Corrected LM tests for AR models with time-varying variance

Raja Ben Hajria

LGM-ENIM, Faculté des Sciences de Monastir, Université de Monastir, Tunisie

Résumé. Dans le présent article, nous étendons l'approche de Breusch-Godfrey (Test des multiplicateurs de Lagrange (LM)) pour contrôler la qualité d'ajustement des modèles autorégressifs dans le cas des processus ayant une variance non conditionnelle non constante au cours du temps. Nous effectuons une comparaison asymptotique de la puissance de test LM et tests portmanteau au sens de l'efficacité de Bahadur contre des alternatives linéaires.

Mots-clés. Modèle AR ; Erreurs non conditionnellement hétéroscédastiques ; Autocorrélations résiduelles ; Tests portmanteau ; Efficacité asymptotique relative ; Tests des multiplicateurs de Lagrange.

Abstract. In this paper the Breusch-Godfrey Lagrange Multiplier (LM) approach for checking the goodness-of-fit of autoregressive models is extended in the case of processes with non constant variance. We provide an asymptotic power comparison between the portmanteau and the LM tests in the Bahadur sense against linear alternatives.

Keywords. AR model ; Unconditionally heteroscedastic errors ; Residual autocorrelations ; Portmanteau tests ; Asymptotic relative efficiency ; Lagrange Multiplier tests.

1 Introduction

Les tests d'autocorrélation des résidus dans les processus autorégressifs ont reçu beaucoup d'attention dans la littérature. Ceci peut être expliqué par le fait qu'un ordre autorégressif bien spécifié est important pour une description appropriée de la dynamique linéaire de la série. Le test LM de Breusch-Godfrey (Breusch (1978) et Godfrey (1978)) et le test portmanteau (voir Box et Pierce (1970), BP ci-après, et Ljung et Box (1978), LB ci-après)) ont été introduits dans le cadre standard des processus ayant une variance constante au cours du temps. L'implémentation de ces tests se fait généralement à l'aide de logiciels tels que R, SAS ou JMulTi. Dans le présent article, nous développons un test des multiplicateurs de Lagrange pour contrôler la qualité d'ajustement des modèles autorégressifs dans le cas des processus ayant une variance non conditionnelle non constante au cours du temps (voir Xu et Phillips (2008)). De plus nous démontrons que le test LM proposé est plus puissant que le test portmanteau au sens de l'efficacité de Bahadur contre des alternatives linéaires.

2 Tests d'autocorrélation sur les Modèles de type AR

Nous considérons un processus autorégressif d'ordre p_0 défini par :

$$\begin{aligned} x_t &= a_{01}x_{t-1} + \cdots + a_{0p_0}x_{t-p_0} + u_t \\ u_t &= h_t\epsilon_t, \quad t = 1, \ldots, n, \end{aligned} \tag{1}$$

où les paramètres a_{0i}, $i \in \{1, \ldots, p_0\}$, sont tels que $\det A(z) \neq 0$ pour tout $|z| \leq 1$, avec $A(z) = 1 - \sum_{i=1}^{p_0} a_{0i}z^i$. Les conditions suivantes sur le processus (u_t) donnent le cadre de notre étude.

Hypothèse A1 : Supposons que (u_t) satisfait $u_t = h_t\epsilon_t$ telles que : **(i)** Les termes h_t satisfont $h_t = g(t/n)$, où $g(r)$ est une fonction déterministe mesurable strictement positive et non constante sur l'intervalle $(0, 1]$, avec $\sup_{r\in(0,1]} |g(r)| < \infty$, et satisfait une condition de Lipschitz par morceaux sur un nombre fini de sous intervalles qui partitionnent $(0, 1]$. Nous supposons que $g(.)$ n'est pas observée. **(ii)** Le processus (ϵ_t) est α-mélangeant (voir Doukhan (1994) pour la définition d'un processus α-mélangeant) tel que $E(\epsilon_t \,|\, \mathcal{F}_{t-1}) = 0$, $E(\epsilon_t^2 \,|\, \mathcal{F}_{t-1}) = 1$ pour la filtration naturelle $\mathcal{F}_t = \sigma(\epsilon_s, s \leq t)$ et $\sup_t ||\epsilon_t||_{4\mu} < \infty$ pour un certain $\mu > 1$.

À partir de nos hypothèses, nous avons $\tilde{x}_t = \sum_{i=0}^{\infty} \psi^i\tilde{u}_{t-i}$, où $\tilde{u}_t = (u_t, 0, \ldots, 0)'$, $\tilde{x}_t = (x_t, x_{t-1}, \ldots, x_{t-p_0+1})'$ sont des vecteurs de dimension p_0 et

$$\psi = \begin{pmatrix} a_{01} & a_{02} & \ldots & a_{0p_0} \\ 1 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Nous supposons que $x_1, \ldots, x_n$ sont observés, et pour calculer les statistiques de test, nous fixons arbitrairement $x_t = 0$ pour tout $t < 1$.

Dans cette partie, nous faisons l'hypothèse non réaliste que la structure de la variance est connue. Par la suite une approximation adaptative de la statistique du test LM sera proposée dans la section suivante. Les paramètres autorégressifs sont estimés par la méthode des moindres carrés généralisés. Pour nos résultats asymptotiques, nous considérons tout d'abord le cas où l'ordre autorégressif est bien ajusté, c'est-à-dire $p = p_0$. Soit le vecteur des vrais paramètres autorégressifs $\theta_0 = (a_{01}, \ldots, a_{0p})'$ et l'estimateur MCG $\hat{\theta}_{MCG} = \left\{\hat{\Sigma}_{\tilde{x}}\right\}^{-1}\left\{\hat{\Sigma}_x\right\}$, où $\hat{\Sigma}_{\tilde{x}} = n^{-1}\sum_{t=2}^{n} h_t^{-2}\tilde{x}_{t-1}\tilde{x}_{t-1}'$ et $\hat{\Sigma}_x = n^{-1}\sum_{t=2}^{n} h_t^{-2}x_t\tilde{x}_{t-1}$. Il est montré dans Xu et Phillips (2008) qu'un tel estimateur est $\sqrt{n}$-asymptotiquement normal. Notons par $\hat{\epsilon}_t := h_t^{-1}(x_t - \hat{\theta}_{MCG}'\tilde{x}_{t-1})$ les résidus MCG. Comme nous avons supposé une structure déterministe pour la variance, le choix du nombre de retard adéquat peut être vérifié en testant les deux hypothèses suivantes : $H_0 : \Gamma_m = 0$ contre l'hypothèse $H_1 : \Gamma_m \neq 0$, avec $\Gamma_m = (\gamma(1), \dots, \gamma(m))'$ pour tout $m \geq 1$ avec $\gamma(i) = E(\epsilon_t \epsilon_{t-i})$. En pratique, plusieurs valeurs de m doivent être testées. Nous rappelons brièvement quelques résultats sur les autocovariances résiduelles et le test portmanteau introduit par Patilea et Raïssi (2013). Définissons les autocovariances résiduelles $\hat{\gamma}(k) := n^{-1} \sum_{t=1+k}^{n} \hat{\epsilon}_t \hat{\epsilon}_{t-k}$. Nous introduisons également le vecteur des m premières autocovariances résiduelles $\hat{\Gamma}_m = (\hat{\gamma}(1), \dots, \hat{\gamma}(m))'$ et le vecteur des m premières autocorrélations résiduelles $\hat{\Phi}_m = (\hat{\gamma}(1)/\hat{\gamma}(0), \dots, \hat{\gamma}(m)/\hat{\gamma}(0))'$. Définissons par $e_p(1)$ le vecteur de dimension p de sorte que la première composante soit un et zero ailleurs. Il peut être montré que $n^{\frac{1}{2}} \hat{\Gamma}_m \Rightarrow \mathcal{N}\left(0, I_m - K_m K_\infty^{-1} K_m'\right)$ et $n^{\frac{1}{2}}(\hat{\Phi}_m - \hat{\Gamma}_m) = o_p(1)$, où $K_m = (\psi^0 e_p(1), \psi^1 e_p(1), \dots, \psi^{m-1} e_p(1))'$ et $K_\infty = \sum_{i=0}^{\infty} \psi^i e_p(1) e_p(1)' \psi^{i'}$. Les statistiques du test de portmanteau sont données par :

$$Q_m^{BP} = n \hat{\Phi}_m' \hat{\Phi}_m, \quad et \quad Q_m^{LB} = n(n+2) \sum_{k=1}^{m} (n-k)^{-1} \hat{\gamma}^2(k) \hat{\gamma}^{-2}(0).$$

Sous l'hypothèse nulle et l'hypothèse **A1**, la distribution asymptotique des statistiques de test portmanteau peuvent être approchées par $\chi^2_{(m-p)}$ pour m assez large.

Le test de Breusch-Godfrey (LM) s'appuie sur le modèle suivant :

$$x_t = \sum_{i=1}^{p} a_{0i} x_{t-i} + \sum_{j=1}^{m} b_{0j} u_{t-j} + u_t. \tag{2}$$

Rappelons que nous supposons que $p = p_0$ et que m est fixé par le praticien. Sous l'hypothèse nulle on a $b_{0j} = 0$ pour tout $j \in \{1, \dots, m\}$. La nullité des b_{0j} est testée en utilisant l'approche LM. Nous pouvons réécrire le modèle (2) sous la forme $x_t = \varphi_0' \zeta_{t-1} + u_t$, où $\varphi_0 = (a_{01}, \dots, a_{0p}, b_{01}, \dots, b_{0m})'$ et $\zeta_{t-1} = (x_{t-1}, \dots, x_{t-p}, u_{t-1}^{m'})'$ avec $u_{t-1}^m = (u_{t-1}, \dots, u_{t-m})'$. Prenons $u_t = 0$ pour tout $t < 1$. En notant par $\mathcal{L}(\varphi)$ la log-vraisemblance de (2) pour tout vecteur $\varphi \in \mathbb{R}^{p+m}$, le vecteur des scores au point $\varphi^c = (\hat{\theta}_{MCG}', 0)'$ est alors donné par

$$R\mathcal{S} = R \frac{\partial \mathcal{L}(\varphi^c)}{\partial \varphi} = \sum_{t=1}^{n} h_t^{-2} \hat{u}_t \hat{u}_{t-1}^m,$$

où $R = (0_{m \times p}, I_m)$ est de dimension $m \times (m+p)$, $\hat{u}_t = x_t - \hat{\theta}_{MCG}' \tilde{x}_{t-1}$ et $\hat{u}_{t-1}^m = (\hat{u}_{t-1}, \dots, \hat{u}_{t-m})'$. Certains calculs montrent que le vecteur des scores peut être approché par les autocovariances résiduelles :

$$n^{-\frac{1}{2}} R\mathcal{S} = n^{\frac{1}{2}} \hat{\Gamma}_m + o_p(1),$$

et que la variance asymptotique de $n^{-\frac{1}{2}} R\mathcal{S}$ est donnée par

$$\mathrm{Var}_{as}(n^{-\frac{1}{2}} R\mathcal{S}) = \left[R J^{-1} R'\right]^{-1} : \; = \; \left[R \begin{pmatrix} K_\infty & K_m' \\ K_m & I_m \end{pmatrix}^{-1} R' \right]^{-1}.$$

La matrice de covariance estimée est définie comme suit

$$\left[R\hat{J}^{-1}R'\right]^{-1} = \left[R\begin{pmatrix} \hat{K}_\infty & \hat{K}'_m \\ \hat{K}_m & I_m \end{pmatrix}^{-1} R'\right]^{-1}$$

où $\hat{K}_m$ est obtenue en remplaçant les a_{0i} par leurs estimateurs MCG dans K_m et $\hat{K}_\infty = n^{-1}\sum_{t=1}^{n} h_t^{-2}\tilde{x}_{t-1}\tilde{x}'_{t-1}$. Il est clair que nous avons $\hat{K}_m = K_m + o_p(1)$, et que $\hat{K}_\infty = K_\infty + o_p(1)$. Considérons maintenant la statistique LM :

$$Q_m^{LM} = n\hat{\Gamma}'_m\left(R\hat{J}^{-1}R'\right)\hat{\Gamma}_m.$$

Sous l'hypothèse nulle et l'hypothèse **A1**, nous avons $Q_m^{LM} \Rightarrow \chi_m^2$. Néanmoins, le test LM présenté ci-dessus s'appuie sur l'hypothèse irréaliste que la structure de la variance est connue. Un test LM asymptotiquement équivalent est proposé dans la section suivante.

3 Dérivation de la statistique de test des multiplicateurs de Lagrange adaptative

Maintenant nous allons construire un test LM des processus AR(p) ayant une variance non constante au cours du temps. Nous proposons un estimateur à noyau $\hat{\theta}_{ALS}$ qui a la même distribution asymptotique que $\hat{\theta}_{MCG}$. Une estimation non paramétrique de la variance est nécessaire pour construire une telle approximation. Soit $\breve{u}_t = x_t - \tilde{x}_{t-1}\hat{\theta}_{MCO}$ les résidus de l'estimation par MCO et $K(z)$ le noyau avec $\int_{-\infty}^{\infty} K(z)dz = 1$. Définissons $\check{h}_t^2 := \sum_{i=1}^{n} w_{ti}\breve{u}_i^2$ où $w_{ti} = \left(\sum_{i=1}^{n} K_{ti}\right)^{-1} K_{ti}$, et

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

Notons que la fenêtre b dépend de la taille de l'échantillon n. La mise en œuvre de l'estimateur $\check{h}_t$ dépend du choix du noyau K et la fenêtre b. D'autres hypothèses sont habituellement faites pour faciliter le développement de résultats asymptotiques (Voir Patilea et Raïssi (2013) pour plus de détails).

Hypothèse A2 : (i) Le noyau $K(.)$ est une fonction de densité de probabilité bornée définie sur $\mathbb{R}$ telle que $K(.)$ est non décroissante sur $(-\infty, 0]$, décroissante sur $[0,\infty)$ et $\int_{\mathbb{R}} v^2K(v)\,\mathrm{d}v < \infty$. La fonction $K(.)$ est différentiable sauf sur un nombre fini de points et la fonction dérivée $K'(.)$ satisfait $\int_{\mathbb{R}} xK'(x)\,\mathrm{d}x < \infty$. De plus la transformée de Fourier $\mathbf{F}[K](.)$ de $K(.)$ satisfait $\int_{\mathbb{R}} |s|^\tau|\mathbf{F}[K](s)|\,\mathrm{d}s < \infty$ pour un certain $\tau > 0$. **(ii)** La fenêtre b est prise dans l'intervalle $\mathcal{B}_n = [c_{min}b_n, c_{max}b_n]$ avec $0 < c_{min} < c_{max} < \infty$ et $nb_n^{4-a} + 1/nb_n^{2+a} \to 0$ quand $n \to \infty$, pour un certain $a > 0$ aussi petit que l'on veut.

Introduisons maintenant les estimateurs (ALS) suivants : $\hat{\theta}_{ALS} = \left\{\check{\Sigma}_{\tilde{x}}\right\}^{-1}\left\{\check{\Sigma}_{x}\right\}$, où $\check{\Sigma}_{\tilde{x}} = n^{-1}\sum_{t=2}^{n}\check{h}_t^{-2}\tilde{x}_{t-1}\tilde{x}_{t-1}'$, et $\check{\Sigma}_{x} = n^{-1}\sum_{t=2}^{n}\check{h}_t^{-2}x_t\tilde{x}_{t-1}$. Les résidus de l'estimation (ALS) sont définis comme suit $\check{\epsilon}_t := \check{h}_t^{-1}(x_t - \hat{\theta}_{ALS}'\tilde{x}_{t-1})$. Nous obtenons alors le vecteur des scores au point $\varphi_{ALS}^c = (\hat{\theta}_{ALS}', 0)'$,

$$R\mathcal{S}_{ALS} = R\frac{\partial\mathcal{L}(\varphi_{ALS}^c)}{\partial\varphi} = \sum_{t=1}^{n}\check{h}_t^{-2}\check{u}_t\check{u}_{t-1}^m.$$

Définissons $\check{\gamma}(k) := n^{-1}\sum_{t=1+k}^{n}\check{\epsilon}_t\check{\epsilon}_{t-k}$ et posons $\check{\Gamma}_m = (\check{\gamma}(1),\ldots,\check{\gamma}(m))'$. Comme il est montré dans Patilea et Raïssi (2013), $\check{\gamma}(k)$ et $\check{\Gamma}(m)$ sont asymptotiquement équivalents à $\hat{\gamma}(k)$ et $\hat{\Gamma}(m)$. Ces résultats nous permettent d'écrire

$$n^{-\frac{1}{2}}R\mathcal{S}_{ALS} = n^{\frac{1}{2}}\check{\Gamma}_m + o_p(1), \tag{3}$$

et de proposer la statistique LM suivante adaptée au cas des séries présentant une variance variable au cours du temps :

$$\check{Q}_m^{LM} = n\check{\Gamma}_m'\left(R\check{J}^{-1}R'\right)\check{\Gamma}_m,$$

où la matrice de covariance estimée est définie comme suit :

$$\left[R\check{J}^{-1}R'\right]^{-1} = \left[R\begin{pmatrix}\check{K}_\infty & \check{K}_m' \\ \check{K}_m & I_m\end{pmatrix}^{-1}R'\right]^{-1},$$

$\check{K}_m$ est obtenu en remplaçant les a_{0i} par leurs estimateurs ALS et $\check{K}_\infty = n^{-1}\sum_{t=1}^{n}\check{h}_t^{-2}\tilde{x}_{t-1}\tilde{x}_{t-1}'$. Ainsi, nous avons le résultat suivant :

Proposition 1 *Sous les hypothèses **A1** et **A2**, nous avons* $\check{Q}_m^{LM} \Rightarrow \chi_m^2$.

Nous considérons également les statistiques adaptatives du test portmanteau (introduites dans Patilea et Raïssi (2013)) dans le cadre univarié : $\check{Q}_m^{BP} = n\check{\Phi}_m'\check{\Phi}_m$, et $\check{Q}_m^{LB} = n(n+2)\sum_{k=1}^{m}(n-k)^{-1}\check{\gamma}^2(k)\check{\gamma}^{-2}(0)$. Dans la section suivante, les puissances du test LM adaptatif introduit plus haut, le test Box-Pierce et le test Ljung-Box adaptatifs sont comparées au sens de l'efficacité de Bahadur.

4 Comparaison des efficacités asymptotiques au sens de Bahadur des tests portmanteau et du test des multiplicateurs de Lagrange

Dans cette partie, nous étudions la puissance asymptotique des tests décrits ci-dessus. Nous considérons l'approche de Bahadur (1960) qui consiste à comparer les vitesses de convergence vers zéro des p-values sous des alternatives fixes $H_1 : \Gamma_m = \varrho \neq 0$ quand $n \to \infty$. Dans un tel cas, nous avons $p < p_0$. Notons par Q^B une statistique de test et pour tout $x > 0$, définissons $q_B(x) = -\log P_0(Q^B > x)$, avec P_0 correspond à la distribution limite de Q^B sous une certaine hypothèse nulle H_0. La pente de Bahadur (approchée) est donnée par $c_B(\varrho) = 2\lim_{n\to\infty} n^{-1} q_B(Q^B)$ sous l'alternative $H_1 : \Gamma_m = \varrho \neq 0$. L'efficacité relative asymptotique du test basé sur Q^B par rapport à celui basé sur Q^D, est donnée par $ARE_{B,D}(\varrho) = c_B(\varrho)/c_D(\varrho)$. Par exemple, une efficacité relative asymptotique $ARE_{B,D}(\varrho) \geq 1$ donne l'avantage au test basé sur Q^B. Pour présenter notre résultat, nous définissons $c_{BP}(\varrho)$, $c_{LB}(\varrho)$ et $c_{LM}(\varrho)$ qui sont basés sur $\check{Q}_m^{BP}$, $\check{Q}_m^{LB}$ et $\check{Q}_m^{LM}$.

Proposition 2 *Sous les hypothèses **A1** et **A2**, l'efficacité relative asymptotique du test basé sur $\check{Q}_m^{LM}$ par rapport à celui basé sur $\check{Q}_m^{BP}$ ou $\check{Q}_m^{LB}$ est supérieure ou égale à 1.*

Par conséquent, il s'avère que le test des multiplicateurs de Lagrange est plus puissant que le test portmanteau sous des alternatives non locales.

Bibliographie

[1] Bahadur, R.R. (1960). Stochastic comparison of tests. *Annals of Mathematical Statistics*, 31, 276–295.

[2] Box, G.E.P., and Pierce, D.A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65, 1509–1526.

[3] Breusch, T.S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, 17, 334–355.

[4] Doukhan, P. (1994). Mixing : Properties and Examples. *Lecture Notes in Statistics*, 85, 15–23.

[5] Godfrey, L.G. (1978). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica*, 46, 1293–1301.

[6] Ljung, G.M. and Box, G.E.P. (1978). On measure of lack of fit in time series models. *Biometrika*, 65, 297–303.

[7] Patilea, V. and Raïssi, H. (2013). Corrected portmanteau tests for VAR models with time-varying variance. *Journal of Multivariate Analysis*, 116, 190-207.

[8] Xu, K.L., and Phillips, P.C.B. (2008). Adaptive estimation of autoregressive models with time-varying variances. *Journal of Econometrics*, 142, 265-280.