



HAL
open science

Estimation adaptative des modeles vectoriels autoregrésifs avec une variance dependant du temps

Valentin Patilea, Hamdi Raïssi

► **To cite this version:**

Valentin Patilea, Hamdi Raïssi. Estimation adaptative des modeles vectoriels autoregrésifs avec une variance dependant du temps. 42èmes Journées de Statistique, 2010, Marseille, France, France. inria-00494847

HAL Id: inria-00494847

<https://hal.inria.fr/inria-00494847>

Submitted on 24 Jun 2010

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.

ESTIMATION ADAPTATIVE DES MODELES VECTORIELS AUTOREGRESSIFS AVEC UNE VARIANCE DEPENDANT DU TEMPS

Valentin Patilea & Hamdi Raïssi*

20, avenue des buttes de Coësmes, CS 70839, F-35708 Rennes Cedex 7, France.

mail: hamdi.raïssi@insa-rennes.fr.

Résumé

Nous analysons les modèles Vectoriels AutoRegressifs (VAR) quand les innovations sont non conditionnellement hétéroscédastiques. La structure de la volatilité est déterministe et générale, incluant des discontinuités ou des tendances comme cas particuliers. Dans ce cadre nous proposons des estimateurs des Moindres Carrés Ordinaires (MCO) et des estimateurs des Moindres Carrés Adaptatifs (MCA). L'estimateur des MCA est calculé en estimant la volatilité de façon non paramétrique. Nous obtenons la distribution asymptotique des estimateurs et comparons leur propriétés. En particulier nous montrons que l'estimateur des MCA est asymptotiquement équivalent à l'estimateur des Moindres Carrés Généralisés (MCG) obtenus en supposant que la volatilité des erreurs est connue.

Abstract

We analyze Vector AutoRegressive (VAR) models when the innovations are unconditionally heteroscedastic. The volatility structure is deterministic and quite general, including breaks or trending variances as special cases. In this framework we propose ordinary least squares and Adaptive Least Squares (ALS). The ALS is obtained using a non parametric estimator of the volatility function. We derive the asymptotic distribution of the proposed estimators for the VAR model coefficients and compare their properties. In particular we show that the ALS estimator is asymptotically equivalent to the infeasible generalized least squares estimator obtained assuming that the volatility is known.

Mots clés: Modèles VAR; Erreurs hétéroscédastiques, Moindres carrés adaptatifs, Moindres carrés ordinaires, Causalité linéaire en moyenne.

1 Introduction

L'étude de modèles linéaires avec erreurs non conditionnellement hétéroscédastiques a trouvé un intérêt croissant ces dernières années (voir entre autres références Boswijk et Zu (2007)). Cet intérêt est motivé par de nombreuses études appliquées qui ont montré que la présence d'erreurs à volatilité non conditionnelle est commune dans des données de nature économique (voir entre autres références Sensier et van Dijk (2004)). Phillips et Xu (2005), Xu et Phillips (2008) ont étudié le problème de l'estimation de processus autorégressifs univariés stables. Ils ont considéré des estimateurs des MCO et MCA. Nous étendons leur résultats dans le cadre des VAR. Nous montrons que dans ce cas la distribution asymptotique de l'estimateur MCA dépend de la structure de la volatilité. De plus nos résultats asymptotiques sont uniformes par rapport à la fenêtre prise dans un intervalle donné. Ces résultats permettent le choix du paramètre de lissage basés sur les données comme par exemple la validation croisée, ainsi que des outils d'analyse pour l'étude de liens linéaires dynamiques entre variables.

2 Estimation des paramètres d'un modèle VAR

Soit les observations $X_{-p+1}, \dots, X_0, X_1, \dots, X_T$ générées par un modèle VAR

$$\begin{aligned} X_t &= A_1 X_{t-1} + \dots + A_p X_{t-p} + u_t \\ u_t &= H_t \epsilon_t, \end{aligned} \tag{1}$$

où les vecteurs X_t sont de dimension d . Les matrices A_i , $i \in \{1, \dots, p\}$, sont telles que $\det A(z) \neq 0$ pour tout $|z| \leq 1$, avec $A(z) = I_d - \sum_{i=1}^p A_i z^i$. Pour une variable aléatoire x nous définissons $\|x\|_r = (E \|x\|^r)^{1/r}$, où $\|x\|$ est la norme euclidienne. Nous définissons également $\mathcal{F}_t = \sigma(\epsilon_s, s \leq t)$ la σ -algèbre générée par $\{\epsilon_s, s \leq t\}$. L'hypothèse sur les matrices H_t et le processus (ϵ_t) donne le cadre de notre étude.

Hypothèse A1: (i) Les matrices H_t sont inversibles et satisfont $H_{[Tr]} = G(r)$, où les composantes de la matrice $G(r) := \{g_{ij}(r)\}$ sont des fonctions mesurables déterministes sur l'intervalle $]0, 1]$, telles que $\sup_{r \in (0, 1]} |g_{ij}(r)| < \infty$, et satisfont une condition de Lipschitz par morceaux. De plus nous supposons que $\Sigma_t := H_t H_t'$ est définie positive pour tout t .

(ii) Le processus (ϵ_t) est α -mélangeant et tel que $E(\epsilon_t | \mathcal{F}_{t-1}) = 0$, $E(\epsilon_t \epsilon_t' | \mathcal{F}_{t-1}) = I_d$ et $\sup_t \|\epsilon_{it}\|_{4\mu} < \infty$ pour un certain $\mu > 1$ et tout $i \in \{1, \dots, d\}$.

Notons que la spécification de la volatilité en **A1** est large et peut décrire des cas communément observés dans les séries notamment économiques. Nous utilisons les approches suivantes pour construire les estimateurs des paramètres autorégressifs. Soit

$\theta_0 = (\text{vec}(A_1)', \dots, \text{vec}(A_p)')' \in \mathbf{R}^{pd^2}$ le vecteur contenant les paramètres. L'équation (1) devient

$$\begin{aligned} X_t &= (\tilde{X}'_{t-1} \otimes I_d)\theta_0 + u_t \\ u_t &= H_t\epsilon_t, \end{aligned} \quad (2)$$

où le produit de Kronecker usuel est noté \otimes . En utilisant (2) nous définissons l'estimateur des MCO

$$\hat{\theta}_{MCO} = \hat{\Sigma}_{\tilde{X}}^{-1} \text{vec}(\hat{\Sigma}_X),$$

où

$$\hat{\Sigma}_{\tilde{X}} = T^{-1} \sum_{t=1}^T \tilde{X}_{t-1} \tilde{X}'_{t-1} \otimes I_d \quad \text{et} \quad \hat{\Sigma}_X = T^{-1} \sum_{t=1}^T X_t \tilde{X}'_{t-1}.$$

Nous considérons également l'estimateur des MCG

$$\hat{\theta}_{MCG} = \hat{\Sigma}_{\underline{X}}^{-1} \text{vec}(\hat{\Sigma}_{\underline{X}}), \quad (3)$$

avec

$$\hat{\Sigma}_{\underline{X}} = T^{-1} \sum_{t=1}^T \tilde{X}_{t-1} \tilde{X}'_{t-1} \otimes \Sigma_t^{-1} \quad \text{et} \quad \hat{\Sigma}_{\underline{X}} = T^{-1} \sum_{t=1}^T \Sigma_t^{-1} X_t \tilde{X}'_{t-1}.$$

Si on suppose que Σ_t est constant, nous avons $\hat{\theta}_{MCG} = \hat{\theta}_{MCO}$. Notons que l'estimateur des MCG n'est pas calculable en pratique puisqu'il dépend de la volatilité inconnue. Dans la section suivante nous étudions les propriétés des estimateurs des MCO et des MCG.

3 Comportement asymptotique des estimateurs

Afin d'énoncer les premiers résultats de cet article nous avons besoin d'introduire les notations suivantes. Puisque nous avons supposé que $\det A(z) \neq 0$ pour tout $|z| \leq 1$, nous avons

$$X_t = \sum_{i=0}^{\infty} \psi_i u_{t-i}, \quad (4)$$

où $\psi_0 = I_d$ et les éléments des ψ_i sont absolument sommables (voir Lütkepohl (2005, pp 14-16)). D'après l'expression (4) nous écrivons également

$$\tilde{X}_t = \sum_{i=0}^{\infty} \tilde{\psi}_i u_{t-i}^p,$$

avec u_t^p est donné par $u_t^p = \mathbf{1}_p \otimes u_t$, et où $\mathbf{1}_p$ est un vecteur de dimension p dont les composantes sont égales à un, et

$$\tilde{\psi}_i = \begin{pmatrix} \psi_i & 0 & 0 & 0 \\ 0 & \psi_{i-1} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \psi_{i-p+1} \end{pmatrix},$$

où on prend $\psi_j = 0$ pour $j < 0$. Définissons $\Sigma(r) = G(r)G(r)'$ et notons par $\mathbf{1}_{p \times p}$ la matrice de dimension $p \times p$ avec des composantes égales à un. La proposition suivante nous donne le comportement asymptotique des estimateurs introduits à la section précédente.

Proposition 3.1 *Sous l'hypothèse A1 nous avons*

$$T^{\frac{1}{2}}(\hat{\theta}_{MCG} - \theta_0) \Rightarrow \mathcal{N}(0, \Lambda_1^{-1}), \quad (5)$$

où

$$\Lambda_1 = \int_0^1 \sum_{i=0}^{\infty} \left\{ \tilde{\psi}_i(\mathbf{1}_{p \times p} \otimes \Sigma(r)) \tilde{\psi}_i' \right\} \otimes \Sigma(r)^{-1} dr.$$

Nous avons également

$$T^{\frac{1}{2}}(\hat{\theta}_{MCO} - \theta_0) \Rightarrow \mathcal{N}(0, \Lambda_3^{-1} \Lambda_2 \Lambda_3^{-1}), \quad (6)$$

avec

$$\Lambda_2 = \int_0^1 \sum_{i=0}^{\infty} \left\{ \tilde{\psi}_i(\mathbf{1}_{p \times p} \otimes \Sigma(r)) \tilde{\psi}_i' \right\} \otimes \Sigma(r) dr$$

et

$$\Lambda_3 = \int_0^1 \sum_{i=0}^{\infty} \left\{ \tilde{\psi}_i(\mathbf{1}_{p \times p} \otimes \Sigma(r)) \tilde{\psi}_i' \right\} \otimes I_d dr.$$

De plus les matrices Λ_1 , Λ_2 et Λ_3 sont définies positives.

Des estimateurs convergents des matrices Λ_2 et Λ_3 sont proposés dans une version longue de cet article. Nous montrons à travers un exemple dans la version longue de cet article (Patilea et Raïssi (2010)) que considérer les résultats standard pour l'analyse d'un modèle VAR alors que le processus d'erreur est non conditionnellement hétéroscédastique peut amener à de fausses conclusions. Notons aussi que contrairement au cas univarié (voir Xu et Phillips (2008)), le comportement asymptotique de l'estimateur des MCG dépend de $\Sigma(r)$. Comme la volatilité des termes d'erreur n'est pas connue, l'estimateur des MCG ne peut être calculé. Nous proposons dans la section suivante un estimateur s'appuyant sur l'estimation non paramétrique de la volatilité.

4 Estimation adaptative

Soit l'estimateur adaptatif de la fonction de volatilité qui généralise l'estimateur de Xu et Phillips (2008) au cas multivarié :

$$\check{\Sigma}_t = \sum_{i=1}^T w_{ti} \hat{u}_i \hat{u}_i',$$

où les \hat{u}_i sont les résidus de l'estimation par MCO du modèle (1). Les poids w_{ti} sont donnés par

$$w_{ti} = w_{ti}(b) = \left(\sum_{i=1}^T K_{ti} \right)^{-1} K_{ti},$$

avec

$$K_{ti} = K_{ti}(b) = \begin{cases} K\left(\frac{t-i}{Tb}\right) & \text{si } t \neq i \\ 0 & \text{si } t = i, \end{cases}$$

où $K(z)$ est le noyau. La fenêtre b appartient à un intervalle $\mathcal{B}_T = [c_{min}b_T, c_{max}b_T]$ avec c_{min}, c_{max} des constantes et $b_T \downarrow 0$ à une vitesse T^{-1} qui sera précisée plus loin. Comme les $\hat{u}_i, i = 1, \dots, T$ sont presque sûrement linéairement indépendents entre eux, $\check{\Sigma}_t$ est presque sûrement définie positive pourvu que T est suffisamment grand. En pratique la fenêtre b peut être choisie par validation croisée en minimisant le critère suivant

$$\widehat{CV}(b) = \sum_{t=1}^T \|\check{\Sigma}_t - \hat{u}_t \hat{u}_t'\|^2,$$

avec $b \in \mathcal{B}_T$, où $\|\cdot\|$ est une norme pour une matrice carrée, par exemple la norme de Frobenius. Nos résultats théoriques sont obtenus uniformément par rapport à $b \in \mathcal{B}_T$ et apportent ainsi une justification pour la sélection par validation croisée communément pratiquée dans notre cadre. Nous pouvons maintenant introduire l'estimateur des MCA suivant

$$\hat{\theta}_{ALS} = \check{\Sigma}_{\check{X}}^{-1} \text{vec} \left(\check{\Sigma}_{\check{X}} \right),$$

avec

$$\check{\Sigma}_{\check{X}} = T^{-1} \sum_{t=1}^T \check{X}_{t-1} \check{X}_{t-1}' \otimes \check{\Sigma}_t^{-1}, \quad \text{and} \quad \check{\Sigma}_{\check{X}} = T^{-1} \sum_{t=1}^T \check{\Sigma}_t^{-1} X_t \check{X}_{t-1}'.$$

Hypothèse A1': Supposons que toutes les conditions dans **A1** (i) sont vérifiées. De plus:

(i) $\inf_{r \in (0,1]} \lambda_{min}(\Sigma(r)) > 0$ où $\lambda_{min}(\Gamma)$ est la plus petite valeur propre d'une matrice symétrique Γ .

(ii) $\sup_t \|\epsilon_{it}\|_8 < \infty$ pour tout $i \in \{1, \dots, d\}$.

Hypothèse A2: (i) Le noyau $K(\cdot)$ est une densité, borné, continu défini sur l'ensemble des réels, croissante sur $] -\infty, 0]$ et décroissante sur $[0, \infty)$.

(ii) La fenêtre b est choisie dans l'intervalle $\mathcal{B}_T = [c_{min}b_T, c_{max}b_T]$ avec $0 < c_{min} < c_{max} < \infty$ et $b_T + 1/Tb_T^2 \rightarrow 0$ quand $T \rightarrow \infty$.

Notons que **A2** (i) est satisfaite par tout les noyaux couramment utilisés. Soit $\Omega_1 := \int_0^1 \Sigma(r) \otimes \Sigma(r)^{-1} dr$. Une suite de matrices aléatoires A_t , $t \geq 1$ est $o_p(1)$ uniformément par rapport à $b \in \mathcal{B}_T$ si $\sup_{b \in \mathcal{B}_T} \|\text{vec}(A_t)\| = o_p(1)$. La proposition suivante nous donne le comportement asymptotique de l'estimateur adaptatif.

Proposition 4.1 *Sous les hypothèses A1' and A2, nous avons uniformément par rapport à $b \in \mathcal{B}_T$*

$$\check{\Lambda}_1 := \check{\Sigma}_{\check{X}} = \Lambda_1 + o_p(1),$$

$$\check{\Omega}_1 := T^{-1} \sum_{t=1}^T \check{\Sigma}_t \otimes \check{\Sigma}_t^{-1} = \Omega_1 + o_p(1)$$

et

$$\sqrt{T}(\hat{\theta}_{ALS} - \hat{\theta}_{GLS}) = o_p(1).$$

D'après la Proposition 4.1 les estimateurs MCA et MCG ont le même comportement asymptotique. Dans une version longue de cet article (Patilea et Raïssi (2010)) nous utilisons les différents résultats présentés dans ce résumé pour développer des tests de causalité de Granger linéaire en moyenne adaptés à notre cadre théorique. Des expériences Monte Carlo illustrent les propriétés des estimateurs étudiés.

Bibliographie

- [1] Boswijk, H.P., et Zu, Y. (2007) Testing for cointegration with nonstationary volatility. Document de travail. University d'Amsterdam.
- [2] Lütkepohl, H. (2005) *New Introduction to Multiple Time Series Analysis*. Springer, Berlin.
- [3] Patilea, V., et Raïssi, H., (2010) Adaptive estimation of vector autoregressive models with time-varying variance: application to testing linear causality in mean. Document de travail. IRMAR-INSA.
- [4] Phillips, P.C.B., et Xu, K.L. (2005) Inference in autoregression under heteroskedasticity. *Journal of Time Series Analysis* 27, 289–308.
- [5] Sensier, M., et van Dijk, D. (2004) Testing for volatility changes in U.S. macroeconomic time series. *The Review of Economics and Statistics* 86, 833–839.
- [6] Xu, K.L., et Phillips, P.C.B. (2008) Adaptive estimation of autoregressive models with time-varying variances. *Journal of Econometrics* 142, 265–280.