

MPRA

Munich Personal RePEc Archive

Forms and Estimation Methods of Panel Recursive Dynamic Systems

Hassan B. Ghassan

Sidi Mohammed Ben Abdullah University

10 March 2000

Online at <https://mpa.ub.uni-muenchen.de/56432/>

MPRA Paper No. 56432, posted 18 November 2014 00:29 UTC

Formes et méthodes d'estimation des systèmes récurrents dynamiques à double indice

Hassan GHASSAN*

Avril 2002

Résumé

L'article traite des méthodes d'estimation des modèles récurrents dynamiques au cas où l'observation porte sur des individus statistiques, chacun étant définie par un système d'équations de type récurrent dynamique. A la dimension temporelle, présente dans l'étude classique de ces systèmes, la dimension transversale est ajoutée pour obtenir des "modèles récurrents dynamiques à double indice". Le traitement des formes de ces modèles conduit à regrouper les variables par individu puis par équation [BG94]. Ce regroupement a permis d'assimiler le modèle au système d'équations apparemment non-liées SUR, avec la particularité que la variance relative aux individus i et j pour l'équation k n'est pas nécessairement nulle. L'estimation est réalisée par les méthodes du maximum de vraisemblance et les SUR-GLS non itér et itéré avec transformation de Taylor. Ces derniers donnent des estimateurs convergents¹ plus viables. L'application de ces méthodes fait l'objet d'un autre article sur un panel sectoriel de l'économie marocaine durant les trois dernières décennies.

Mots clés : *Causalité, système récurrent, estimation, asymptotique.*

Abstract

The purpose of this paper is to study the model belongs to the family of structural equation models with data varying both accross individuals (sectors) and in time. A complete theoretical analysis is developed in this work for the case of a dynamic recursive structure. Maximum likelihood estimation and SUR-GLS "*Seemingly Unrelated Regressions-Generalized Least Square*" estimators (iterated or not, with proper instruments and with Taylor's transformation) are carefully used. These last convergents estimators are most efficient. The application of these methods to panel sector of morocco's economy is treated in another paper.

Key words : *Causality, recursive system, estimation, asymptotic.*

*Professeur Assistant à la Faculté des Sciences Juridiques, Économiques et Sociales de Fès. Tel : (212.55)614322 E-mail address : h_ghassan@yahoo.fr

¹La convergence en probabilité des estimateurs purs est établie de manière originale par des lemmes, qui ne figurent pas dans ce papier.

1 Introduction

Le but de l'article est de traiter des méthodes d'estimation des modèles récurrents dynamiques au cas où l'observation porte sur des individus statistiques, chacun étant définie par un système d'équations de type récurrent dynamique. A la dimension temporelle, présente dans l'étude classique de tels systèmes, la dimension transversale est ajoutée pour obtenir les "modèles récurrents dynamiques à double indice".

Dans ce papier, il est démontré que la fonction log-vraisemblance du système complet est la somme de K log-vraisemblances indépendantes une pour chaque équation. L'avantage de cette décomposition est que le maximum de la fonction de vraisemblance peut se faire séparément pour chaque équation, mais simultanément pour les N individus. Ce résultat important permet de développer les méthodes d'estimation appropriées pour un système de N individus en présence de variables retardées et d'un processus vectoriel autorégressif d'ordre 1 des erreurs.

Comparativement à Hatanaka [Hat76] et à Fuller et Alii [FW80], l'estimation réalisable des paramètres du modèle complet est faite de manière simple à partir du développement de Taylor autour de coefficients de corrélation des erreurs obtenus par la méthode des variables instrumentales.

2 Le modèle individuel

La formalisation proposée généralise l'étude des systèmes récurrents dynamiques au cas où l'observation porte sur un ensemble de N unités statistiques. Un système récurrent dynamique à K équations est défini pour chaque individu, qui est observé sur T périodes. Le modèle est écrit sous sa forme individuelle (2.1), puis sous forme globale (2.2).

2.1 La forme de base du modèle

Au niveau de l'individu i ($i = 1, \dots, N$) pour la période d'observation t ($t = 1, \dots, T$), la forme structurelle du modèle est définie par le système à K équations suivant :

$$\phi^{i'} y_{t-1}^i + \beta^{i'} y_t^i + \gamma^{i'} x_t^i + u_t^i = 0 \quad (1)$$

où :

- y_t^i est le vecteur $K \times 1$ des observations concernant les K variables endogènes (le système étant alors par définition complet)
- x_t^i est le vecteur $L \times 1$ des observations concernant les L variables exogènes
- $\beta^{i'}$ est la matrice $K \times K$ des coefficients des variables endogènes, supposée non singulière
- $\phi^{i'}$ est la matrice $K \times K$ des coefficients des variables endogènes retardées, supposée non singulière. Les éléments de la diagonale principale sont notés α_d^i
- $\gamma^{i'}$ est la matrice $K \times L$ des coefficients des variables exogènes
- u_t^i est le vecteur $K \times 1$ d'erreurs.

2.1.1 Les hypothèses

Le modèle 1 est muni d'un corps d'hypothèses :

(i1) H1a : Hypothèse de comportements

β^i , ϕ^i et γ^i sont des matrices de constantes (inconnues), variant d'individu à individu. On admet par conséquent que chaque individu est caractérisé par un comportement qui lui est propre (mais qui reste invariant dans le temps).

(i2) H1b : Hypothèse de restrictions

Pour des raisons de commodité, on supposera également que les restrictions théoriques a-priori d'exclusion (les éléments nuls a-priori de β^i , ϕ^i et γ^i) sont les mêmes pour tous les individus.

(ii) H2 : Récursivité

Les matrices non-singulières $\beta^{i'}$ et $\phi^{i'}$ sont triangulaire inférieure².

(iii) H3 : Variables exogènes

Les variables exogènes sont fixes, non stochastiques, indépendantes des erreurs et telles que :

(a) $\sum_t x_t^i x_t^{i'}$ est une matrice de rang complet L

(b) $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_t x_t^i = \bar{x}^i$, fini

(c) $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_t x_t^i x_t^{i'} = M^i$, définie positive

(iv) H4 : Hypothèses des erreurs

Les erreurs possèdent des moments finis au moins jusqu'à l'ordre deux. En particulier, en désignant par u_{kt}^i , $k = 1, \dots, K$, l'élément typique du vecteur u_t^i , on postule que

(a) l'effet des perturbations est nul en moyenne :

$$E(u_{kt}^i) = 0 \quad \forall i, k, t$$

(b) la structure des variances-covariances est telle que :

$$E(u_{kt}^i u_{ls}^j) = \begin{cases} \sigma_k^{ij} & \text{si } k = l \text{ et } t = s \\ 0 & \text{sinon}^3 \end{cases}$$

La nullité de la covariance pour deux équations différentes ($k \neq l$) définit (conjointement à H2) la récursivité du système. La covariance contemporaine σ_k^{ij} entre l'erreur de l'individu i et celle de l'individu j pour l'équation k n'est pas nécessairement nulle. Cette possibilité donne au modèle sa première spécificité et exprime les liaisons aléatoires entre les individus.

²Leurs éléments au-dessus de la diagonale principale sont tous nuls et ceux de la diagonale principale de $\beta^{i'}$ sont tous égaux à -1 (du fait de la normalisation).

De l'hypothèse H4, on tire les deux formulations suivantes :

$$E(u_t^i) = 0 \quad \text{et}$$

$$E(u_t^i u_s^j) = \begin{cases} \text{diag}(\sigma_1^{ii}, \dots, \sigma_K^{ii}) = D^{ii} & \text{si } i = j \text{ et } t = s \\ \text{diag}(\sigma_1^{ij}, \dots, \sigma_K^{ij}) = D^{ij} & \text{si } i \neq j \text{ et } t = s \\ 0 & \text{sinon} \end{cases}$$

La seconde spécificité du modèle est que le processus des erreurs aléatoires u_{kt} d'ordre $N \times 1$ est un processus vectoriel auto-régressif "VAR(1)"⁴ :

$$u_{kt} = R_k u_{kt-1} + \epsilon_{kt} \quad \text{où } R_k = (\rho_{ij}) \quad \forall i, j = 1, 2, \dots, N \quad (2)$$

2.2 Forme de base d'une équation

La k -ème équation pour l'individu i (la k -ème ligne du modèle (1)) s'écrit :

$$(\phi_k^i)' y_{t-1}^i + (\beta_k^i)' y_t^i + (\gamma_k^i)' x_t^i + u_{kt}^i = 0 \quad (3)$$

où $(\phi_k^i)'$, $(\beta_k^i)'$ et $(\gamma_k^i)'$ sont respectivement la k -ème ligne de $(\phi^i)'$, $(\beta^i)'$ et $(\gamma^i)'$.

En raison de la triangularité de $(\beta^i)'$ et de $(\phi^i)'$, de la normalisation (le k -ème coefficient de (β_k^i) étant égal à -1) et en tenant compte d'éventuelles autres restrictions nulles sur certains paramètres, l'équation (2) prend la forme suivante :

$$y_{kt}^i = \alpha_{kk}^i y_{kt-1}^i + (\phi^{i/k})' y_{t-1}^{i/k} + (\beta^{i/k})' y_t^{i/k} + (\gamma^{i/k})' x_t^{i/k} + u_{kt}^i \quad (4)$$

où

- $y_t^{i/k}$ est le vecteur des variables endogènes explicatives dans cette équation (les variables endogènes y_{pt}^i d'indice $p \geq k$ ne figurent pas dans le vecteur $y_t^{i/k}$)
- $y_{t-1}^{i/k}$ est le vecteur des variables endogènes retardées explicatives dans cette équation (les variables endogènes y_{pt-1}^i d'indice $p \geq k$ ne figurent pas dans le vecteur $y_{t-1}^{i/k}$)
- $x_t^{i/k}$ est le vecteur des variables exogènes figurant dans l'équation (si toutes les variables exogènes interviennent dans l'équation, alors $x_t^{i/k}$ est égal à x_t^i). $y^{i/k}$ est l'ensemble des variables endogènes explicatives dans la k -ème équation, il sera noté par y_{*}^i pour alléger l'écriture.
- $(\phi^{i/k})'$, $(\beta^{i/k})'$ et $(\gamma^{i/k})'$ sont les parties correspondantes de $(\phi_k^i)'$, $(\beta_k^i)'$ et $(\gamma_k^i)'$.

Le nombre total de variables explicatives de la k -ème équation est désigné par s_k avec s_k inférieur ou égale à $L + 2k - 1$. θ_k^i est un vecteur colonne de paramètres d'ordre s_k ou encore $K_* + L_\Delta$ avec $(K_* < k)$.

⁴ Pour un scalaire, on écrit $u_{kt}^i = \sum_{j=1}^N \rho_{ij} u_{kt-1}^j + \epsilon_{kt}^i$ où ϵ_{kt}^i est un bruit blanc pour $i = 1, 2, \dots, N$. Ce processus est stationnaire si les racines du polynôme de $|I_N - Rz| = 0$ sont plus grandes que 1 en module.

Il est commode de réunir toutes les variables explicatives dans un seul vecteur ainsi que les paramètres correspondants. En posant :

$$Z_t^{i/k} = \begin{bmatrix} y_{t-1}^{i/k} \\ y_t^{i/k} \\ x_t^{i/k} \end{bmatrix} \quad (\theta_k^i)' = [(\phi^{i/k})' \quad (\beta^{i/k})' \quad (\gamma^{i/k})']$$

il vient alors :

$$y_{kt}^i = \alpha_{kk}^i y_{kt-1}^i + (\theta_k^i)' Z_t^{i/k} + u_{kt}^i = \alpha_{kk}^i y_{kt-1}^i + (Z_t^{i/k})' (\theta_k^i) + u_{kt}^i \quad (5)$$

on utilisera l'écriture suivante au niveau des méthodes d'estimation :

$$\varphi_k^i = \begin{bmatrix} \alpha_{kk}^i \\ \theta_k^i \end{bmatrix} \quad z_{kt}^i = [z_{kt-1}^i \quad (Z_t^{i/k})']$$

Par la récursivité du système individuel de base⁵, l'indépendance des erreurs d'une équation à l'autre fait que les erreurs u_{kt}^i ne sont pas corrélées avec $y_t^{i/k}$ ni avec $y_{t-1}^{i/k}$, car :

$$E(u_{kt}^i y_{lt-1}^i) = 0 \quad \text{pour } k \neq l \text{ pour tout } i, t$$

Ainsi, la matrice $Z_t^{i/k}$ est constituée de variables prédéterminées. En revanche, la variable y_{kt-1}^i est corrélée avec l'erreur aléatoire u_{kt}^i .

3 Forme compacte pour la période t

3.1 Regroupement des observations

Le regroupement des observations pour la période t peut s'opérer de deux manières différentes, chacune présente un avantage particulier. Le regroupement des individus consiste à empiler le modèle de base pour les N individus. On obtient :

$$\beta^l y_t + \phi^l y_{t-1} + \gamma^l x_t + u_t = 0 \quad (6)$$

⁵En considérant le modèle pour toutes les observations, on obtient :

$$y_k^i = \alpha_{kk}^i y_{kR}^i + Z^{i/k} \theta_k^i + u_k^i \quad \text{avec } Z^{i/k} = (y_R^{i/k} \quad y^{i/k} \quad x^{i/k})$$

La corrélation entre la variable endogène retardée explicative et les erreurs u_k^i s'écrit par :

$$E(u_k^i y_{kR}^i) \neq 0 \quad \text{pour tout } k, i$$

alors que les erreurs u_k^i ne sont pas corrélées avec y_*^i . Ces caractéristiques font que la méthode des moindres carrés ordinaires donne des estimateurs non convergents.

où β' et ϕ' sont triangulaire inférieure et diagonale par bloc, d'ordre $NK \times NK$, seulement la matrice ϕ' ne contient pas forcément des éléments unitaires sur la diagonale principale, qui est notée α_d . γ' est d'ordre $NK \times NL$, et y_t, y_{t-1}, u_t et x_t sont des vecteurs respectivement de dimension NK et NL . Pour le vecteur de perturbations u_t , à partir de H4 nous déduisons :

$$E(u_t) = 0, \text{ et}$$

$$E(u_t u'_s) = \begin{cases} [D^{ij}] := \Omega = R\Omega R' + \Sigma & \text{si } t = s \\ 0 & \text{sinon} \end{cases}$$

avec R une matrice bloc-diagonale. La matrice Ω n'est pas diagonale, mais bloc-diagonale. Ses propriétés ne sont pas faciles à déduire. Mais, l'équation (6) permet d'extraire immédiatement la fonction de vraisemblance du vecteur y en partant de la fonction de densité de ϵ_t , en raison de la triangularité de β' et de ϕ' .

Le regroupement des équations s'opère à partir de la k -ème équation pour l'individu i (4). Pour l'ensemble des K équations on arrive au modèle suivant pour l'observation t :

$$\tilde{y}_t = \tilde{y}_{t-1}\alpha + Z_t\theta + \tilde{u}_t \quad (7)$$

Pour l'observation t , l'équation (5) possède les propriétés suivantes par l'hypothèse H4 :

$$E(u_{kt}) = 0, \text{ et}$$

$$E(u_{kt} u'_{lt'}) = \begin{cases} [\sigma_k^{ij}] := \tilde{\Omega}_k & \text{si } k = l \\ 0 & \text{sinon} \end{cases} \quad (8)$$

la matrice des variances-covariances du vecteur \tilde{u}_t est une matrice diagonale par bloc :

$$E(\tilde{u}_t \tilde{u}'_t) = \text{diag}(\tilde{\Omega}_1, \dots, \tilde{\Omega}_K) := \tilde{\Omega} = \tilde{R}\tilde{\Omega}\tilde{R}' + \tilde{\Sigma} \quad (9)$$

la matrice \tilde{R} est diagonale par bloc.

En admettant l'absence de restrictions linéaires sur les éléments \tilde{u}_t , la matrice $\tilde{\Omega}$ doit être définie positive. Cette propriété est satisfaite dès lors que chaque bloc $\tilde{\Omega}_k$ est une matrice définie positive. Le déterminant et l'inverse de $\tilde{\Omega}$ se déduisent et sont donnés par :

$$\bullet \quad |\tilde{\Omega}| = \prod_{k=1}^K |\tilde{\Omega}_k| \quad \text{et} \quad \tilde{\Omega}^{-1} = \text{diag}(\tilde{\Omega}_1^{-1}, \dots, \tilde{\Omega}_K^{-1})$$

3.2 Matrice de passage de u_t à \tilde{u}_t

Les vecteurs u_t et \tilde{u}_t contiennent les mêmes éléments au nombre total de NK . Il est possible de les ordonner tous dans une matrice $K \times N$. A cet effet, nous posons :

$$U_t = [u_t^1, \dots, u_t^N] = \begin{bmatrix} u'_{1t} \\ \vdots \\ u'_{Kt} \end{bmatrix}$$

Dans cette matrice, les N colonnes sont les vecteurs u_t^i de l'équation (1), pour $i = 1, \dots, N$. En revanche, les K lignes représentent les vecteurs $(u_{kt})'$ de l'équation (6), avec $k = 1, \dots, K$.

Par les propriétés de l'opération `vec` (qui empile les colonnes d'une matrice) [Bal76] [Tur00] :

$$u_t = \text{vec } U_t \quad \text{et} \quad \tilde{u}_t = \text{vec } U_t'$$

$\text{vec } U_t' = P_{N,K} \text{vec } U_t$ où $P_{N,K}$ est la matrice unitaire permutée, on obtient :

$$\tilde{u}_t = P_{N,K} u_t \quad \text{et} \quad \tilde{\epsilon}_t = P_{N,K} \epsilon_t \quad (10)$$

La matrice orthogonale $P_{N,K}$ est donc la matrice de passage de u_t à \tilde{u}_t . Elle permet de passer indifféremment, par une transformation linéaire non-singulière, de la formulation (6) à la formulation (5) et surtout de dériver les propriétés de la matrice Ω de celles de $\tilde{\Omega}$.

En pré-multipliant l'équation (6) par $-P_{N,K}$, il vient :

$$-P_{N,K}(\beta' y_t + \phi' y_{t-1} + \gamma' x_t) = P_{N,K} u_t = \tilde{u}_t$$

En comparant avec (5), on a la relation suivante :

$$\tilde{y}_t - \tilde{y}_{t-1} \alpha - Z_t \theta = -P_{N,K}(\beta' y_t + \phi' y_{t-1} + \gamma' x_t) \quad (11)$$

L'équivalence ci-dessus sera exploitée pour dériver la fonction de vraisemblance du modèle.

En partant de (10), on a :

$$\tilde{\Omega} = P_{N,K} \Omega P_{K,N} \quad \text{et} \quad \widetilde{\Sigma} = P_{N,K} \Sigma P_{K,N} \quad (12)$$

par les propriétés de la matrice unitaire permutée⁶, on déduit immédiatement :

$$\Omega = P_{K,N} \tilde{\Omega} P_{N,K}$$

par conséquent :

$$|\Omega| = |P_{K,N} \tilde{\Omega} P_{N,K}| = |\tilde{\Omega}| |P_{K,N} P_{N,K}| = |\tilde{\Omega}|$$

d'où :

$$|\Omega| = \prod_{k=1}^K |\tilde{\Omega}_k| \quad \text{et} \quad |\Sigma| = \prod_{k=1}^K |\widetilde{\Sigma}_k| \quad (13)$$

$$\Omega^{-1} = P_{K,N} \tilde{\Omega}^{-1} P_{N,K} \quad \text{et} \quad \Sigma^{-1} = P_{K,N} \widetilde{\Sigma}^{-1} P_{N,K} \quad (14)$$

ou encore en utilisant la matrice de passage P et à partir de la relation (7), on a :

$$\begin{aligned} \Omega &= P_{K,N} \tilde{\Omega} P_{N,K} = P_{K,N} \tilde{R} \tilde{\Omega} \tilde{R}' P_{N,K} + P_{K,N} \widetilde{\Sigma} P_{N,K} \\ &= P_{K,N} \tilde{R} P_{N,K} \Omega P_{K,N} \tilde{R}' P_{N,K} + \Sigma = R \Omega R' + \Sigma \end{aligned}$$

⁶ $P_{N,K}' = P_{K,N} = P_{N,K}^{-1}$.

4 Factorisation de la vraisemblance

Soit ϵ_t un processus gaussien [IG84], la fonction de densité des erreurs est :

$$f(\epsilon_t) = (2\pi)^{-\frac{NK}{2}} |\Sigma|^{-\frac{1}{2}} \exp - \frac{1}{2} \epsilon_t' \Sigma^{-1} \epsilon_t$$

La fonction de densité de y_t est obtenue de celle de ϵ_t par changement des variables :

$$f(y_t) = (2\pi)^{-\frac{NK}{2}} |\Sigma|^{-\frac{1}{2}} \left\| \frac{\partial \epsilon}{\partial y_t'} \right\| \exp - \frac{1}{2} \epsilon_t' \Sigma^{-1} \epsilon_t$$

et la fonction de vraisemblance logarithmique, à une constante additive près, est la suivante⁷ :

$$L(y_t | \beta, \phi, \gamma, R, \Sigma) = -\frac{1}{2} \ln |\Sigma| + \ln |\beta' + \phi'| - \frac{1}{2} \epsilon_t' \Sigma^{-1} (\beta' y_t + \phi' y_{t-1} + \gamma' x_t + R u_{t-1}) \quad (15)$$

à l'aide de (11), (13) et (14), on obtient successivement :

$$\begin{aligned} \epsilon_t' \Sigma^{-1} \epsilon_t &= \epsilon_t' P'_{N,K} \widetilde{\Sigma}^{-1} P_{N,K} \epsilon_t \\ &= (\tilde{y}_t - \tilde{y}_{t-1} \alpha - Z_t \theta - \tilde{R} \tilde{u}_{t-1})' \widetilde{\Sigma}^{-1} (\tilde{y}_t - \tilde{y}_{t-1} \alpha - Z_t \theta - \tilde{R} \tilde{u}_{t-1}) \\ &= \text{tr} \widetilde{\Sigma}^{-1} (\tilde{y}_t - \tilde{y}_{t-1} \alpha - Z_t \theta - \tilde{R} \tilde{u}_{t-1}) (\tilde{y}_t - \tilde{y}_{t-1} \alpha - Z_t \theta - \tilde{R} \tilde{u}_{t-1})' \\ &= \text{tr} \text{diag} [\widetilde{\Sigma}_k^{-1} \epsilon_{kt} \epsilon_{kt}'] \\ &= \text{tr} \text{diag} (\epsilon_{1t} \widetilde{\Sigma}_1^{-1} \epsilon_{kt}', \dots, \epsilon_{kt} \widetilde{\Sigma}_k^{-1} \epsilon_{kt}') \\ &= \sum_k \epsilon_{kt} \widetilde{\Sigma}_k^{-1} \epsilon_{kt}' \end{aligned}$$

on a aussi :

$$\frac{1}{2} \ln |\Sigma| = \sum_{k=1}^K \frac{1}{2} \ln |\widetilde{\Sigma}_k|$$

comme les matrices β et ϕ sont triangulaires, on obtient :

$$|\beta' + \phi'| = |\text{diag} (\beta^{1'} + \phi^{1'}, \dots, \beta^{N'} + \phi^{N'})| = |\text{diag} [(\prod_k (1 + \phi_{kk}^1), \dots, \prod_k (1 + \phi_{kk}^N))]|$$

d'où :

$$\ln |\beta' + \phi'| = \ln \prod_i \prod_k (1 + \phi_{kk}^i) = \sum_k \ln \prod_i (1 + \phi_{kk}^i)$$

L'introduction de ces résultats dans (15) donne :

$$L(y_t) = \sum_{k=1}^K \left[-\frac{1}{2} \ln |\widetilde{\Sigma}_k| + \ln \prod_i (1 + \phi_{kk}^i) - \frac{1}{2} \epsilon_{kt}' \widetilde{\Sigma}_k^{-1} \epsilon_{kt} \right]$$

⁷ Le jacobien de transformation de ϵ_t à y_t est $\frac{\partial \epsilon}{\partial y_t'} = -\beta' - \phi'$.

ce qui revient à écrire :

$$L(y_t|\beta, \phi, \gamma, R, \sum) = \sum_{k=1}^K \mathcal{L}(y_{kt}|\theta_k, \alpha_k, R_k, \widetilde{\sum}_k) = \sum_{k=1}^K \mathcal{L}(y_{kt}|\varphi_k, \widetilde{\Omega}_k) \quad \square \quad (16)$$

La log-vraisemblance globale est donc la somme de K log-vraisemblances indépendantes, une pour chaque équation. L'avantage de cette décomposition est que la maximisation de la fonction de vraisemblance peut se faire séparément pour chaque équation, mais simultanément pour les N individus.

5 Forme compacte pour l'équation k

5.1 Regroupement des individus

En empilant les N individus, le modèle (5) devient :

$$y_k = y_{kR}\alpha_{kk} + Z_k\theta_k + u_k \quad (17)$$

en posant

$$u_k = \begin{bmatrix} u_k^1 \\ \vdots \\ u_k^N \end{bmatrix} \quad \alpha_{kk} = \begin{bmatrix} \alpha_{kk}^1 \\ \vdots \\ \alpha_{kk}^N \end{bmatrix} \quad \theta_k = \begin{bmatrix} \theta_k^1 \\ \vdots \\ \theta_k^N \end{bmatrix} \quad \text{et}$$

$$y_{kR} = \begin{bmatrix} y_{kR}^1 & & \\ & \ddots & \\ & & y_{kR}^N \end{bmatrix} \quad Z_k = \begin{bmatrix} Z^{1/k} & & \\ & \ddots & \\ & & Z^{N/k} \end{bmatrix}$$

où u_k est un vecteur aléatoire d'ordre $NT \times 1$ avec :

$$E(u_k) = 0 \quad \text{et} \quad E(u_k u_k') = \Omega_k \otimes I_T$$

Ce modèle ne correspond pas exactement à un système d'équations apparemment non-liées de type de Zellner [Mae80]. Le vecteur d'erreurs u_k ne possède pas la structure classique d'un tel modèle. En plus, de la présence de variables endogènes parmi les variables explicatives surtout retardées, qui sont corrélées avec les erreurs u_k .

5.2 Structure des erreurs et matrice de transformation

Du fait que u_k est un processus vectoriel auto-régressif d'ordre 1, la matrice Ω_k des variances-covariances complète de u_k est d'une forme compliquée [JGR⁺85]. En supposant que u_{kt} est

stationnaire⁸, nous écrivons ce qui suit :

$$\begin{aligned} V_{k0} &= R_k V_{k0} R_k' + \Sigma_k & \text{où } V_{k0} &:= E(u_{kt} u_{kt}') \text{ et } \Sigma_k := E(\epsilon_{kt} \epsilon_{kt}') = (\sigma_{\epsilon_k}^{ij}) \\ V_{ks} &= R_k^s V_{k0} & \text{où } V_{ks} &:= E(u_{kt} u_{k,t-s}') \quad 1 \leq s \leq t-1 \quad t = 1, 2, \dots, T \end{aligned}$$

dont la résolution en terme des éléments de V_{k0} donne :

$$vec(V_{k0}) = (I - R_k \otimes R_k)^{-1} vec \Sigma_k$$

A l'aide d'une matrice de transformation P carrée d'ordre NT tel que $P u_k = \epsilon_k$, nous n'aurons pas besoin de la spécifier explicitement pour estimer les paramètres du modèle. La matrice P est choisie de façon à satisfaire la condition suivante :

$$P \Omega_{u_k} P' = \Sigma_k \otimes I_T$$

En utilisant une matrice R_k non diagonale, chaque variable transformée par la matrice P sera exprimée en fonction de l'ensemble des variables du modèle, ce qui consomme beaucoup de degré de liberté. Aussi, le choix de P dépend de la structure de la matrice R_k . En considérant R_k diagonale⁹, il en résulte que : $u_{kt}^i = \rho_k^{ii} u_{kt-1}^i + \epsilon_{kt}^i$ avec $E(u_{kt}^i \epsilon_{kt}^i) = 0$ et $E(u_{kt}^i u_{kt}^j) = 0$ pour $i \neq j$, qui montre que y_{kt-1}^i est corrélée avec l'élément u_{kt}^i .

A partir du modèle (17) :

$$\begin{aligned} y_k &= y_{kR} \alpha_k + Z_k \theta_k + u_k \\ Z_k &= diag(y_{k*R} \quad y_{k*} \quad x_k) \\ u_k &= (R_k' \otimes I_T) u_{kR} + \epsilon_k \end{aligned}$$

avec

$$E(\epsilon_k \epsilon_k') = \Sigma_k \otimes I_T \quad E(u_k u_k') = \Omega_k \otimes I_T \quad E(u_k) = E(\epsilon_k) = 0 \quad \text{et} \quad E(u_k \epsilon_k') = 0$$

l'application des moindres carrés ordinaires donne des estimateurs non convergents, car :

$$plim z_k' u_k \neq 0 \quad \text{où} \quad z_k = (y_{kR} \quad Z_k) \quad (18)$$

Deux problèmes majeurs sont posés : celui de la corrélation entre y_{kt-1}^i et u_{kt}^i et ensuite le fait que les matrices R_k et Σ_k soient inconnues [Spe79]. Ils sont liés et traités en même temps.

⁸ Si la matrice Ω_k est générée par (Ω_k^{ij}) , celle-ci n'est pas nécessairement symétrique à moins que $i = j$. Les éléments de la diagonale de la matrice Ω_k^{ij} sont identiques à l'élément (i, j) de la matrice V_{k0} . La matrice V_{k1} fournit les éléments adjacents à la diagonale dans chaque matrice Ω_k^{ij} en respectant l'ordre des indices (i, j) . La matrice V_{k2} donne les éléments de la deuxième position par rapport à la diagonale de chaque matrice Ω_k^{ij} , et ainsi de suite.

⁹ Avec une matrice R_k non-diagonale y_{kt-1}^i est corrélée avec u_{kt}^i i.e. toutes les erreurs individuelles du modèle.

6 Méthodes d'estimation

Plusieurs techniques d'estimation sont envisageables pour briser ou du moins diminuer fortement la dépendance entre y_{it-1} et u_{it} pour tout i [Ham94]. Pour l'estimation des paramètres des résidus retardés, on utilise soit la méthode des variables instrumentales (*IV*), soit celle des doubles moindres carrés (*2SLS*). Pour l'estimation du modèle complet, on utilise soit la méthode des moments généralisés (*GMM*), soit celle du maximum de vraisemblance (*ML*), soit différentes variantes des techniques d'estimation¹⁰ (*GLS-SUR*).

6.1 Estimation par le maximum de vraisemblance

6.1.1 Fonction de vraisemblance

La log-vraisemblance pour la k -ème équation (voir section 3 et 4) s'écrit comme suit :

$$\begin{aligned}\mathcal{L}(y_k|\varphi_k, \tilde{\Omega}_k) &= -\frac{T}{2}\ln|\tilde{\Omega}_k| - \frac{1}{2}(y_k - z_k\varphi_k)'(\tilde{\Omega}_k^{-1} \otimes I)(y_k - z_k\varphi_k) \\ &= -\frac{T}{2}\ln|\tilde{\Omega}_k| - \frac{1}{2}tr\tilde{\Omega}_k^{-1}U_k'U_k\end{aligned}\quad (19)$$

avec $u_k = y_k - z_k\varphi_k$, $U_k = [u_k^1 \dots u_k^N]$ de dimension $T \times N$ et $u_k = \text{vec } U_k$.

Nous étudions d'abord la solution du problème de maximisation de la log-vraisemblance et dérivons ensuite la distribution limite des estimateurs du maximum de vraisemblance.

6.1.2 Les estimateurs du maximum de vraisemblance

Les conditions de premier ordre pour la maximisation de \mathcal{L} s'écrivent par :

$$\begin{aligned}\text{(i)} \quad \frac{\partial \mathcal{L}}{\partial \tilde{\Omega}_k} &= -\frac{T}{2}\tilde{\Omega}_k^{-1} + \frac{1}{2}\tilde{\Omega}_k^{-1}U_k'U_k\tilde{\Omega}_k^{-1} = 0 \\ \text{(ii)} \quad \frac{\partial \mathcal{L}}{\partial \varphi_k} &= z_k'(\tilde{\Omega}_k^{-1} \otimes I)(y_k - z_k\varphi_k) = 0\end{aligned}$$

ce qui donne les deux résultats suivants :

$$\tilde{\Omega}_k^* = \frac{1}{T}U_k'U_k \quad (20)$$

$$\varphi_k^* = [z_k'(\tilde{\Omega}_k^{-1} \otimes I)z_k]^{-1}z_k'(\tilde{\Omega}_k^{-1} \otimes I)y_k \quad (21)$$

¹⁰ Si R est diagonale et $\alpha = 0$, nous obtenons un modèle *SUR* avec auto-corrélation des erreurs intra-individuelles.

La matrice de transformation P tel que : $\Omega_u^{-1} = P'(\sum^{-1} \otimes I_T)P$ permet d'obtenir les estimateurs GLS-SUR de θ à partir du modèle (17) :

$$\hat{\theta} = [Z^{*'}(\sum^{-1} \otimes I)Z^*]^{-1}Z^{*'}(\sum^{-1} \otimes I)y^*$$

où $Z^* = PZ$, $y^* = Py$ et $\epsilon = Pu$.

Les estimateurs du maximum de vraisemblance de φ_k et $\tilde{\Omega}_k$ sont donnés par la solution simultanée des équations (20) et (21). Le système est non-linéaire et sans solution analytique, mais la structure en “zig-zag” des deux équations suggère une procédure numérique de solution simple et efficace. Elle comporte les pas suivants :

- Pas 1 : condition initiale $\tilde{\Omega}_k = I$;
- Pas 2 : on calcule φ_k par (21) et une nouvelle valeur de $\tilde{\Omega}_k$ par (20) ;
- Pas 3 : on retourne au pas 2 jusqu’à convergence numérique.

6.2 Estimation et matrice de transformation appropriée

Au vu des hypothèses du modèle et lorsque R est diagonale, la matrice de transformation P d’ordre $NT \times NT$, triangulaire inférieure $P = (P_{ij})$ pour $i, j = 1, \dots, N$, satisfait la condition de base $P\Omega_u P = \sum_\epsilon \otimes I$ avec :

$$P_{ii} = \begin{bmatrix} \alpha_{ii} & 0 & & 0 \\ -\rho_{ii} & 1 & & \\ & & \ddots & \ddots \\ 0 & & & -\rho_{ii} & 1 \end{bmatrix} \quad \text{et} \quad P_{ij} = \begin{bmatrix} \alpha_{ij} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \quad i \neq j \quad i, j = 1, \dots, N$$

La détermination de la matrice P permet de réaliser la transformation des variables du modèle. En considérant le cas où R est diagonale et α non nul :

$$\begin{aligned} y_{i0}^* &= \sum_{j=1}^i \alpha_{ij} y_{j0}, & y_{i1}^* &= \sum_{j=1}^i \alpha_{ij} y_{j1} & \text{et} & x_{i1}^* = \sum_{j=1}^i \alpha_{ij} x_{j1} & \text{pour } t = 1 \\ y_{it-1}^* &= y_{it-1} - \rho_{ii} y_{it-2}, & y_{it}^* &= y_{it} - \rho_{ii} y_{it-1}, & \text{et} & x_{it}^* = x_{it} - \rho_{ii} x_{it-1} & \text{pour } t > 1 \end{aligned}$$

Les paramètres α_{ij} sont calculés à partir de la matrice A , dont les éléments sont choisis tels que :

$$\sum = AV_0 A' \quad (22)$$

où la matrice $A = (\alpha_{ij}) \quad i, j = 1, \dots, N$ est une matrice de transformation d’ordre $N \times N$ triangulaire inférieure.

Le calcul des paramètres α_{ij} s’effectue à partir de la relation (22). Comme les matrices \sum et V_0 sont définies positives, il existe une matrice H et une matrice B qui sont triangulaires inférieures avec des éléments positifs sur leurs diagonales. Les éléments de A peuvent être trouvés à l’aide de la décomposition de Cholesky : $\sum = HH'$ et $V_0 = BB'$, on en déduit : $H = AB$ d’où $A = HB^{-1}$.

Le modèle transformé global est déterminé comme suit :

$$y^* = y_R^* \alpha + Z^* \theta + \epsilon \quad \text{avec} \quad E(\epsilon \epsilon') = \sum \otimes I_T \quad (23)$$

on n’est pas en présence d’un modèle *SUR* qui permet d’appliquer la méthode *GLS* même sur des variables transformées, en raison de la corrélation intra-individuelle entre y_{it-1} et ϵ_{it} .

6.3 Estimation réalisable et développement de Taylor

Pour définir de manière réalisable le modèle transformé, il faut réaliser un développement de Taylor au premier ordre autour d'une valeur initiale estimée de ρ_{ii} . Pour cela, il faut briser efficacement la corrélation entre y_{it-1} et u_{it} avec la méthode des variables instrumentales.

Soit un vecteur d'erreurs aléatoires $u = \text{vec}(u_1, \dots, u_i, \dots, u_N)$ où u_i est un $VAR(1)$ d'ordre $T \times 1$:

$$\text{si } R \text{ est diagonale } E(u_i' y_{jR}) \neq 0 \text{ pour } i = j = 1, \dots, N$$

En collectionnant toutes les variables explicatives prédéterminées avec les instruments dans une matrice w_i d'ordre $T \times r_i$ (nous supposons $r_i = r$). La condition nécessaire est qu'au moins il faut autant d'instruments que de variables endogènes explicatives : $r \geq L + 1$.

Habituellement, on retient Z_{iR} comme instrument¹¹ pour y_{iR} :

$$(Z_{iR} \quad Z_i) := w_i \quad \text{avec } \text{plim} \frac{1}{T} Z_{iR}' u_i = 0$$

le modèle global (17) s'écrit :

$$y = \begin{pmatrix} y_R & Z \end{pmatrix} \begin{pmatrix} \alpha \\ \theta \end{pmatrix} + u := z\varphi + u \quad (24)$$

où z est d'ordre $T \times N(L + 1)$ et z_i d'ordre $T \times (L + 1)$. Soit w la matrice d'instruments¹² :

$$w = \text{diag}(w_1, \dots, w_i, \dots, w_N)$$

par le principe des *GLS*, nous obtenons :

$$\hat{\varphi}_{i(IV)} = (z_i' w_i (w_i' M_i w_i) w_i' z_i)^{-1} z_i' w_i (w_i' M_i w_i)^{-1} w_i' y_i$$

puisque $r = L + 1$ ¹³, alors :

$$\hat{\varphi}_{i(VI)} = (w_i' z_i)^{-1} w_i' y_i \quad \square$$

¹¹ ξ_{iR} est un instrument ssi $E(y_{jR}' \xi_{iR}) \neq 0$ et $E(u_{iR}' \xi_{iR}) = 0$.

¹² Par la convergence quadratique en probabilité, on montre que $\text{plim} \frac{1}{T} w_i' u_i \rightarrow 0$ pour tout i :

$$V\left(\frac{1}{T} w_i' u_i\right) = \frac{1}{T^2} w_i' V(u_i) w_i = \frac{1}{T^2} w_i' \Omega_i w_i \quad \text{avec } \Omega_i = \frac{\sigma_{ii}}{1 - \rho_{ii}^2} M(1, \rho_{ii})$$

où $M(1, \rho_{ii})$ contient 1 sur la diagonale principale et des puissances de ρ_{ii} par ailleurs ;

$$\text{plim} \frac{1}{T^2} w_i' \Omega_i w_i = \left(\text{plim} \frac{1}{T} \frac{\sigma_{ii}}{1 - \rho_{ii}^2} \right) \text{plim} \frac{1}{T} w_i' M(1, \rho_{ii}) w_i = 0 \times Q_{w_i} = 0.$$

¹³ Lorsque $r = L + 1$, la méthode des IV est équivalente exactement à la méthode 2SLS :

A cette étape, utilisant les résidus \tilde{u}_i pour fournir des coefficients de corrélation $\tilde{\rho}_{ii}$:

$$\begin{aligned} \tilde{u}_i &= y_i - \hat{y}_i & \text{où } \hat{y}_i &= z_i \hat{\varphi}_i(VI) \\ \tilde{u}_{iR} &= y_{iR} - z_{iR} \hat{\varphi}_i(VI) & \text{pour } i &= 1, 2, \dots, N \end{aligned}$$

ainsi,

$$\tilde{R} = (\tilde{U}'_R \tilde{U}_R)^{-1} \tilde{U}'_R \tilde{U} \quad \text{ou} \quad \tilde{\rho}_i = (\tilde{u}'_{iR} \tilde{u}_{iR})^{-1} \tilde{u}'_{iR} \tilde{u}_i, \quad \text{car} \quad U = U_R R + \Upsilon$$

où \tilde{U}_R et Υ sont d'ordre $T \times N$, ce qui donne :

$$\tilde{U}_R = (\tilde{u}_{1R}, \dots, \tilde{u}_{iR}, \dots, \tilde{u}_{NR}) \quad \Upsilon = (\epsilon_1, \dots, \epsilon_i, \dots, \epsilon_N) \quad \text{avec} \quad E(\Upsilon \Upsilon') = \sum_{i=1}^N \sigma_{ii}^2 I_T.$$

En partant du modèle global, nous pouvons le réécrire comme suit :

$$y - z\varphi = (R' \otimes I)u_R + \epsilon = (R' \otimes I)(y_R - z_R\varphi) + \epsilon \quad (26)$$

ou encore :

$$y - (R' \otimes I)y_R = (z - (R' \otimes I)z_R)\varphi + \epsilon \quad (27)$$

Implicitement, ce modèle n'utilise pas la première observation. Il serait équivalent au modèle transformé global, si nous omettons de ce dernier la première observation.

Pour démarrer le processus d'estimation, il faut réaliser un développement de Taylor d'ordre 1 autour du paramètre $\tilde{\rho}_i$ pour le modèle individuel suivant :

$$\begin{aligned} y_i - \rho_i y_{iR} &= (z_i - \rho_i z_{iR})\varphi_i + \epsilon_i \\ y_i^* &= z_i^* \varphi_i + \epsilon_i \end{aligned}$$

par définition, le développement de Taylor d'ordre 1 autour de $\tilde{\rho}_i$ donne :

$$\begin{aligned} y_i^*(\rho_i) &= y_i^*(\tilde{\rho}_i) - \Delta \rho_i y_{iR} \\ z_i^*(\rho_i) \varphi_i &= z_i^*(\tilde{\rho}_i) \varphi_i - \Delta \rho_i z_{iR} \hat{\varphi}_i(IV) \\ \epsilon_i(\rho_i) &= \epsilon_i(\tilde{\rho}_i) - \Delta \rho_i u_{iR} \end{aligned}$$

$$\begin{cases} y_i = z_i \varphi_i + u_i & E(z' u_i) \neq 0 & z_i = (y_{iR} \quad Z_i) & w_i := (Z_{iR} \quad Z_i) \\ z_i = w_i \delta_i + e_i & E(w_i' e_i) = 0 \\ u_i = \rho_{ii} u_{iR} + \epsilon_i \end{cases}$$

nous avons :

$$\begin{aligned} \hat{z} &= w \hat{\delta} & \text{avec} & \quad \hat{\delta} = (w' w)^{-1} w' z & \text{et} & \quad \hat{z} := (\hat{y}_R \quad Z) \\ \hat{e} &= z - \hat{z} = M_w z & \hat{z}' \hat{e} &= \hat{\delta}' w' \hat{e} = 0 & \hat{z}' \hat{z} &= \hat{z}' (\hat{z} + \hat{e}) = \hat{z}' z \end{aligned}$$

d'où :

$$\hat{\varphi}_{2SLS} = (\hat{z}' z)^{-1} \hat{z}' y = (\hat{\delta}' w' z)^{-1} \hat{\delta}' w' y = (z' w (w' w)^{-1} w' z)^{-1} z' w (w' w)^{-1} w' y \quad (25)$$

si de plus $r = L + 1$, alors : $\hat{\varphi}_{2SLS} = (w' z)^{-1} w' y$.

par conséquent, le modèle individuel se réécrit par :

$$\begin{aligned} y_i^*(\tilde{\rho}_i) &= z_i^*(\tilde{\rho}_i)\varphi_i + \Delta\rho_i(y_{iR} - z_{iR}\widehat{\varphi}_{i(VI)} - u_{iR}) + \epsilon_i(\tilde{\rho}_i) \\ &= z_i^*(\tilde{\rho}_i)\varphi_i + \Delta\rho_i(\tilde{u}_{iR} - u_{iR}) + \Delta\rho_i u_{iR} + \epsilon_i \\ &= z_i^*(\tilde{\rho}_i)\varphi_i + \Delta\rho_i\tilde{u}_{iR} + \epsilon_i \end{aligned}$$

pour tous les individus, nous obtenons :

$$y - (\tilde{R}' \otimes I)y_R = (z - (\tilde{R}' \otimes I)z_R)\varphi + \text{diag}(\tilde{u}_{1R}, \dots, \tilde{u}_{NR})\Delta\rho + \epsilon \quad (28)$$

$$\text{avec}(\Delta\rho)' = (\Delta\rho_1, \dots, \Delta\rho_N) \quad \text{et} \quad \Delta\rho_i = \rho_i - \tilde{\rho}_i$$

Soit $h := [z - (\tilde{R}' \otimes I)z_R \quad \text{diag}(\tilde{u}_{1R}, \dots, \tilde{u}_{NR})]$, le modèle (13) devient :

$$y - (\tilde{R}' \otimes I)y_R = h\theta^* + \epsilon \quad \text{avec} \quad \theta^* = \begin{pmatrix} \varphi \\ \Delta\rho \end{pmatrix} \quad \text{et} \quad E(\epsilon\epsilon') = \sum_{\epsilon} \otimes I_{T-1} \quad (29)$$

Pour trouver des estimateurs efficaces réalisables pour φ et $\Delta\rho$, il faut estimer les variances-covariances de l'échantillon de la matrice \sum_{ϵ} . Elles peuvent être formées en se servant des résidus \tilde{u}_i et des coefficients de corrélation estimés $\tilde{\rho}_i$:

$$\tilde{\sigma}_{ij} = (T^{-1} - L_i^{-1})^{-\frac{1}{2}}(T^{-1} - L_j^{-1})^{-\frac{1}{2}}\tilde{\epsilon}_i\tilde{\epsilon}_j \quad \text{pour } i, j = 1, \dots, N$$

où $\tilde{\epsilon}_i = \tilde{u}_i - \tilde{\rho}_i \tilde{u}_{iR}$ et $L_i + 1$ représente la taille des variables explicatives de z_i . Ainsi, nous obtenons :

$$\hat{\theta}^* = \begin{pmatrix} \widehat{\varphi} \\ \widehat{\rho} - \tilde{\rho} \end{pmatrix} = (h'(\widetilde{\sum}^{-1} \otimes I_{T-1})h)^{-1}h'(\widetilde{\sum}^{-1} \otimes I_{T-1})(y - (\tilde{R}' \otimes I_{T-1})y_R) \quad (30)$$

L'estimation finale est donnée par :

$$\hat{\theta} = \begin{pmatrix} \widehat{\varphi} \\ \widehat{\rho} \end{pmatrix} = \hat{\theta}^* + \begin{pmatrix} 0 \\ \tilde{\rho} \end{pmatrix} \quad \square \quad (31)$$

Sous certaines conditions de régularité, la matrice des variances asymptotiques est la même que l'inverse de la matrice d'information associée à la méthode du maximum de vraisemblance. L'estimateur $\hat{\theta}$ ignore N observations, une par individu, il pourrait être moins efficace que $\hat{\theta}_0$ qui tiendrait compte de la première observation. La matrice des variances-covariances de l'estimateur $\hat{\theta}$ est donnée par :

$$\sum_{\hat{\theta}} = (h'(\widetilde{\sum}^{-1} \otimes I_{T-1})h)^{-1}$$

Les deux estimateurs $\hat{\theta}$ et $\hat{\theta}_0$ doivent avoir les mêmes propriétés asymptotiques et que dans des échantillons de taille finie $\hat{\theta}_0$ n'est pas nécessairement plus efficace que $\hat{\theta}$.

En ignorant la première observation dans chaque équation, la procédure de transformation est grandement simplifiée, car P_{ij}^0 devient une matrice nulle d'ordre $(T-1) \times T$, et donc la matrice de transformation P^0 devient diagonale d'ordre $N(T-1) \times NT$ sans dépendre des paramètres α_{ii} , éléments de la diagonale de la matrice A .

En tenant compte de toutes les observations, il faut intégrer dans la procédure d'estimation l'équation individuelle correspondante à la première observation :

$$\sum_{j=1}^i \alpha_{ij} y_{j1} = \sum_{j=1}^i \alpha_{ij} y_{j0} a_j + \sum_{r=1}^{L_i} \sum_{j=1}^i \alpha_{ij} x_{jr1} \beta_{jr} + \sum_{j=1}^i \alpha_{ij} u_{j1}$$

avec

$$E\left(\sum_{j=1}^i \alpha_{ij} u_{j1}\right)^2 = \sigma_{ij}, \quad \text{car} \quad E(u_1^* u_1^{*'}) = AE(u_1 u_1')A' = AV_0 A' = \sum \epsilon$$

La décomposition de Cholesky des matrices $\widetilde{\Sigma}$ et \widetilde{V}_0 permet de déduire des estimateurs pour $\widetilde{\alpha}_{ij}$:

$$\widetilde{A} = \widetilde{H} \widetilde{B}^{-1} \quad \text{où} \quad \widetilde{H} \widetilde{H}' = \widetilde{\Sigma} \quad \text{et} \quad \widetilde{B} \widetilde{B}' = \widetilde{V}_0$$

avec

$$\widetilde{\Sigma} = (\widetilde{\sigma}_{ij}) \quad \text{et} \quad \widetilde{V}_0 = \frac{\widetilde{\sigma}_{ij}}{(1 - \widetilde{\rho}_i \widetilde{\rho}_j)} \quad \widetilde{A} = (\widetilde{\alpha}_{ij})_{j \leq i}$$

En dépit du fait que A peut être obtenu numériquement d'une façon relativement facile, il est clair que la transformation de la première observation pour chaque équation individuelle est une procédure lourde.

L'estimation de \widetilde{A} permet d'obtenir entièrement la matrice de transformation estimée \widetilde{P} , après transformation appropriée des premières équations, et d'en déduire $\widehat{\theta}_0$.

Références

- [Bal76] P. Balestra. La dérivation matricielle (techniques et résultats pour économistes). Technical report, IME, dijon, 1976.
- [BG94] P. Balestra and H. Ghassan. Modèles récursifs à double indice : théorie et application. Technical report, LATEC-IME. Université de Bourgogne, 1994.
- [FHW80] W.A. Fuller, M. Hidiroglou, and J.H.K. Wang. Estimation of seemingly unrelated regressions with lagged dependent variables and autocorrelation errors. *Journal of statistical computation and simulation*, 10, 1980.
- [Ham94] J.D. Hamilton. *Time series analysis*. Princeton, New Jersey, 1994.
- [Hat76] M. Hatanaka. Several efficient two-step estimators for the dynamic simultaneous equations model with autoregressive disturbances. *Journal of econometrics*, IV, May 1976.
- [IG84] M.D. Intriligator and Z. Griliches. *Handbook of econometrics*. Intriligator, North Holland, 1984.

- [JGR⁺85] G. Judge, W.E. Griffiths, Carter Hill R., Lutkepohl H., and Lee T. Ch. *The theory and practice of econometrics*. John Wiley and Sons, New-York, 1985.
- [Mae80] A. Maeshiro. New evidence on the small properties of estimators of sur models with autocorrelated disturbances. *Journal of econometrics*, 12, 1980.
- [Spe79] D.E. Spencer. Estimation of a dynamic system of seemingly unrelated regressions with autoregressive disturbances. *Journal of econometrics*, 10, 1979.
- [Tur00] D. Turkington. Generalized vec operators and the seemingly unrelated regression equations model with vector correlated disturbances. *Journal of econometrics*, 99, 2000.