REVUE DE STATISTIQUE APPLIQUÉE

FARIB E. ABDEL-BADIE
MOHAMED F.M. ABDEL-FATAH

Les effets de corrélation contemporaine et non contemporaine sur l'estimation d'un système d'équations de régression

Revue de statistique appliquée, tome 22, nº 2 (1974), p. 69-80
<http://www.numdam.org/item?id=RSA_1974__22_2_69_0>

LES EFFETS DE CORRÉLATION CONTEMPORAINE ET NON CONTEMPORAINE SUR L'ESTIMATION D'UN SYSTÈME D'ÉQUATIONS DE REGRESSION

Farib E. ABDEL-BADIE et Mohamed F. M. ABDEL-FATAH

RESUME

Le problème considéré dans cet article consiste à obtenir des estimateurs efficaces des paramètres d'un système de M équations de régression. On suppose que les termes perturbants de ce système sont liés à la fois par des corrélations sériales et contemporaines. Sous l'hypothèse supplémentaire que la corrélation sériale est un processus auto-régressif de premier ordre, on ébauche un estimateur convergent et ayant la même distribution asymptotique normale que l'estimateur de Aitken, qui suppose connue la matrice de covariance. L'article évalue les résultats dans le cas d'un système de deux équations de régression.

1 – INTRODUCTION

Le problème de l'estimation efficace d'un système d'équations de régression dans le cas où les "perturbations" du système ont des corrélations contemporaines a été étudié par Zellner [8]. Dans cet article, et dans d'autres : Zellner [9], Zellner et Theil [10] et Telser [7], le concept de corrélation sériale entre les perturbations avait été exclu. Il y a cependant beaucoup de situations où les perturbations aléatoires d'un système d'équations de régression peuvent présenter aussi bien des corrélations contemporaines qu'en série. En particulier, dans l'analyse du comportement de demande du consommateur, analyse utilisant des données en séries temporelles et qui est un cas cité par Zellner comme un domaine où sa technique pourrait être utile, la présence des deux phénomènes est très fréquente [4].

Voici le plan de l'article :

– dans la section 2, le modèle est posé et une technique efficace d'estimation est esquissée sous les hypothèses les plus générales sur la covariance,

– dans la section 3, on donne les propriétés des estimateurs,

– dans la section 4, on considère l'évaluation du modèle dans un système de deux équations de régression.

2 – ESTIMATION D'UN SYSTEME D'EQUATIONS DE REGRESSION

a) Le modèle

Considérons un système de M équations de régression, dont la $j^{\text{ème}}$ équation est :

$$\underline{Y}_j = X_j \, \underline{\beta}_j + \underline{U}_j, \qquad j = 1, 2, \ldots\ldots\ldots\ldots, M \tag{2.1}$$

où Y_j est le $(T \times 1)$ vecteur des observations sur la $j^{\text{ème}}$ variable dépendante.

X_j est une $(T \times K_j)$ matrice non stochastique des observations de K_j variables indépendantes, de rang $K_j < T$.

β_j est un $(K_j \times 1)$ vecteur des coefficients de régression non stochastiques, inconnus, à estimer.

U_j est un $(T \times 1)$ vecteur des termes de perturbation, non mesurables, chacun d'espérance nulle.

Le système dont (2.1) est une équation peut être noté :

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_M \end{bmatrix} = \begin{bmatrix} X_1 & 0 & 0 \ldots\ldots & 0 \\ 0 & X_2 & 0 \ldots\ldots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & & X_M \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_M \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ \cdot \\ \cdot \\ \cdot \\ U_M \end{bmatrix} \tag{2.2}$$

ou :

$$\underline{Y} = X \, \underline{\beta} + \underline{U} \tag{2.3}$$

où les variables dans (2.3) représentent les vecteurs et les matrices correspondants dans (2.2) c'est-à-dire

$$Y = [Y_1' \; Y_2' \ldots\ldots Y_M']', \qquad \beta = [\beta_1' \; \beta_2' \ldots\ldots \beta_M']',$$

et X représente la matrice diagonale en blocs avec des 0 comme matrices nulles. Considérons aussi les hypothèses très générales sur la covariance :

$$E(\underline{U} \, \underline{U}') = \underset{(MT \,.\, MT)}{B} = \{B_{ij.}\} \tag{2.4}$$

où B_{ij} est la $(T \times T)$ matrice de covariance $E(\underline{U}_i \, \underline{U}_j)'$. L'hypothèse (2.4) permet la corrélation des termes de perturbation, contemporains et non contemporains ; c'est-à-dire que les perturbations sont corrélées d'une équation à l'autre et en même temps auto-corrélées. Le modèle (2.3) est identique à celui considéré par Zellner [8], mais Zellner suppose de plus la matrice de covariance :

$$E(\underline{U}\,\underline{U}') = W$$
$$= W_c \oplus I \tag{2.5}$$

Cette hypothèse sur la covariance n'autorise que des corrélations contemporaines entre les perturbations : la covariance des perturbations non contemporaines est supposée nulle.

Considérons maintenant le système (2.3) avec les hypothèses (2.4). On peut regarder ce système comme un seul modèle d'équations de régression, et c'est un résultat bien connu que le meilleur estimateur linéaire sans biais b de β est donné par l'estimateur des moindres carrés de Aitken généralisé

$$\underline{\approx{b}} = (X' B^{-1} X)^{-1} X' B^{-1} \underline{Y} \tag{2.6}$$

De plus la matrice de variance-covariance de $\underline{\approx{b}}$ est

$$\mathrm{Var}(\underline{\approx{b}}) = (X' B^{-1} X)^{-1} \tag{2.7}$$

Si on suppose (2.4) vraie, l'estimateur $\underline{\approx{b}}$ (2.6) est plus efficace que l'estimateur

$$\underline{b} = (X' W^{-1} X)^{-1} X' W^{-1} Y$$

donné par Zellner. Ce résultat découle immédiatement du fait que l'estimateur de Aitken (2.6) a une plus petite variance généralisée que tout autre estimateur linéaire sans biais de β.

La matrice de covariance B est généralement inconnue et les résultats formels ci-dessus ont peu d'intérêt pratique, à moins qu'il n'y ait une méthode convenable pour estimer B. En général on doit mettre certaines contraintes sur la forme de B, lorsqu'il faut l'estimer.

b) **Les contraintes sur la forme de B**

Considérons de nouveau la $j^{\text{ème}}$ équation de régression du système (2.1). Il est supposé que les éléments $\underline{U}_j$ du vecteur des perturbations sont générés par le processus stationnaire, auto-regressif :

$$u_{jt} = \rho_j\, u_{jt-1} + v_{jt} \;;\; \rho_j < 1 \;;\; j = 1, 2, \dots, M \tag{2.8}$$

où les v_{jt} sont des variables aléatoires satisfaisant aux conditions :

$$E(v_{jt}) = 0 \;;\; j = 1, 2, \dots, M \;;\; t = 1, 2, \dots, T$$

$$E(v_{it}\, v_{jt'}) = \begin{cases} \sigma_{ij} \text{ pour } i, j = 1, 2, \dots, M \text{ et } t = t' \\ 0 \text{ pour } i, j = 1, 2, \dots, M \text{ et } t \neq t' \end{cases} \tag{2.9}$$

Le traitement de la valeur initiale de la perturbation exige une certaine attention qui manque souvent dans la littérature : la perturbation initiale a la même variance que les suivantes, pourvu que u_{jt} soit déterminé par (2.8) pour $t = 2, 3, \dots, T$, et u_{j1} donné par

$$u_{j1} = (1 - \rho_j^2)^{1/2} \, v_1 \qquad \text{pour } t = 1 \tag{2.10}$$

Il y a quelque avantage à utiliser (2.10) dans le développement qui suit. L'équation (2.1) peut être écrite :

$$\underline{Y}_j = X_j \, \underline{\beta}_j + P_j \underline{V}_j \tag{2.11}$$

$\underline{U}_j = P_j \underline{V}_j$, où $\underline{V}_j$ est un (T x 1) vecteur aléatoire avec

$$E(\underline{V}_j) = 0 \qquad \text{et} \qquad E(\underline{V}_j, \underline{V}_j') = \sigma_{jj} \, I_T \, , \; j = 1, 2, \dots, M,$$

et où

$$\underset{(T.T.)}{P_j =} \begin{bmatrix} (1 - \rho_j^2)^{-1/2} & 0 & 0 & \dots & 0 \\ \rho_j \, (1 - \rho_j^2)^{-1/2} & 1 & 0 & \dots & 0 \\ \rho_j^2 \, (1 - \rho_j^2)^{-1/2} & \rho_j & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \rho_j^{T-1} \, (1 - \rho_j^2)^{-1/2} & \rho_j^{T-2} & \rho_j^{T-3} & \dots & 1 \end{bmatrix} \tag{2.12}$$

P_j^{-1} prend la forme

$$P_j^{-1} = \begin{bmatrix} (1 - \rho_j)^{1/2} & 0 & 0 & \dots & 0 & 0 \\ -\rho_j & 1 & 0 & \dots & 0 & 0 \\ 0 & -\rho_j & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & \cdot & \dots & -\rho_j & 1 \end{bmatrix} \tag{2.13}$$

Le système complet d'équations (2.3) est maintenant

$$\underline{Y} = X \underline{\beta} + P \, \underline{V} \tag{2.14}$$

où P est la matrice diagonale en blocs $\{P_j\}$, j = 1, 2, . . . , M et

$$\underline{V} = [\underline{V}_1', \underline{V}_2', \dots, \underline{V}_M'].$$

Les hypothèses (2.9) peuvent s'écrire :

$$E(\underline{V}\,\underline{V}') = W$$
$$= W_c \oplus I \tag{2.15}$$

où $W = \{\sigma_{ij}\ I\}$ est la (MT x MT) matrice de covariances, $W_c = \{\sigma_{ij}\}$ est une (M x M) matrice, I est la (T x T) matrice identité, $\oplus$ représente le produit de Kronecker, et i, j = 1, 2, . . . , M. Alors en termes du modèle de base (2.3) la structure de covariance est :

$$E(\underline{U}\,\underline{U}') = B$$
$$= E(P\,\underline{V}\,\underline{V}'\,P')$$
$$= P\,E(\underline{V}\,\underline{V}')\,P'$$
$$= P\,(W_c \oplus I)\,P'$$
$$= \begin{bmatrix} \sigma_{11}\,P_1\,P_1' & \sigma_{12}\,P_1\,P_2' & \ldots & \sigma_{1M}\,P_1\,P_M' \\ \sigma_{21}\,P_2\,P_1' & \sigma_{22}\,P_2\,P_2' & \ldots & \sigma_{2M}\,P_2\,P_M' \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \sigma_{M1}\,P_M P_1' & \sigma_{M2}\,P_M\,P_2' & \ldots & \sigma_{MM} P_M\,P_M' \end{bmatrix} \tag{2.16}$$

c) **Estimation des coefficients de régression**

Dans la situation pratique où B est inconnue, l'estimateur de Aitken (2.6) n'est utile que si l'estimation de B peut être faite avec celle des coefficients de régression β. Nous allons donner une technique d'estimation convergente et asymptotiquement efficace pour les coefficients de régression β. Une telle estimation se fait en trois étapes. La première étape utilise une seule équation de régression pour estimer les paramètres du modèle auto-régressif. La seconde étape utilise une seule équation de régression sur les équations transformées, pour estimer les covariances contemporaines. Enfin l'estimateur de Aitken est formé grâce aux covariances estimées. Pour estimer les paramètres du processus auto-régressif, considérons

$$\hat{\underline{U}}_j = \underline{Y}_j - X_j\,\hat{\underline{\beta}}_j$$

où $\hat{\underline{\beta}}_j$ est l'estimateur usuel par les moindres carrés de $\underline{\beta}_j$, donné par les équations (2.11). Les paramètres ρ_j du processus auto-régressif peuvent être estimés de manière consistante sur la base d'une régression sur ces résidus et l'équation de régression

$$u_{jt} = \rho_j\,u_{jt-1} + v_{jt}$$

donne

$$\hat{\rho}_j = \frac{\sum_{t=2}^{T} \hat{u}_{jt}\hat{u}_{jt-1}}{\sum_{t=2}^{T} \hat{u}^2_{jt-1}} \tag{2.17}$$

La matrice $\hat{P}_j$ est obtenue en substituant $\hat{\rho}_j$ dans (2.12). Prémultipliant l'équation (2.11) par $\hat{R}_j = \hat{P}_j^{-2}$ donne :

$$\hat{R}_j \underline{Y}_j = \hat{R}_j X_j \underline{\beta}_j + \hat{R}_j P_j \underline{V}_j \quad \text{ou}$$

$$\underline{\tilde{Y}}_j = \tilde{X}_j \underline{\beta}_j + \tilde{V}_j, \qquad j = 1, 2, \ldots, M \tag{2.18}$$

Les résidus estimés à partir des équations de régression transformées peuvent être utilisés pour estimer les éléments de la matrice des covariances contemporaines $W_c = \{\sigma_{ij}\}$, par

$$\begin{aligned} s_{ij} &= \frac{\hat{\hat{V}}'_i \hat{\hat{V}}_j}{(T - K_i)^{1/2} (T - K_j)^{1/2}} \\ &= \frac{(\underline{\tilde{Y}}_i - \tilde{X}_i \hat{\underline{\beta}}_i)' (\underline{\tilde{Y}}_j - \tilde{X}_j \hat{\underline{\beta}}_j)}{(T - K_i)^{1/2} (T - K_j)^{1/2}} \end{aligned} \tag{2.19}$$

où $\hat{\hat{\underline{\beta}}}_j$ est maintenant l'estimateur usuel par les moindres carrés des coefficients de régression des $j^{èmes}$ équations transformées (2.18). La procédure d'estimation décrite ci-dessus donne un estimateur pour la matrice de covariance B, c'est-à-dire :

$$\begin{aligned} \hat{B} &= \hat{P}\, \hat{W}\, \hat{P}' \\ &= \hat{P}(\hat{W}_c \oplus I)\, \hat{P}' \end{aligned} \tag{2.20}$$

où $\hat{P}$ est la (MT x MT) matrice diagonale en blocs des P_j, et $\hat{W}_c = \{S_{ij}\}$. Pour l'estimateur de Aitken (2.6) on a besoin d'inverser B. L'estimateur de cette inverse est donné par :

$$\begin{aligned} \hat{B}^{-1} &= \hat{P}'^{-1}\, \hat{W}^{-1}\, \hat{P}^{-1} \\ &= \hat{R}'\, \hat{W}^{-1}\, \hat{R} \\ &= \hat{R}'\, (\hat{W}^{-1} \oplus I)\, \hat{R} \end{aligned} \tag{2.21}$$

Enfin dans la troisième étape, l'estimateur

$$\underline{\approx{\beta}} = (X'\, \hat{B}^{-1}\, X)^{-1}\, X'\, \hat{B}^{-1}\, \underline{Y} \tag{2.22}$$

est calculé. La dernière étape est équivalente à la technique de Zellner [8] appliquée au modèle transformé, c'est-à-dire

$$\underline{\approx{\beta}} = [\tilde{X}'\, (\hat{W}_0^{-1} \oplus I)\, \tilde{X}]^{-1}\, [\tilde{X}'\, (\hat{W}_0^{-1} \oplus I)\, \underline{\tilde{Y}}]$$

et Var $(\underline{\approx{\beta}}) = (X'\, \hat{B}^{-1}\, X)^{-1}$, c'est-à-dire que la variance asymptotique est la même que celle de $\underline{\approx{b}}$ (2.6), d'où $\underline{\approx{\beta}}$ est, au moins asymptotiquement, plus efficient que l'estimateur de Zellner :

$$\underset{\sim}{\tilde{b}} = [X' (\hat{W}_0^{-1} \oplus I) X]^{-1} [X' (W_0^{-1} \oplus I) \underline{Y}] \tag{2.23}$$

quand seules sont présentes les corrélations contempóraines entre les erreurs.

3 – PROPRIETES DES ESTIMATEURS

Pour le système des équations de régression (2.3) où le vecteur des perturbations a une espérance nulle et une matrice de covariance connue (2.4), l'estimateur de Aitken (2.6) est l'estimateur linéaire, sans biais de $\underline{\beta}$ à variance minimum. L'erreur d'échantillonnage de cet estimateur peut être écrite :

$$\begin{aligned} d &= \underline{b} - \underline{\beta} \\ &= (X' B^{-1} X)^{-1} X' B^{-1} \underline{U} \end{aligned} \tag{3.1}$$

Une généralisation du théorème limite central pour des variables aléatoires dépendantes peut être appliquée pour montrer que $\sqrt{T}$ d converge en distribution vers une distribution normale d'espérance nulle et de matrice de covariance asymptotique $T(X' B^{-1} X)^{-1}$. Considérons l'estimateur (2.22) où $\hat{B}^{-1}$ est donné par (2.21) ; les résultats suivants seront prouvés :

i) $\hat{B}$ est un estimateur convergent pour B.

ii) $\underline{\approx{\beta}}$ est un estimateur convergent pour $\underline{\beta}$.

iii) $\sqrt{T}(\underline{\approx{\beta}} - \underline{\beta})$ a la même distribution asymptotique normale que $\sqrt{T}(\underline{\approx{b}} - \underline{\beta})$.

Pour prouver que $\hat{B}$ est convergent, il faut montrer que $\hat{R}$ est un estimateur convergent de P^{-1} et que $\hat{W}_c$ est un estimateur convergent de W_c, en employant le théorème de Slutsky (2, pp. 252-253).

Il est bien connu que les éléments ρ_j de P_j sont estimés de manière convergente par (2.17). La matrice P est donc estimée de manière convergente par $\hat{P}$, par le théorème de Slutsky :

$$P \lim \hat{R} = P \lim \hat{P}^{-1} = (P \lim \hat{P})^{-1} = P^{-1} \tag{3.2}$$

Les résidus estimés de la $j^{ème}$ équation transformée peuvent être écrits :

$$\begin{aligned} \hat{\underline{V}}_j &= \hat{R}_j \underline{Y}_j - \hat{R}_j X_j \hat{\hat{\underline{\beta}}}_j \\ &= [I - \hat{R}_j X_j (X_j' \hat{R}_j \hat{R}_j X_j)^{-1} X_j' \hat{R}_j'] \underline{V}_j \end{aligned} \tag{3.3}$$

Soit

$$N_j = \hat{R}_j X_j (X_j' \hat{R}_j' \hat{R}_j X_j)^{-1} X_j \hat{R}_j.$$

Alors par (2.19) :

$$S_{ij} = \frac{\underline{V}_i' (I - N_i)' (I - N_j) \underline{V}_j}{(T - K_i)^{1/2} (T - K_j)^{1/2}} \tag{3.4}$$

Considérons maintenant

$$\text{P lim } S_{ij} = \text{P lim } \left\{ \frac{\underline{V}_i' \underline{V}_j}{T} - \frac{\underline{V}_i' N_3 \underline{V}_j}{T} - \frac{\underline{V}_i' N_i' \underline{V}_j}{T} - \frac{\underline{V}_i' N_i' N_j \underline{V}_j}{T} \right\} \tag{3.5}$$

Le premier terme, P lim $\frac{\underline{V}_j \underline{V}_j}{T} = \sigma_{ij}$ est le moment de la population.

Le deuxième et le troisième termes peuvent s'écrire :

$$\text{P lim } \underline{V}_i' N_j \underline{V}_j/T = \text{P lim } \frac{\underline{V}_i' [\hat{R}_j X_j (X_j' \hat{R}_j' \hat{R}_j X_j)^{-1} X_j' \hat{R}_j'] \underline{V}_j}{T}$$

$$= \text{P lim } \underline{V}_i' \hat{R}_j X_j/T . \text{P lim } (T^{-1} X_j' \hat{R}_j' \hat{R}_j' X_j)^{-1}.$$

$$\text{P lim } X_j' \hat{R}_j' \underline{V}_j/T \tag{3.6}$$

grâce au théorème de Slutsky.

Puisque P lim $\hat{R}_j = P_j^{-1}$, et $E(\underline{V}_j) = 0$, il s'ensuit que P lim $\underline{V}_j' \hat{R}_j X_j/T = 0$. Puisque X_j est non stochastique, le terme du milieu de (3.6) converge vers une matrice $L_{x_j x_j}$.

De même, le dernier terme de (3.5) peut s'écrire :

$$\text{P lim } \underline{V}_i' N_i' N_J \underline{V}_j/T = 0 \,.\, L_{x_i x_i} L_{x_i x_j} L_{x_j x_j} \,.\, 0 = 0 \tag{3.7}$$

Donc (3.5) devient P lim $S_{ij} = \sigma_{ij} + 0 + 0 + 0 = \sigma_{ij}$, d'où P lim $W_c = W_c$.

L'erreur d'échantillonnage pour l'estimation $\underline{\approx\beta}$ de (2.22) est

$$\tilde{d} = \underline{\approx\beta} - \underline{\beta}$$

$$= (X' \hat{B}^{-1} X) \quad X' \hat{B}^{-1} \underline{U} \tag{3.8}$$

et nous devions montrer que

$$\sqrt{T}\ d\ (3.1) \sim N\ [0, T\ (X'\ B^{-1} X)^{-1}]$$

$$\sqrt{T}\ (d - \tilde{d}) = \sqrt{T}\ [(X'\ B^{-1}\ X)^{-1}\ X'\ B^{-1}\ \underline{U} - (X'\ \hat{B}^{-1}\ X)^{-1}\ X'\ \hat{B}^{-1}\ \underline{U}]$$

$$\text{P lim} \sqrt{T}\ (d - \tilde{d}) = \text{p lim} \sqrt{T}\ [(X' B^{-1} X)^{-1}\ X' B^{-1}\ \underline{U} - (X'\ \hat{B}^{-1} X)^{-1}\ X' \hat{B}^{-1}\ \underline{U}]$$

$$= \text{P lim} \sqrt{T}\ (X'\ B^{-1}\ X)^{-1}\ [X' (B^{-1} - \hat{B}^{-1})\ \underline{U}] \tag{3.9}$$

par le théorème de Slutsky. Mais alors

$$\text{P lim } (B^{-1} - \hat{B}^{-1}) = 0$$

c'est pourquoi

$$\text{P lim } T(d - \tilde{d}) = 0$$

d'où $\sqrt{T}$ d, $\sqrt{T}\ \tilde{d}$ ont la même distribution normale limite. De plus, puisque

$$\text{P lim} \sqrt{T}\ (d - \tilde{d}) = \text{P lim} \sqrt{T}\ (\underline{\tilde{b}} - \underline{\approx\beta}) = 0$$

nous avons prouvé ce qu'il fallait.

Nous avons montré que $\sqrt{T}\ \tilde{\tilde{\beta}}$ converge vers la même distribution asymptotique que $\sqrt{T}\ \tilde{\tilde{b}}$. Donc $\tilde{\tilde{\beta}}$ est asymptotiquement plus efficace que $\tilde{b}$ (2.23), qui ignore les corrélations non contemporaines entre les termes perturbants.

4 – ESSAI DU MODELE DANS UN CAS PARTICULIER

Pour évaluer les résultats ci-dessus, considérons le cas où (2.3) consiste en deux équations de régression, nommément :

$$\underline{Y}_1 = X_1\ \beta_1 + \underline{U}_1 \tag{4.1}$$

$$\underline{Y}_2 = X_2\ \beta_2 + \underline{U}_2 \tag{4.2}$$

où (4.1) et (4.2) satisfont toutes les hypothèses émises plus haut. Le modèle peut s'écrire :

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$$

ou brièvement

$$\underline{Y} = X\,\underline{\beta} + \underline{U} \tag{4.3}$$

Pour estimer le vecteur $\underline{\beta}$ des coefficients de régression par la procédure exposée, nous ferons les trois pas suivants :

On estime premièrement $\underline{\beta}_1$ et $\underline{\beta}_2$ [(4.1), (4.2)] par la technique habituelle des moindres carrés, c'est-à-dire :

$$\underline{\beta}_1 = (X_1'\ X_1)^{-1}\ X_1'\ \underline{Y}_1$$

$$\underline{\beta}_2 = (X_2'\ X_2)^{-1}\ X_2'\ \underline{Y}_2$$

puis les paramètres auto-régressifs ρ_1 et ρ_2 d'après la formule (2.17), c'est-à-dire:

$$\hat{\rho}_1 = \frac{\sum_{t=2}^{T} \hat{u}_{1t}\ \hat{u}_{1t-1}}{\sum_{t=2}^{T} \hat{u}_{1t-1}^2}\ ;\ \hat{\rho}_2 = \frac{\sum_{t=2}^{T} \hat{u}_{2t}\ \hat{u}_{2t-1}}{\sum_{t=2}^{T} \hat{u}_{2t-1}^2}$$

Ensuite les matrices $\hat{P}_1$ et $\hat{P}_2$ sont obtenues en substituant $\hat{\rho}_1$ et $\hat{\rho}_2$ dans (2.12). La prémultiplication de (4.1) par $\hat{P}_1^{-1}$ et de (4.2) par $\hat{P}_2^{-1}$ donne :

$$\hat{P}^{-1}\ \underline{Y}_1 = \hat{P}^{-1}\ X_1\ \underline{\beta}_1 + \hat{P}_1^{-1}\ \underline{U}_1$$

$$\underline{\tilde{Y}}_1 = \tilde{X}_1\ \underline{\beta}_1 + \underline{\tilde{V}}_1 \tag{4.4}$$

et

$$\underline{\tilde{Y}}_2 = \tilde{X}_2\ \underline{\beta}_2 + \underline{\tilde{V}}_2 \tag{4.5}$$

où $\underset{\sim}{\tilde{V}}_1$ et $\underset{\sim}{\tilde{V}}_2$ sont les vecteurs des résidus sur les nouvelles variables transformées, d'après nos contraintes sur la matrice de covariance B, (2.8) – (2.16).

Deuxièmement, on estime $\underline{\beta}_1$ et $\underline{\beta}_2$ de (4.4) et (4.5) en utilisant de nouveau la technique des moindres carrés, c'est-à-dire :

$$\hat{\hat{\underline{\beta}}}_1 = (\tilde{X}_1' \tilde{X}_1)^{-1} \tilde{X}_1' \tilde{\underline{Y}}_1$$

$$\hat{\hat{\underline{\beta}}}_2 = (\tilde{X}_2' \tilde{X}_2)^{-1} \tilde{X}_2' \tilde{\underline{Y}}_2$$

puis les éléments de la matrice $W_c = \{\sigma_{ij}\}$ des covariances contemporaines grâce aux résidus estimés des équations de régression transformées, (4.4) et (4.5), d'après la formule (2.19), c'est-à-dire :

$$S_{11} = \frac{\hat{\tilde{\underline{V}}}_1' \hat{\tilde{\underline{V}}}_1}{T-K_1} \quad , \quad S_{12} = S_{21} = \frac{\hat{\tilde{\underline{V}}}_1' \hat{\tilde{\underline{V}}}_2}{(T-K_1)^{1/2}\,(T-K_2)^{1/2}}$$

$$S_{22} = \frac{\underline{V}_2' \underline{V}_2}{T-K_2}$$

Nous avons donc, par ces deux étapes, estimé les expressions suivantes :

$$\underset{(2T\,.\,2T)}{\hat{P}} = \begin{bmatrix} \underset{(T.T)}{\hat{P}_1} & 0 \\ 0 & \underset{(T.T)}{\hat{P}_2} \end{bmatrix} \tag{4.6}$$

et

$$\underset{(2T\,.\,2T)}{\hat{W}} = \begin{bmatrix} S_{11}\, I_T & S_{12}\, I_T \\ S_{21}\, I_T & S_{22}\, I_T \end{bmatrix} = \underset{(2.2)}{\hat{W}_c I_t} \tag{4.7}$$

Si bien que l'estimée B de la matrice de covariance est :

$$\hat{B} = \hat{P}\,\hat{W}\,\hat{P}'$$

$$= \begin{bmatrix} S_{12}\, \hat{P}_1\, \hat{P}_1' & S_{12}\, \hat{P}_1\, \hat{P}_2' \\ S_{21}\, \hat{P}_2\, \hat{P}_1' & S_{22}\hat{P}_2\, \hat{P}_2' \end{bmatrix}$$

$$= \begin{bmatrix} \hat{B}_{11} & \hat{B}_{12} \\ \hat{B}_{21} & \hat{B}_{22} \end{bmatrix},$$

et nous pouvons écrire $\hat{B}^{-1}$ comme :

$$\hat{B}^{-1} = \begin{bmatrix} \hat{B}^{11} & \hat{B}^{12} \\ \hat{B}^{21} & \hat{B}^{22} \end{bmatrix} \tag{4.8}$$

où :

$$\hat{B}^{11} = [I - S_{12}^2/S_{11}\ S_{22}\{(\hat{P}_1\hat{P}_1')^{-1}(\hat{P}_1\hat{P}_2')(\hat{P}_2\hat{P}_2')^{-1}(\hat{P}_2\hat{P}_1')\}]^{-1}(\hat{P}_1\hat{P}_1')^{-1}/S_{11}$$

$$\hat{B}^{21} = -S_{12}/S_{22}\ [(\hat{P}_2\ \hat{P}_2')^{-1}\ (\hat{P}_2\ \hat{P}_1')]\ B^{11}$$

$$\hat{B}^{22} = [I - S_{12}^2/S_{11}\ S_{12}\ \{(\hat{P}_2\hat{P}_2')^{-1}(\hat{P}_2\hat{P}_1')(\hat{P}_1\hat{P}_1')^{-1}(\hat{P}_1\hat{P}_2')\}]^{-1}(\hat{P}_2\hat{P}_2')^{-1}/S_{22}$$

$$\hat{B}^{12} = -S_{12}/S_{11}\ (\hat{P}_1\hat{P}_1')^{-1}\ (\hat{P}_1\hat{P}_2')\ \hat{B}^{22}$$

en appliquant la règle sur l'inversion des matrices partitionnées.

Finalement, dans la troisième étape, on estime le vecteur $\underline{\beta}$ (4.4), par $\underline{\approx\beta}$ disons, d'après la formule (2.22), c'est-à-dire

$$(X'\ \hat{B}^{-1}\ X) = \begin{bmatrix} X_1'\ \hat{B}^{11}\ X_1 & X_1'\ \hat{B}^{12}\ X_2 \\ X_2'\ \hat{B}^{21}\ X_1 & X_2'\ \hat{B}^{22}\ X_2 \end{bmatrix} \qquad (4.9)$$

(4.9) est une $(K_1 + K_2) \times (K_1 + K_2)$ matrice définie, positive et dont l'inverse est nécessaire pour les calculs. On peut l'écrire :

$$(X'\ B^{-1}\ X)^{-1} = \begin{bmatrix} C^{11} & C^{12} \\ C^{21} & C^{22} \end{bmatrix}$$

où

$$C^{11} = [I - (X_1'\ \hat{B}^{11}X_1)^{-1}\ (X_1'\hat{B}^{12}X_2)\ (X_2'\hat{B}^{22}X_2)^{-1}\ (X_2'\ \hat{B}^{21}\ X_1)]^{-1}(X_1'\hat{B}^{11}\ X_1)^{-1}$$

$$C^{21} = -(X_2'\ \hat{B}^{22}\ X_2)^{-1}\ (X_2'\ \hat{B}^{21}\ X_1)\ C^{11}$$

$$C^{22} = [I - (X_2'\hat{B}^{22}\ X_2)^{-1}(X_2'\hat{B}^{21}\ X_1)\ (X_1'\hat{B}^{11}\ X_1)^{-1}\ (X_1'\hat{B}^{12}\ X_2)]^{-1}(X_2'\hat{B}^{22}\ X_2)^{-1}$$

$$C^{12} = -(X_1'\ \hat{B}^{11}\ X_1)^{-1}\ (X_1'\ \hat{B}^{12}\ X_2)\ C^{22}$$

en appliquant la régle d'inversion des matrices partitionnées.

De même :

$$X'\ \hat{B}^{-1}\ \underline{Y} = \begin{bmatrix} (X_1'\ \hat{B}^{11}\ \underline{Y}_1 + X_1'\ \hat{B}^{12}\ \underline{Y}_2) \\ (K_1\ .\ 1) \\ (X_2'\ \hat{B}^{21}\ \underline{Y}_1 + X_2'\ \hat{B}^{22}\ \underline{Y}_2) \\ (K_2\ .\ 1) \end{bmatrix}$$

et enfin

$$\underline{\approx\beta} = \begin{bmatrix} \underline{\approx\beta_1} \\ \underline{\approx\beta_2} \end{bmatrix} = \begin{bmatrix} C^{11}(X_1'\hat{B}^{11}\ \underline{Y}_1 + X_1'\hat{B}^{12}\underline{Y}_2) + C^{12}(X_2\hat{B}^{21}\underline{Y}_1 + X_2'\hat{B}^{22}\underline{Y}_2) \\ C^{21}\ (X_1'\hat{B}^{11}\ \underline{Y}_1 + X_1'\hat{B}^{12}\ \underline{Y}_2) + C^{22}(X_2'\hat{B}^{21}\ \underline{Y}_1 + X_2'\hat{B}^{22}\underline{Y}_2) \end{bmatrix}$$

(4.12)

Ce dernier vecteur des coefficients de régression estimés d'après la technique décrite ci-dessus est un estimateur convergent de $\underline{\beta}$, et a les propriétés que nous avons données.

REFERENCES

[1] CHERNOFF, H., — Large-Sample Theory : Parametric Case, Annals of Mathematical Statistics, Volume 27, 1965, pp. 1-22.

[2] CRAMER H., – Mathematical Methods of Statistics, Princeton University, Princeton, N. J., 1946.

[3] GOLDBERGER, A.S., – Econometric Theory, John Wiley & Sons, New-York, 1964.

[4] HOUTHAKKER H.S., – Additive Preferences, Econometrica, Volume 28, 1960, pp. 244-257.

[5] JOHNSTON, J., – Econometric Methods, Mc Graw Hill Book Gs., New-York, 1963.

[6] LOEVE, M., – Probability Theory, D. Van Nostrand Co, Princeton, 1955.

[7] TELSER, L.G., – Iterative Estimation of a set of Linear Regression Equations, Journal of the American Statistical Association, Volume 9, 1964, pp. 854-862.

[8] ZELLNER, A., – An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias, Journal of the American Statistical Association, Volume 57, 1962, pp. 348-368.

[9] ZELLNER, A., – Estimators of Seemingly Unrelated Regressions : Some Exact Finite Sample Results, Journal of the American Statistical Association, Volume 58, 1963, pp. 977-992.

[10] ZELLNER, A., and H. THEIL, Three Stage Least Squares – Simultaneous Estimation of Simultaneous Equations, Econometrica, Volume 30, 1960, pp. 54-78.