



HAL
open science

Estimateurs récursifs 1

Zaka Ratsimalahelo

► **To cite this version:**

Zaka Ratsimalahelo. Estimateurs récursifs 1. [Rapport de recherche] Institut de mathématiques économiques (IME). 1991, 27 p., ref. bib. : 10 ref. hal-01534418

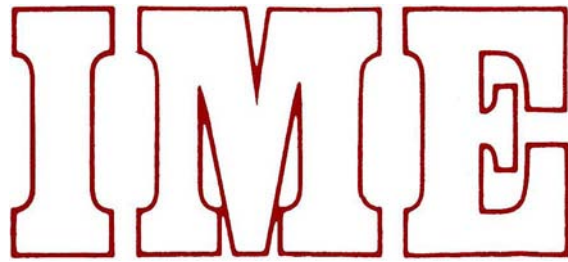
HAL Id: hal-01534418

<https://hal.science/hal-01534418>

Submitted on 7 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



INSTITUT DE MATHÉMATIQUES ÉCONOMIQUES

LATEC C.N.R.S. URA 342

DOCUMENT de TRAVAIL



UNIVERSITE DE BOURGOGNE

FACULTE DE SCIENCE ECONOMIQUE ET DE GESTION

4, boulevard Gabriel - 21000 DIJON - Tél. 80 39 54 30 - Fax 80 39 56 48

ISSN : 0292-2002

9106

ESTIMATEURS RECURSIFS 1

Zaka RATSIMALAHELO*

juin 1991

** Institut de Mathématiques Economiques*

ESTIMATEURS RECURSIFS 1

RESUME

Dans cet article, nous présentons d'abord l'estimation récursive des paramètres d'un modèle de régression et établissons la relation entre la matrice de projection et la matrice de gain. Deux cas sont envisagés : le cas de paramètres constants et le cas de paramètres obéissant à une loi d'évolution. Ensuite nous donnons une généralisation des formulations récursives pour les moindres carrés doubles.

ABSTRACT

In this paper, we present first the recursive estimation of parameters of linear regression models and we show the link between a matrix of projection and a matrix of gain. Two cases are examined : constant parameters and parameters changing over time. In the second part the recursion formulae for the two stage least squares are given.

MOTS-CLES: estimation récursive, les moindres carrés ordinaires récursifs, coefficients aléatoires, Filtre de Kalman, les moindres carrés doubles récursifs.

KEYWORDS: recursive estimation, recursive least squares, time varying parameters, Kalman filter, two stage least squares.

Introduction.

Nous assistons depuis quelques années à un intérêt marqué pour l'utilisation des méthodes d'estimation récursives comme les moindres carrés récursifs, le filtre de Kalman et ses variantes et ceci dans des domaines très variés : économétrie, systèmes utilisés par les ingénieurs, etc

Il faut souligner que la procédure récursive présente en économétrie un certain avantage tant au niveau théorique que pratique.

Sur le plan théorique, on a utilisé par exemple les résidus récursifs pour développer le test de changement de structure (Brown, Durbin et Evans, 1975), le test d'autocorrélation (Harvey et Phillips, 1974), la réduction du biais; le Jackknife est basé sur la structure récursive de l'estimateur.

Quant au niveau pratique, la formulation récurrente offre une structure efficiente de calcul parce qu'elle ne nécessite pas d'inversion supplémentaire de matrice.

Les développements les plus récents des méthodes récursives ou méthodes d'estimation en ligne des paramètres concernent davantage l'identification des systèmes, Ljung (1985), Young (1985).

Nous verrons dans un premier temps les estimateurs récurrents des paramètres constants d'un modèle de régression, nous montrerons le lien existant entre la matrice de gain et la matrice de projection. Nous présenterons ensuite les estimateurs récurrents des paramètres variant dans le temps et quelques modèles particuliers.

Dans la deuxième partie, nous donnerons deux formulations récursives des moindres carrés doubles : l'une nécessite pour son calcul l'inversion d'une matrice d'ordre deux, le calcul de la seconde s'effectue en deux étapes sans inversion de matrice.

I- Les moindres carrés récurrents.

1.1. Paramètres constants.

Soit le modèle de régression classique

$$Y_n = X_n \beta + \varepsilon_n \quad (1)$$

où Y_n : vecteur des n observations des variables endogènes,
 X_n : matrice non stochastique ($n \times k$) des variables exogènes,
 β : vecteur des paramètres inconnus de \mathbb{R}^k ,
 ε_n : vecteur d'erreurs d'espérance nulle et de variance $\sigma^2 I_n$.

L'indice n est utilisé pour désigner la taille de l'échantillon.
Pour une observation individuelle, la i -ème, le modèle s'écrit :

$$y_i = x_i' \beta + \varepsilon_i$$

y_i étant un scalaire et x_i' un vecteur ligne à k composantes.
L'estimateur des moindres carrés ordinaires (m.c.o.) pour un échantillon de taille $n+1$ est défini par :

$$\hat{\beta}_{n+1} = (X_{n+1}' X_{n+1})^{-1} X_{n+1}' Y_{n+1} \quad (2)$$

Le but est d'exprimer cet estimateur en fonction de l'estimateur pour un échantillon de taille n .

En écrivant

$$X_{n+1} = \begin{bmatrix} X_n \\ x_{n+1}' \end{bmatrix} \quad \text{et} \quad Y_{n+1} = \begin{bmatrix} Y_n \\ y_{n+1} \end{bmatrix}$$

l'estimateur donné en (2) peut se mettre sous la forme suivante :



$$\hat{\beta}_{n+1} = \left[X'_n X_n + x_{n+1} x'_{n+1} \right]^{-1} \left[X'_n Y_n + x_{n+1} y_{n+1} \right] \quad (3)$$

En appliquant le lemme d'inversion matricielle A1¹ (voir Annexe) avec $A = X'_n X_n$; $C = x'_{n+1}$; $D = x_{n+1}$ et $B = 1$, il vient

$$(X'_{n+1} X_{n+1})^{-1} = (X'_n X_n)^{-1} - K_n x'_{n+1} (X'_n X_n)^{-1} \quad (4)$$

avec
$$K_n = (X'_n X_n)^{-1} x_{n+1} [1 + x'_{n+1} (X'_n X_n)^{-1} x_{n+1}]^{-1}$$

qui est la matrice de gain. Elle peut également être exprimée par :

$$K_n = (X'_{n+1} X_{n+1})^{-1} x_{n+1}$$

d'après le lemme d'inversion matricielle A2.

En utilisant la relation (4), on obtient l'estimateur récursif

$$\hat{\beta}_{n+1} = \hat{\beta}_n + K_n (y_{n+1} - x'_{n+1} \hat{\beta}_n) \quad (5)$$

où $y_{n+1} - x'_{n+1} \hat{\beta}_n$ représente l'erreur de prévision d'espérance nulle et de variance $\sigma^2(1 + x'_{n+1} (X'_n X_n)^{-1} x_{n+1})$.

Nous avons ainsi une formulation récurrente sans inversion supplémentaire de matrice car $x'_{n+1} (X'_n X_n)^{-1} x_{n+1}$ est un scalaire et $(X'_n X_n)^{-1}$ est connu à l'étape précédente.

Au vu de ces deux formulations, l'estimateur des m.c.o. (2) nécessite pour son calcul toutes les informations passées, tandis que pour l'estimateur récursif (5), les informations passées sont résumées dans l'état $\hat{\beta}_n$ et seules les données pour $n+1$ permettent la mise à jour de l'estimateur (en plus de la connaissance de $(X'_n X_n)^{-1}$).

¹voir BALESTRA (1978) pour des formulations plus générales et diverses applications économiques.

Interprétation de la matrice de gain.

La matrice de gain est le rapport de la variance de $\hat{\beta}_n$ sur la variance de l'erreur de la prévision a priori ($V(y_{n+1} - x_{n+1}'\hat{\beta}_n)$), x_{n+1} qui est la nouvelle observation des variables exogènes jouant le rôle de facteur de proportionnalité.

Ces équations correspondent à la prise en compte d'une information supplémentaire à l'instant $n+1$. On peut en revanche établir la procédure inverse c'est-à-dire la suppression de cette information supplémentaire, les équations associées sont :

$$\left\{ \begin{array}{l} \text{la matrice de gain} \\ \quad \tilde{K}_{n+1} = (X'_{n+1} X_{n+1})^{-1} x_{n+1} [x'_{n+1} (X'_{n+1} X_{n+1})^{-1} x_{n+1} - 1]^{-1} \quad (6.a) \\ \quad = - (X' X_n)^{-1} x_{n+1} \quad (6.b) \\ \text{la matrice de variance-covariance} \\ \quad \sigma^2 (X' X_n)^{-1} = \sigma^2 (X'_{n+1} X_{n+1})^{-1} - \sigma^2 \tilde{K}_{n+1} x'_{n+1} (X'_{n+1} X_{n+1})^{-1} \quad (6.c) \\ \text{l' estimateur récurrent} \\ \quad \hat{\beta}_n = \hat{\beta}_{n+1} + \tilde{K}_{n+1} (y_{n+1} - x'_{n+1} \hat{\beta}_{n+1}) \quad (6.d) \\ \text{où } y_{n+1} - x'_{n+1} \hat{\beta}_{n+1} \text{ est l'erreur de prédiction a posteriori} \end{array} \right.$$

La preuve de ces relations est immédiate, en utilisant le lemme A1 d'inversion matricielle :

$$\begin{aligned} (X' X_n)^{-1} &= \left[X'_{n+1} X_{n+1} - x_{n+1} x'_{n+1} \right]^{-1} \\ &= (X'_{n+1} X_{n+1})^{-1} - \tilde{K}_{n+1} x'_{n+1} (X'_{n+1} X_{n+1})^{-1} \end{aligned}$$

avec $\tilde{K}_{n+1} = (X'_{n+1} X_{n+1})^{-1} x_{n+1} [x'_{n+1} (X'_{n+1} X_{n+1})^{-1} x_{n+1} - 1]^{-1}$

L'estimateur de β_n est défini par :

$$\hat{\beta}_n = \left[X'_{n+1} X_{n+1} - x_{n+1} x'_{n+1} \right]^{-1} \left[X'_{n+1} Y_{n+1} - x'_{n+1} y_{n+1} \right].$$

D'après la relation ci-dessus et après simplification, on obtient :

$$\hat{\beta}_n = \hat{\beta}_{n+1} + \tilde{K}_{n+1} (y_{n+1} - x'_{n+1} \hat{\beta}_{n+1}) \quad \blacksquare$$

Relation entre les deux matrices de gain.

D'après la relation (6.a), le terme $(X'_{n+1} X_{n+1})^{-1} x_{n+1}$ n'est autre que K_n tandis que le reste de l'expression s'écrit :

$$\begin{aligned} x'_{n+1} (X'_{n+1} X_{n+1})^{-1} x_{n+1} - 1 &= x'_{n+1} (X'_n X_n)^{-1} x_{n+1} [1 + x'_{n+1} (X'_n X_n)^{-1} x_{n+1}]^{-1} - 1 \\ &= - [1 + x'_{n+1} (X'_n X_n)^{-1} x_{n+1}]^{-1} \end{aligned}$$

On obtient alors une relation de proportionnalité entre les deux matrices de gain

$$\tilde{K}_{n+1} = - K_n d$$

avec $d = 1 + x'_{n+1} (X'_n X_n)^{-1} x_{n+1}$.

On peut montrer également la relation entre l'erreur de prédiction a priori et l'erreur de prédiction a posteriori.

Définissons l'erreur de prédiction a priori par :

$$\tilde{e}_t^{\circ} = y_{n+1} - x'_{n+1} \hat{\beta}_n$$

et l'erreur de prédiction a posteriori ou erreur d'estimation par :

$$\begin{aligned} \tilde{e}_t &= y_{n+1} - x'_{n+1} \hat{\beta}_{n+1} \\ &= \tilde{e}_t^{\circ} - x'_{n+1} (\hat{\beta}_{n+1} - \hat{\beta}_n) \end{aligned}$$

En utilisant (5) pour évaluer $\hat{\beta}_{n+1} - \hat{\beta}_n$, on obtient

$$\tilde{e}_t = \tilde{e}_t^\circ - x'_{n+1} (X'_n X_n)^{-1} x_{n+1} [1 + x'_{n+1} (X'_n X_n)^{-1} x_{n+1}]^{-1} \tilde{e}_t^\circ.$$

Après simplification, nous avons la relation fondamentale liant l'erreur a priori et l'erreur a posteriori :

$$\tilde{e}_t = \tilde{e}_t^\circ / (1 + x'_{n+1} (X'_n X_n)^{-1} x_{n+1}).$$

Ce qui nous permet de réécrire la relation (5) sous la forme :

$$\hat{\beta}_{n+1} = \hat{\beta}_n + (X'_n X_n)^{-1} x_{n+1} \tilde{e}_t.$$

Le vecteur d'erreurs calculées est défini par :

$$\hat{\varepsilon}_{n+1} = Y_{n+1} - X_{n+1} \hat{\beta}_n = M_{n+1} Y_{n+1} = M_{n+1} \varepsilon_{n+1}$$

où
$$M_{n+1} = I_{n+1} - X_{n+1} (X'_{n+1} X_{n+1})^{-1} X'_{n+1}$$

est la matrice de projection.

En tenant compte de la partition de la matrice X_{n+1} et de l'inverse de la matrice $(X'_{n+1} X_{n+1})$, nous pouvons écrire successivement la matrice de projection comme suit :

$$M_{n+1} = \begin{bmatrix} I_n - X_n (X'_n X_n)^{-1} X'_n & -X_n (X'_{n+1} X_{n+1})^{-1} x_{n+1} \\ -x'_{n+1} (X'_{n+1} X_{n+1})^{-1} X'_n & 1 - x'_{n+1} (X'_{n+1} X_{n+1})^{-1} x_{n+1} \end{bmatrix}$$

$$= \begin{bmatrix} I_n - X_n (X'_n X_n)^{-1} X'_n + E & -X_n (X'_{n+1} X_{n+1})^{-1} x_{n+1} \\ -x'_{n+1} (X'_{n+1} X_{n+1})^{-1} X'_n & 1 - x'_{n+1} (X'_{n+1} X_{n+1})^{-1} x_{n+1} \end{bmatrix} \quad (7)$$

avec
$$E = X_n (X'_n X_n)^{-1} x_{n+1} x'_{n+1} (X'_{n+1} X_{n+1})^{-1} X'_n / 1 + x'_{n+1} (X'_{n+1} X_{n+1})^{-1} x_{n+1}$$

Relation entre la matrice de gain et la matrice de projection.

En utilisant les expressions de la matrice de gain et en rappelant que

$$1 - x'_{n+1} (X'_{n+1} X_{n+1})^{-1} x_{n+1} = [1 + x'_{n+1} (X'_n X_n)^{-1} x_{n+1}]^{-1} = d^{-1}$$

la matrice de projection devient

$$M_{n+1} = \begin{bmatrix} M_n & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} X_n K_n K'_n X'_n d & - X_n K_n \\ - K'_n X'_n & d^{-1} \end{bmatrix} \quad (8)$$

Cette formulation nous indique le lien entre la matrice de projection et la matrice de gain.

L'équation (8) peut se mettre sous la forme suivante :

$$M_{n+1} = L M_n L' + d^{-1} \tilde{J}' \tilde{J} \quad (9)$$

avec $L' = [I_n \quad 0]$ et $\tilde{J} = [- K'_n X'_n d \quad 1]$.

La dernière expression (9) nous permet de calculer facilement la somme des carrés des erreurs.

Proposition 1.

La formulation récurrente de la somme des carrés des erreurs notée SS est :

$$SS_{n+1} = SS_n + e_n^2$$

où $e_n = (y_{n+1} - x_{n+1} \hat{\beta}_n) / \sqrt{d}$

avec $d = 1 + x'_{n+1} (X'_n X_n)^{-1} x_{n+1}$

e_n est l'erreur de prédiction (innovation) d'espérance nulle et de variance σ^2 .

Preuve.

La somme des carrés des erreurs pour les $n+1$ observations est définie par :

$$\begin{aligned}
 SS_{n+1} &= Y'_{n+1} M_{n+1} Y_{n+1} \\
 SS_{n+1} &= \begin{bmatrix} Y_n \\ y_{n+1} \end{bmatrix}' \begin{bmatrix} LM_n L' + d^{-1} \check{J}' \check{J} \end{bmatrix} \begin{bmatrix} Y_n \\ y_{n+1} \end{bmatrix} \\
 &= Y'_{n+1} M_{n+1} Y_{n+1} + d^{-1} (y_{n+1} - dK'_n X'_n Y_n)' (y_{n+1} - dK'_n X'_n Y_n)
 \end{aligned}$$

En remplaçant la matrice de gain par son expression

$$K'_n = d^{-1} x'_{n+1} (X'_n X_n)^{-1},$$

nous avons

$$SS_{n+1} = SS_n + d^{-1} (y_{n+1} - x'_{n+1} \hat{\beta}_n)^2 \quad \blacksquare$$

Les résidus récurrents e_i pour $i = k+1 \dots, n$ ont servi pour les tests statistiques tels que les CUSUM (sommes cumulées des résidus récurrents), les CUSMSQ (sommes cumulées des carrés des résidus récurrents) dues à Brown, Durbin et Evans (1975).

1.2 Paramètres variables dans le temps.

La généralisation du modèle précédent consiste à supposer que les paramètres sont eux-mêmes des variables aléatoires. Une structure générale des paramètres variant dans le temps est donnée par :

$$\beta_{n+1} = A\beta_n + \gamma_{n+1} + u_{n+1} \tag{10}$$

où A est la matrice de transition ($k \times k$),
 γ est le vecteur des paramètres fixes,
 u est le vecteur d'erreurs d'espérance nulle et de variance $\sigma_u^2 I$.

Dans ce qui suit, nous supposerons d'une part que les vecteurs aléatoires ε_t et u_s sont indépendants pour toutes valeurs de t et s et d'autre part que la matrice A , le vecteur γ et les variances-covariances σ_u^2 , σ^2 sont connus.

L'espérance conditionnelle des relations (10) et (1) sachant les observations à l'instant n est donnée respectivement par :

$$\tilde{\beta}_{n+1} = A\hat{\beta}_n + \gamma_{n+1} \quad (11)$$

$$\tilde{y}_{n+1} = x'_{n+1}\tilde{\beta}_{n+1} \quad (12)$$

où $\tilde{\beta}_{n+1}$ et \tilde{y}_{n+1} sont respectivement les espérances conditionnelles de β_{n+1} et y_{n+1} , étant donné les observations au temps n .

L'estimateur linéaire récursif est défini par :

$$\hat{\beta}_{n+1} = \tilde{\beta}_{n+1} + K_n(y_{n+1} - \tilde{y}_{n+1}) \quad (13)$$

où $\hat{\beta}_{n+1}$ est l'espérance conditionnelle de β_{n+1} sachant les observations à l'instant $n+1$; $y_{n+1} - \tilde{y}_{n+1}$ représente l'erreur de prévision d'horizon 1 faite à l'instant n .

Détermination de la matrice de gain.

Notons $v_{n+1} = \hat{\beta}_{n+1} - \beta_{n+1}$, l'erreur d'estimation au temps $n+1$. La matrice de gain minimise la variance conditionnelle de l'erreur d'estimation.

D'après les relations (1), (10), (11), (12) et (13), nous pouvons exprimer l'erreur d'estimation par :

$$\begin{aligned}
v_{n+1} &= (I - K_n x'_{n+1}) A \hat{\beta}_n + (K_n x'_{n+1} - I)(A \beta_n + u_{n+1}) + K_n \varepsilon_{n+1} \\
&= (I - K_n x'_{n+1}) A v_n - (I - K_n x'_{n+1}) u_{n+1} + K_n \varepsilon_{n+1} \quad (14)
\end{aligned}$$

Posons

$$\begin{aligned}
D_{n+1} &= E(v_{n+1} v'_{n+1}) \\
&= (I - K_n x'_{n+1}) A D_n A' (I - K_n x'_{n+1})' \\
&\quad + \sigma_u^2 (I - K_n x'_{n+1}) (I - K_n x'_{n+1})' + \sigma^2 K_n K_n'.
\end{aligned}$$

Le problème consiste à minimiser la variance conditionnelle de v_{n+1} , choisissons le critère de la minimisation de la trace de la matrice de variances-covariances D_{n+1} .

La condition nécessaire s'écrit :

$$\frac{\partial}{\partial K_n} (\text{tr } D_{n+1}) = 0$$

où tr désigne l'opérateur trace

$$\Rightarrow - (I - K_n x'_{n+1}) [A D_n A' + \sigma_u^2] x_{n+1} + \sigma^2 K_n = 0.$$

En résolvant par rapport à K_n , il vient

$$K_n = \left(A D_n A' + \sigma_u^2 \right) x_{n+1} \left[x'_{n+1} \left(A D_n A' + \sigma_u^2 \right) x_{n+1} + \sigma^2 \right]^{-1} \quad (15)$$

le gain est optimal parce que le Hessien

$$\frac{\partial^2}{\partial K_n^2} (\text{tr } D_{n+1}) = x'_n \left(A D_n A' + \sigma_u^2 \right) x_n + \sigma^2$$

est défini positif.

Interprétation de la matrice de gain.

C'est le rapport de la variance conditionnelle de β_{n+1} sur la variance conditionnelle de l'observation y_{n+1} ; x'_{n+1} joue le rôle de facteur de proportionnalité.

Quelques cas particuliers du modèle.

1) Le vecteur β évolue suivant une marche aléatoire.

Nous avons le modèle suivant :

$$Y_n = X_n \beta_n + \varepsilon_n$$

$$\beta_{n+1} = \beta_n + \gamma_{n+1} + u_{n+1}$$

le paramètre γ peut être interprété comme un trend.

En posant $A = I$ dans les relations (11) et (15), on obtient :

$$\hat{\beta}_{n+1} = \hat{\beta}_n + \gamma_{n+1} + K_n (y_{n+1} - \tilde{y}_{n+1}).$$

La matrice de gain devient

$$K_n = \left(D_n + \sigma^2 \right) x_{n+1} \left[x'_{n+1} \left(D_n + \sigma_u^2 \right) x_n + \sigma^2 \right]^{-1}$$

et la matrice de covariance s'écrit :

$$D_{n+1} = (I - K_n x'_{n+1}) D_n (I - K_n x'_{n+1})' + \sigma_u^2 (I - K_n x'_{n+1}) (I - K_n x'_{n+1})' \\ + \sigma^2 K_n K_n'$$

Dans le cas où γ_{n+1} est nul, l'estimateur récursif devient

$$\hat{\beta}_{n+1} = \hat{\beta}_n + K_n (y_{n+1} - \tilde{y}_{n+1}).$$

La matrice de gain et la matrice de covariance restent inchangées.

Ce cas particulier a été étudié plus précisément par Garbade (1977) et Chow (1981).

2) Le vecteur β est constant au cours du temps : $A = I$; $\gamma = 0$ et $\sigma^2 = 0$ (i.e $u = 0$ parce que sa moyenne est nulle).

De la relation (11), on a $\tilde{\beta}_{n+1} = \hat{\beta}_n$. Par conséquent on obtient l'estimateur des m.c.o récursifs (5).

3) Supposons que seuls les k_1 premiers paramètres évoluent dans le temps, nous avons alors

$$\beta_{in+1} = \beta_{in} \quad i = k_1+1, \dots, k$$

(les $k - k_1$ derniers paramètres sont fixes).

L'équation dynamique des paramètres devient :

$$\beta_{n+1} = \begin{bmatrix} A_1 & 0 \\ 0 & I \end{bmatrix} \beta_n + \begin{bmatrix} I_{k_1} \\ 0 \end{bmatrix} \gamma_{n+1} + \begin{bmatrix} I_{k_1} \\ 0 \end{bmatrix} u_{n+1}$$

où A_1 est une matrice $k_1 \times k_1$.

Nous venons de présenter les formulations récursives du modèle linéaire dans les cas où les paramètres sont d'une part constants et d'autre part variables dans le temps. Dans chaque cas, le rôle de la matrice de gain a été mis en évidence et son interprétation a été donnée. Considérons maintenant un cadre plus général : les moindres carrés doubles.

PHILLIPS G.D.A (1977) a généralisé la formulation récursive de l'estimateur d'un modèle de régression pour les moindres carrés doubles, en vue d'estimer les paramètres d'un modèle à équations simultanées. Néanmoins ses formulations semblent peu pratiques et difficiles à réaliser sur des programmes numériques, ainsi nous allons suggérer des formulations récursives beaucoup plus pratiques.

II- Estimateur récurrent des moindres carrés doubles.

L'estimateur des moindres carrés doubles est souvent présenté dans la littérature comme un estimateur à variables instrumentales optimal.

De plus, de nombreux modèles économétriques peuvent être estimés en utilisant le formalisme des variables instrumentales. Nous considérons dans ce cas le modèle suivant :

$$\begin{cases} Y_n = X_n \beta_n + \varepsilon_n \\ X_n = Z_n \gamma_n + U_n \end{cases}$$

avec $E(\varepsilon) = 0$, $V(\varepsilon) = \sigma^2 I$ et $E(\varepsilon / X) \neq 0$,

où Z est une matrice de n observations sur p variables instrumentales, γ_n est une matrice de coefficients inconnus, de dimension (p,k) et U une matrice de perturbations.

Cette formulation est assez générale, elle est particulièrement adaptée à l'étude des équations simultanées.

Les hypothèses.

- i) $\text{plim } \frac{1}{n} Z' \varepsilon = 0$
- ii) $\text{plim } \frac{1}{n} Z' U = 0$
- iii) $\text{plim } \frac{1}{n} Z' Z = Q_Z$ définie positive
- iv) γ de rang k .

Sous ces hypothèses l'estimateur des moindres carrés doubles de β défini par :

$$\hat{\beta}_n = (\hat{X}'_n \hat{X}_n)^{-1} \hat{X}'_n Y_n.$$

est convergent.

Pour un échantillon de taille $n+1$ nous avons :

$$\hat{X}'_{n+1} \hat{X}_{n+1} = \hat{\gamma}'_{n+1} Z'_{n+1} Z_{n+1} \hat{\gamma}_{n+1}$$

où $\hat{\gamma}_{n+1} = (Z'_{n+1} Z_{n+1})^{-1} X_{n+1}$ est une matrice $(p \times k)$ (16)

de même nous pouvons écrire

$$\hat{X}'_{n+1} Y_{n+1} = \hat{\gamma}'_{n+1} Z'_{n+1} Z_{n+1} \tilde{Y}_{n+1}$$

où $\tilde{Y}_{n+1} = (Z'_{n+1} Z_{n+1})^{-1} Z'_{n+1} Y_{n+1}$. (17)

Nous pouvons remarquer que les relations (16) et (17) ressemblent à la forme classique de résolution des problèmes par les moindres carrés. Elles nous permettent de formaliser la récursivité de l'estimateur

$$\hat{\beta}_{n+1} = \left(\hat{\gamma}'_{n+1} Z'_{n+1} Z_{n+1} \hat{\gamma}_{n+1} \right)^{-1} \hat{\gamma}'_{n+1} Z'_{n+1} Z_{n+1} \tilde{Y}_{n+1} .$$

Pour avoir une relation récurrente de l'estimateur ainsi défini, il faut d'abord formuler la récursivité de l'inverse de la matrice

$\left(\hat{\gamma}'_{n+1} Z'_{n+1} Z_{n+1} \hat{\gamma}_{n+1} \right)$ et ensuite compléter par le développement du terme $\hat{\gamma}'_{n+1} Z'_{n+1} Z_{n+1} \tilde{Y}_{n+1}$. Chaque étape de la procédure demande une démonstration assez longue et fastidieuse, il est donc nécessaire de donner dans un premier temps quelques propositions.

Proposition 2.

La formulation réursive de la matrice $(Z'Z_n)^{-1}$ est :

$$(Z'_{n+1} Z_{n+1})^{-1} = (Z'Z_n)^{-1} - B_n Z'_{n+1} (Z'Z_n)^{-1}. \quad (18.a)$$

Les relations récurrentes des deux transformations (16) et (17) sont respectivement :

$$\hat{\gamma}_{n+1} = \hat{\gamma}_n + B_n (x'_{n+1} - z'_{n+1} \hat{\gamma}_n) \quad (18. b)$$

$$\tilde{Y}_{n+1} = \tilde{Y}_n + B_n (y_{n+1} - z'_{n+1} \tilde{Y}_n) \quad (18. c)$$

$$\text{où } B_n = (Z'_n Z_n)^{-1} z_{n+1} (1 + z'_{n+1} (Z'_n Z_n)^{-1} z_{n+1})^{-1} \quad (18. d)$$

est la matrice de gain.

Preuve.

Définissons la matrice instrumentale Z_{n+1} par :

$$Z_{n+1} = \begin{bmatrix} Z_n \\ z'_{n+1} \end{bmatrix}$$

$$\Rightarrow Z'_{n+1} Z_{n+1} = Z'_n Z_n + z'_{n+1} z_{n+1}. \quad (19)$$

La démonstration de la relation (18.a) est immédiate, il suffit d'appliquer le résultat du lemme A1 d'inversion matricielle sur la relation (19).

Les relations (18.b) et (18.c) sont obtenues en utilisant la formulation récurrente des m.c.o.

Proposition 3.

La formulation récurrente de la matrice $P_{n+1} = \hat{\gamma}'_{n+1} (Z'_{n+1} Z_{n+1})^{-1} \hat{\gamma}_{n+1}$ s'écrit :

$$P_{n+1} = P_n + x_{n+1} x'_{n+1} - u_{n+1} u'_{n+1} \quad (20. a)$$

$$\text{avec } u_{n+1} = (x'_{n+1} - z'_{n+1} \hat{\gamma}_n) / \sqrt{b} \quad (20. b)$$

$$\text{où } b = (1 + z'_{n+1} (Z'_n Z_n)^{-1} z_{n+1}). \quad (20. c)$$

Preuve.

D'après la proposition 2, la matrice P_{n+1} s'écrit :

$$P_{n+1} = [\hat{\gamma}'_n + B_n(x'_{n+1} - z'_{n+1}\hat{\gamma}_n)]' [Z'_n Z_n + z_{n+1}z'_{n+1}] [\hat{\gamma}_n + B_n(x'_{n+1} - z'_{n+1}\hat{\gamma}_n)]$$

En remplaçant la matrice de gain B_n par son expression (18.d) et après quelques manipulations, on obtient :

$$\begin{aligned} P_{n+1} &= \hat{\gamma}'_n (Z'_n Z_n) \hat{\gamma}_n + \hat{\gamma}'_n z_{n+1} z'_{n+1} \hat{\gamma}_n \\ &+ (x'_{n+1} - z'_{n+1} \hat{\gamma}_n)' z'_{n+1} \hat{\gamma}_n + \hat{\gamma}'_n z'_n (x'_{n+1} - z'_{n+1} \hat{\gamma}_n) \\ &+ (x'_{n+1} - z'_{n+1} \hat{\gamma}_n)' (Z'_n Z_n)^{-1} z_{n+1} (x'_{n+1} - z'_{n+1} \hat{\gamma}_n) / b. \end{aligned}$$

En développant les 3^{ème} et 4^{ème} termes, il vient :

$$\begin{aligned} P_{n+1} &= \hat{\gamma}'_n (Z'_n Z_n) \hat{\gamma}_n - \hat{\gamma}'_n z_{n+1} z'_{n+1} \hat{\gamma}_n \\ &+ x_{n+1} z'_{n+1} \hat{\gamma}_n + \hat{\gamma}'_n z_{n+1} x'_{n+1} \\ &+ (x'_{n+1} - z'_{n+1} \hat{\gamma}_n)' z'_{n+1} (Z'_n Z_n)^{-1} z_{n+1} (x'_{n+1} - z'_{n+1} \hat{\gamma}_n) / b. \end{aligned}$$

En ajoutant et en soustrayant le terme $x_{n+1} x'_{n+1}$ au côté droit, nous avons :

$$\begin{aligned} P_{n+1} &= \hat{\gamma}'_n (Z'_n Z_n) \hat{\gamma}_n + x_{n+1} x'_{n+1} - (x'_{n+1} - z'_{n+1} \hat{\gamma}_n)' (x'_{n+1} - z'_{n+1} \hat{\gamma}_n) \\ &+ (x'_{n+1} - z'_{n+1} \hat{\gamma}_n)' z'_{n+1} (Z'_n Z_n)^{-1} z_{n+1} (x'_{n+1} - z'_{n+1} \hat{\gamma}_n) / b. \end{aligned}$$

Après simplification et en reprenant la définition de P_n , nous avons alors :

$$P_{n+1} = P_n + x_{n+1}x'_{n+1} - (x'_{n+1} - z'_{n+1}\hat{\gamma}_n)'(x'_{n+1} - z'_{n+1}\hat{\gamma}_n) / b$$

avec $b = 1 + z'_{n+1}(Z'_n Z_n)^{-1}z_{n+1}$ ■

A partir de cette dernière proposition, nous pouvons définir une relation récurrente de l'inverse de la matrice P_{n+1} .

Ecrivons la relation (20.a) comme suit :

$$\begin{aligned} P_{n+1} &= P_n + [x_{n+1}, u_{n+1}] \begin{bmatrix} x'_{n+1} \\ -u'_{n+1} \end{bmatrix} \\ &= P_n + g_{n+1}h_{n+1} \end{aligned} \quad (21)$$

En appliquant le résultat du lemme A1 d'inversion matricielle, nous avons la proposition suivante.

Proposition 4.

La relation récurrente de la matrice $P_{n+1}^{-1} = [\hat{\gamma}'_{n+1}(Z'_{n+1}Z_{n+1})\hat{\gamma}_n]^{-1}$ s'écrit :

$$P_{n+1}^{-1} = P_n^{-1} - C_n h_{n+1} P_n^{-1} \quad (22.a)$$

avec $C_n = P_n^{-1} g_{n+1} [I + h_{n+1} P_n^{-1} g_{n+1}]^{-1}$. (22.b)

De cette relation, à la n-ième itération, la matrice P_n^{-1} a été supposée calculée, il suffit alors de calculer l'inverse de la matrice

d'ordre deux $\left[I + h_{n+1} P_n^{-1} g_{n+1} \right]$.

Ces différentes propositions nous permettent d'établir la formulation récurrente de l'estimateur de β , donnée par le théorème suivant :

théorème 1.

L'estimateur $\hat{\beta}$ de β s'écrit sous la forme récurrente suivante :

$$\hat{\beta}_{n+1} = \hat{\beta}_n + C_n (r_{n+1} - h_{n+1} \hat{\beta}_n)$$

où
$$r_{n+1} = \begin{bmatrix} y_{n+1} \\ -v_{n+1} \end{bmatrix}; \quad v_{n+1} = (y_{n+1} - z'_{n+1} \tilde{Y}_n) / \sqrt{b}$$

$$h_{n+1} = \begin{bmatrix} x'_{n+1} \\ -u'_{n+1} \end{bmatrix}; \quad u'_{n+1} = (x'_{n+1} - z'_{n+1} \hat{\gamma}_n) / \sqrt{b} \quad \square$$

Preuve.

L'estimateur de β à la période $n+1$ est défini par :

$$\hat{\beta}_{n+1} = P_{n+1}^{-1} \hat{\gamma}'_{n+1} (Z'_{n+1} Z_{n+1}) \tilde{Y}_{n+1}$$

Compte tenu de la proposition 2, il s'écrit :

$$\hat{\beta}_{n+1} = P_{n+1}^{-1} \left\{ \left(\hat{\gamma}_n + B_n (x'_{n+1} - z'_{n+1} \hat{\gamma}_n) \right)' \left(Z'_n Z_n + z'_{n+1} z_{n+1} \right) \right. \\ \left. \times \left(\tilde{Y}_n + B_n (y_{n+1} - z'_{n+1} \tilde{Y}_n) \right) \right\}$$

Après quelques manipulations, on obtient :

$$\hat{\beta}_{n+1} = P_{n+1}^{-1} \left\{ \hat{\gamma}'_n (Z'_n Z_n) \tilde{Y}_n - \hat{\gamma}'_n z'_{n+1} z'_{n+1} \tilde{Y}_n + \hat{\gamma}'_n z'_{n+1} y_{n+1} + x'_{n+1} z'_{n+1} \tilde{Y}_n \right\}$$

$$\left. (x'_{n+1} - z'_{n+1}\hat{\gamma}'_n)' z'_{n+1} (Z'_n Z_n)^{-1} z_{n+1} (y_{n+1} - z'_{n+1}\hat{\gamma}'_n) / b \right\}.$$

Après simplification, nous avons :

$$\begin{aligned} \hat{\beta}_{n+1} &= P_{n+1}^{-1} \left\{ \hat{\gamma}'_n (Z'_n Z_n) \tilde{Y}_n + x_{n+1} y_{n+1} \right. \\ &\quad \left. - (x'_{n+1} - z'_{n+1}\hat{\gamma}'_n)' (y_{n+1} - z'_{n+1}\tilde{Y}_n) / b \right\} \\ &= P_{n+1}^{-1} \left\{ \hat{\gamma}'_n (Z'_n Z_n) \tilde{Y}_n + g_{n+1} r_{n+1} \right\} \end{aligned}$$

où $g_{n+1} = (x_{n+1}, u_{n+1})$ (23)

$$r_{n+1} = \begin{bmatrix} y_{n+1} \\ -v_{n+1} \end{bmatrix} \quad \text{avec} \quad v_{n+1} = (y_{n+1} - z'_{n+1}\tilde{Y}_n) / \sqrt{b}$$

or $\hat{\gamma}'_n (Z'_n Z_n) \tilde{Y}_n = P_n \hat{\beta}_n$

$$= (P_{n+1} - g_{n+1} h_{n+1}) \hat{\beta}_n$$

d'où $\hat{\beta}_{n+1} = \hat{\beta}_n + P_{n+1}^{-1} g_{n+1} (r_{n+1} - h_{n+1} \hat{\beta}_n)$ (24)

et la matrice de gain s'écrit :

$$\begin{aligned} P_{n+1}^{-1} g_{n+1} &= \left(P_n + g_{n+1} h_{n+1} \right)^{-1} g_{n+1} \\ &= P_n^{-1} g_{n+1} \left(I + h_{n+1} P_n^{-1} g_{n+1} \right)^{-1} = C_n \quad \blacksquare \end{aligned}$$

Nous venons de démontrer une formulation récurrente dont la mise à jour de l'estimateur s'obtient en calculant l'inverse de la matrice

$[I + hP_n^{-1}g]$ de dimension (2×2) .

D'après les relations (23) et (24), l'estimateur de β s'écrit aussi :

$$\hat{\beta}_{n+1} = \hat{\beta}_n + P_{n+1}^{-1} \left\{ x_{n+1}(y_{n+1} - x'_{n+1}\hat{\beta}_n) - u_{n+1}(v_{n+1} - u'_{n+1}\hat{\beta}_n) \right\} \quad (25)$$

Néanmoins nous pouvons définir une formulation beaucoup plus pratique sans inversion de matrice.

Posons

$$\tilde{P}_n = P_n + x_{n+1}x'_{n+1} \quad (26. a)$$

$$\Rightarrow P_{n+1} = \tilde{P}_n - u_{n+1}u'_{n+1} \quad (26. b)$$

la matrice \tilde{P}_n correspond à la prise en compte de la nouvelle observation.

Proposition 5.

La récursivité de la matrice $P_{n+1}^{-1} = [\hat{\gamma}'_{n+1}(Z'_{n+1}Z_{n+1})\hat{\gamma}_{n+1}]^{-1}$ s'effectue en deux étapes :

1^{ère} étape :

$$\tilde{P}_n^{-1} = P_n^{-1} - M_n x_{n+1} P_n^{-1} \quad (27. a)$$

$$\text{avec } M_n = P_n^{-1} x_{n+1} (1 + x'_{n+1} P_n^{-1} x_{n+1})^{-1} \quad (27. b)$$

$$= \tilde{P}_n^{-1} x_{n+1} \quad (27. b)'$$

2^{ème} étape :

$$P_{n+1}^{-1} = \tilde{P}_n^{-1} - N_n u_{n+1} \tilde{P}_n^{-1} \quad (27. c)$$

$$\text{avec } N_n = -P_{n+1}^{-1}u_{n+1} \quad (27.d)$$

$$= \tilde{P}_n^{-1}u_{n+1}(u'_{n+1}\tilde{P}_n^{-1}u_{n+1} - 1)^{-1} \quad (27.d)'$$

Preuve.

Appliquons le résultat du lemme d'inversion matricielle A1 à la relation (26.a), on obtient :

$$\tilde{P}_n^{-1} = P_n^{-1} - P_n^{-1}x_{n+1}(1 + x'_{n+1}P_n^{-1}x_{n+1})^{-1}x'_{n+1}P_n^{-1}.$$

Définissons la matrice de gain de dimension $(k \times 1)$ par :

$$M_n = P_n^{-1}x_{n+1}(1 + x'_{n+1}P_n^{-1}x_{n+1})^{-1}.$$

D'après le résultat du lemme A2 d'inversion matricielle, elle s'écrit également :

$$M_n = \left(P_n + x_{n+1}x'_{n+1} \right)^{-1} x_{n+1} = \tilde{P}_n^{-1}x_{n+1}.$$

De la même manière, à partir de la relation (26.b), l'utilisation du résultat du lemme d'inversion matricielle A1 nous donne :

$$P_{n+1}^{-1} = \tilde{P}_n^{-1} - \tilde{P}_n^{-1}u_{n+1}(u'_{n+1}\tilde{P}_n^{-1}u_{n+1} - 1)^{-1}u'_{n+1}\tilde{P}_n^{-1}.$$

La matrice de gain est définie par :

$$N_n = \tilde{P}_n^{-1}u_{n+1}(u'_{n+1}\tilde{P}_n^{-1}u_{n+1} - 1)^{-1}$$

que l'on écrit aussi d'après le résultat du lemme A2 d'inversion matricielle

$$N_n = - \left(\tilde{P}_n - u_{n+1} u'_{n+1} \right)^{-1} u_{n+1} = - P_{n+1}^{-1} u_{n+1} \quad \blacksquare$$

Cette dernière proposition nous permet de définir une formulation récurrente sans inversion de matrice supplémentaire donnée par le théorème suivant :

théorème 2.

La récurrence de l'estimateur s'effectue en deux étapes :

1^{ère} étape :

$$\tilde{\beta}_n = \hat{\beta}_n + M_n (y_{n+1} - x'_{n+1} \hat{\beta}_n)$$

2^{ème} étape : correction du biais de l'estimation réalisée.

$$\hat{\beta}_{n+1} = \tilde{\beta}_n + N_n (v_{n+1} - u_{n+1} \tilde{\beta}_n).$$

Preuve.

$$\begin{aligned} \hat{\beta}_{n+1} &= P_{n+1}^{-1} \left\{ P_n \hat{\beta}_n + x_{n+1} y_{n+1} - u_{n+1} v_{n+1} \right\} \\ &= P_{n+1}^{-1} \left\{ \tilde{P}_n \hat{\beta}_n + x_{n+1} (y_{n+1} - x'_{n+1} \hat{\beta}_n) \right\} - P_{n+1}^{-1} u_{n+1} v_{n+1} \end{aligned}$$

D'après la proposition 4, nous avons :

$$\begin{aligned} \hat{\beta}_{n+1} &= [\tilde{P}_n^{-1} - N_n u_{n+1} \tilde{P}_n^{-1}] [\tilde{P}_n \hat{\beta}_n + x_{n+1} (y_{n+1} - x'_{n+1} \hat{\beta}_n)] + N_n v_{n+1} \\ &= \hat{\beta}_n + M_n (y_{n+1} - x'_{n+1} \hat{\beta}_n) - N_n u'_{n+1} [\hat{\beta}_n + M_n (y_{n+1} - x'_{n+1} \hat{\beta}_n)] + N_n v_{n+1} \end{aligned}$$

Définissons un vecteur intermédiaire $\tilde{\beta}_n$ tel que

$$\tilde{\beta}_n = \hat{\beta}_n + M_n (y_{n+1} - x'_{n+1} \hat{\beta}_n)$$

$$\text{d'où} \quad \hat{\beta}_{n+1} = \tilde{\beta}_n + N_n (v_{n+1} - u'_{n+1} \tilde{\beta}_n) \quad \text{C.Q.F.D.} \quad \blacksquare$$

A la première étape, on reconnaît la forme récursive des moindres carrés ordinaires, la deuxième étape consiste à corriger le biais de l'estimation réalisée.

Conclusion.

Nous avons démontré une formulation récursive pour les paramètres constants d'une part et variables d'autre part d'un modèle de régression, nous avons établi une relation entre la matrice de projection et la matrice de gain et entre l'erreur de prédiction a priori et l'erreur de prédiction a posteriori. Quelques cas particuliers du modèle ont été présentés.

Dans la deuxième partie, nous avons donné des formulations récursives des moindres carrés doubles qui peuvent être réalisées en pratique à l'aide de programmes numériques.

Annexe

Lemme d'inversion matricielle.

$$\left[A + CBD \right]^{-1} = A^{-1} - A^{-1}C \left[B^{-1} + DA^{-1}C \right]^{-1} DA^{-1} \quad A1$$

$$\left[A + CBD \right]^{-1} CB = A^{-1}C \left[B^{-1} + DA^{-1}C \right]^{-1} \quad A2$$

Preuve.

$$\text{Posons } R = \left[A + CBD \right]^{-1} \quad (1)$$

où les matrices A et B sont supposées régulières,

$$R^{-1} = A + CBD \quad (2)$$

multiplions les deux membres de cette relation par A^{-1} à droite et par R à gauche, il vient

$$A^{-1} = R \left[I + CBDA^{-1} \right] \quad (3)$$

multiplions la relation (3) par la matrice C

$$\begin{aligned} A^{-1}C &= RC \left[I + BDA^{-1}C \right] \\ &= RCB \left[B^{-1} + DA^{-1}C \right] \end{aligned} \quad (4)$$

la matrice $B^{-1} + DA^{-1}C$ est supposée régulière

$$A^{-1}C \left[B^{-1} + DA^{-1}C \right]^{-1} = RCB, \quad (5)$$

multiplions la relation (5) par DA^{-1}

$$A^{-1}C \left[B^{-1} + DA^{-1}C \right]^{-1} DA^{-1} = RCBDA^{-1}, \quad (6)$$

compte tenu de l'expression (3), on obtient

$$R = A^{-1} - RCBDA^{-1} \quad (7)$$

nous avons d'après les relations (6), (7), et (1)

$$\left[A + CBD \right]^{-1} = A^{-1} - A^{-1}C \left[B^{-1} + DA^{-1}C \right]^{-1} DA^{-1} \quad (8)$$

qui vérifie A1.

De même l'expression (5) vérifie A2

$$A^{-1}C \left[B^{-1} + DA^{-1}C \right]^{-1} = \left[A + CBD \right]^{-1} CB. \quad \blacksquare$$

BIBLIOGRAPHIE

- BALESTRA P. (1978) Determinant and inverse of a sum of matrices with applications in economics and statistics. *Document de L'IME N°24*.
- BROWN R.L., DURBIN J., et EVANS J.M., (1975), Techniques for testing the constancy of regression relationships over time, *Journal of the royal statistical society series B* 37, pp 149-192.
- CHOW G. (1981), *Econometric analysis by control methods*, John wiley, New-York.
- DUCAN D.B. et HORN J.D., (1972), Linear dynamic recursive estimation from the wiepoint of regression analysis, *Journal of American Statistical Association*, Vol 67, n° 340, pp 815-821.
- DE LARMINAT P. (1986) Une commande adaptative basée sur l'identification par variable instrumentale, *R.A.I.R.O, Automatique Productique Informatique Industrielle (APII)*, Vol 20, N°2, pp 99 - 118.
- GARBADE K. (1977), Two methods for examining the stability of regression coefficients, *Journal of American Statistical Association*, Vol 72, pp 54-63.
- HARVEY A.C. et PHILLIPS G.D.A. (1974), A simple test for serial correlation in regression analysis, *Journal of American Statistical Association*, Vol 69, pp 935-939.
- LJUNG L. (1985) Estimation of parameters in dynamical systems in HANNAN E.J et RAO M.M. eds *Handbook of Statistics*, Vol 5.
- PHILLIPS G.D.A (1977), Recursions for the two stage least squares estimators, *Journal of Econometrics*, vol 6, N°1, pp 65 - 77.
- YOUNG P. (1985) Recursive identification, estimation and control, in HANNAN E.J et RAO M.M eds *Hanbook of Statistics*, Vol 5.

