REVUE DE STATISTIQUE APPLIQUÉE

A. HOLLY

Sur l'estimation des paramètres d'un modèle linéaire lorsque la matrice des variances-covariances des résidus est singulière

Revue de statistique appliquée, tome 22, n° 2 (1974), p. 5-27
<http://www.numdam.org/item?id=RSA_1974__22_2_5_0>

SUR L'ESTIMATION DES PARAMÈTRES D'UN MODÈLE LINÉAIRE LORSQUE LA MATRICE DES VARIANCES-COVARIANCES DES RÉSIDUS EST SINGULIÈRE [2]

A. HOLLY [1]
Université Paris IX – Dauphine
U.E.R. de Mathématique que de la décision

INTRODUCTION

Soit le modèle linéaire

$$
\text{(I)} \qquad \begin{array}{l} \underset{(T,1)}{y} = \underset{(T,K)}{X}\; \underset{(T,1)}{b} + \underset{(T,1)}{u} \quad \text{avec} \\ \\ r(X) = K \end{array} \quad \left\{ \begin{array}{l} E(u) = 0 \\ Var(u) = \Omega \\ \Omega \text{ connue, définie non négative} \\ \text{de rang} \\ r(\Omega) = h < T \end{array} \right.
$$

On sait, lorsque Ω est régulière et connue, que l'estimateur des moindres carrés généralisé, ou encore "l'estimateur d'Aitken" de $\tilde{b}$ est

$$\tilde{b} = (X'\Omega^{-1} X)^{-1} X'\Omega^{-1} y \tag{1}$$

L'estimateur $\tilde{b}$ est linéaire par rapport aux observations y_t $(t = 1, \ldots, T)$ de la variable endogène. De plus, $\tilde{b}$ est un estimateur sans biais de b. Une des propriétés les plus importantes de b est sans doute que cet estimateur vérifie le théorème de Gauss-Markov, c'est-à-dire que si l'on désigne par $\ddot{b}$ un autre estimateur linéaire et sans biais de b, de la forme

$$\ddot{b} = G\, y \tag{2}$$

---

(1) Je tiens à remercier tout particulièrement P. Mazodier (de l'I.N.S.E.E.), instigateur de ce travail, pour ses nombreuses suggestions. J'ai aussi tiré grand profit de discussions avec MM. C. Fourgeaud, B. Lenclud et D. Piet. Je remercie également, Mlle J. Ulmo et Mr. P. Thionet de leurs critiques formulées sur une première version de ce travail.

(2) Article remis le 1/3/73, révisé le 22/3/73

on montre que la matrice $G - (X'\Omega^{-1}X)^{-1}$ est définie non négative. En particulier, d'après ce théorème on a,

$$\mathrm{Var}(\ddot{b}_k) - \mathrm{Var}(\tilde{b}_k) \geqslant 0 \tag{3}$$

On dit en anglais que le vecteur $\tilde{b}$ est le "best linear unbiased estimator" ce qui s'abrège en b.l.u.e.. On dit aussi dans la littérature économétrique que $\tilde{b}$ est l'estimateur de "Gauss-Markov" de b. Le b.l.u.e. de b est par ailleurs unique.

Cependant, dans certains problèmes économétriques il peut exister des relations linéaires entre les résidus u_t, qui rendent la matrice Ω singulière. Empruntons un exemple à H. Theil ([10] pp. 274-275). Supposons que l'on étudie dans une "coupe instantannée" la consommation des ménages. On peut distinguer dans l'ensemble des biens consommés, les biens alimentaires, les biens durables et les autres biens. Supposons que l'on a formulé un modèle où on a lié chaque catégorie de biens consommés à la consommation totale des ménages, indicés par n (n = 1, . . ., N). On a par conséquent les relations suivantes,

$$(\text{biens alimentaires})_n = a_0 + a_1 \,(\text{consommation totale})_n + v_n \tag{4}$$

$$(\text{biens durables})_n = a'_0 + a'_1 \,(\text{consommation totale})_n + v'_n \tag{5}$$

$$(\text{autres biens})_n = a''_0 + a''_1 \,(\text{consommation totale})_n + v''_n \tag{6}$$

Le vecteur u des résidus de ce modèle a 3N composantes et la matrice des variances-covariances de u est de format (3N, 3N). Il est cependant clair que pour tout n les résidus aléatoires v_n, v'_n et v''_n, ne sont pas linéairement indépendants. En effet, desrrelations (4), (5) et (6) on déduit que

$$[1 - (a_1 + a'_1 + a''_1)]\,(\text{consom. tot.})_n = (a_0 + a'_0 + a''_0) + v_n + v'_n + v''_n \tag{7}$$

Si on suppose que la variable exogène de ce modèle est la consommation totale de chaque ménage, celle-ci ne peut être considérée comme une variable aléatoire. Par conséquent la somme des trois résidus, $v_n + v'_n + v''_n$, est une quantité certaine. Si on suppose que

$$E(v_n) = E(v'_n) = E(v''_n) = 0 \tag{8}$$

il est clair que

$$E(v_n + v'_n + v''_n) = 0 \tag{9}$$

ce qui montre que

$$v_n + v'_n + v''_n = 0 \tag{10}$$

La matrice des variances-covariances du résidu u de ce modèle a donc un rang au plus égal à 2N, d'après la relation (10). On a donc un exemple de matrice de variances-covariances des résidus singulière.

Lorsque la matrice Ω est singulière et connue, un certain nombre d'auteurs ont proposé des méthodes d'estimation du vecteur b. Mais on peut dire qu'il existe dans la littérature sur ce sujet trois principales approches qui sont les suivantes.

Une première approche est proposée par Rao et Mitra ([9], pp. 147-150 (*). Elle consiste à transformer le modèle (1) en un modèle avec contrainte sur les coefficients et avec une matrice de variances-covariances régulière. Cette transformation s'effectue de la manière suivante. Puisque Ω est singulière, il existe au moins une matrice R de rang maximum telle que

$$\left\{ \begin{array}{c} \Omega\, R = 0 \\ R \text{ de format } (T,\ T - h), \text{ et de rang } T - h \end{array} \right. \tag{11}$$

Par ailleurs, puisque Ω est d.n.n. il existe une matrice J de format (T, h) telle que

$$\Omega = J\, J' \tag{12}$$

avec

$$r(\Omega) = r(J) \tag{13}$$

Soit P une inverse à gauche de J, de format (h, T), c'est-à-dire telle que

$$P\, J = I_h \tag{14}$$

Rao et Mitra imposent la condition supplémentaire

$$P\, R = 0$$

On verra à la section 2 que l'on peut se dispenser de cette condition et qu'il existe une seule matrice P qui la vérifie.

La transformation du modèle (I) conduit au modèle (II)

$$\text{(II)} \quad \left\{ \begin{array}{l} Py = P\, X\, b + Pu \\ \\ R'y = R'\, X\, b \end{array} \right. \quad \text{avec} \quad \left\{ \begin{array}{l} E(P\, u) = 0 \\ \\ \mathrm{Var}(P\, u) = I_h \end{array} \right.$$

On aboutit de cette manière à un modèle avec contrainte sur les coefficients, vérifiée presque sûrement puisque $E(R'u) = 0$ et $\mathrm{Var}(R'u) = 0$ avec $u = y - X\, b$. C'est-à-dire que l'on est conduit à résoudre le problème suivant

$$\text{(P 1)} \quad \left\{ \begin{array}{l} \min\limits_{b} (y - Xb)'P'P(y - Xb) \\ \text{sous contrainte} \\ R'X\, b = R'y \end{array} \right.$$

---

(1) Récemment Rao ([7], [8]) a proposé une théorie unifiée de l'estimation de b, comprenant le cas où la matrice X est de rang inférieur à K. Comme dans le présent travail on n'étudie pas ce cas, seule l'approche proposée dans l'ouvrage de Rao et Mitra est exposée.

ou encore

$$
(\mathrm{P'1}) \begin{cases} \min\limits_{b} (y - Xb)'P'P(y - Xb) \\ \text{sous contrainte} \\ R'u = 0 \\ \quad u = y - Xb \end{cases}
$$

où la matrice P'P est une matrice singulière, de format (T, T) et de rang h.

Malinvaud ([3], chap. 5) montre, par une approche géométrique, que l'estimateur de Gauss-Markov, ou encore le b.l.u.e. de b, est obtenu en résolvant le problème suivant (voir prop. 4 p. 179).

$$
(\mathrm{P\ 2}) \begin{cases} \min\limits_{b} (y - Xb)' (\Omega + RR')^{-1} (y - Xb) \\ \text{sous contrainte} \\ R'u = 0 \\ \quad u = y - Xb \end{cases}
$$

pour toute matrice R vérifiant (11).

Theil ([10], pp. 273-293) utilise une approche faisant intervenir la pseudo-inverse de Ω, notée Ω^+ (unique), étudiée par Moore [4] et Penrose [5]. On montre ainsi que le b.l.u.e. de b est obtenu en résolvant le problème suivant

$$
(\mathrm{P\ 3}) \begin{cases} \min\limits_{b} (y - Xb)' \Omega^+ (y - Xb) \\ \text{sous contrainte} \\ R'_o u = 0 \\ \quad u = y - Xb \end{cases}
$$

L'approche de Theil constitue un cas particulier de celle de Rao et Mitra. La matrice R_o est une matrice R particulière formée de vecteurs propres orthogonaux et unitaires, associés à la valeur propre nulle de Ω. D'autre part, pour une matrice P particulière, notée ici H (et présentée plus loin) on a $H'H = \Omega^+$.

Dans un travail récent, Fourgeaud et Lenclud [1] aboutissent, par une approche générale de l'estimation dans les modèles linéaires fondée sur l'inverse de Moore-Penrose à la formulation de Theil.

L'objet de ce travail est de montrer que ces trois approches fournissent le même estimateur du vecteur des paramètres b.

Le plan adopté est le suivant. Dans la section 1 on rappelle certains résultats relatifs aux g-inverses de Rao (voir par exemple [6]) et la définition de la pseudo-inverse de Moore-Penrose. Dans la section 2 on montre, en ne supposant pas que P R = 0, que l'on peut estimer le vecteur b en résolvant le problème suivant :

$$
(P\ 4)\ \begin{cases} \min\limits_{b}(y - Xb)'\,(P'P + RR')\,(y - Xb) \\ \text{sous-contrainte} \\ R'u = 0 \\ \quad u = y - Xb \end{cases}
$$

On montre aussi que ce problème, dans le cas particulier où $P = H$ et où la matrice R est égale à R_0, se ramène au problème de Malinvaud.

On vérifie aussi, facilement, que les quatre problèmes précédents sont un cas particulier du problème suivant, où Ω^- désigne une g-inverse quelconque de Ω :

$$
(P\ 5)\quad \begin{cases} \min\limits_{b}(y - Xb)'\,\Omega^-(y - Xb) \\ \text{sous contrainte} \\ R'u = 0 \\ \quad u = y - Xb \end{cases}
$$

Dans la section 3, on démontre le résultat essentiel de ce travail, à savoir que la solution de (P 5) est indépendante du choix de Ω^-. Le résultat montre l'unicité des estimateurs proposés par Rao et Mitra, Malinvaud et Theil. Il en résulte en particulier que l'estimateur de b obtenu en résolvant (P 5) est le b.l.u.e. de b.

Finalement, dans la section 4, on donne une justification du problème (P 5) par une approche géométrique identique à celle de Malinvaud.

Une dernière remarque doit être faite. Comme il a été précisé plus haut, on suppose ici que la matrice Ω est connue. Bien entendu, il s'agit d'un cas exceptionnel. Le cas où la matrice Ω est inconnue est très important à étudier. Les quelques indications que l'on peut donner sur ce cas nous conduiraient trop loin. C'est la raison pour laquelle on se propose d'examiner différents aspects de ce problème, d'une manière approfondie, dans un travail ultérieur.

1. QUELQUES RESULTATS RELATIFS AUX G-INVERSES DE RAO ET DEFINITION DE Ω^+

On énonce dans ce paragraphe quelques résultats relatifs aux g-inverses définis par C.R. Rao (voir par exemple [6]). On ne donne pas de démonstration des résultats énoncés, mais on indique en référence les ouvrages où ces démonstrations peuvent se trouver.

On a la définition suivante d'une g-inverse (de Rao) :

Définition 1

Soit A une matrice de format (m, n) et de rang quelconque. Une g-inverse de A est une matrice de format (n, m), notée A^-, vérifiant la propriété suivante : pour chaque y, tel que $Ax = y$ est possible, $x = A^- y$ est une des solutions.

On a le lemme suivant :

Lemme 1

Une condition nécessaire et suffisante pour que A^- soit une g-inverse de A est que :

$$A A^- A = A \tag{16}$$

Il est possible de donner l'expression générale de toutes les g-inverses de A en connaissant une g-inverse particulière A^-. Le lemme suivant, dont on peut trouver la démonstration dans Lentner [2] et dans Rao et Mitra ([9] ; pp. 26-27) sera d'une grande utilité.

Lemme 2

Soit A^- une g-inverse quelconque de A. On peut caractériser toutes les g-inverses de A par la relation :

$$\text{g-inverse de } A = A^- + U(I - A A^-) + (I - A^- A) W$$

où U et W sont des matrices quelconques ayant des formats convenables.

Parmi les g-inverses de A, il existe une g-inverse particulière qui est l'unique à posséder certaines propriétés. C'est l'inverse généralisée de Moore-Penrose.

Définition 2

On note A^+ la g-inverse de A qui vérifie les quatre propriétés suivantes :

$$\begin{aligned} A A^+ A &= A \\ A^+ A A^+ &= A^+ \\ (A A^+)^+ &= A A^+ \\ (A^+ A)' &= A^+ A \end{aligned}$$

On montre que la g-inverse A^+ qui vérifie ces propriétés est unique. Pour la démonstration, voir par exemple Penrose [5], Rao-Mitra ([9] p. 51) et Theil ([10] pp. 269-270).

On aura besoin par la suite du résultat suivant :

Lemme 3

Une condition nécessaire et suffisante pour que l'équation matricielle en X.

$$A X B = C$$

admette une solution est que

$$A A^- C B^- B = C$$

où A^- et B^- sont des g-inverses quelconques de A et B. La solution générale est donnée par :

$$X = A^- C B^- + Z - A^- A Z B B^-$$

où Z est une matrice arbitraire de format convenable.

Pour la démonstration de ce lemme voir Rao et Mitra ([9] p. 24)

2. PASSAGE DU MODELE (I) A DIFFERENTS MODELES AVEC CONTRAINTES SUR LES COEFFICIENTS

On montre dans cette section comment on aboutit à l'étude des problémes (P 4) et (P 5).

Soit R une matrice quelconque vérifiant (11). On peut remarquer que l'application du lemme 3 donne la forme générale de R. On a le lemme suivant :

Lemme 4

La solution de l'équation

$$\Omega \; R = 0$$

est telle que

$$R' = Z (I_T - \Omega \; \Omega^-)$$

où Z est une matrice quelconque de format $(T - h, T)$ et Ω^- une g-inverse quelconque de Ω.

Démonstration

En effet, l'application du lemme 3 à l'équation transposée de (11)

$$R' \; \Omega = 0$$

montre que la solution en R' existe et est donnée par :

$$R' = Z (I_T - \Omega \; \Omega^-) \tag{17}$$

■

La proposition suivante servira à démontrer que les résultats obtenus par la suite sont indépendants du choix de R.

Proposition 1

Pour toute matrice R vérifiant $\Omega \; R = 0$, *on a presque sûrement :*

$$R'u = 0 \tag{18}$$

Démonstration

En effet,

$$E(R'u) = 0$$

et

$$\mathrm{Var} \; (R'u) = R' \; \Omega \; R$$

$$\mathrm{Var} \; (R'u) = 0$$

■

Remarque

Cette proposition explique la formulation de la proposition 4 de Malinvaud ([3] p. 179) sous la forme (P 2). En effet, Ω R = 0 entraîne que R'u = 0. Ceci revient à dire que, presque sûrement, y appartient à la variété linéaire L définie par R' X b = R'y.

Parmi les matrices R qui vérifient (11), on peut choisir une matrice R_o formée de T – h vecteurs propres orthogonaux et normés associés à la valeur propre de Ω. Une telle matrice R_o jouera un rôle important dans différentes démonstrations. Précisons que Ω étant définie non négative, de rang $h < T$, on peut la diagonaliser et écrire :

$$\Omega = [Q : R_o] \begin{bmatrix} \Lambda_h & : 0 \\ \cdot\ \cdot & \cdot\ \cdot \\ 0 & : 0 \end{bmatrix} \begin{bmatrix} Q' \\ \cdot\ \cdot \\ R'_o \end{bmatrix} = Q \Lambda_h Q' \tag{19}$$

où

$$\Lambda_h = \begin{bmatrix} \lambda_1 & & & & \\ & \diagdown & & & 0 \\ & & \lambda_j & & \\ & & & \diagdown & \\ 0 & & & & \lambda_h \end{bmatrix} (\lambda > 0\ j = 1, -, h) \tag{20}$$

et $[Q : R_o]$ orthogonale, c'est-à-dire que les conditions suivantes sont vérifiées :

$$\left.\begin{array}{l} Q'R_o = 0 \\ R'_o R_o = I_{T-h} \\ Q'Q = I_h \\ Q\,Q' + R_o R'_o = I_T \end{array}\right\} \begin{array}{r} (21) \\ (22) \\ (23) \\ (24) \end{array}$$

La relation (21) entraîne que :

$$\Omega\, R_o = Q\, \Lambda_h\, Q'R_o = 0$$

Comme R_o est de rang T – h, elle vérifie (11).

On pose :

$$J = Q\, \Lambda_h^{1/2} \tag{25}$$

La matrice J est de format (T, h).

On a :

$$J\, J' = Q\, \Lambda_h\, Q' = \Omega \tag{26}$$

Il est clair que l'on a :

$$r(\Omega) = r(J\,J') = r(J) = h$$

Soit P une inverse à gauche quelconque de J, de format (h, T). Elle vérifie :

$$P\,J = I_h \tag{28}$$

Il est clair que, quel que soit P, on a :

$$r(P) = h \tag{29}$$

Il est utile de remarquer qu'une inverse à gauche particulière de J est la matrice H définie par :

$$H = \Lambda_h^{-1/2}\,Q' \tag{30}$$

On a alors les résultats suivants :

Proposition 2

1) H'H est l'inverse de Moore-Penrose de Ω notée Ω^+

2) $$(H'H + R_o R_o) = (\Omega + R_o R_o')^{-1}$$

Démonstration

En effet, la première partie de cette proposition résulte du fait que :

$$H'H = Q\,\Lambda_h^{-1}\,Q, \tag{31}$$

Les propriétés suivantes sont alors vérifiées :

a) $$\Omega\,(H'H)\,\Omega = \Omega$$

b) $$(H'H)\,\Omega\,(H'H) = H'H$$

car

$$(Q\,\Lambda^{-1}\,Q')\,Q\,\Lambda_h\,Q'\,(Q\,\Lambda_h^{-1}\,Q') = Q\,\Lambda_h{}^{-1}\,q'$$

en tenant compte de (23).

c) $$[(H'H)\Omega]' = (H'H)\Omega$$

En effet

$$H'H\ \Omega = (Q\ \Lambda^{-1} Q')\ (Q\ \Lambda_h Q') = Q\ Q' = I_T - R_o\ R'_o \tag{32}$$

qui est bien symétrique

d) $$[\Omega\,(H'H)]' = \Omega\,(H'H)$$

En effet, on montre aussi que :

$$\Omega\,(H'H) = Q\ Q' = I_T - R_o\ R'_o \tag{33}$$

Pour démontrer la deuxième partie de la proposition, on peut écrire que :

$$H'H + R_o R'_o = \Omega^+ + R_o R'_o \tag{34}$$

mais on a :

$$(\Omega^+ + R_o R'_o)\ (\Omega + R_o R'_o) = \Omega^+ \Omega + \Omega^+ R_o R'_o + R_o R'_o R_o R'_o \tag{35}$$

du fait que

$$R'_o\ \Omega = 0$$

mais $\quad \Omega^+ \Omega = Q\ Q'$ d'après (c)

et $\quad \Omega^+ R_o R'_o = Q\ \Lambda_h^{-1}\ Q' R_o R'_o = 0$ d'après (21)

D'autre part,

$$R'_o R_o = I_{T-h} \text{ d'après (22)}$$

d'où

$$(\Omega^+ + R_o R'_o)\ (\Omega + R_o R'_o) = Q\ Q' + R_o R'_o = I_T \tag{36}$$

d'après (24) ; $H'H + R_o R'_o$ est donc une inverse à gauche de $\Omega + R_o R'_o$.

En prenant le transposé de (36) et en remarquant que toutes les matrices sont symétriques, on voit que $H'H + R_o R'_o$ est une inverse à droite de $\Omega + R_o R'_o$. On a donc bien :

$$(\Omega + R_o\ R'_o)^{-1} = H'H + R_o\ R'_o \quad \blacksquare \tag{37}$$

Le lemme suivant sera très utile par la suite.

Lemme 5

Pour toute matrice R vérifiant (11), il existe une matrice S régulière telle que :

$$R' = S\ R'_o \tag{38}$$

Démonstration

La démonstration que l'on donne de ce lemme n'est sans doute pas la plus rapide, mais elle fournit l'expression explicite de S.

D'après le lemme 4, on peut trouver une matrice Z de format (T – h, T) telle que :

$$R' = Z(I - \Omega\Omega^{+}) \tag{39}$$

puisque Ω^{+} est une Ω^{-} particulière. C'est-à-dire, d'après (33), pour toute matrice R, vérifiant $\Omega R = O$ et de rang T – h, on peut trouver une matrice Z de format (T – h, T) telle que :

$$R' = Z\, R_o R_o' \tag{40}$$

La matrice $Z\, R_o$ est une matrice carrée de format (T – h, T – h). Elle vérifie :

$$R'R = Z\, R_o\, R_o'\, R_o\, (Z\, R_o)' \tag{41}$$

$$R'R = (Z\, R_o)\,(Z\, R_o)' \text{ (d'après 21))} \tag{42}$$

or

$$r(R) = r(R'R) = T - h$$

d'où

$$r[(ZR_o)\,(ZR_o)'] = r(Z\, R_o) = T - h$$

Ceci montre que $Z\, R_o$ est régulière ∎

On posera par la suite :

$$S = Z\, R_o \tag{43}$$

Posons

$$\underset{(T,T)}{M} = \begin{bmatrix} P \\ .. \\ R' \end{bmatrix} \tag{44}$$

et

$$M_o = \begin{bmatrix} P \\ .. \\ R_o' \end{bmatrix}$$

Il n'est pas utile d'émettre l'hypothèse (15), PR = O comme le font Rao et Mitra ([9], p. 149). On montre en fait que, sous cette hypothèse, P est unique et égale à H.

Proposition 3

La seule inverse à gauche P de J, qui vérifie PR = 0 est la matrice H.

Démonstration

En effet, d'après la définition de J et de P, on a :

$$P\,Q\,\Lambda_h^{1/2} = I_h \tag{46}$$

ou encore
$$P\,Q\,Q' = \Lambda_h^{-1/2}\,Q' = H \tag{47}$$

et d'après (24)
$$P(I_T - R_o R_o') = H \tag{48}$$

ou encore
$$P = P\,R_o\,R'_o + H \tag{49}$$

Or, P R = 0 entraîne :

$$P\,R_o S' = 0 \tag{50}$$

Comme S est régulière, il en résulte que :

$$P\,R_o = 0 \tag{51}$$

et, d'après (49), on voit que P = H ■

On peut montrer que, quels que soient P et R, la matrice M définie par (44) est régulière. On a besoin pour cela du lemme suivant, que l'on énonce sans démonstration.

Lemme 6

Soit D une matrice symétrique décomposée en blocs de la manière suivante :

$$D = \begin{bmatrix} A & : & C \\ .. & & .. \\ C' & : & B \end{bmatrix}$$

où B est régulière.

Une condition nécessaire et suffisante pour que D soit régulière est que la matrice $(A - CB^{-1}C')$ soit régulière.

On a alors le lemme suivant.

Lemme 7

La matrice M_o est régulière.

Démonstration

En effet, on a :

$$M_o M_o' = \begin{bmatrix} PP' & : & PR_o \\ .. & & .. \\ R_o' P_o' & : & R_o' R_o \end{bmatrix} \tag{52}$$

et, en tenant compte de (22)

$$M_o M_o' = \begin{bmatrix} PP' & : & PR_o \\ .. & & .. \\ R'P' & : & I_{T-h} \end{bmatrix} \tag{53}$$

Considérons la matrice

$$PP' - PR_o R_o' P' \tag{54}$$

On a, d'après (24)

$$PP' - P R_o R_o' P' = P(I_T - R_o R_o')P' = P Q Q'P' \tag{55}$$

et d'après la définition de P, on a

$$P Q = Q'P' = \Lambda_h^{-1/2} \tag{56}$$

ce qui montre que

$$P P' - (P R_o)(R_o' p') = \Lambda_h^{-1} \tag{57}$$

et que cette matrice est inversible. En appliquant le lemme 6, on voit que $M_o M_o'$ est régulière. Mais :

$$r(M_o) = r(M_o M_o') = T \tag{58}$$

ce qui montre que la matrice M_o est régulière. ■

Il en résulte alors le lemme suivant.

Lemme 8

La matrice M est régulière.

Il suffit de remarquer en effet que

$$M = \begin{bmatrix} I_h & : & 0 \\ .. & & .. \\ 0 & & S \end{bmatrix} M_o \qquad (59)$$

et que M est le produit de M_o par une matrice régulière. ∎

On est alors en mesure de transformer le modèle initial (I) en un modèle strictement équivalent en écrivant que :

$$y = Xb + u$$

ce qui équivaut à :

$$My = M\,Xb + Mu \qquad (60)$$

ce qui conduit au modèle III suivant :

$$\text{III} \begin{cases} Py = Xb + Pu \\ R'y = R'Xb \text{ (p.s.)} \end{cases} \quad \begin{cases} E\ (Pu) = 0 \\ \text{Var}\ (Pu) = I_h \end{cases}$$

La différence entre le modèle (III) et le modèle (II) est que l'on n'a pas forcément PR = 0.

On est donc ramené à l'étude d'un problème des moindres carrés ordinaires sous contrainte linéaire sur. les coefficients.

$$(P\ 6) \begin{cases} \min\limits_{b}\ (y - Xb)'\ P'P\ (y - Xb) \\ \text{sous contrainte} \\ R'Xb = R'y \end{cases}$$

où P'P n'est pas régulière.

On a le résultat suivant :

Lemme 9

P'P est une g-inverse de Ω

Démonstration

En effet, on a :

$$\begin{aligned} \Omega\ P'P\ \Omega &= (J\ J')\ (P'P)\ J\ J' \\ &= J\,(J'P')\ (P\ J)\ J' \\ &= J\,(I)\ (I)\ J' \\ &= J\ J' \\ &= \Omega \end{aligned}$$

∎

Il résulte de ce lemme que le problème (P 6) est un cas particulier du problème (P 5).

D'autre part, lorsque $P = H$ et $R = R_o$, le problème (P 6) est identique au problème (P 3) de Theil.

La résolution du problème (P 6) conduit aux conditions du premier ordre, où $\hat{\mu}$ est un vecteur de multiplicateurs de Lagrange.

$$\left.\begin{aligned} X'P'P\, X\, \hat{b} + X'R\, \hat{\mu} &= X'P'Py \\ R'X\, \hat{b} &= R'y \end{aligned}\right\} \qquad (61)$$

qui deviennent, après prémultiplication de la 2e équation par X'R et en l'additionnant à la première :

$$\left.\begin{aligned} X'(P'P + R\, R')\, X\, \hat{b} + X'R\, \hat{\mu} = X'(P'P + R\, R')y \\ R'X\, \hat{b} = R'y \end{aligned}\right\} \qquad (62)$$

Il est clair que les solutions du système (61) sont solution du système (62) qui, à leur tour, sont solution du problème suivant :

$$(P_4) \quad \left\{\begin{array}{l} \min\limits_{b} \ (y - Xb)\, (P'P + R\, R')\, (y - Xb) \\ \text{sous contrainte} \\ R'Xb = R'y \end{array}\right.$$

On a obtenu ainsi une nouvelle méthode d'estimation de b. Il est possible de montrer directement que le b.l.u.e. de b est la solution de (P 4). On fournit plus loin la raison pour laquelle cette démonstration n'est pas donnée.

La deuxième partie du lemme 10 nous montre que (P 4) est un cas particulier de (P 5).

D'autre part, les propriétés de la matrice (P'P + R R') sont intéressantes à étudier. On a les résultats suivants :

Lemme 10

1) La matrice (P'P + R R') est une matrice régulière définie positive.

2) (P'P + R R') est une g-inverse de Ω.

Démonstration

La première partie de ce lemme résulte du lemme 8. En effet,

$$r(M) = r(M'M) = T$$

(M'M) est donc une matrice symétrique, inversible.

D'autre part, puisque M est régulière, si x est un vecteur non nul, on a :

$$Mx = 0 \text{ et } x'M'Mx > 0$$

ce qui montre que M'M est définie positive.

Pour démontrer la deuxième partie de ce lemme il suffit de remarquer que $\Omega R = 0$ entraîne que :

$$\Omega (P'P + R R') \Omega = \Omega P'P \Omega \tag{63}$$

qui est égal à Ω d'après le lemme 9. ∎

Il résulte de ce lemme qu'il existe un matrice Σ définie positive telle que

$$(P'P + R R') = \Sigma^{-1} \tag{64}$$

Dans le cas où $P = H$ et $R = R_o$, la deuxième partie de la proposition 2 montre que :

$$\Sigma = \Omega + R_o R'_o$$

Le problème (P 2) de Malinvaud est, sous sa forme générale, un cas particulier du problème (P 5). Il suffit de remarquer que l'on a le lemme suivant.

Lemme 11

Soit Ω une matrice de format (T, T) d.n.n. et de rang h et R une matrice de format $(T, T - h)$ de rang $(T - h)$ telle que $\Omega R = 0$. Alors :

1) La matrice $\Omega + RR'$ est régulière, définie positive.

2) $(\Omega + R R')^{-1}$ est une g-inverse de Ω.

On trouve une démonstration de ce résultat dans Malinvaud ([3] p. 168).

Ainsi les différents problèmes de Rao-Mitra, Malinvaud. Theil et (P 4) sont des cas particuliers du problème (P 5). Il importe de démontrer que leur résolution conduit au même estimateur de b. Ce sera l'objet du paragraphe suivant.

Remarque importante.

Revenons un instant sur le sens de la proposition 3, avant de passer au paragraphe suivant.

Ce que l'on montre en fait, grâce à cette proposition, c'est que la condition supplémentaire $PR = 0$ qu'imposent Rao et Mitra rend leur problème identique à celui de Theil.

En ce sens, la proposition 3 constitue une sorte de parenthèse.

Il est clair, cependant, que la matrice J définie par (25) n'est pas la seule qui vérifie (26) et (27). Cependant les résultats obtenus dans ce paragraphe dépendent de la seule matrice J.

En effet, l'ensemble des matrices F qui vérifient

$$\Omega = FF' \tag{65}$$

et

$$r(\Omega) = r(FF') = h \tag{66}$$

est obtenu en postmultipliant J par une matrice B orthogonale de format (h,h), c'est-à-dire que l'on a

$$F = JB \qquad (67)$$

Les inverses à gauche de F, notées G, se déduisent des inverses à gauche de J par la relation

$$G = B'P \qquad (68)$$

En effet, les matrices G qui vérifient

$$GF = I_h \qquad (69)$$

satisfont à

$$GJB = I_h \qquad (70)$$

c'est-à-dire, en tenant compte du fait que B est orthogonale, à

$$BGJ = I_h \qquad (71)$$

Autrement dit,

$$BG = P \qquad (72)$$

ce qui montre la relation (68).

Il est alors clair, d'après l'orthogonalité de B que la relation PR = 0 se traduit par

$$GR = 0 \qquad (73)$$

Finalement on peut remarquer que, pour toute matrice G, résoudre le problème

$$\begin{cases} \min\limits_{b} (y - Xb)' G' G (y - Xa) \\ \text{sous contraintes} \\ R' (y - Xb) = 0 \end{cases}$$

équivaut, du fait que

$$G' G = P' P \qquad (74)$$

à résoudre le problème (P1).

De même, le problème suivant

$$\begin{cases} \min\limits_{b} (y - Xb)'(G'G + RR') (y - Xb) \\ \text{sous contraintes} \\ R' (y - Xb) = 0 \end{cases}$$

est identique au problème (P 4)

3. ETUDE DU PROBLEME (P 5)

On commencera par démontrer le résultat essentiel suivant.

Proposition 4

Soit Ω une matrice d.n.n. de format (T, T) et de rang h $(h < T)$. Soit une matrice quelconque R, de format $(T, T - h)$ et rang $(T - h)$ vérifiant $\Omega R = 0$. Alors la solution du problème (P 5) :

$$(P\ 5)\left\{\begin{array}{l} \min\limits_{b}\ (y - Xb)'\ \Omega^{-}(y - Xb) \\ \text{sous contrainte} \\ R'Xb = R'y \end{array}\right.$$

est indépendante du choix de Ω^{-} et de R.

Démonstration

Ω^- est une g-inverse particulière de Ω. On sait, d'après le lemme 2, que si Ω^- est une g-inverse quelconque de Ω, il existe deux matrices U et W telles que :

$$\Omega^{-} = \Omega^{+} + U(I - \Omega\,\Omega^{+}) + (I - \Omega^{+}\Omega)\,W \tag{75}$$

Or, d'après (c) et (d) de la proposition 3, on a :

$$\Omega\,\Omega^{+} = \Omega^{+}\Omega = Q\,Q' \tag{76}$$

c'est-à-dire, d'après (24)

$$\Omega\,\Omega^{+} = \Omega^{+}\Omega = I_T - R_o R_o' \tag{77}$$

Il en résulte que, pour toute g-inverse de Ω, on peut trouver deux matrices U et W telles que :

$$\Omega^{-} = \Omega^{+} + U\,R_o R_o' + R_o R_o'\,W \tag{78}$$

Par conséquent, on a :

$$\begin{aligned}(y - Xb)'\ \Omega^{-}(y - Xb) = {} & (y - Xb)'\ \Omega^{+}\ (y - Xb) \\ & + (y - Xb)'\ U\,R_o R_o'\ (y - Xb) \\ & + (y - Xb)'\ R_o R_o' W\ (y - Xb)\end{aligned} \tag{79}$$

Or, R_o vérifie $\Omega\, R_o = 0$. D'après la proposition 1, on a presque sûrement

$$R_o'\ u = u'R_o = 0 \text{ avec } y - Xb = u$$

Il en résulte que, quel que soit Ω^-

$$(y - Xb)' \Omega^-(y - Xb) = (y - Xb)' \Omega^+ (y - Xb) \qquad (80)$$

Ainsi, pour toute g-inverse Ω^- de Ω, la solution de (P 5) est donnée par la solution du problème

$$\begin{cases} \min\limits_{b} (y - Xb)' \Omega^+ (y - Xb) \\ \text{sous contrainte} \\ R'Xb = R'y \end{cases}$$

Il reste à démontrer que la contrainte est indépendante de R.

La contrainte

$$R'Xb = R'y \qquad (81)$$

se traduit par :

$$S R'_o Xb = S R'_o y \qquad (82)$$

c'est-à-dire par :

$$R'_o Xb = R'_o y \qquad (83)$$

La solution de (P 4) est donc donnée par la solution de

$$(P\ 3) \begin{cases} \min\limits_{b} (y - Xb)' \Omega^+ (y - Xb) \\ \text{sous contrainte} \\ R'_o Xb = R'_o y \end{cases}$$

qui est bien indépendant du choix de Ω^- et de R. ■

Il est facile de vérifier que si on ajoute des contraintes sur les coefficients, de la forme $Lb = l$, la proposition suivante est vraie.

Proposition 5

La solution du problème

$$(P\ 6) \begin{cases} \min\limits_{b} (y - Xb)' \Omega^- (y - Xb) \\ \text{sous contraintes} \\ R' Xb = R'y \\ Lb = l \end{cases}$$

Ω et R vérifiant les hypothèses de la proposition (4), est indépendante du choix de Ω^- et de R.

On voit facilement en effet que la solution de ce problème est obtenue en résolvant :

$$(P\ 7)\begin{cases} \min\limits_{b} (y - Xb)\ \Omega^+\ (y - Xb) \\ \text{sous contraintes} \\ R'_o Xb = R'y \\ Lb = l \end{cases}$$

On retrouve sous cette forme le problème étudié par Theil ([8], pp. 282-289) qui démontre que l'estimateur obtenu en résolvant (P 7) est le b.l.u.e. de b. C'est la raison pour laquelle on ne démontre pas, dans le présent travail, que la solution de (P 4) est le b.l.u.e. de b.

Jusqu'à présent nous avons simplement constaté que les problèmes (P 1), (P 2), (P 3) et (P 4) sont des cas particuliers de (P 5). Il est vrai qu'une étude directe du modèle I conduit à ces cas particuliers. On peut cependant se demander pourquoi l'étude du problème (P 5) s'impose dans le cas général avec une matrice Ω^- quelconque. On peut donner une justification géométrique de ce problème en utilisant l'approche de Malinvaud.

4. JUSTIFICATION GEOMETRIQUE DU PROBLEME (P 5)

Il est possible de justifier par une approche géométrique la recherche de l'estimateur de b en résolvant le problème (P 5). En fait, on peut montrer que l'approche de Malinvaud est vérifiée pour toute matrice Ω^-g-inverse de Ω. Plus précisément, l'approche de Malinvaud s'applique en remplaçant $(\Omega + RR')^{-1}$par Ω^-et conduit au problème (P 5).

L'approche de Malinvaud repose fortement sur trois notions géométriques : la notion d'ellipsoïde indicateur E d'un vecteur aléatoire, la notion de support d'un vecteur aléatoire et la notion de conjugaison de deux vecteurs, appartenant à ce support, par rapport à l'ellipsoïde E.

On applique directement ces notions au vecteur u.

On peut montrer que la partie (ii) de la proposition 1 de Malinvaud ([3] p. 168) se généralise de la manière suivante :

Proposition 6

Soit Ω une matrice d.n.n. de format (T, T) et de rang h. Si le support de u est défini par $R'u = 0$ avec une matrice R de format $(T, T - h)$, de rang $T - h$, telle que $0\ R = 0$, alors l'ellipsoïde indicateur de u est l'ensemble des vecteurs ϵ satisfaisant $R'\epsilon = 0$ et :

$$\epsilon'\ \Omega^-\epsilon \leqslant 1 \tag{84}$$

Cet ellipsoïde est indépendant du choix de Ω^-et de R.

Démonstration

Malinvaud a démontré que cet ellipsoïde est l'ensemble des vecteurs ϵ vérifiant :

$$R'\epsilon = 0 \tag{85}$$

et

$$\epsilon'(\Omega + R\,R')^{-1}\,\epsilon \leqslant 1 \tag{86}$$

Il faut alors remarquer que, d'après (40), $R'\epsilon = 0$ entraîne, du fait que $Z\,R_o$ est régulière :

$$R'_o\,\epsilon = 0 \tag{87}$$

D'autre part, on sait, d'après le lemme 11, que $(\Omega + R\,R')^{-1}$ est une g-inverse particulière de Ω. On peut donc trouver deux matrices A et B telles , que, d'après le lemme 2, et la première partie de la proposition 2,

$$(\Omega + R\,R')^{-1} = \Omega^{+} + A\,R_o R_o' + R_o R_o' B \tag{88}$$

Il résulte de (87) que :

$$\epsilon'\,(\Omega + RR')^{-1}\epsilon = \epsilon'\,\Omega^{+}\,\epsilon \tag{89}$$

On sait aussi, d'après (41), que :

$$\Omega^{-} = \Omega^{+} + U\,R_o R_o' + R_o\,R_o'\,V$$

qui conduit, en tenant compte de (87), à :

$$\epsilon'\Omega^{-}\epsilon = \epsilon'\,\Omega^{+}\,\epsilon \tag{90}$$

qui est bien indépendant du choix de Ω^{-}. ∎

Rappelons la définition de conjugaison par rapport à un ellipsoïde.

Définition 3

Deux vecteurs ϵ et η, du support d'un vecteur aléatoire u, sont dits conjugués par rapport à l'ellipsoïde indicateur E de u si :

$$v'\Omega\,\nu = 0 \tag{91}$$

pour tout couple de vecteurs v *et* ν *tels que* $\epsilon = \Omega\,v$ *et* $\eta = \Omega\,\nu$

On a alors le résultat suivant.

Proposition 7

Si deux vecteurs ϵ et η sont conjugués par rapport à l'ellipsoïde indicateur E du vecteur aléatoire u, ils vérifient

$$\epsilon'\Omega^{-}\eta = 0 \tag{92}$$

quelle que soit la g-inverse Ω- de Ω.

Inversement, si deux vecteurs ϵ et η du support de u vérifient (92), alors les vecteurs v *et ν correspondants sont conjugués par rapport à l'ellipsoïde indicateur E du vecteur aléatoire u.*

Démonstration

On sait que $\epsilon = \Omega$ v et $\eta = \Omega\ \nu$ entraînent que :

$$v'\Omega\ \eta = v'\ \Omega\Omega^{-}\Omega\ \eta \tag{93}$$

pour tout choix de Ω^{-}.

C'est-à-dire :

$$\epsilon'\Omega^{-}\eta = 0 \tag{94}$$

Inversement, si deux vecteurs ϵ et η du support de u vérifient (77) on a :

$$v'\Omega\ \Omega^{-}\eta = 0$$

c'est-à-dire

$$v'\Omega\ \eta = 0$$

■

Les propositions 6 et 7 conduisent, par une approche géométrique, au problème (P 5). On ne donne pas les détails de la démonstration d'un tel résultat. Il suffit de refaire la démarche de Malinvaud en remplaçant $(\Omega + R\ R')^{-1}$ par Ω^{-}.

REFERENCES

[1] FOURGEAUD, C. et LENCLUD, B., (1972) – L'estimation dans les modèles linéaires. Paris, miméo.

[2] LENTNER, M.M. (1970) – "A characterization of all generalized inverses of an arbitrary real matrix". *Technical Report Series n° 19. September 1970,* Statistical Laboratory– New Mexico State University.

[3] MALINVAUD, E. (1969) – *Méthodes statistiques de l'économétrie.* 2e édition. Paris. Dunod.

[4] MOORE, E.H. (1920) – "On the reciprocal of general algebraic matrix (abstract). *Bull. Amer. Math. Soc., 26,* pp. 394-395.

[5] PENROSE, R, (1955) – "A generalized inverse for matrices". *Proc. Cambridge Philos. Soc.,* 51, pp. 17-19.

[6] RAO, C.R. (1965) – *Linear statistical interference and its applications.* J. Wiley, New-York.

[7] RAO, C.R. (1971) – "Unified theory of linear estimation". *Sankhya, series A*, vol. 33, Dec. 1971, pp. 371-194.

[8] RAO, C.R. (1972) – "Unified theory of least-squares". Résumé paru dans : *The Institute of Mathematical Statistics. Bulletin, vol. 1, n°. 4,* July 1972, p. 198.

[9] RAO, C,R. et MITRA, S.K. (1971) – *Generalized inverse of matrices and its applications.* J. Wiley. New-York.

[10] THEIL, H. (1971) – *Principles of Econometrics.* J. Wiley. North Holland.