

TESTS DE RANGS ADAPTATIFS POUR LES MODÈLES DE RÉGRESSION LINÉAIRE AVEC ERREURS ARMA

JELLOUL ALLAL ET SAÏD EL MELHAOUI

RÉSUMÉ. Les auteurs montrent comment tester, dans un cadre semi-paramétrique, des hypothèses sous contraintes linéaires sur le paramètre d'un modèle de régression linéaire avec erreurs ARMA. Afin de construire des tests adaptatifs, ils modifient les statistiques de rangs de Hallin et Puri (1994) en remplaçant la fonction score efficace, qui dépend de la densité sous-jacente, par une suite d'estimateurs convergents et mesurables par rapport au vecteur d'ordre des résidus. Les tests ainsi construits sont valides et asymptotiquement les plus stringents pour tester diverses hypothèses linéaires pour une classe étendue de densités. Puisque leurs tests adaptatifs sont des tests de permutation alignés, ils sont à la fois sans biais, α -semblables et conditionnellement libres.

ABSTRACT. The authors solve, in a semiparametric setting, the problem of testing hypotheses without linear constraints on the parameter of a linear regression model with ARMA innovations. In order to construct adaptative tests, they modify the rank statistics of Hallin and Puri (1994) by replacing the efficient score function, which depends on an underlying density, by a sequence of consistent estimators that are measurable with respect to the vector of ordered residuals. This leads to tests that are valid and asymptotically most stringent for testing various linear hypotheses on a large class of densities. As their adaptative tests are aligned permutation tests, they are also unbiased, α -similar and conditionally free.

1. Introduction. Les modèles de régression linéaire avec erreurs ARMA, dits aussi modèles ARMA à tendance linéaire, constituent une large et riche classe de modèles statistiques ayant de nombreuses applications. Vu son importance, cette classe de modèles a été largement étudiée ; voir par exemple Anderson [2], Box et Pierce [6], Durbin [10], Pierce [27], Judge et coll. [19], Swensen [29] ou Hallin et Puri [13].

Nous nous intéressons ici aux modèles ARMA avec tendance linéaire, mais dans un contexte semi-paramétrique. En général, les modèles semi-paramétriques incluent, d'une façon simultanée, un paramètre d'intérêt $\theta \in \mathbb{R}^{k_1}$ et un paramètre nuisible de dimension infinie f . Souvent, f représente la densité du bruit blanc (ou innovation) associé au modèle. Il est clair que ces modèles sont mieux adaptés aux situations pratiques que les modèles paramétriques où la densité f est supposée connue.

Dans les modèles semi-paramétriques, l'indétermination de f induit généralement une perte d'efficacité par rapport aux situations paramétriques. L'efficacité considérée

Reçu le 30 mai 2002 et, sous forme définitive, le 26 janvier 2005.

ici est prise à la Le Cam, au sens local et asymptotique sous une structure de normalité locale asymptotique (LAN). Cette perte d'efficacité est principalement due au fait que les éventuelles perturbations locales de θ ou de f ont un effet similaire sur la loi des observations. Lorsque cette perte est nulle, c'est-à-dire lorsqu'il existe une méthode uniformément efficace en f dans chaque sous-modèle paramétrique régi par f (un f -sous-modèle), le modèle semi-paramétrique est alors dit *adaptatif* et la méthode utilisée est dite *adaptative*. Ainsi, un test est adaptatif s'il atteint uniformément en f la puissance réalisée par le test optimal dans chaque f -sous modèle paramétrique.

L'approche répandue dans l'inférence semi-paramétrique est celle qui consiste à estimer f et à réduire localement et asymptotiquement le problème semi-paramétrique à un simple problème paramétrique, via l'argument du "*sous modèle le moins favorable*" du modèle semi-paramétrique. Dans ce cadre se situe l'ouvrage de Bickel et coll. [5], consacré aux modèles semi-paramétriques avec observations indépendantes, ainsi qu'une série d'articles traitant les modèles semi-paramétriques des séries-temporelles : Kreiss [22, 23], Jeganathan [18], Drost et coll. [7], Koul et Schick [21].

Une seconde approche à l'inférence adaptative consiste aussi à réduire le problème semi-paramétrique à un problème paramétrique, mais cette fois-ci à travers l'emploi du principe d'invariance et des statistiques mesurables en fonction des rangs des résidus. Dans ce sens, Hájek et Šidák [11] ont proposé des tests adaptatifs pour les modèles de régression, tandis que Beran [4] et Hušková et Sen [16, 17] ont respectivement étudié l'adaptativité dans les modèles de localisation et dans les modèles de régression. Par ailleurs, Allal [1] a proposé des tests adaptatifs pour le modèle ARMA et Hallin et Werker [14, 15] ont abordé l'inférence semi-paramétrique dans les modèles de séries temporelles. En particulier, ces derniers ont montré une certaine complémentarité entre deux méthodologies statistiques différentes, d'une part l'inférence semi-paramétrique et adaptative, d'autre part l'inférence basée sur les statistiques invariantes.

Dans ce travail, nous adoptons la deuxième approche pour construire des tests adaptatifs pour les modèles linéaires avec erreurs ARMA. Notre point de départ est le travail de Hallin et Puri [13], qui propose des tests de rangs alignés pour tester des hypothèses de contraintes linéaires sur le paramètre θ . Ces tests sont asymptotiquement libres et donc valides uniformément en f . Cependant, ils ne sont optimaux que si la fonction f est spécifiée. En effet, pour atteindre l'optimalité, il faut que la fonction score associée à la statistique test soit issue de la densité sous-jacente f , ce qui est impossible dans le cadre semi-paramétrique. En revanche, le même travail assure que les modèles ARMA avec tendance linéaire sont LAN avec une suite centrale libre, ce qui prouve, d'après les résultats de Hallin et Werker [15], que ces modèles sont adaptatifs.

Afin de construire des statistiques de rangs adaptatives, nous modifions les statistiques de rangs en remplaçant la fonction score efficace, qui dépend de la densité sous-jacente, par une suite d'estimateurs convergents et mesurables en terme du vecteur d'ordre des résidus. Ce choix assure l'indépendance entre l'inférence sur θ , fondée sur les rangs, et l'inférence sur f , basée sur le vecteur d'ordre. Nous évitons donc le recours à la technique usuelle mais peu commode du *partage de l'échantillon*. Nous montrons ensuite l'équivalence asymptotique entre les statistiques de rangs adaptatives et les statistiques efficaces dans chaque sous-modèle paramétrique.

Les résidus alignés, où le paramètre inconnu θ est estimé par une suite d'estimateurs convergents à l'ordre $\sqrt{n}$, sont considérés à la place des résidus exacts. Nous montrons que cette substitution n'a pas d'effet sur le comportement asymptotique de

nos statistiques adaptatives. Enfin, nous proposons une suite de tests de rangs adaptatifs, asymptotiquement et localement optimaux pour une large classe de densités f . Notons que même si nos tests adaptatifs ne sont plus de véritables tests de rangs alignés, ils restent des tests de rangs alignés conditionnels. Aussi sont-ils conditionnellement libres. De plus, ces tests sont des tests de permutations alignés et par suite, ils sont asymptotiquement sans biais et α -semblables au sens de Neyman.

Notre présentation est organisée comme suit : dans le §2, nous définissons le modèle et nous formulons les hypothèses techniques. Dans le §3, nous définissons les statistiques de rangs et nous rappelons la propriété LAN. Dans le §4, nous définissons les statistiques de rangs adaptatives, puis nous montrons l'équivalence asymptotique entre les statistiques de rangs adaptatives et les statistiques de rangs efficaces. Dans le §5, nous définissons les hypothèses linéaires, les statistiques de rangs alignés adaptatives, et nous montrons que l'alignement ne modifie pas le comportement asymptotique des statistiques de rangs adaptatives. Nous proposons ensuite des tests de rangs adaptatifs, dont nous calculons les puissances asymptotiques locales ainsi que les efficacités relatives asymptotiques par rapport aux tests gaussiens. Dans le §6, nous développons une brève étude de simulations pour deux importants exemples. Les démonstrations techniques sont reportées en annexe.

2. Notations et hypothèses techniques.

2.1. Modèle linéaire avec erreurs ARMA. Considérons le modèle linéaire

$$Y^{(n)} = X^{(n)}\beta + e^{(n)}, \quad (1)$$

où $Y^{(n)} = (Y_1^{(n)}, \dots, Y_n^{(n)})^\top$ est une série observée de taille n ,

$$X^{(n)} = \begin{pmatrix} x_{1,1}^{(n)} & \cdots & x_{1,k}^{(n)} \\ \vdots & & \vdots \\ x_{n,1}^{(n)} & \cdots & x_{n,k}^{(n)} \end{pmatrix} = \begin{pmatrix} (x_1^{(n)})^\top \\ \vdots \\ (x_n^{(n)})^\top \end{pmatrix}$$

est une suite de matrices de plein rang de dimension $n \times k$ dont les éléments sont les constantes de régression supposées connues, $\beta = (\beta_1, \dots, \beta_k)^\top$ est le vecteur des paramètres de régression, et

$$e^{(n)} = (e_1^{(n)}, \dots, e_n^{(n)})^\top$$

désigne le vecteur des erreurs (non observables). Pour $t \leq 0$, posons $x_t^{(n)} = 0$.

Contrairement à l'hypothèse classique d'indépendance, nous supposons que $e^{(n)}$ est la réalisation de longueur n de la solution stationnaire de l'équation aux différences stochastiques :

$$e_t - \sum_{i=1}^p A_i e_{t-i} = \varepsilon_t + \sum_{i=1}^q B_i \varepsilon_{t-i}, \quad t \in \mathbb{Z} \quad (2)$$

où $A_p \neq 0$, $B_q \neq 0$ et $\{\varepsilon_t : t \in \mathbb{Z}\}$ est une suite de variables aléatoires indépendantes identiquement distribuées. Les ordres p et q dans (2) sont pour le moment non spécifiés, mais nous supposons que $p \leq p_1$ et $q \leq q_1$, où p_1 et q_1 sont deux entiers donnés. Posons

$$A^\top = (A_1, \dots, A_p, 0, \dots, 0) \in \mathbb{R}^{p_1}, \quad B^\top = (B_1, \dots, B_q, 0, \dots, 0) \in \mathbb{R}^{q_1}$$

et

$$\theta^\top = (\beta^\top, A^\top, B^\top) \in \mathbb{R}^{k_1}, \quad k_1 = k + p_1 + q_1.$$

Notons $H_f^{(n)}(\theta)$ l'hypothèse selon laquelle $Y^{(n)}$ satisfait à (1) et (2), θ étant la valeur du paramètre et f étant la densité de probabilité du bruit blanc ε_t . La notation $H^{(n)}(\theta)$ est utilisée quand la densité f n'est pas spécifiée.

2.2. Hypothèses techniques. La propriété LAN, les résultats de la linéarité asymptotique ainsi que la normalité des statistique de rangs et des statistique de rangs adaptatives requièrent quelques hypothèses de régularité concernant la densité des innovations. Notons $\mathcal{F}$ la classe des fonctions de densité vérifiant les hypothèses suivantes.

Hypothèses A1

- (i) $\int x f(x) dx = 0$ et $0 < \int x^2 f(x) dx = \sigma_f^2 < \infty$.
- (ii) La densité f est absolument continue, c'est-à-dire qu'il existe une fonction $\dot{f}$ telle que pour tous $0 < a < b < \infty$,

$$f(b) - f(a) = \int_a^b \dot{f}(x) dx.$$

- (iii) Posons $\phi_f = -\dot{f}/f$ et supposons que

$$\int \{\phi_f(x)\}^4 f(x) dx < \infty;$$

cela implique que l'information de Fisher

$$\mathcal{I}(f) = \int \phi_f^2(x) f(x) dx$$

est finie. Posons de plus $\mathcal{I}(f_1) = \sigma_f^2 \mathcal{I}(f)$.

- (iv) La densité f est fortement unimodale ; autrement dit, ϕ_f est croissante.
- (v) On a $f(x) > 0$ pour tout $x \in \mathbb{R}$. De plus, $(\varepsilon_{-q_1+1}, \dots, \varepsilon_0, e_{-p_1+1}, \dots, e_0)$ possède une densité jointe non nulle $g^0(\cdot, \theta)$ qui vérifie

$$g^0(\cdot, \theta^{(n)}) - g^0(\cdot, \theta) = o_P(1)$$

sous $H_f^{(n)}(\theta)$ quand $\theta^{(n)} \rightarrow \theta$.

- (vi) La fonction ϕ_f est lipschitzienne par morceaux.

Étant donné une densité $f \in \mathcal{F}$, notons F sa fonction de répartition, F^{-1} l'inverse généralisé de F et posons $\varphi_f(u) = \phi_f \circ F^{-1}(u)$ pour tout $u \in (0, 1)$.

Concernant la partie ARMA du modèle, on considère une deuxième classe d'hypothèses qui assurent la causalité du modèle (2) sous les hypothèses nulles. Soit L l'opérateur retard,

$$A(L) = 1 - \sum_{i=1}^{p_0} A_i L^i \quad \text{et} \quad B(L) = 1 + \sum_{i=1}^{q_0} B_i L^i.$$

Notons Θ le sous ensemble des vecteurs θ de $\mathbb{R}^{k_1}$ vérifiant les hypothèses suivantes.

Hypothèses A2

- (i) $A(L)$ et $B(L)$ sont respectivement de degré p_0 et q_0 .
- (ii) Toutes les racines des polynômes caractéristiques $A(z)$ et $B(z)$, $z \in \mathbb{C}$, sont en dehors du disque unité.
- (iii) Les polynômes caractéristiques $A(z)$ et $B(z)$ n'ont pas de racines communes.

Finalement, pour la tendance linéaire du modèle, on considère les hypothèses suivantes, qui sont typiques de la théorie asymptotique des modèles linéaires.

Hypothèses A3

Définissons les matrices

$$C_i^{(n)} = (n-i)^{-1} \sum_{t=i+1}^n x_t^{(n)} (x_{t-i}^{(n)})^\top, \quad i = 0, \dots, n-1$$

et notons $D^{(n)}$ la matrice diagonale d'éléments $(C_0^{(n)})_{11}, \dots, (C_0^{(n)})_{kk}$.

- (i) Tous les éléments de la diagonale $(C_0^{(n)})_{jj}$, $j = 0, \dots, k$, sont strictement positifs.
- (ii) Posons

$$R_i^{(n)} = (D^{(n)})^{-1/2} C_i^{(n)} (D^{(n)})^{-1/2}, \quad i = 0, \dots, n-1.$$

On suppose que pour tout $i = 0, \dots, n-1$, $R_i^{(n)} - R_i \rightarrow 0$ quand $n \rightarrow \infty$; R_0 est une matrice définie positive et peut donc être mise sous la forme $R_0 = (KK^\top)^{-1}$, où K est une matrice $k \times k$ de plein rang. Posons

$$K^{(n)} = (D^{(n)})^{-1/2} K.$$

- (iii) Pour tout n et $j = 1, \dots, k$, $\bar{x}_j^{(n)} = (x_{1,j}^{(n)} + \dots + x_{n,j}^{(n)})/n = 0$.
- (iv) Les conditions classiques de Noether sont vérifiées :

$$\lim_{n \rightarrow \infty} \frac{\max_{1 \leq t \leq n} (x_{t,j}^{(n)})^2}{\sum_{t=1}^n (x_{t,j}^{(n)})^2} = 0, \quad j = 1, \dots, k.$$

- (v) $D^{(n)}/\text{tr } D^{(n)} \rightarrow D^2$ quand $n \rightarrow \infty$ et la limite est une matrice diagonale définie positive.

Les hypothèses ci-dessus sont pratiquement les mêmes que celles requises pour l'inférence paramétrique dans les modèles en question. Ainsi, même si notre approche est semi-paramétrique, elle ne nécessite aucune condition de régularité de plus.

3. Normalité asymptotique locale.

3.1. Préliminaires. Notons $Z^{(n)}(\theta) = (Z_1^{(n)}(\theta), \dots, Z_n^{(n)}(\theta))$ la série des résidus obtenue à partir des observations $Y^{(n)}$ par récurrence linéaire :

$$Z_t^{(n)}(\theta) = Y_t^{(n)} - (x_t^{(n)})^\top \beta - \sum_{i=1}^{p_0} A_i(Y_{t-i}^{(n)} - (x_{t-i}^{(n)})^\top \beta) - \sum_{i=1}^{q_0} B_i Z_{t-i}^{(n)}, \quad t = 1, \dots, n, \quad (3)$$

avec les valeurs initiales

$$Z_t^{(n)} = \varepsilon_t, \quad t = -q_0 + 1, \dots, 0, \quad Y_t^{(n)} - (x_t^{(n)})^\top \beta = e_t, \quad t = -p_0 + 1, \dots, 0.$$

Il est clair que sous l'hypothèse $H^{(n)}(\theta)$, $Z_t^{(n)}(\theta)$ coïncide avec $\varepsilon_t^{(n)}$. En pratique, les valeurs initiales $e_0, \dots, e_{-p_0+1}, \varepsilon_0, \dots, \varepsilon_{-q_0+1}$ ne sont pas observables. Des résidus approximatifs seront donc utilisés et cela ne change en rien les résultats asymptotiques ; voir le lemme 4.2 de Hallin et Puri [13].

Notons $R_t^{(n)}(\theta)$ ou simplement $R_t^{(n)}$ le rang de $Z_t^{(n)}(\theta)$ parmi $Z_1^{(n)}(\theta), \dots, Z_n^{(n)}(\theta)$. Considérons maintenant les vecteurs de dimension k des statistiques de rangs non sérielles introduites dans Hallin et Puri [13] :

$$t_{i,f}^{(n)}(\theta) = \frac{1}{n-i} \{ \mathcal{J}(f) \}^{-1/2} \sum_{t=i+1}^n \left\{ \varphi_f \left(\frac{R_t^{(n)}}{n+1} \right) - \bar{\varphi}_f^{(n)} \right\} (K^{(n)})^\top x_{t-i}^{(n)},$$

pour $i = 0, \dots, n-1$, où

$$\bar{\varphi}_f^{(n)} = \frac{1}{n} \sum_{t=1}^n \varphi_f \left(\frac{t}{n+1} \right),$$

Considérons aussi les statistiques de rang sérielles, dites *coefficients d'autocorrélation de rangs*, introduites dans Hallin et coll. [12]. Comme le montrent Dufour et Roy [8, 9], celles-ci jouent le même rôle que les coefficients d'autocorrélation usuels dans le contexte gaussien. Pour $i = 0, \dots, n-1$, on pose donc

$$r_{i,f}^{(n)}(\theta) = \frac{1}{\sigma_{i,f}^{(n)}} \left\{ \frac{1}{n-i} \sum_{t=i+1}^n \varphi_f \left(\frac{R_t^{(n)}}{n+1} \right) F^{-1} \left(\frac{R_{t-i}^{(n)}}{n+1} \right) - m_f^{(n)} \right\},$$

où

$$m_f^{(n)} = \frac{1}{n(n-1)} (S_{10}^{(n)} S_{01}^{(n)} - S_{11}^{(n)}) \quad (4)$$

et

$$\begin{aligned} (\sigma_{i,f}^{(n)})^2 = & \frac{1}{n(n-1)} (S_{20}^{(n)} S_{02}^{(n)} - S_{22}^{(n)}) + \frac{2(n-2i)^+}{n(n-1)(n-2) \times (n-i)} \times \\ & \{ S_{10}^{(n)} S_{01}^{(n)} S_{11}^{(n)} - S_{21}^{(n)} S_{01}^{(n)} - S_{12}^{(n)} S_{10}^{(n)} - (S_{11}^{(n)})^2 + 2S_{22}^{(n)} \} \\ & + \frac{(n-i)(n-i-1) - 2(n-2i)^+}{n(n-1)(n-2) \times (n-3)(n-i)} \times \\ & \{ (S_{10}^{(n)} S_{01}^{(n)})^2 + 2(S_{11}^{(n)})^2 + S_{20}^{(n)} S_{02}^{(n)} - 6S_{22}^{(n)} - 4S_{11}^{(n)} S_{10}^{(n)} S_{01}^{(n)} - S_{20}^{(n)} (S_{01}^{(n)})^2 \\ & - S_{02}^{(n)} (S_{10}^{(n)})^2 + 4S_{21}^{(n)} S_{01}^{(n)} + 4S_{12}^{(n)} S_{10}^{(n)} \} - (n-i)(m_f^{(n)})^2, \end{aligned} \quad (5)$$

avec $(\cdot)^+ = \max(\cdot, 0)$ et

$$S_{\ell m}^{(n)} = \sum_{i=1}^n \left\{ \phi_f \left(\frac{i}{n+1} \right) \right\}^\ell \left\{ F^{-1} \left(\frac{i}{n+1} \right) \right\}^m, \quad \ell, m = 0, 1, 2.$$

Soient g_u , h_u et c_u , $u = 0, 1, \dots$ les fonctions de Green respectivement associées aux opérateurs $A(L)$, $B(L)$ et $A(L)(B(L))^{-1}$. Soit $\{\psi_t^{(1)}, \dots, \psi_t^{(p_0+q_0)}\}$ un système fondamental des solutions de l'équation homogène

$$A(L)B(L)\psi_t = 0, \quad t \in \mathbb{Z},$$

des choix convenables de ce système sont donnés dans le §4 de Hallin et Puri [13].

Posons $\pi = \max(p_1 - p_0, q_1 - q_0)$, $\pi_0 = \pi + p_0 + q_0$, et définissons les matrices suivantes :

$$C_\psi = \begin{pmatrix} \psi_{\pi+1}^{(1)} & \cdots & \psi_{\pi+1}^{(p_0+q_0)} \\ \vdots & & \vdots \\ \psi_{\pi_0}^{(1)} & \cdots & \psi_{\pi_0}^{(p_0+q_0)} \end{pmatrix},$$

$$M(\theta) = \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ g_1 & 1 & \cdots & & h_1 & 1 & \cdots & \\ \vdots & & & \vdots & \vdots & & & \vdots \\ g_{p_1-1} & \cdots & \cdots & 1 & h_{q_1-1} & \cdots & \cdots & 1 \\ g_{p_1} & \cdots & \cdots & g_1 & h_{q_1} & \cdots & \cdots & h_1 \\ \vdots & & & \vdots & \vdots & & & \vdots \\ g_{\pi_0-1} & \cdots & \cdots & g_{\pi_0-p_1} & h_{\pi_0-1} & \cdots & \cdots & h_{\pi_0-q_1} \end{pmatrix}$$

et w_ψ^2 la matrice de taille $(p_0 + q_0) \times (p_0 + q_0)$ ayant pour éléments

$$(w_\psi^2)_{i,j} = \sum_{t=\pi+1}^{\infty} \psi_t^{(i)} \psi_t^{(j)}, \quad i, j = 1, \dots, p_0 + q_0.$$

Finalement, considérons la statistique de rangs de dimension $\pi_0 + k$ définie par

$$n^{1/2} T_{\psi,f}^{(n)}(\theta) = \begin{pmatrix} \sum_{i=1}^{n-1} (n-i)^{1/2} c_{i,f}^{(n)}(\theta) \\ (n-1)^{1/2} r_{1,f}^{(n)}(\theta) \\ \vdots \\ (n-\pi)^{1/2} r_{\pi,f}^{(n)}(\theta) \\ \sum_{i=\pi+1}^{n-1} (n-i)^{1/2} \psi_i^{(1)} r_{i,f}^{(n)}(\theta) \\ \vdots \\ \sum_{i=\pi+1}^{n-1} (n-i)^{1/2} \psi_i^{(p_0+q_0)} r_{i,f}^{(n)}(\theta) \end{pmatrix} = n^{1/2} \begin{pmatrix} T_{I,f}^{(n)}(\theta) \\ T_{II,\psi,f}^{(n)}(\theta) \end{pmatrix}. \quad (6)$$

Notons que $T_{\psi,f}(\theta)$ est la statistique sur laquelle se basent les tests de rangs pour les modèles ARMA avec tendance linéaire. La statistique $T_{I,f}(\theta)$ de dimension k est une statistique non sérielle liée à la partie linéaire du modèle, tandis que la statistique

$T_{II,f}(\theta)$ de dimension π_0 est une statistique sérielle associée à la partie ARMA du modèle. Nous montrons dans la suite qu'une statistique adaptative provient d'une transformation adéquate de la statistique (6).

3.2. Propriété LAN basée sur les rangs. Soient $\eta^{(n)}, \gamma^{(n)}$ et $\delta^{(n)}$ des suites de vecteurs réels ayant respectivement pour dimension k, p_1 et q_1 . Posons

$$\tau^{(n)} = \begin{pmatrix} \eta^{(n)} \\ \gamma^{(n)} \\ \delta^{(n)} \end{pmatrix} \quad \text{et} \quad \nu(n) = n^{-1/2} \begin{pmatrix} \sigma_f K^{(n)} & 0_{k \times (p_1+q_1)} \\ 0_{(p_1+q_1) \times k} & I_{(p_1+q_1) \times (p_1+q_1)} \end{pmatrix}.$$

De plus, soient

$$\theta^{(n)} = \theta + \nu(n) \tau^{(n)}$$

et $\Lambda_{f;\theta^{(n)}/\theta}^{(n)}$ le logarithme du rapport de vraisemblance de $H_f^{(n)}(\theta^{(n)})/H_f^{(n)}(\theta)$. Sous quelques hypothèses de régularité, Hallin et Puri [13] ont établi la propriété suivante.

Propriété LAN

Pour toute suite bornée $\tau^{(n)}$, sous l'hypothèse $H_f^{(n)}(\theta)$,

$$\Lambda_{f;\theta^{(n)}/\theta}^{(n)}(Y^{(n)}) = \tau^{(n)\top} \Delta_f^{(n)}(\theta) - 1/2 \tau^{(n)\top} \Gamma_f(\theta) \tau^{(n)} + o_P(1)$$

quand $n \rightarrow \infty$ et

$$De_f^{(n)}(\theta) \xrightarrow{\mathcal{L}} \mathcal{N}[0, \Gamma_f(\theta)],$$

avec

$$\Delta_f^{(n)}(\theta) = \{n\mathcal{J}(f_1)\}^{1/2} \begin{pmatrix} T_{I;f}^{(n)}(\theta) \\ M^\top(\theta) \begin{pmatrix} I_{\pi \times \pi} & 0 \\ 0 & C_\psi^\top \end{pmatrix}^{-1} T_{II;\psi,f}^{(n)}(\theta) \end{pmatrix},$$

$$\Gamma_f(\theta) = \mathcal{J}(f_1) \begin{pmatrix} \Gamma_I(\theta) & 0 \\ 0 & \Gamma_{II}(\theta) \end{pmatrix}$$

où

$$\Gamma_I(\theta) = K^\top \left(\sum_{u=0}^{\infty} \sum_{v=0}^{\infty} c_u c_v R_{|u-v|} \right) K,$$

$$\Gamma_{II}(\theta) = M^\top(\theta) \begin{pmatrix} I_{\pi \times \pi} & 0 \\ 0 & (C_\psi^{-1})^\top w_\psi^2 C_\psi^{-1} \end{pmatrix} M(\theta).$$

Parmi les importants conséquences de cette propriété, citons :

- (i) La contiguïté mutuelle des deux suites d'hypothèses $H_f^{(n)}(\theta^{(n)})$ et $H_f^{(n)}(\theta)$; voir le Théorème 1 de Le Cam [25].
- (ii) Sous la suite de contre-hypothèses locales $H_f^{(n)}(\theta^{(n)})$, lorsque $n \rightarrow \infty$,

$$\Delta_f^{(n)}(\theta) - \Gamma_f(\theta) \tau^{(n)} \xrightarrow{\mathcal{L}} \mathcal{N}[0, \Gamma_f(\theta)];$$

voir le troisième lemme de Le Cam dans Hájek et Šidák [11].

- (iii) La convergence faible des suites des expériences statistiques étudiées vers l'expérience de position gaussienne ; voir Le Cam [25]. Cette convergence a une importante signification décisionnelle puisqu'elle implique que les procédures statistiques (tests, estimation, etc.) ayant de bonnes propriétés dans des expériences gaussiennes possèdent également ces bonnes propriétés dans les expériences considérées initialement.

Pour les problèmes de tests, les résultats (i)–(iii) aident à la construction d'une suite de tests de rangs $\Delta_f^{(n)}(\theta)$ -mesurables. Ces tests sont libres et par conséquent ils sont valides sous les hypothèses semi-paramétriques $H^{(n)}(\theta)$. Toutefois, ils ne sont asymptotiquement optimaux que sous la contre-hypothèse paramétrique particulière $H_f^{(n)}(\theta^{(n)})$. En effet, leur efficacité est affectée quand la densité f utilisée dans la construction de la fonction score φ_f ne coïncide pas avec la vraie densité g non spécifiée qui régit la distribution des innovations. Par conséquent, ces tests de rangs ne sont pas adaptatifs.

En revanche, la propriété LAN ci-dessus jouit d'une caractéristique déterminante pour l'inférence semi-paramétrique. En effet, pour tout $(\theta, f) \in \Theta \times \mathcal{F}$ le f -sous-modèle paramétrique est LAN avec une suite centrale $\Delta_f^{(n)}(\theta)$ mesurable par rapport aux rangs, et donc libre sous l'hypothèse semi-paramétrique $H^{(n)}(\theta)$. D'après la proposition 2.4 de Hallin et Werker [15], ceci est suffisant pour affirmer que le modèle semi-paramétrique est adaptatif.

En conclusion, il est possible de construire des tests qui seront valides et optimaux pour tester l'hypothèse nulle $H^{(n)}(\theta)$ par rapport à une contre-hypothèse semi-paramétrique locale $H^{(n)}(\theta^{(n)})$.

4. Statistiques de rangs adaptatives. Bien qu'elle soit inconnue sous l'hypothèse semi-paramétrique $H^{(n)}(\theta)$, densité sous-jacente f intervient dans la statistique test efficace ; celle-ci est $\Delta_f^{(n)}$ -mesurable et donc $T_{\psi, f}^{(n)}(\theta)$ -mesurable, à travers la fonction score $\varphi_f(\cdot)$. Pour cette raison, nous envisageons de remplacer cette dernière par une suite d'estimateurs construits à partir des observations. Considérons donc $\tilde{\varphi}^{(n)}(\cdot)$ une suite d'estimateurs de la fonction $\varphi_f(\cdot)$ vérifiant les hypothèses suivantes.

Hypothèses A4

- (i) la suite d'estimateurs $\tilde{\varphi}^{(n)}(\cdot)$ est convergente ; en d'autres termes, pour tous $f \in \mathcal{F}$ et $\varepsilon > 0$, sous $H_f^{(n)}(\theta)$,

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left[\int_0^1 \{ \tilde{\varphi}^{(n)}(u) - \varphi_f(u) \}^2 du > \varepsilon \right] = 0; \quad (7)$$

- (ii) La suite d'estimateurs $\tilde{\varphi}^{(n)}(\cdot)$ est $Z_{(\cdot)}$ -mesurable, où $Z_{(\cdot)} = (Z_{(1)}, \dots, Z_{(n)})^\top$ est le vecteur d'ordre associée à la série des résidus $Z^{(n)}$.

L'hypothèse (i), qui assure la convergence de la suite d'estimateurs $\tilde{\varphi}^{(n)}$ vers φ_f , est naturelle dans l'inférence semi-paramétrique et adaptative ; voir par exemple la définition VII.1.5 de Hájek et Šidák [11]. L'hypothèse (ii) assure l'indépendance entre l'inférence sur θ , basée sur les rangs, et l'inférence sur f , basée sur le vecteur d'ordre. Cette indépendance est habituellement acquise par le recours peu commode à la technique du *partage de l'échantillon* ; voir Bickel et coll. [5] ou Koul et Schik [21]. Cette technique

artificielle est évitée ici. En fait, (ii) n'est pas vraiment très restrictive, puisqu'elle est vérifiée par la plupart des estimateurs d'une fonctionnelle de densité. En effet, le vecteur d'ordre des innovations contient l'information complète sur leur distribution ; voir Lehmann [26, p. 143].

Des formes explicites de $\tilde{\varphi}^{(n)}$ sont fournies par Hájek et Šidák [11, p. 260], Beran [4] et Kim et Van Ryzin [20]. En général, d'autres exemples de $\tilde{\varphi}^{(n)}$ peuvent être construits, notamment en fonction des estimateurs à noyau.

Dans la suite, les statistiques de rangs dans lesquelles les fonctions scores sont estimées seront dites *statistiques de rangs adaptatives*. Pour $i = 0, \dots, n-1$, considérons les vecteurs des statistiques de rangs adaptatives :

$$\tilde{r}_i^{(n)} = \frac{1}{n-i} (\tilde{I}^{(n)})^{-1} \sum_{t=i+1}^n \left\{ \tilde{\varphi}^{(n)} \left(\frac{R_t^{(n)}}{n+1} \right) - \bar{\varphi}^{(n)} \right\} (K^{(n)})^\top x_{t-i}^{(n)},$$

où

$$\bar{\varphi}^{(n)} = \frac{1}{n} \sum_{t=1}^n \tilde{\varphi}^{(n)} \left(\frac{t}{n+1} \right),$$

$$(\tilde{I}^{(n)})^2 = \frac{1}{n-1} \sum_{t=1}^n \left\{ \tilde{\varphi}^{(n)} \left(\frac{t}{n+1} \right) - \bar{\varphi}^{(n)} \right\}^2.$$

Les statistiques de rangs sérielles adaptatives sont définies par

$$\tilde{r}_i^{(n)} = \frac{1}{\tilde{\sigma}_i^{(n)}} \left\{ \frac{1}{n-i} \sum_{t=i+1}^n \tilde{\varphi}_n \left(\frac{R_t^{(n)}}{n+1} \right) Z_{t-i}^{(n)} - \tilde{m}^{(n)} \right\}, \quad i = 0, \dots, n-1,$$

où

$$\tilde{m}^{(n)} = \frac{1}{n(n-1)} (\tilde{S}_{10}^{(n)} \tilde{S}_{01}^{(n)} - \tilde{S}_{11}^{(n)}), \quad (8)$$

$$\begin{aligned} (\tilde{\sigma}_i^{(n)})^2 = & \frac{1}{n(n-1)} (\tilde{S}_{20} \tilde{S}_{02}^{(n)} - \tilde{S}_{22}^{(n)}) + \frac{2(n-2i)^+}{n(n-1)(n-2)(n-i)1} \times \\ & \{ \tilde{S}_{10}^{(n)} \tilde{S}_{01}^{(n)} \tilde{S}_{11}^{(n)} - \tilde{S}_{21}^{(n)} \tilde{S}_{01}^{(n)} - \tilde{S}_{12}^{(n)} (\tilde{S}_{10}^{(n)})^2 - (\tilde{S}_{11}^{(n)})^2 + 2\tilde{S}_{22}^{(n)} \| \\ & + \frac{(n-i)(n-i-1) - 2(n-2i)^+}{n(n-1)(n-2)(n-3)(n-i)} \times \\ & \{ (\tilde{S}_{10}^{(n)} \tilde{S}_{01}^{(n)})^2 - 2(\tilde{S}_{11}^{(n)})^2 + \tilde{S}_{20}^{(n)} \tilde{S}_{02}^{(n)} - 6\tilde{S}_{22}^{(n)} - 4\tilde{S}_{11}^{(n)} \tilde{S}_{10}^{(n)} \tilde{S}_{01}^{(n)} - \tilde{S}_{20}^{(n)} (\tilde{S}_{01}^{(n)})^2 \\ & - \tilde{S}_{02}^{(n)} (\tilde{S}_{10}^{(n)})^2 + 4\tilde{S}_{21}^{(n)} \tilde{S}_{01}^{(n)} + 4\tilde{S}_{12}^{(n)} \tilde{S}_{10}^{(n)} \} - (n-i)(\tilde{m}^{(n)})^2, \end{aligned} \quad (9)$$

avec

$$\tilde{S}_{lm}^{(n)} = \sum_{i=1}^n \left\{ \tilde{\varphi}^{(n)} \left(\frac{i}{n+1} \right) \right\}^\ell \{Z_{(i)}\}^m, \quad \ell, m = 0, 1, 2.$$

Pour prouver l'équivalence asymptotique entre les statistiques de rangs adaptatives et les statistiques paramétriquement efficaces, nous avons d'abord besoin du lemme suivant.

Lemme 1. *Supposons que les hypothèses A1–A4 soient satisfaites. Alors, sous $H_f^{(n)}(\theta)$, lorsque $n \rightarrow \infty$,*

- (i) $\bar{\Phi}^{(n)} - \bar{\Phi}_f^{(n)} = o_P(1)$;
- (ii) $(\tilde{I}^{(n)})^2 - \mathcal{J}(f) = o_P(1)$;
- (iii) $n^{1/2}(\tilde{m}^{(n)} - m_f^{(n)}) = o_P(1)$;
- (iv) $\tilde{\sigma}_i^{(n)} - \sigma_{i,f}^{(n)} = o_P(1)$, $i = 1, \dots, n-1$.

Maintenant, énonçons le résultat concernant l'équivalence asymptotique entre les statistiques de rangs adaptatives et les statistiques de rangs.

Proposition 1. *Supposons que les hypothèses A1–A4 soient satisfaites. Alors, sous $H_f^{(n)}(\theta)$, lorsque $n \rightarrow \infty$,*

- (i) $n^{1/2}\{t_{i,f}^{(n)}(\theta) - \tilde{t}_i^{(n)}(\theta)\} = o_P(1)$, $i = 1, \dots, n-1$;
- (ii) $n^{1/2}\{r_{i,f}^{(n)}(\theta) - \tilde{r}_i^{(n)}(\theta)\} = o_P(1)$, $i = 1, \dots, n-1$.

Les démonstrations du lemme et de la proposition sont fournies en annexe.

À présent, nous définissons la statistique de rangs adaptative $\tilde{T}_\psi^{(n)}$. Pour ce faire, nous remplaçons dans (6) les statistiques $t_{i,f}^{(n)}$ et $r_{i,f}^{(n)}$ par leur version adaptative respective $\tilde{t}_i^{(n)}$ et $\tilde{r}_i^{(n)}$. Nous sommes maintenant en mesure d'énoncer le résultat clé pour la dérivation des tests adaptatifs.

Proposition 2. *Supposons que les hypothèses A1–A4 soient satisfaites. Alors, sous $H_f^{(n)}(\theta)$, lorsque $n \rightarrow \infty$,*

$$n^{1/2}\{T_{\psi,f}^{(n)}(\theta) - \tilde{T}_\psi^{(n)}(\theta)\} = o_P(1).$$

Démonstration. Montrons d'abord que

$$n^{1/2}\{T_{I,f}^{(n)}(\theta) - \tilde{T}_I^{(n)}(\theta)\} = o_P(1). \quad (10)$$

Pour cela on procède par composante, soit $t_{i,f,\ell}^{(n)}$ la ℓ -ième composante de $t_{i,f}^{(n)}$, $\tilde{t}_{i,\ell}^{(n)}$ la ℓ -ième composante de $\tilde{t}_i^{(n)}$, avec $1 \leq \ell \leq k$ et soit p un entier naturel tel que $2 \leq p < n-1$. On considère la décomposition suivante :

$$\begin{aligned} & \sum_{i=1}^{n-1} (n-i)^{1/2} c_i(t_{i,f,\ell}^{(n)} - \tilde{t}_{i,\ell}^{(n)}) \\ &= \sum_{i=1}^p (n-i)^{1/2} c_i(t_{i,f,\ell}^{(n)} - \tilde{t}_{i,\ell}^{(n)}) + \sum_{i=p+1}^{n-1} (n-i)^{1/2} c_i(t_{i,f,\ell}^{(n)} - \tilde{t}_{i,\ell}^{(n)}). \end{aligned} \quad (11)$$

D'après la proposition 1(i), on sait que quand $n \rightarrow \infty$,

$$\sum_{i=1}^p (n-i)^{1/2} c_i(t_{i,f,\ell}^{(n)} - \tilde{t}_{i,\ell}^{(n)}) = o_P(1).$$

Pour montrer que le terme de droite dans (11) est un $o_P(1)$ lorsque $n \rightarrow \infty$, il suffit de montrer que

$$\sum_{i=p+1}^{n-1} (n-i)^{1/2} c_i(t_{i,f,\ell}^{(n)} - \tilde{t}_{i,\ell}^{(n)}) = o_P(1) \quad (12)$$

uniformément en n quand $p \rightarrow \infty$ et d'appliquer le théorème d'Anderson [3] dans sa version non dégénérée. Pour cela, nous calculons les moyennes quadratiques conditionnelles sachant $Z_{(\cdot)}$:

$$\begin{aligned} \mathbb{E} \left[\frac{1}{Z_{(i)}} \left\{ \sum_{i=p+1}^{n-1} (n-i)^{1/2} c_i \tilde{t}_{i,l}^{(n)} \right\}^2 \right] &= \sum_{i=p+1}^{n-1} c_i^2 \mathbb{E} \{ (n-i) (\tilde{t}_{i,l}^{(n)})^2 / Z_{(\cdot)} \} \\ &\quad + \sum_{i \neq j}^{n-1} c_i c_j \mathbb{E} \{ (n-i)^{1/2} \tilde{t}_{i,l}^{(n)} (n-j)^{1/2} \tilde{t}_{j,l}^{(n)} / Z_{(\cdot)} \} \\ &\leq \sum_{i=p+1}^{n-1} c_i^2 + \sum_{i \neq j}^{n-1} |c_i c_j| = \left(\sum_{i=p+1}^{n-1} |c_i| \right)^2. \end{aligned}$$

Or, la suite (c_i) décroît d'une façon exponentielle vers 0, de sorte que $\sum_{i=p+1}^{n-1} |c_i|$ converge vers 0 quand $p \rightarrow \infty$ uniformément en n . Ainsi,

$$\sum_{i=p+1}^{n-1} (n-i)^{1/2} c_i \tilde{t}_{i,l}^{(n)} = o_P(1) \quad (13)$$

lorsque $p \rightarrow \infty$ uniformément en n . De la même façon, on montre que

$$\sum_{i=p+1}^{n-1} (n-i)^{1/2} c_i t_{i,f,l}^{(n)} = o_P(1) \quad (14)$$

lorsque $p \rightarrow \infty$ uniformément en n . Le résultat (12) découle de (13) et (14), ce qui prouve (10). À présent montrons que lorsque $n \rightarrow \infty$,

$$n^{1/2} \{ T_{II,\psi,f}^{(n)}(\theta) - \tilde{T}_{II,\psi}^{(n)}(\theta) \} = o_P(1). \quad (15)$$

La proposition 1(ii) assure que pour tout $j \in \{1, \dots, \pi\}$,

$$(n-j)^{1/2} (r_{j,f}^{(n)} - \tilde{r}_j^{(n)}) = o_P(1) \quad \text{lorsque } n \rightarrow \infty. \quad (16)$$

Aussi, par une décomposition similaire à la décomposition (11), on montre que pour tout $j \in \{1, \dots, p_0 + q_0\}$,

$$\sum_{i=\pi+1}^{n-1} (n-i)^{1/2} \psi_i^{(j)} \{ r_{i,f}^{(n)}(\theta) - \tilde{r}_i^{(n)}(\theta) \} = o_P(1) \quad \text{lorsque } n \rightarrow \infty. \quad (17)$$

Le résultat (15) découle de (16) et (17), ce qui achève la démonstration. $\square$

5. Tests de rangs adaptatifs.

5.1. Hypothèses linéaires et problème d'alignement. Notre approche aux problèmes de tests dans la classe des modèles linéaires avec erreurs ARMA serait d'un intérêt modeste pour la pratique si nous nous limitons aux tests d'hypothèses nulle $H^{(n)}(\theta)$, où la valeur du paramètre θ est complètement spécifiée. Les vrais problèmes rencontrés dans la pratique consistent généralement à tester des contraintes linéaires sur θ , comme par exemple dans les problèmes d'identification de l'ordre de la part ARMA du modèle (voir Pötscher [28]) ou dans les problèmes liés à la détection d'autocorrélation entre les erreurs d'un modèle de régression linéaire. Pour cela, nous considérons des hypothèses

nulles plus générales dites *hypothèses linéaires*. Sous ces hypothèses, le paramètre θ est supposé vérifier un nombre fini de restrictions linéaires.

Notons $\mathcal{M}(E)$ le sous-espace vectoriel de $\mathbb{R}^n$ engendré par les colonnes d'une matrice E d'ordre $n \times m$, $\mathcal{M}(E)_\perp$ son espace orthogonal par rapport à $\mathbb{R}^n$ muni du produit scalaire usuel et $E_\perp$ une base arbitraire de $\mathcal{M}(E)_\perp$.

Soit Ω une matrice de dimension $k_1 \times (k_0 + \pi_0 - r)$, où $\pi \leq r \leq k + \pi_0$, qui contient des zéros dans les lignes $k + p_0 + 1, \dots, k + p_1$ et $k + p_1 + q_0 + 1, \dots, k + p_1 + q_1 = k_1$. Nous supposons que Ω est de rang $k + \pi_0 - r$. Si θ_0 est un élément particulier de Θ , on note $H_f^{(n)}(\Omega; \theta_0)$ l'hypothèse nulle selon laquelle $Y^{(n)}$ est générée par (1) et (2), avec f est la densité du bruit blanc et θ vérifiant

$$\theta - \theta_0 \in \mathcal{M}(\Omega). \quad (18)$$

Lorsque la densité f est non spécifiée, nous utilisons la notation $H^{(n)}(\Omega; \theta_0)$.

Nous supposons que les matrices Ω et $\Omega'_\perp$ ont une forme diagonale spéciale et nous reprenons les hypothèses A5 de Hallin et Puri [13].

Hypothèse A5

Les matrices Ω et $\Omega'_\perp$ sont de la forme

$$\Omega = \begin{pmatrix} \Omega_I & 0 \\ 0 & \Omega_{II} \end{pmatrix} \text{ et } \Omega'_\perp = \begin{pmatrix} \Omega'_{I,\perp} & 0 \\ 0 & \Omega'_{II,\perp} \end{pmatrix},$$

où Ω_I , Ω_{II} , $\Omega'_{I,\perp}$ et $\Omega'_{II,\perp}$ sont respectivement de dimension $k \times (k - r_1)$, $(p_1 + q_1) \times (\pi_0 - r_2)$, $r_1 \times k$ et $\{r_2 + \min(p_1 - p_0, q_1 - q_0)\} \times (p_1 + q_1)$, où $\pi \leq r_2 \leq \pi_0$, $r_1 + r_2 = r$.

Comme sous l'hypothèse nulle $H^{(n)}(\Omega; \theta_0)$, la valeur de θ est non spécifié, on utilise une suite $\hat{\theta}^{(n)}$ d'estimateurs de θ tels que

Hypothèse A6

- (i) Pour tout $n \in \mathbb{N}$, $\hat{\theta}^{(n)}$ satisfait à (18);
- (ii) $\hat{\theta}^{(n)}$ est $v(n)$ -convergent; autrement dit, pour tout $f \in \mathcal{F}$, tout θ vérifiant (18) et tout $\varepsilon > 0$, il existe $b(f, \theta, \varepsilon)$ et $N = N(f, \theta, \varepsilon)$ tels que sous $H_f^{(n)}(\theta)$

$$P(\|v^{(n)}(\hat{\theta}^{(n)} - \theta) > b\|) < \varepsilon, \quad \forall n \geq N.$$

- (iii) $\hat{\theta}^{(n)}$ est localement asymptotiquement discret au sens de Kreiss [22].

L'hypothèse (ii) est une condition de taux de convergence; elle remplace l'hypothèse usuelle (faite par exemple, dans les tests de multiplicateurs de Lagrange) où les estimateurs de θ sont les estimateurs à vraisemblance maximale. Il existe plusieurs estimateurs qui vérifient l'hypothèse (ii); à titre d'exemple citons les estimateurs des moindres carrés, les M-estimateurs et les estimateurs de Yule-Walker, etc. L'hypothèse (iii) est typique dans le contexte LAN; voir Kreiss [22]. Elle est cependant sans aucune conséquence pratique, puisque tout estimateur peut être transformé par discrétisation en un estimateur localement asymptotiquement discret; voir Le Cam [24].

En remplaçant le paramètre θ par $\hat{\theta}$ dans (3), nous obtenons les résidus alignés $\hat{Z}_t$, dont les rangs $\hat{R}_t$ sont dits *les rangs alignés*. La statistique $\tilde{T}_{\psi, f}(\hat{\theta})$ est la statistique calculée, cette fois-ci, à partir des résidus et des rangs alignés au lieu des résidus et

des rangs exacts. La proposition suivante montre que la substitution de $\hat{\theta}$ par θ n'a pas d'effet sur le résultat de la proposition 2.

Proposition 3. *Supposons que les hypothèses A1–A3 soient satisfaites et que $\tilde{\varphi}^{(n)}(\cdot)$, Ω et $\hat{\theta}^{(n)}$ vérifient respectivement les hypothèses A4, A5 et A6. Alors, sous $H_f^{(n)}(\Omega, \theta_0)$, lorsque $n \rightarrow \infty$,*

$$n^{1/2}\{T_{\psi,f}^{(n)}(\hat{\theta}^{(n)}) - \tilde{T}_{\psi}^{(n)}(\hat{\theta}^{(n)})\} = o_P(1).$$

Démonstration. La proposition 2 implique que sous l'hypothèse $H_f^{(n)}(\theta^{(n)})$,

$$n^{1/2}\{T_{\psi,f}^{(n)}(\theta^{(n)}) - \tilde{T}_{\psi}^{(n)}(\theta^{(n)})\} = o_P(1)$$

lorsque $n \rightarrow \infty$. Or puisque $H_f^{(n)}(\theta^{(n)})$ et $H_f^{(n)}(\theta)$ sont mutuellement contiguës alors, sous l'hypothèse $H_f^{(n)}(\theta)$,

$$n^{1/2}\{T_{\psi,f}^{(n)}(\theta^{(n)}) - \tilde{T}_{\psi}^{(n)}(\theta^{(n)})\} = o_P(1).$$

quand $n \rightarrow \infty$. Le lemme 4.4 de Kreiss [22] permet de conclure. $\square$

5.2. Tests de rangs alignés adaptatifs. Les tests de rangs alignés définis dans la proposition 5.10 de Hallin et Puri [13] jouissent asymptotiquement des bonnes propriétés suivantes : l'invariance, la liberté, l'absence du biais, la α -similarité au sens de Neyman et l'optimalité dans les modèles paramétriques. Cependant, ces tests dépendent de la densité sous-jacente f ; aussi perdent-ils leur efficacité lorsque la densité sous-jacente est non spécifiée.

Les tests que nous proposons ne dépendent pas de la densité sous-jacente et ils sont uniformément optimaux pour une large classe de densité $f \in \mathcal{F}$. Notons que même si nos tests ne sont plus de véritables tests de rangs alignés, ils restent des tests de rangs alignés conditionnels et ils appartiennent à la classe des tests de permutation alignés. Ils sont donc libres conditionnellement au vecteur d'ordre, sans biais et α -semblables. La proposition suivante résume les résultats principaux de ce papier.

Proposition 4. *Supposons que les hypothèses A1–A3 soient satisfaites sous $H^{(n)}(\Omega, \theta_0)$ et que $\tilde{\varphi}^{(n)}(\cdot)$, Ω et $\hat{\theta}^{(n)}$ vérifient respectivement les hypothèses A4, A5 et A6. Notons $\mathring{M}(\theta)$ la matrice $(\pi + p_0 + q_0) \times (p_0 + q_0)$ de plein rang obtenue à partir de $M(\theta)$ par la suppression des colonnes $p_0 + 1, \dots, p_1$ et $p_1 + q_0 + 1, \dots, p_1 + q_1$. De même, notons $\mathring{\Omega}_{II}$ la matrice $(p_0 + q_0) \times (p_1 + q_1 - r_2)$ obtenue à partir de Ω_{II} par la suppression des lignes $p_0 + 1, \dots, p_1$ et $p_1 + q_0 + 1, \dots, p_1 + q_1$. Soit*

$$Q^{(n)}(\theta) = \begin{pmatrix} Q_I^{(n)}(\theta) & 0 \\ 0 & Q_{II}^{(n)}(\theta) \end{pmatrix}$$

avec

$$Q_I^{(n)}(\theta) = \Gamma_I^{-1}(\theta) - (K^{(n)})^{-1} \Omega_I [\Omega_I' \{ (K^{(n)})^{-1} \}^\top \Gamma_I(\theta) (K^{(n)})^{-1} \Omega_I]^{-1} \Omega_I' \{ (K^{(n)})^{-1} \}^\top,$$

et

$$Q_{II}^{(n)}(\theta) = \begin{pmatrix} I & 0 \\ 0 & w_\psi^2 \end{pmatrix}^{-1} - \begin{pmatrix} I & 0 \\ 0 & C_\psi \end{pmatrix}^{-1} \mathring{M}(\theta) \mathring{\Omega}_{II}$$

$$\times \left[\ddot{\Omega}'_{II} \ddot{M}'(\theta) \begin{pmatrix} I & 0 \\ 0 & C_{\psi} w_{\psi}^{-2} C_{\psi'} \end{pmatrix}^{-1} \ddot{M}(\theta) \ddot{\Omega}_{II} \right]^{-} \ddot{\Omega}'_{II} \ddot{M}'(\theta) - \begin{pmatrix} I & 0 \\ 0 & C'_{\psi} \end{pmatrix}^{-1}.$$

Considérons la suite de tests de rangs alignés adaptatifs $\tilde{\phi}^{(n)}$ rejetant $H^{(n)}(\Omega, \theta_0)$ quand

$$n \tilde{T}_{\tilde{\psi}}^{(n)}(\hat{\theta}^{(n)})^{\top} Q^{(n)}(\hat{\theta}^{(n)}) \tilde{T}_{\tilde{\psi}}^{(n)}(\hat{\theta}^{(n)}) > \chi_{r;1-\alpha}^2, \quad (19)$$

où $\chi_{r;1-\alpha}^2$ désigne le quantile d'ordre $1 - \alpha$ de la loi du khi-deux à $r = r_1 + r_2$ degrés de liberté. Alors

(i) Les tests $\tilde{\phi}^{(n)}$ confrontant les hypothèses $H^{(n)}(\Omega, \theta_0)$ et

$$\bigcup \left\{ H_f^{(n)}(\theta) / \theta - \theta_0 \notin \mathcal{M}(\Omega) : f \in \mathcal{F} \right\}$$

sont localement asymptotiquement les plus stringents au seuil α .

(ii) Les tests $\tilde{\phi}^{(n)}$ sont localement asymptotiquement α -semblables sous $H^{(n)}(\Omega, \theta_0)$.

(iii) Pour des valeurs données de p_0 et q_0 , la statistique de test (19) ne dépend de p_1 et q_1 qu'à travers $\pi = \max(p_1 - p_0, q_1 - q_0)$.

Démonstration. La condition A3(v) entraîne que la suite des matrices $Q^{(n)}(\theta)$ converge, quand $n \rightarrow \infty$, vers une matrice $Q(\theta)$. Or puisque $Q^{(n)}(\theta)$ est une fonction continue en θ , $Q^{(n)}(\hat{\theta}^{(n)})$ converge en probabilité vers la matrice $Q(\theta)$, d'où

$$Q^{(n)}(\hat{\theta}^{(n)}) = o_P(1) \quad (20)$$

lorsque $n \rightarrow \infty$. La proposition 2 et la relation (20) impliquent que pour toute fonction f appartenant à la classe $\mathcal{F}$,

$$n T_{\psi,f}^{(n)}(\hat{\theta}^{(n)})^{\top} Q^{(n)}(\hat{\theta}^{(n)}) T_{\psi,f}^{(n)}(\hat{\theta}^{(n)}) - n \tilde{T}_{\tilde{\psi}}^{(n)}(\hat{\theta}^{(n)})^{\top} Q^{(n)}(\hat{\theta}^{(n)}) \tilde{T}_{\tilde{\psi}}^{(n)}(\hat{\theta}^{(n)}) = o_P(1), \quad (21)$$

sous $H_f^{(n)}(\Omega, \theta_0)$, lorsque $n \rightarrow \infty$. Notre proposition résulte donc de (21) et la proposition 5.10 de Hallin et Puri [13]. $\square$

Nous concluons cette section par le calcul de la puissance asymptotique locale, c'est-à-dire sous des contre-hypothèses locales, des tests adaptatifs de la proposition 4, ainsi que leur efficacité relative asymptotique (ERA) par rapport aux tests paramétriques gaussiens décrits dans la proposition 6.1 de Hallin et Puri [13].

Pour un vecteur $\tau = (\eta^{\top}, \gamma^{\top}, \delta^{\top})^{\top}$, associons les paramètres locaux

$$a_i = \sum_{j=1}^{\min(p_1, i+p_0-1)} \gamma_j g_{i-j}, \quad b_i = \sum_{j=1}^{\min(q_1, i+q_0-1)} \delta_j h_{i-j}.$$

Corollaire 1. Supposons que les hypothèses de la proposition 4 sont vérifiées. Alors

(i) la puissance $1 - \beta_f^{(n)}\{\theta + v(n)\tau\}$ des tests $\tilde{\phi}^{(n)}$ sous $H_f^{(n)}\{\theta + v(n)\tau\}$ avec $\theta - \theta_0 \in \mathcal{M}(\Omega)$ et $\tau \notin \mathcal{M}(v(n)^{-1}\Omega)$ vérifie

$$\lim_{n \rightarrow \infty} [1 - \beta_f^{(n)}\{\theta + v(n)\tau\} - \{1 - F_{\chi}^r(\chi_{r;1-\alpha}^2; \rho_f^{(n)}(\theta; \tau))\}] = 0,$$

où $F_{\chi}^r(\cdot; \rho)$ représente la fonction de répartition de la loi du khi-deux à r degrés de liberté et de paramètre d'excentricité ρ , avec

$$\rho_f^{(n)}(\theta; \tau) = \mathcal{J}(f_1) \rho^{(n)}(\theta; \tau) = \mathcal{J}(f_1) \Xi Q^{(n)}(\theta) \Xi^{\top},$$

où

$$\Xi = \left(\Gamma_I(\theta) \eta a_1 + b_1 \cdots a_\pi + b_\pi \sum_{i=\pi+1}^{\infty} (a_i + b_i) \psi_i^{(1)} \cdots \sum_{i=\pi+1}^{\infty} (a_i + b_i) \psi_i^{(p_0+q_0)} \right).$$

(ii) l'efficacité relative asymptotique des tests $\tilde{\phi}^{(n)}$ par rapport aux tests gaussiens $\phi^{(n)}$, sous $H_f^{(n)}\{\theta + v(n)\tau\}$, est

$$\text{ERA}_f(\tilde{\phi}^{(n)}, \phi^{(n)}) = \lim_{n \rightarrow \infty} \frac{\rho_f^{(n)}(\theta; \tau)}{\rho^{(n)}(\theta; \tau)} = \mathcal{J}(f_1).$$

Démonstration. La contiguïté mutuelle des hypothèses

$$H_f^{(n)}(\theta) \quad \text{et} \quad H_f^{(n)}\{\theta + v(n)\tau\}$$

entraîne que le résultat (21) reste vrai sous la contre-hypothèse locale $H_f^{(n)}\{\theta + v(n)\tau\}$. Ce résultat et la proposition 5.14 de Hallin et Puri [13] assurent (i). L'énoncé (ii) est une conséquence immédiate de (i) et de la proposition 6.1 (iii) de Hallin et Puri [13]. $\square$

Le tableau 1 contient les valeurs numériques des ERA sous les densités normale, logistique et double exponentielle. Remarquons que ces valeurs sont toujours supérieures à un, ce qui signifie que les tests adaptatifs sont asymptotiquement plus puissants que les tests gaussiens. Notons aussi que chaque valeur est maximale dans sa colonne correspondante dans le tableau 6.1 de Hallin et Puri [13]. Ainsi sous chacune des trois densités, les tests adaptatifs réalisent la meilleure efficacité, par rapport aux tests gaussiens, parmi les efficacités réalisées par les tests de van der Waerden, Wilcoxon et Laplace.

Densité sous-jacente f	normale	logistique	double exponentielle
$\text{ERA}_f(\tilde{\phi}^{(n)}, \phi^{(n)})$	1.00	1.098	2.00

TAB. 1. Les efficacités relatives asymptotiques pour les tests de rangs adaptatifs par rapport aux tests gaussiens, sous les densités normale, logistique et double exponentielle.

6. Exemples. Afin d'illustrer nos résultats théoriques, nous avons réalisé des simulations en S-PLUS sur deux exemples importants. Nous avons étudié le problème de test d'une dépendance $\text{AR}(p_0)$ contre une dépendance $\text{AR}(p_0 + 1)$, avec une tendance linéaire simple non spécifiée ($k = 1$), et ce dans deux cas : $p_0 = 0$ et $p_0 = 1$. Notons que pour toutes les simulations nous avons construit la suite $\tilde{\phi}^{(n)}(\cdot)$ de la façon suivante.

Pour $x \in \mathbb{R}$,

$$k(x) = c/(1+x^6), \quad \text{où} \quad c = 1 / \int_{-\infty}^{+\infty} (1+x^6)^{-1} dx \approx 2.094,$$

$$\hat{\phi}_n(x) = \frac{-\frac{1}{n0.5^2} \sum_{t=1}^n k' \left(\frac{x - \hat{Z}_t}{0.5} \right)}{0.02 + \frac{1}{n0.5} \sum_{t=1}^n k \left(\frac{x - \hat{Z}_t}{0.5} \right)}, \quad F_n(x) = \frac{1}{n} \sum_{t=1}^n \mathbf{I}(\hat{Z}_t \leq x)$$

et finalement

$$\tilde{\phi}^{(n)}(u) = \hat{\phi}_n \circ F_n^{-1}(u), \quad \forall u \in (0, 1).$$

Exemple 1. Le cas particulier $p_0 = 0$, est le problème de tester l'indépendance des erreurs $e^{(n)}$ contre une dépendance AR(1). Alors $p_1 = \pi = 1$, $\theta_0 = 0$, $\hat{\theta} = (\hat{\beta}, 0)^\top$, $r_1 = 0$, $r_2 = 1$, $r = 1$, $\Omega_I = 1$, $\Omega_{II} = 0$, $Q_I = 0$, $Q_{II} = 1$ et les tests $\tilde{\phi}^{(n)}$ rejettent l'hypothèse d'indépendance $H^{(n)}(\Omega, 0)$ quand

$$(n-1)\{\tilde{r}_1^{(n)}(\hat{\theta})\}^2 > \chi_{1;1-\alpha}^2.$$

Dans le but de comparer les puissances des tests $\tilde{\phi}^{(n)}$ aux puissances des tests de rangs optimaux $\phi_f^{(n)}$ et des tests gaussiens $\phi^{(n)}$, nous avons simulé $N = 1000$ échantillons $Y^{(n)}$ de taille n , sous une contre-hypothèse locale comme suit :

$$\begin{cases} e_t = 0.1 e_{t-1} + \varepsilon_t; \\ Y_t = 0.75 x_t + e_t, \quad \text{où } x_t = (2t - n^2 - n)/2n. \end{cases}$$

Nous avons considéré trois tailles des échantillons $n = 50$, $n = 100$ et $n = 200$ et trois types de densités f de ε_t , à savoir

$$\begin{aligned} f_{\mathcal{N}}(x) &= 1/\sqrt{2\pi} \exp(-x^2/2), \\ f_{\mathcal{L}}(x) &= \frac{\pi \exp(-\pi x/\sqrt{3})}{\sqrt{3} \{1 + \exp(-\pi x/\sqrt{3})\}^2}, \\ f_{\mathcal{D}e}(x) &= \sqrt{2} \exp(-\sqrt{2}|x|). \end{aligned}$$

Dans les neuf cas et pour chaque échantillon, nous avons utilisé les estimateurs des moindres carrées

$$\hat{\beta} = \sum_{t=1}^n Y_t x_t / \sum_{t=1}^n x_t^2$$

pour calculer les résidus alignés $\hat{Z}_t^{(n)} = Y_t - \hat{\beta}^{(n)} x_t$, desquels nous avons déduit les trois tests au seuil $\alpha = 0.05$. Pour chaque test, nous avons calculé le nombre κ de rejet de l'hypothèse nulle pour déduire enfin sa puissance empirique $\xi = \kappa/N$.

Le tableau 2 rapporte les puissances empiriques des tests $\phi^{(n)}$, $\phi_f^{(n)}$ et $\tilde{\phi}^{(n)}$. Comme prévu, les tests de rangs qui utilisent la densité sous-jacente réelle sont les plus performants. Malheureusement, ces tests ne peuvent être bâtis que si la densité f est connue. Pour les densités, logistique et double exponentielle, les tests adaptatifs sont plus performants que les tests gaussiens, ce qui confirme les valeurs du tableau 1. Dans le cas d'une densité normale, les tests adaptatifs rivalisent quand même avec les tests gaussiens, qui sont bien sur optimaux dans le cas gaussien. En conclusion, les tests adaptatifs sont recommandés pour les situations non standards (non gaussiennes).

Exemple 2. Le cas $p_0 = 1$ est le problème de test d'une hypothèse nulle sous laquelle les erreurs $e^{(n)}$ réalisent une dépendance autorégressive d'ordre un, avec un paramètre d'autorégression $\rho \in (-1, 1)$ non spécifié, contre une dépendance autorégressive d'ordre deux. Alors $\pi = 1$, $p_1 = 2$, $\hat{\theta} = (\hat{\beta}, \hat{\rho}, 0)^\top$, $r_1 = 0$, $r_2 = 1$, $r = 1$, $\Omega_I = 1$, $\Omega_{II} = (1, 0)^\top$, $\psi_t = \rho^{t-1}$, $w_\psi^2 = \rho^2/(1 - \rho^2)$, $C_\psi = \rho$, $\hat{M} = 1$,

$$Q_I = 0, \quad Q_{II} = \begin{pmatrix} \rho^2 & \rho^2 - 1 \\ \rho^2 - 1 & (1 - \rho^2)^2 / \rho^2 \end{pmatrix},$$

densité f	n	$\xi(\phi^{(n)})$	$\xi(\phi_f^{(n)})$	$\xi(\tilde{\phi}^{(n)})$
$f_{\mathcal{N}}$	50	0.076	0.081	0.051
	100	0.107	0.123	0.088
	200	0.242	0.243	0.209
$f_{\mathcal{L}}$	50	0.075	0.083	0.064
	100	0.137	0.146	0.138
	200	0.259	0.276	0.271
$f_{\mathcal{D}e}$	50	0.061	0.094	0.080
	100	0.110	0.150	0.120
	200	0.212	0.398	0.379

TAB. 2. Les puissances empiriques des tests gaussiens, tests de rangs et tests adaptatifs sous les densités, normale, logistique et double exponentielle.

et les tests $\tilde{\phi}^{(n)}$ rejettent l'hypothèse nulle $H^{(n)}(\Omega, 0)$ lorsque

$$\left((n-1)^{1/2} \hat{\rho} \tilde{r}_1^{(n)}(\hat{\theta}) - \frac{1-\hat{\rho}^2}{\hat{\rho}} \sum_{i=2}^{n-1} (n-1)^{1/2} \hat{\rho}^{i-1} \tilde{r}_i^{(n)}(\hat{\theta}) \right)^2 > \chi_{1;1-\alpha}^2.$$

Nous avons simulé $N = 100$ échantillons $Y^{(n)}$ de taille n , sous une contre-hypothèse locale, comme suit :

$$\begin{cases} e_t = 0.5 e_{t-1} + 0.1 e_{t-2} + \varepsilon_t; \\ Y_t = 0.75 x_t + e_t. \end{cases}$$

Pour les trois types de densités $f_{\mathcal{N}}(x)$, $f_{\mathcal{L}}(x)$, et $f_{\mathcal{D}e}(x)$, pour $n = 50$, $n = 100$ et $n = 200$ et pour chaque échantillon, nous avons calculé les estimateurs des moindres carrés de β et ρ ,

$$\hat{\beta} = \frac{\sum_{t=1}^n Y_t x_t}{\sum_{t=1}^n x_t^2}, \quad \hat{\rho} = \frac{\sum_{t=1}^n \hat{e}_t \hat{e}_{t-1}}{\sum_{t=1}^n \hat{e}_t^2},$$

où $\hat{e}_t = Y_t - \hat{\beta}^{(n)} x_t$ et $\hat{e}_0 = 0$. Enfin, nous avons construit les trois tests à partir des résidus alignés $\hat{Z}_t = \hat{e}_t - \hat{\rho} \hat{e}_{t-1}$.

Le tableau 3 donne les puissances empiriques des tests $\phi^{(n)}$, $\phi_f^{(n)}$ et $\tilde{\phi}^{(n)}$, et il confirme les conclusions tirées du tableau 2.

7. Annexe. Démonstration du lemme 1. (i) En utilisant (7), on montre que

$$\int_0^1 \left\{ \tilde{\phi}^{(n)}(u) - \phi_f \left(\frac{1 + [nu]}{n+1} \right) \right\}^2 du \xrightarrow{P} 0. \quad (22)$$

densité f	n	$\phi^{(n)}$	$\phi_f^{(n)}$	$\tilde{\phi}^{(n)}$
$f_{\mathcal{N}}$	50	0.04	0.05	0.06
	100	0.14	0.13	0.12
	200	0.19	0.19	0.17
$f_{\mathcal{L}}$	50	0.04	0.08	0.08
	100	0.07	0.07	0.06
	200	0.21	0.25	0.22
$f_{\mathcal{D}e}$	50	0.05	0.10	0.05
	100	0.11	0.16	0.14
	200	0.25	0.43	0.32

TAB. 3. Les puissances empiriques des tests gaussiens, tests de rangs et tests adaptatifs sous les densités normale, logistique et double exponentielle.

L'énoncé (i) découle alors de (22) et de l'inégalité suivante :

$$\begin{aligned}
\left| \tilde{\phi}^{(n)} - \bar{\phi}_f^{(n)} \right| &= \left| n^{-1} \sum_{t=1}^n \tilde{\phi}^{(n)} \left(\frac{t}{n+1} \right) - \phi_f \left(\frac{t}{n+1} \right) \right| \\
&= \left| \int_0^1 \left\{ \tilde{\phi}^{(n)}(u) - \phi_f \left(\frac{1+[nu]}{n+1} \right) \right\} du \right| \\
&\leq \left[\int_0^1 \left\{ \tilde{\phi}^{(n)}(u) - \phi_f \left(\frac{1+[nu]}{n+1} \right) \right\}^2 du \right]^{1/2}.
\end{aligned}$$

(ii) D'une part, on a

$$\begin{aligned}
(\tilde{I}^{(n)})^2 &= (n-1)^{-1} \sum_{t=1}^n \left\{ \tilde{\phi}^{(n)} \left(\frac{t}{n+1} \right) - \bar{\phi}^{(n)} \right\}^2 \\
&= \frac{n}{n-1} \int_0^1 (\tilde{\phi}^{(n)}(u) - \bar{\phi}^{(n)})^2 du \\
&= \frac{n}{n-1} \left\{ \int_0^1 \tilde{\phi}^{(n)}(u) du - (\bar{\phi}^{(n)})^2 \right\}.
\end{aligned} \tag{23}$$

D'autre part, (i) et le fait que $\tilde{\phi}_f^{(n)} \rightarrow \int \phi_f(u) f(u) du = 0$ impliquent que

$$\tilde{\phi}^{(n)} = o_P(1), \quad n \rightarrow \infty. \tag{24}$$

De (23) et (24) on déduit que

$$(\tilde{I}^{(n)})^2 - \int_0^1 \{\tilde{\phi}^{(n)}(u)\}^2 du = o_P(1), \quad n \rightarrow \infty. \tag{25}$$

D'après (22), on a

$$\int_0^1 \{\tilde{\phi}^{(n)}(u)\}^2 du - \int_0^1 \phi_f^2 \left(\frac{1+[nu]}{n+1} \right) du = o_P(1), \quad n \rightarrow \infty; \tag{26}$$

le théorème V.1.4.b de Hájek et Šidák [11] implique alors que

$$\int_0^1 \phi_f^2 \left(\frac{1 + [nu]}{n+1} \right) du \rightarrow \mathcal{J}(f), \quad n \rightarrow \infty. \quad (27)$$

En définitive (25), (26) et (27) assurent (ii).

(iii) D'après (4) et (8) on peut écrire

$$\begin{aligned} n^{1/2}(\tilde{m}^{(n)} - m^{(n)}) &= \frac{n^{1/2}}{n(n-1)} (\tilde{S}_{10}^{(n)} \tilde{S}_{01}^{(n)} - S_{10}^{(n)} S_{01}^{(n)} + S_{11}^{(n)} - \tilde{S}_{11}^{(n)}) \\ &= \frac{n^{1/2}}{n(n-1)} \{ (\tilde{S}_{10}^{(n)} - S_{10}^{(n)}) \tilde{S}_{01}^{(n)} + (\tilde{S}_{01}^{(n)} - S_{01}^{(n)}) S_{10}^{(n)} - (\tilde{S}_{11}^{(n)} - S_{11}^{(n)}) \} \\ &= I_1^{(n)} + I_2^{(n)} - I_3^{(n)}. \end{aligned}$$

Il est facile de voir que chaque terme $I_j^{(n)}$, $j = 1, 2, 3$, converge en probabilité vers zéro.

(iv) D'une façon similaire à (iii), on montre que chaque terme de (9) est asymptotiquement équivalent en probabilité à son analogue dans (5). $\square$

Démonstration de la proposition 1.

(i) Posons $k_{t-i,j}^{(n)}$ la j ème composante du vecteur $(K^{(n)})^\top x_{t-i}^{(n)}$ et soit

$$\bar{k}^{(n)} = \frac{1}{n-i} \sum_{t=i+1}^n k_{t-i,j}^{(n)}.$$

Puisque $(\tilde{I}^{(n)})^2$ converge en probabilité vers $\mathcal{J}(f)$ (lemme 1(i)), alors montrer (i) revient à montrer que

$$L^{(n)} = \sum_{t=i+1}^n \left[\left\{ \tilde{\phi}^{(n)} \left(\frac{R_t^{(n)}}{n+1} \right) - \phi_f \left(\frac{R_t^{(n)}}{n+1} \right) \right\} - (\tilde{\phi}^{(n)} - \phi_f^{(n)}) \right] k_{t-i,j}^{(n)} = o_P(n^{1/2}), \quad (28)$$

lorsque $n \rightarrow \infty$. Pour cela, on calcule la moyenne quadratique de $n^{-1/2} L^{(n)}$ sachant $Z_{(\cdot)}$. Ce calcul est bien possible puisque le vecteur de rangs et le vecteur d'ordre de la série résiduelle sont indépendants. En effet,

$$\begin{aligned} E_\theta \left\{ \frac{(n^{-1/2} L^{(n)})^2}{z_{(\cdot)}} \right\} &= \frac{1}{n-1} \sum_{t=i+1}^n \left(k_{t-i,j}^{(n)} - \bar{k}^{(n)} \right)^2 \times \\ &\quad \left[\frac{1}{n} \sum_{t=1}^n \left\{ \tilde{\phi}^{(n)} \left(\frac{t}{n+1} \right) - \phi_f \left(\frac{t}{n+1} \right) \right\}^2 - (\tilde{\phi}^{(n)} - \phi_f^{(n)})^2 \right]. \end{aligned}$$

Aussi est-il clair que sous l'hypothèse $H_f^{(n)}(\theta)$, lorsque $n \rightarrow \infty$,

$$(\tilde{\phi}^{(n)} - \phi_f^{(n)})^2 = o_P(1), \quad (29)$$

$$\frac{1}{n} \sum_{t=1}^n \left\{ \tilde{\phi}^{(n)} \left(\frac{t}{n+1} \right) - \phi_f \left(\frac{t}{n+1} \right) \right\}^2 = o_P(1), \quad (30)$$

$$\frac{1}{n-1} \sum_{t=i+1}^n \left(k_{t-i,j}^{(n)} - \bar{k}^{(n)} \right)^2 = O(1). \quad (31)$$

Les résultats (29), (30) et (31) impliquent que lorsque $n \rightarrow \infty$,

$$\mathbb{E}_\theta \left\{ \frac{(n^{-1/2} L^{(n)})^2}{z_{(\cdot)}} \right\} = o_P(1),$$

d'où (28), ce qui assure (i).

(ii) Posons

$$S_{n,i} = \frac{1}{n-i} \sum_{t=i+1}^n \varphi_f \left(\frac{R_t}{n+1} \right) F^{-1} \left(\frac{R_{t-i}}{n+1} \right)$$

et

$$\tilde{S}_{n,i} = \frac{1}{n-i} \sum_{t=i+1}^n \tilde{\varphi}^{(n)} \left(\frac{R_t}{n+1} \right) Z_{t-i}.$$

En vertu du lemme 1(iii)–(iv), pour prouver (ii) il suffit de montrer que sous $H_f^{(n)}(\theta)$,

$$S_{n,i} - \tilde{S}_{n,i} = o_P(n^{-1/2}), \quad \text{lorsque } n \rightarrow \infty. \quad (32)$$

Pour cela, considérons la décomposition suivante :

$$\begin{aligned} \frac{1}{n^{1/2}} (S_{n,i} - \tilde{S}_{n,i}) &= \frac{n^{1/2}}{n-i} \sum_{t=i+1}^n \left\{ \tilde{\varphi}^{(n)} \left(\frac{R_t^{(n)}}{n+1} \right) - \varphi_f \left(\frac{R_t^{(n)}}{n+1} \right) \right\} Z_{t-i} \\ &\quad + \frac{n^{1/2}}{n-i} \sum_{t=i+1}^n \varphi_f \left(\frac{R_t^{(n)}}{n+1} \right) \left\{ Z_{t-i} - F^{-1} \left(\frac{R_{t-i}^{(n)}}{n+1} \right) \right\} \\ &= A^{(n)} + B^{(n)}. \end{aligned}$$

Calculons les moyennes quadratiques conditionnelles de $A^{(n)}$ et $B^{(n)}$ sachant $Z_{(\cdot)}$. En particulier, nous obtenons pour $A^{(n)}$

$$\begin{aligned} \mathbb{E}_\theta \left\{ A^{(n)2} / Z_{(\cdot)} \right\} &= \frac{1}{n} \mathbb{E}_\theta \left[\left[\sum_{t=i+1}^n \left\{ \tilde{\varphi}^{(n)} \left(\frac{R_t^{(n)}}{n+1} \right) - \varphi_f \left(\frac{R_t^{(n)}}{n+1} \right) \right\} Z_{t-i} \right]^2 / Z_{(\cdot)} \right] \\ &= \frac{1}{n} \sum_{t=i+1}^n \mathbb{E}_{\theta_0} \left[\left\{ \tilde{\varphi}^{(n)} \left(\frac{R_t^{(n)}}{n+1} \right) - \varphi_f \left(\frac{R_t^{(n)}}{n+1} \right) \right\}^2 Z_{t-i}^2 / Z_{(\cdot)} \right] \\ &\quad + \frac{1}{n} \sum_{i+1 \leq t, s \leq n} \mathbb{E}_{\theta_0} \left[\left\{ \tilde{\varphi}^{(n)} \left(\frac{R_t^{(n)}}{n+1} \right) - \varphi_f \left(\frac{R_t^{(n)}}{n+1} \right) \right\} \right. \\ &\quad \times \left. \left\{ \tilde{\varphi}^{(n)} \left(\frac{R_s^{(n)}}{n+1} \right) - \varphi_f \left(\frac{R_s^{(n)}}{n+1} \right) \right\} Z_{t-i} Z_{s-i} / Z_{(\cdot)} \right] \\ &= A_1^{(n)} + A_2^{(n)}. \end{aligned}$$

Il est facile de voir que

$$\begin{aligned}
 A_1^{(n)} &= \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \left\{ \tilde{\varphi}^{(n)}\left(\frac{i}{n+1}\right) - \varphi_f\left(\frac{i}{n+1}\right) \right\}^2 Z_{(j)}^2 \\
 &= \frac{1}{n(n-1)} \left[\sum_{i=1}^n \left\{ \tilde{\varphi}^{(n)}\left(\frac{i}{n+1}\right) - \varphi_f\left(\frac{i}{n+1}\right) \right\}^2 \sum_{i=1}^n Z_i^2 \right. \\
 &\quad \left. - \sum_{i=1}^n \left\{ \tilde{\varphi}^{(n)}\left(\frac{i}{n+1}\right) - \varphi_f\left(\frac{i}{n+1}\right) \right\}^2 Z_{(i)}^2 \right] \\
 &\leq \int_0^1 \left\{ \tilde{\varphi}^{(n)}(u) - \varphi_f\left(\frac{1+[nu]}{n+1}\right) \right\}^2 du \times \frac{1}{n+1} \sum_{i=1}^n Z_i^2.
 \end{aligned}$$

Puisque sous $H_f^{(n)}(\theta)$, $(n+1)^{-1} \sum_{t=1}^n Z_t^2$ converge presque sûrement vers σ_f^2 et en vertu de (22), alors le terme $A_1^{(n)}$ converge en probabilité vers zéro. Posons maintenant

$$T_{\ell m} = \sum_{i=1}^n \left\{ \tilde{\varphi}^{(n)}\left(\frac{i}{n+1}\right) - \varphi_f\left(\frac{i}{n+1}\right) \right\}^\ell \{Z_{(i)}\}^m, \quad \ell, m = 0, 1, 2.$$

Le terme $A_2^{(n)}$ peut être décomposé comme suit

$$\begin{aligned}
 A_2^{(n)} &= 2 \{n(n-1)(n-2)(n-i)\}^{-1} (T_{11}T_{01}T_{10} - T_{20}T_{01} - 2T_{11}T_{01}T_{10} \\
 &\quad + 4T_{12}T_{10} - 2T_{11}T_{10}T_{01} - T_{10}^2T_{02} + T_{20}T_{02} + 2T_{11}^2 + 6T_{22}).
 \end{aligned}$$

En utilisant (7), il est facile de vérifier que chaque terme de cette décomposition converge en probabilité vers zéro, sous $H_f^{(n)}(\theta)$. Par conséquent, $A_2^{(n)}$ converge en probabilité vers zéro, ce qui prouve que $A^{(n)}$ converge en probabilité vers zéro.

À présent, montrons que $B^{(n)}$ converge en probabilité vers zéro. Soit

$$\begin{aligned}
 E_\theta (B^{(n)2}/Z_{(\cdot)}) &= \frac{1}{n} E_\theta \left[\left[\sum_{t=i+1}^n \left\{ Z_{t-i} - F^{-1}\left(\frac{R_t^{(n)}}{n+1}\right) \right\} \varphi_f\left(\frac{R_t^{(n)}}{n+1}\right) \right]^2 / Z_{(\cdot)} \right] \\
 &= \frac{1}{n} \sum_{t=i+1}^n E_\theta \left[\left\{ Z_{t-i} - F^{-1}\left(\frac{R_t^{(n)}}{n+1}\right) \right\}^2 \left[\varphi_f\left(\frac{R_t^{(n)}}{n+1}\right) \right]^2 / Z_{(\cdot)} \right] \\
 &\quad + \frac{1}{n} \sum_{i+1 \leq t \neq s \leq n} E_\theta \left[\left\{ Z_{t-i} - \varphi_f\left(\frac{R_t^{(n)}}{n+1}\right) \right\} \right. \\
 &\quad \times \left. \left\{ Z_{s-i} - \varphi_f\left(\frac{R_s^{(n)}}{n+1}\right) \right\} \varphi_f\left(\frac{R_t^{(n)}}{n+1}\right) \varphi_f\left(\frac{R_s^{(n)}}{n+1}\right) / Z_{(\cdot)} \right] \\
 &= B_1^{(n)} + B_2^{(n)},
 \end{aligned}$$

avec

$$\begin{aligned}
 B_1^{(n)} &= \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \left\{ \varphi_f\left(\frac{i}{n+1}\right) \right\}^2 \left\{ Z_{(j)} - F^{-1}\left(\frac{j}{n+1}\right) \right\}^2 \\
 &\leq \frac{1}{n(n-1)} \left[\sum_{i=1}^n \left\{ \varphi_f\left(\frac{i}{n+1}\right) \right\}^2 \sum_{i=1}^n \left\{ Z_{(i)} - F^{-1}\left(\frac{i}{n+1}\right) \right\}^2 \right].
 \end{aligned}$$

En utilisant le fait que

$$\frac{1}{n} \sum_{i=1}^n \left\{ \varphi_f \left(\frac{i}{n+1} \right) \right\}^2 \rightarrow \int_0^1 \{ \varphi_f(u) \}^2 du = \mathcal{J}(f)$$

et

$$\frac{1}{n-1} \sum_{i=1}^n \left\{ Z_{(i)} - F^{-1} \left(\frac{i}{n+1} \right) \right\}^2 = o_P(1)$$

quand $n \rightarrow \infty$, on déduit que $B_1^{(n)}$ converge en probabilité vers 0, sous l'hypothèse $H_f^{(n)}(\theta)$. Une décomposition de $B_2^{(n)}$ similaire à celle donnée pour $A_2^{(n)}$ montre que $B_2^{(n)}$ converge en probabilité vers zéro, sous $H_f^{(n)}(\theta)$, ce qui assure (32) et achève donc la démonstration. $\square$

English extended abstract. ARMA models with a regression trend constitute a very large and very rich class of statistical models. Their importance has long been recognized; see e.g. Anderson [2], Durbin [10], Pierce [27], Judje *et al.* [19], Swensen [29] or Hallin and Puri [13].

In this paper, we consider the ARMA model with linear trend from a semiparametric point of view, i.e., we suppose that the innovation density, however, will be considered unspecified. This situation is by far more realistic than the classical parametric one. Semiparametric models typically involve a finite-dimensional parameter of interest $\theta \in \mathbb{R}^{k_1}$, along with an infinite-dimensional nuisance parameter f , where f is an unspecified density of some white noise underlying the data generating process.

The fact that f , in the semiparametric model, remains unspecified induces, in general, a loss of efficiency with respect to the parametric situation. Such loss is essentially caused by local changes in f that may have identical effects on the distribution of observations as local changes in θ . Whenever this loss is nil, i.e., whenever there exists an efficient method, uniformly on f , for all parametric f -models, the semiparametric model is called *adaptive* and the method used is called *adaptive method*. For example, the adaptive test is the test which achieves—uniformly on f —the power of each optimal test in every parametric f -model.

Approaches that take into account the semiparametric nature of the models at hand mainly belong to one of two types : either to estimate the nuisance f from the sample using split technique, or to take care of it by means of some adequate statistical principle as invariance argument. We adopte the second approach to construct adaptive procedures for testing hypothesis of linear constraints on a some parameter θ related to the linear model with errors ARMA.

Our starting point is the LAN property proposed in Hallin and Puri [13] which claims that the ARMA models with linear trend are locally asymptotically normal with a free central sequence, this is sufficient, in view of results of Hallin and Werker [14], to prove the adaptiveness of the considered models. Moreover, this same paper proposes some aligned rank tests which are asymptotically distribution-free hence they are valid uniformly on f , but unfortunately they are not optimal when the innovation density f is unknown. To extract tests which are valid and optimal for testing a variety of linear hypothesis under a wide class of innovation densities f , we modify the rank-statistics defined in Hallin and Puri [13] by replacing the efficient scores functions, which depend on unknown error density f , by a consistent estimate measurable on the

order statistics of innovations (which form a sufficient statistic for the density function). As ranks and order statistics are independent, we automatically obtain the independence needed between the inference of θ based on ranks and the inference of f based on order statistics.

The method we are using allows for bypassing the usual steps related to establishing “joint” LAN results and avoids some unpleasant technicalities, such as sample-splitting. Moreover, adaptive rank tests belong to the class of (aligned) permutation tests, consequently they are unbiased, α -similar and conditionally distribution-free. For the notation in the results below, see the French text.

Proposition 1. *Assume that hypotheses A1–A4 hold. Then, under $H_f^{(n)}(\theta)$, as $n \rightarrow \infty$,*

- (i) $n^{1/2}\{t_{i,f}^{(n)}(\theta) - \tilde{t}_i^{(n)}(\theta)\} = o_P(1)$, $i = 1, \dots, n-1$;
- (ii) $n^{1/2}\{r_{i,f}^{(n)}(\theta) - \tilde{r}_i^{(n)}(\theta)\} = o_P(1)$, $i = 1, \dots, n-1$.

Proposition 2. *Assume that hypotheses A1–A4 hold. Then, under $H_f^{(n)}(\theta)$, as $n \rightarrow \infty$,*

$$n^{1/2}\{T_{\psi,f}^{(n)}(\theta) - \tilde{T}_{\psi}^{(n)}(\theta)\} = o_P(1).$$

Proposition 3. *Assume that hypotheses A1–A3 hold, and that $\tilde{\phi}^{(n)}(\cdot)$, Ω and $\hat{\theta}^{(n)}$ satisfy hypothesis A4, A5 and A5 respectively. Then, under $H_f^{(n)}(\Omega, \underline{\theta}_0)$, as $n \rightarrow \infty$,*

$$n^{1/2}\{T_{\psi,f}^{(n)}(\hat{\theta}^{(n)}) - \tilde{T}_{\psi}^{(n)}(\hat{\theta}^{(n)})\} = o_P(1).$$

Proposition 4. *Assume that hypotheses A1–A3 hold, and that $\tilde{\phi}^{(n)}(\cdot)$, Ω and $\hat{\theta}^{(n)}$ satisfy hypothesis A4, A5 and A5 respectively. Consider the sequence of adaptive aligned rank tests $\tilde{\phi}^{(n)}$ defined in (19). Then,*

- (i) $\tilde{\phi}^{(n)}$ is locally asymptotically most stringent, for $H^{(n)}(\Omega, \theta_0)$, against

$$\bigcup \left\{ H_f^{(n)}(\theta) / \theta - \theta_0 \notin \mathcal{M}(\Omega) : f \in \mathcal{F} \right\}$$

at probability level α .

- (ii) $\tilde{\phi}^{(n)}$ is locally asymptotically α -similar, under $H^{(n)}(\Omega, \theta_0)$;
- (iii) for given values of p_0 et q_0 , the test statistics in (19) depend on p_1 and q_1 only through $\pi = \max(p_1 - p_0, q_1 - q_0)$.

Corollary 1. *Assume that the assumptions of Proposition 4 hold. Then,*

- (i) the power $1 - \beta_f^{(n)}(\underline{\theta} + \underline{v}(n)\underline{\tau})$ of tests $\tilde{\phi}^{(n)}$ at $H_f^{(n)}(\underline{\theta} + \underline{v}(n)\underline{\tau})$, with $\underline{\theta} - \underline{\theta}_0 \in \mathcal{M}(\Omega)$ and $\underline{\tau} \notin \mathcal{M}(\underline{v}(n)^{-1}\Omega)$, satisfies

$$\lim_{n \rightarrow \infty} \left\{ \left[1 - \beta_f^{(n)}(\underline{\theta} + \underline{v}(n)\underline{\tau}) \right] - \left[1 - F_{\chi^2}^r \left(\chi_{r;1-\alpha}^2; \rho_f^{(n)}(\underline{\theta}; \underline{\tau}) \right) \right] \right\} = 0;$$

- (ii) the asymptotic relative efficiencies of the tests $\tilde{\phi}^{(n)}$ with respect to the Gaussian tests $\phi^{(n)}$, under $H_f^{(n)}(\underline{\theta} + \underline{v}(n)\underline{\tau})$, is

$$ARE_f(\tilde{\phi}^{(n)}, \phi^{(n)}) = \lim_{n \rightarrow \infty} \left[\frac{\rho_f^{(n)}(\underline{\theta}; \underline{\tau})}{\rho^{(n)}(\underline{\theta}; \underline{\tau})} \right] = \mathcal{J}(f_1).$$

BIBLIOGRAPHIE

1. J. Allal, *R-estimation dans le modèle autorégressif d'ordre un*. Thèse, Université libre de Bruxelles, Belgique, 1991.
2. R. L. Anderson, *The problem of autocorrelation in regression analysis*, J. Amer. Statist. Assoc. **49** (1954), 113–129.
3. T. W. Anderson, *The Statistical Analysis of Time Series*, Wiley, New York, 1971.
4. R. Beran, *Asymptotically efficient adaptive rank estimates in location models*, Ann. Statist. **2** (1974), 63–74.
5. P. J. Bickel, C. A. J. Klaassen, Y. Ritov et J. A. Wellner, *Efficient and Adaptive Statistical Inference for Semiparametric Models*, Johns Hopkins University Press, Baltimore, MD, 1993.
6. G. E. P. Box et D. A. Pierce, *Distribution of residual autocorrelations in autoregressive-integrated moving average time series models*, J. Amer. Statist. Assoc. **65** (1970), 1509–1526.
7. F. C. Drost, C. A. J. Klaassen et B. J. M. Werker, *Adaptive estimation in time series models*, Ann. Statist. **25** (1997), 786–817.
8. J.-M. Dufour et R. Roy, *Some robust exact results on sample autocorrelations and tests of randomness*, J. Econom. **29** (1985), 257–273. [Corr. : J. Econom. **41**, 279–281.]
9. J.-M. Dufour et R. Roy, *Generalized portmanteau statistics and tests for randomness*, Comm. Statist. Theory Meth. **15** (1986), 2953–2972.
10. J. Durbin, *Estimation of parameters in time-series regression models*, J. Roy. Statist. Soc. B **22** (1960), 139–153.
11. J. Hájek et Z. Šidák, *Theory of Rank Tests*, Academic Press, New York, 1967.
12. M. Hallin, J.-F. Ingenbleek et M. L. Puri, *Linear and quadratic serial rank tests for randomness against serial dependence*, J. Time Ser. Anal. **8** (1987), 409–424.
13. M. Hallin et M. L. Puri, *Aligned rank tests for linear models with autocorrelated error terms*, J. Multivariate Anal. **50** (1994), 175–237.
14. M. Hallin et B. J. M. Werker, *Optimal testing for semi-parametric AR models : From Gaussian Lagrange multipliers to autoregression rank scores and adaptive tests*, in Asymptotics, Nonparametrics and Time Series (S. Ghosh, Ed.), pp. 295–350, Dekker, New York, 1999.
15. M. Hallin et B. J. M. Werker, *Semi-parametric efficiency, distribution-freeness, and invariance*, Bernoulli **9** (2003), 137–165.
16. M. Hušková et P. K. Sen, *On sequentially adaptive asymptotically efficient rank statistics*, Sequen. Anal. **4** (1985), 125–151.
17. M. Hušková et P. K. Sen, *On sequentially adaptive signed-rank statistics*, Sequen. Anal. **5** (1986), 237–252.
18. P. Jeganathan, *Some aspects of asymptotic theory with applications to time series models*, Econom. Th. **11** (1995), 818–887.
19. G. G. Judge, W. E. Griffiths, R. C. Hill, H. Lütkepohl et T. C. Lee, *The Theory and Practice of Econometrics*, Wiley, New York, 1985.
20. B. K. Kim et J. Van Ryzin, *Estimation of the derivative of a density and related functions*, Comm. Statist. Theory Meth. **9** (1985), 697–709.
21. H. L. Koul et A. Schick, *Efficient estimation in nonlinear autoregressive time-series models*, Bernoulli **3** (1997), 247–277.

- 22. J.-P. Kreiss, *On adaptive estimation in stationary ARMA processes*, Ann. Statist. **15** (1987), 112–133.
- 23. J.-P. Kreiss, *Testing linear hypotheses in autoregressions*, Ann. Statist. **18** (1990), 1470–1482.
- 24. L. Le Cam, *Locally asymptotically normal families of distributions*, University of California Publications in Statistics **3** (1960), 37–98.
- 25. L. Le Cam, *Asymptotic Methods in Statistical Decision Theory*, Springer, New York, 1986.
- 26. E. L. Lehmann, *Testing Statistical Hypotheses*, 2^e édition, Wiley, New York, 1986.
- 27. D. A. Pierce, *Distribution of residual autocorrelations in the regression model with autoregressive-moving average errors*, J. Roy. Statist. Soc. B **33** (1971), 140–146.
- 28. B. M. Pötscher, *The behaviour of the Lagrangian multiplier test in testing the orders of an ARMA-model*, Metrika **32** (1985), 129–150.
- 29. A. R. Swensen, *The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend*, J. Multivariate Anal. **16** (1985), 54–70.

J. ALLAL
DÉPARTEMENT DE MATHÉMATIQUES
UNIVERSITÉ MOHAMMED PREMIER
OUJDA, MAROC
allal@sciences.univ – oujda.ac.ma

S. EL MELHAOUI
DÉPARTEMENT DE MATHÉMATIQUES
UNIVERSITÉ MOHAMMED PREMIER
OUJDA, MAROC