 705-710.

P. Thionet

29 avril 1972