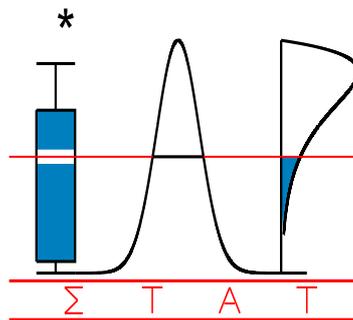


T E C H N I C A L  
R E P O R T

0319

PROCEDURES OPTIMALES FONDEES SUR  
LES RANGS MULTIVARIES

D. PAINDAVEINE



I A P S T A T I S T I C S  
N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

<http://www.stat.ucl.ac.be/IAP>

# PROCÉDURES OPTIMALES FONDÉES SUR LES RANGS MULTIVARIÉS

Davy Paindaveine\*

I.S.R.O, E.C.A.R.E.S et Département de Mathématique  
Université Libre de Bruxelles, Belgique

## Abstract

Les méthodes fondées sur les rangs et/ou les signes répondent au besoin de s'affranchir de l'hypothèse classique de normalité. Les procédures de test semi paramétriques qu'elles engendrent sont en effet valides (car libres) sous une large classe de distributions. Possédant de bonnes propriétés d'efficacité, les procédures de rangs sont restées longtemps confinées aux modèles à observations indépendantes univariées. Bien que le problème de la généralisation au cadre multivarié, que nous considérons dans cet article, se pose depuis plus de quarante ans, les premières extensions multivariées satisfaisantes des notions de signes et de rangs n'émergèrent qu'au début des années 90 (à titre de comparaison, l'extension aux modèles de séries chronologiques a essentiellement eu lieu dans les années 80). Dans cet article, nous rappelons le rôle fondamental joué par les signes et les rangs dans les problèmes univariés jouissant de bonnes propriétés d'invariance, passons brièvement en revue les diverses extensions multivariées de ces notions, puis décrivons nos procédures optimales (au sens de Le Cam) de rangs signés multivariés. A notre connaissance, ces dernières sont les seules qui maintiennent (et même parfois améliorent) dans le cadre multivarié, l'efficacité des procédures optimales de rangs univariés, ainsi que l'ensemble des propriétés d'invariance, qui constituent la justification théorique du recours aux signes et aux rangs.

## 1 Introduction : les rangs, pourquoi, et pour quoi faire?

Nous expliquons tout d'abord brièvement pourquoi les procédures de rangs sont fondamentales dans une large classe de problèmes d'inférence non paramétrique univariée. Dans la suite, nous nous concentrerons sur les procédures de rangs signés, car ce sont ces dernières que nous avons pour objectif de généraliser au cas multivarié.

Les procédures non paramétriques (ou semi paramétriques) permettent de faire l'économie d'une spécification—généralement artificielle et discutable—de la loi des erreurs du modèle statistique considéré. Cette spécification, qui, par opposition, débouche sur les procédures paramétriques, trouve en effet plus souvent son origine dans un besoin de commodité analytique que dans un réel souci de modélisation. Les procédures de test semi paramétriques sont ainsi valides (dans le sens qu'elles ont le niveau nominal  $\alpha$ ) quelle que soit la densité des erreurs dans une certaine collection  $\mathcal{F}$  de densités admissibles. De plus, ces procédures ont de bonnes propriétés d'efficacité sous une large gamme de lois, tout en cédant très peu (et même parfois rien) aux procédures paramétriques sous la densité auxquelles ces dernières sont adaptées.

D'autre part, il est courant que le sous-modèle identifié par l'hypothèse nulle d'un problème de test fixé soit invariant sous l'action d'un groupe  $\mathcal{G}$  de transformations, tel que le groupe des transformations monotones croissantes continues. A titre d'illustration, considérons la situation où on observe un échantillon de  $n$  variables aléatoires univariées  $Y_1, \dots, Y_n$  i.i.d., de loi absolument continue par rapport à la mesure de Lebesgue et symétrique par rapport à  $\theta$ . Plus précisément nous supposons que la fonction de densité

---

\*Recherche encouragée par un contrat P.A.I. du Gouvernement fédéral belge, et un contrat A.R.C. de la Communauté française de Belgique.

commune de ces variables est du type  $x \mapsto f(x - \theta)$ , où  $\theta \in \mathbb{R}$  et où  $f$  appartient à la collection  $\mathcal{F}$  des densités strictement positives et symétriques par rapport à l'origine de la droite réelle (cette hypothèse de symétrie permet d'identifier  $\theta$  même lorsque la moyenne de cette distribution n'existe pas; si on se restreint à des lois ayant des moments finis d'ordre un, le centre de symétrie  $\theta$  coïncide bien entendu avec la moyenne, de sorte que le problème considéré ci-dessous est le problème de position classique sur une moyenne). Considérons alors le problème de test

$$\begin{cases} \mathcal{H}_0 : \theta = 0 \\ \mathcal{H}_1 : \theta \neq 0, \end{cases} \quad (1)$$

dans lequel le paramètre infinidadimensionnel  $f \in \mathcal{F}$  est clairement une nuisance. Une approche paramétrique pour le problème (1) consiste à spécifier la forme analytique de  $f$  (à un paramètre d'échelle près) et d'en déduire la procédure optimale correspondante (typiquement, une procédure localement et asymptotiquement optimale). Le cas particulier de l'hypothèse gaussienne—pour laquelle  $f \in \{x \mapsto f_\sigma(x) := (1/\sigma\sqrt{2\pi}) \exp(-x^2/2\sigma^2), \sigma \in \mathbb{R}_0^+\}$ —débouche sur les procédures gaussiennes habituelles, c'est-à-dire, dans ce cas précis, sur les tests de Student (et, dans d'autres situations, sur les tests de Fisher, de Hotelling, les méthodes fondées sur les corrélogrammes, etc).

Dans l'approche semi paramétrique, par contre, la densité inconnue  $f$  n'est pas spécifiée au-delà du fait qu'elle est un certain élément de  $\mathcal{F}$ . Une des stratégies permettant d'éliminer cette nuisance infinidadimensionnelle consiste à recourir à un argument d'invariance. Considérons le groupe  $\mathcal{G}_+^{(n)} = \{g_g^{(n)}\}$  des transformations symétriques définies par

$$g_g^{(n)}(Y_1, \dots, Y_n) := (g(Y_1), \dots, g(Y_n)),$$

où  $g: \mathbb{R} \rightarrow \mathbb{R}$  est une fonction continue, monotone croissante et impaire, qui vérifie  $\lim_{x \rightarrow \infty} g(x) = \infty$ . Il est clair que si  $Y_1, \dots, Y_n$  "vérifient" l'hypothèse nulle  $\mathcal{H}_0$  en (1), il en va de même pour  $g(Y_1), \dots, g(Y_n)$ . On dit dans ce cas que l'hypothèse nulle du problème de test considéré est *invariante sous l'action de ce groupe*.

Dans ce type de situation, le *principe d'invariance* préconise le recours exclusif à des procédures dont les règles de décision ne varient pas le long des orbites du groupe  $\mathcal{G}_+^{(n)}$ , ce qui est le cas si et seulement si ces procédures ne font intervenir les observations qu'à travers l'invariant maximal associé à  $\mathcal{G}_+^{(n)}$  (c'est-à-dire à travers une statistique ne variant pas le long des orbites de  $\mathcal{G}_+^{(n)}$ , mais variant bien d'orbite en orbite; en d'autres termes, une statistique engendrant la même tribu que les orbites de  $\mathcal{G}_+^{(n)}$ ). En désignant respectivement par  $S_i := \text{sign}(Y_i)$  et  $R_i^+ := \text{rank}(|Y_i|; |Y_1|, \dots, |Y_n|)$ , le signe de  $Y_i$  et le rang de  $|Y_i|$  parmi  $|Y_1|, \dots, |Y_n|$ , il est facile de vérifier qu'un invariant maximal de  $\mathcal{G}_+^{(n)}$  est le vecteur des rangs signés

$$(S_1, \dots, S_n, R_1^+, \dots, R_n^+). \quad (2)$$

Le principe d'invariance conduit donc de manière naturelle à l'utilisation des procédures de rangs signés en pareille situation.

Il est également facile de vérifier que le groupe  $\mathcal{G}_+^{(n)}$  engendre  $\mathcal{H}_0$  (ou est un groupe générateur de  $\mathcal{H}_0$ ), ce qui signifie que

$$\forall f_1, f_2 \in \mathcal{F}, \exists g_g^{(n)} \in \mathcal{G}_+^{(n)}, \text{ tel que } Y_1, \dots, Y_n \text{ i.i.d. } P_{0, f_1} \Rightarrow g(Y_1), \dots, g(Y_n) \text{ i.i.d. } P_{0, f_2},$$

où  $P_{\theta, f}$  désigne la loi jointe des observations lorsque la fonction de densité de  $Y_1$  est  $x \mapsto f(x - \theta)$ .

En d'autres termes, le groupe  $\mathcal{G}_+^{(n)}$  est un groupe générateur si l'action du groupe  $\mathcal{G}_+^{(n)}$  induit, dans  $\mathcal{H}_0$ , une et une seule orbite. Conséquence importante : puisque leurs lois ne varient pas le long des orbites de  $\mathcal{G}_+^{(n)}$ , les procédures invariantes (c'est-à-dire les procédures de rangs signés) sont en fait *libres* sous l'hypothèse  $\mathcal{H}_0$  tout entière; il est donc possible de construire des tests dont la dimension sous l'hypothèse nulle est uniformément égale au niveau nominal  $\alpha$ , quelle que soit  $f \in \mathcal{F}$ .

A titre d'exemple, le test de Wilcoxon (de niveau asymptotique  $\alpha$ ) consiste à rejeter l'hypothèse nulle  $\mathcal{H}_0$  si

$$\left| \frac{3}{n^{3/2}} \sum_i S_i R_i^+ \right|^2 > \chi_{1,1-\alpha}^2, \quad (3)$$

où  $\chi_{1,1-\alpha}^2$  désigne le quantile d'ordre  $1 - \alpha$  d'une variable chi-deux à un degré de liberté.

L'invariance, la liberté qui en résulte, ainsi que leur robustesse intrinsèque (comparativement aux concurrents paramétriques usuels) sont autant de bonnes raisons qui conduisent à recourir aux procédures de rangs signés. Néanmoins, on pourrait croire que ne garder, de l'information résidant dans l'échantillon original, que l'information, à première vue relativement réduite, qui est contenue dans le vecteur des rangs signés se traduira par un déficit d'efficacité. Il est remarquable que, dans de nombreux problèmes (dont celui considéré ci-dessus), il n'en soit pas ainsi, l'information laissée de côté dans l'échantillon initial étant en fait de l'information sur la nuisance  $f$ , et non pas sur le paramètre d'intérêt  $\theta$ .

Les indices les plus clairs de cette bonne efficacité (asymptotique) des procédures de rangs signés sont les fameux résultats de Hodges et Lehmann (1956) et de Chernoff et Savage (1958). Hodges et Lehmann (1956) ont montré que la borne inférieure, en les  $f \in \mathcal{F}$  de variance finie (afin que le test de Student soit valide), de l'efficacité asymptotique relative du test de Wilcoxon  $\phi_W$  par rapport au test de Student  $\phi_T$  est de 0.864, i.e.,

$$\inf_f \text{ARE}_f[\phi_W/\phi_T] = 0.864.$$

Ce résultat signifie qu'asymptotiquement, le test de Wilcoxon demande *dans le pire des cas* 14.6% d'observations de plus que le test de Student pour atteindre la même puissance que ce dernier. Et il n'y a pas de "meilleur des cas", dans le sens où il est possible d'exhiber des suites de densités  $f_k$ ,  $k = 1, 2, \dots$  ( $f_k \in \mathcal{F}$  et de variance finie), telles que

$$\lim_{k \rightarrow \infty} \text{ARE}_{f_k}[\phi_W/\phi_T] = \infty.$$

Pour tout nombre  $M$  indéceimment grand, on peut donc trouver des lois sous lesquelles la procédure de Student consomme plus de  $M$  fois plus d'observations que la procédure de rangs signés de Wilcoxon pour discriminer entre l'hypothèse nulle et des contre-hypothèses locales avec la même puissance.

Le résultat de Chernoff et Savage (1958) concerne la procédure de rangs signés *de type van der Waerden*, c'est-à-dire celle qui est obtenue lorsqu'on pondère les rangs dans (3) avec des *fonctions scores gaussiennes*. Le test  $\phi_{\text{vdW}}$  de van der Waerden consiste à rejeter  $\mathcal{H}_0$  si

$$\left| \frac{1}{\sqrt{n}} \sum_i S_i \Phi_+^{-1} \left( \frac{R_i^+}{n+1} \right) \right|^2 > \chi_{1,1-\alpha}^2, \quad (4)$$

où  $\Phi_+$  désigne la fonction de répartition de la valeur absolue d'une variable aléatoire de loi normale centrée réduite. Chernoff et Savage (1958) ont montré que

$$\inf_f \text{ARE}_{f_k}[\phi_{\text{vdW}}/\phi_T] = 1,$$

et que cette borne inférieure n'est atteinte que sous des lois gaussiennes. Autrement dit, le test de rangs signés de van der Waerden est toujours meilleur—strictement meilleur, dès que la loi sous-jacente n'est pas normale—que le test classique de Student. Comme dans le cas du test de Wilcoxon, le gain relatif tend vers l'infini le long de suites de densités bien choisies.

## 2 Vers le multivarié.

Nous avons vu ci-dessus comment, dans le cas du problème de position symétrique à un échantillon, les signes et les rangs (des valeurs absolues) émergent naturellement en tant qu'invariants maximaux d'un certain groupe de transformations. Nous considérons maintenant la version multivariée du problème de

position symétrique à un échantillon, et présentons deux groupes de transformations sous l'action desquels ce problème de test est invariant.

Le modèle de position multivarié à un échantillon est défini par

$$\mathbf{Y}_i = \boldsymbol{\theta} + \boldsymbol{\varepsilon}_i, \quad i = 1, \dots, n, \quad (5)$$

où  $\boldsymbol{\varepsilon}_i$ ,  $i = 1, \dots, n$ , est un *bruit blanc indépendant*, c'est-à-dire un processus i.i.d. centré à l'origine;  $\boldsymbol{\theta}$  joue donc dans ce modèle le rôle d'un paramètre de position. Comme dans le cas univarié, nous désirons couvrir les cas où la moyenne de la loi sous-jacente n'existe pas. Nous recourrons donc à une hypothèse de symétrie, qui permet d'identifier *le centre de symétrie*  $\boldsymbol{\theta}$  (notons aussi que c'est cette symétrie qui, dans le cas univarié, justifie le recours aux rangs signés, les rangs signés constituant l'invariant maximal d'un groupe de transformations qui engendre une classe de lois *symétriques*). Une notion de symétrie souvent utilisée est celle de *symétrie centrale*, pour laquelle  $\boldsymbol{\varepsilon}_i$  et  $-\boldsymbol{\varepsilon}_i$  sont égaux en loi. Pour des raisons qui deviendront claires plus loin, nous considérerons plutôt le concept plus fort de *symétrie elliptique*, qui veut que la densité de probabilité de  $\boldsymbol{\varepsilon}_i$  soit de la forme

$$f_{\boldsymbol{\Sigma}, f}(\mathbf{z}) = c_{k, f} \frac{1}{(\det \boldsymbol{\Sigma})^{1/2}} f\left(\|\boldsymbol{\Sigma}^{-1/2} \mathbf{z}\|\right), \quad \mathbf{z} \in \mathbb{R}^k, \quad (6)$$

où  $c_{k, f}$  est une constante de normalisation,  $\|\boldsymbol{\Sigma}^{-1/2} \mathbf{z}\| = (\mathbf{z}' \boldsymbol{\Sigma}^{-1} \mathbf{z})^{1/2}$  désigne la distance entre  $\mathbf{z}$  et l'origine de  $\mathbb{R}^k$  dans la métrique associée à la matrice symétrique et définie positive  $\boldsymbol{\Sigma}$ , et où  $f$  est une fonction de  $\mathbb{R}_0^+$  dans  $\mathbb{R}_0^+$ . Les ensembles de niveau (ou *contours*) de (6) sont des hyper-ellipsoïdes—ce qui justifie bien entendu la terminologie—dont le centre commun est l'origine de  $\mathbb{R}^k$ , et l'orientation et la forme (rapports des tailles des axes principaux) sont déterminées par la matrice  $\boldsymbol{\Sigma}$ , dite *matrice de forme*. Le “paramètre” infindimensionnel  $f$  gouverne quant à lui, le comportement aléatoire de la distance à l'origine  $\|\boldsymbol{\Sigma}^{-1/2} \boldsymbol{\varepsilon}_t\|$ . C'est donc de  $f$  que dépend l'existence d'éventuels moments finis d'ordre  $r$  pour le bruit  $\boldsymbol{\varepsilon}_i$ . En particulier, si  $f$  a des queues—il n'y a en fait bien entendu qu'une seule queue (radiale)—trop lourdes, la moyenne de  $\boldsymbol{\varepsilon}_i$  n'existe pas. Si, par contre, cette queue est suffisamment légère pour que le bruit ait des moments finis d'ordre deux, il est clair que la variance de  $\boldsymbol{\varepsilon}_i$  est dans ce cas un multiple positif de la matrice de forme  $\boldsymbol{\Sigma}$ . Les paramètres  $\boldsymbol{\Sigma}$  et  $f$  sont généralement inconnus, et seront dès lors constamment considérés comme des paramètres de nuisance.

La version multivariée du problème de test (1) est bien entendu

$$\begin{cases} \mathcal{H}_0 : \boldsymbol{\theta} = \mathbf{0} \\ \mathcal{H}_1 : \boldsymbol{\theta} \neq \mathbf{0}. \end{cases} \quad (7)$$

En désignant par  $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$  l'hypothèse sous laquelle les observations sont engendrées par (5), où les  $\boldsymbol{\varepsilon}_i$  sont i.i.d. de fonction de densité (6), la version elliptique du même problème s'écrit

$$\begin{cases} \mathcal{H}_0 : \bigcup_{\boldsymbol{\Sigma}} \bigcup_f \mathcal{H}^{(n)}(\mathbf{0}, \boldsymbol{\Sigma}, f) \\ \mathcal{H}_1 : \bigcup_{\boldsymbol{\theta} \neq \mathbf{0}} \bigcup_{\boldsymbol{\Sigma}} \bigcup_f \mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f). \end{cases} \quad (8)$$

Considérons le *groupe des transformations linéaires* agissant sur l'échantillon  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ , c'est-à-dire le groupe  $\mathcal{G}^{(n)} = \{\mathcal{G}_{\mathbf{M}}^{(n)}, \mathbf{M} \in GL_k(\mathbb{R})\}$ , défini par

$$\mathcal{G}_{\mathbf{M}}^{(n)}(\mathbf{Y}_1, \dots, \mathbf{Y}_n) := (\mathbf{M}\mathbf{Y}_1, \dots, \mathbf{M}\mathbf{Y}_n),$$

où  $GL_k(\mathbb{R})$  désigne la collection des matrices  $k \times k$  non singulières à coefficients réels. Clairement, si  $(\mathbf{Y}_1, \dots, \mathbf{Y}_n)$  vérifie  $\mathcal{H}_0$ , alors  $(\mathbf{M}\mathbf{Y}_1, \dots, \mathbf{M}\mathbf{Y}_n)$  vérifie  $\mathcal{H}_0$  également. Autrement dit, l'hypothèse nulle du problème (8) est invariante sous l'action du groupe  $\mathcal{G}^{(n)}$ . Le principe d'invariance préconise donc de se restreindre à des statistiques de test *affine-invariantes*, i.e., qui sont invariantes le long des orbites de  $\mathcal{G}^{(n)}$ .

Un argument en faveur du principe d'invariance dans cette situation est le suivant. Considérons deux statisticiens, qui observent le même échantillon, mais utilisent des systèmes (des dispositifs expérimentaux,

par exemple) de mesure différents. Le premier “voit” l’échantillon  $(\mathbf{Y}_1, \dots, \mathbf{Y}_n)$ , tandis que le second voit  $(\mathbf{M}\mathbf{Y}_1, \dots, \mathbf{M}\mathbf{Y}_n)$ . La matrice  $\mathbf{M}$  est donc celle du changement de base qui permet de passer d’un système de mesure à l’autre. L’invariance de  $\mathcal{H}_0$  se traduit par le fait que celle-ci s’exprime de la même manière pour les deux statisticiens : “les observations sont centrées en l’origine”. Le principe d’invariance conduit à des tests fondés sur des statistiques qui prennent la même valeur pour les deux statisticiens, ce qui est naturellement désirable.

Le second groupe de transformations que nous utiliserons est l’extension multivariée (radiale) de celui qui a été considéré dans le cas univarié : le *groupe des transformations monotones radiales*. Ce groupe est défini de la façon suivante. Pour chaque  $\Sigma$ , décomposons la  $i$ ème observation  $\mathbf{Y}_i$  en  $\mathbf{Y}_i = d_i(\Sigma)\Sigma^{1/2}\mathbf{U}_i(\Sigma)$ , où

$$d_i(\Sigma) := \|\Sigma^{-1/2}\mathbf{Y}_i\| \quad \text{et} \quad \mathbf{U}_i(\Sigma) := \frac{\Sigma^{-1/2}\mathbf{Y}_i}{\|\Sigma^{-1/2}\mathbf{Y}_i\|}.$$

Sous  $\bigcup_f \mathcal{H}^{(n)}(\mathbf{0}, \Sigma, f)$ , c’est-à-dire, sous l’hypothèse que  $\boldsymbol{\theta} = \mathbf{0}$  et que  $\Sigma$  est la “vraie” matrice de forme, les vecteurs aléatoires  $\mathbf{U}_1(\Sigma), \dots, \mathbf{U}_n(\Sigma)$  sont i.i.d., de loi uniforme sur la sphère unité  $\mathcal{S}^{k-1} \subset \mathbb{R}^k$ , et les variables aléatoires  $d_1(\Sigma), \dots, d_n(\Sigma)$  sont les distances entre les observations sphéricisées et l’origine de  $\mathbb{R}^k$ . Considérons alors, pour  $\Sigma$  fixé, le groupe de transformations  $\mathcal{G}_\Sigma^{(n)} = \{g_g^{(n)}\}$ , agissant sur  $\mathbb{R}^k \times \mathbb{R}^k \times \dots \times \mathbb{R}^k$  ( $n$  fois), défini par

$$g_g^{(n)}(\mathbf{Y}_1, \dots, \mathbf{Y}_n) := (g(d_1(\Sigma))\Sigma^{1/2}\mathbf{U}_1(\Sigma), \dots, g(d_n(\Sigma))\Sigma^{1/2}\mathbf{U}_n(\Sigma)),$$

où  $g: \mathbb{R}^+ \rightarrow \mathbb{R}^+$  est une fonction continue monotone croissante, telle que  $g(0) = 0$  et  $\lim_{r \rightarrow \infty} g(r) = \infty$ . La transformation  $g_g^{(n)}$  est une transformation radiale, puisque, sous l’action de  $g_g^{(n)}$ , les observations  $\mathbf{Y}_i = d_i(\Sigma)\Sigma^{1/2}\mathbf{U}_i(\Sigma)$  sont déplacées le long de la demi-droite issue de l’origine de  $\mathbb{R}^k$  et passant par  $\mathbf{Y}_i$ ; c’est ce groupe (en fait, ces groupes, puisqu’il existe un tel groupe pour chaque valeur de la matrice de forme) que nous appellerons le *groupe des transformations radiales monotones continues*. De nouveau,  $\mathcal{H}_0$  est clairement invariante sous l’action de  $\mathcal{G}_\Sigma^{(n)}$ .

Pour ce qui est du caractère générateur de ces deux groupes, on vérifie facilement que

- (i) pour chaque densité radiale fixée  $f$ , le groupe  $\mathcal{G}^{(n)}$  est un groupe générateur de la sous-hypothèse  $\bigcup_\Sigma \mathcal{H}^{(n)}(\mathbf{0}, \Sigma, f)$ , et que
- (ii) pour chaque matrice de forme  $\Sigma$  fixée (à une constante multiplicative près), le groupe  $\mathcal{G}_\Sigma^{(n)}$  engendre la sous-hypothèse  $\bigcup_f \mathcal{H}^{(n)}(\mathbf{0}, \Sigma, f)$ .

Il est donc intuitivement clair qu’une procédure de test invariante par rapport à ces deux groupes de transformations sera libre (et donc valide) sous la classe complète des lois elliptiques introduites en (6).

### 3 Les signes et rangs multivariés.

Au vu des arguments d’invariance décrits plus haut, nous serons particulièrement attentifs aux propriétés d’équivariance/invariance des extensions multivariées des notions de signes et de rangs dont nous présentons un aperçu ci-dessous.

#### 3.1 L’approche historique.

Assez naturellement, le premier essai de généralisation multivariée des signes et des rangs a consisté à considérer les signes et les rangs composante par composante. Plus précisément, le signe  $\mathbf{S}_i$  et le rang  $\mathbf{R}_i$  de  $\mathbf{Y}_i$  sont respectivement définis par

$$\mathbf{S}_i := \begin{pmatrix} \text{sign}(Y_{i1}) \\ \vdots \\ \text{sign}(Y_{ik}) \end{pmatrix} \quad \text{et} \quad \mathbf{R}_i := \begin{pmatrix} \text{rank}(Y_{i1}; Y_{11}, \dots, Y_{n1}) \\ \vdots \\ \text{rank}(Y_{ik}; Y_{1k}, \dots, Y_{nk}) \end{pmatrix},$$

où  $Y_{ij}$  désigne la  $j$ ème composante de  $\mathbf{Y}_i$  (on peut bien entendu définir de façon similaire le rang  $\mathbf{R}_i^+$  de  $\mathbf{Y}_i$ , en remplaçant tous les  $Y_{ij}$  par leur valeur absolue).

Ces signes et rangs vectoriels souffrent malheureusement d'un manque d'invariance important. En particulier, ils ne sont pas invariants (ni équivariants) sous l'action du groupe des transformations linéaires et donc dépendent fortement du choix, souvent arbitraire, d'un système de coordonnées (cfr l'exemple des deux statisticiens utilisant des dispositifs de mesure différents). L'invariance étant l'essence même des concepts de signes et de rangs, cette extension multivariée est donc insatisfaisante. En outre, le  $n$ -uple  $(\mathbf{R}_1, \dots, \mathbf{R}_n)$  n'est pas libre. De nombreuses procédures de test ont néanmoins été basées sur ces signes et rangs en composantes (en fait, des tests fondés sur la loi permutationnelle de  $(\mathbf{R}_1, \dots, \mathbf{R}_n)$  : voir Puri et Sen (1971) pour un catalogue assez exhaustif de ces procédures dans les modèles à observations indépendantes, ainsi que Hallin et al. (1989) pour une application de ces signes et rangs dans un problème sériel simple).

## 3.2 L'approche d'Oja et de ses collaborateurs.

Développer des signes et des rangs multivariés ayant de bonnes propriétés d'invariance/équivariance fut la principale motivation des autres approches, et notamment celle d'Oja et de ses collaborateurs, que nous nous attachons maintenant à décrire. Ceux-ci ont principalement introduit les deux généralisations multivariées de signes et de rangs suivantes (voir Oja (1999) pour un rappel général sur l'ensemble de ces méthodes).

### 3.2.1 Les signes et les rangs spatiaux d'Oja.

Le signe d'un réel  $y$  peut être vu comme le vecteur unitaire pointant dans la direction de  $y$ . La généralisation naturelle au cas  $k$ -varié de cette fonction "signe" est bien entendu

$$\mathbf{S}(\mathbf{y}) := \begin{cases} \mathbf{y}/\|\mathbf{y}\| & \text{si } \mathbf{y} \neq \mathbf{0} \\ \mathbf{0} & \text{si } \mathbf{y} = \mathbf{0}. \end{cases}$$

Dans l'échantillon  $k$ -varié  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ , le *signe spatial*  $\mathbf{S}_i$  de  $\mathbf{Y}_i$  est défini par  $\mathbf{S}_i := \mathbf{S}(\mathbf{Y}_i)$ .

Il est clair que les signes spatiaux sont équivariants sous transformations orthogonales—plus précisément, on a  $\mathbf{S}(\mathbf{O}\mathbf{Y}_i) = \mathbf{O}\mathbf{S}(\mathbf{Y}_i)$ , pour toute matrice orthogonale  $\mathbf{O}$ —de sorte qu'il est possible de fonder, sur ces signes multivariés, des procédures de test pour le problème de position à un échantillon qui soient invariantes sous transformations orthogonales (voir Möttönen et Oja (1995)). Ces signes ne sont par contre pas équivariants sous transformations affines.

Considérons un échantillon univarié  $Y_1, \dots, Y_n$ . Le rang  $R_i$  de  $Y_i$  parmi  $Y_1, \dots, Y_n$  peut s'exprimer en fonction des signes des différences  $Y_i - Y_j$  :

$$R_i = \frac{n}{2} R_i^{(c)} + \frac{n+1}{2} := \frac{n}{2} \left[ \frac{1}{n} \sum_{j=1}^n \text{sign}(Y_i - Y_j) \right] + \frac{n+1}{2},$$

où  $R_i^{(c)}$  est le "rang centré" de  $Y_i$ . Par analogie, Möttönen et Oja (1995) définissent le rang centré de  $\mathbf{Y}_i$  dans l'échantillon  $k$ -varié  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$  par

$$\mathbf{R}_i^{(c)} := \frac{1}{n} \sum_{j=1}^n \mathbf{S}(\mathbf{Y}_i - \mathbf{Y}_j).$$

De la même manière, à une standardisation près, le rang  $R_i^+$  de  $|Y_i|$  parmi  $|Y_1|, \dots, |Y_n|$  peut s'écrire  $R_{i;\text{sym}}^+$ , où  $R_{i;\text{sym}}^+$  désigne le rang centré de  $Y_i$  dans l'échantillon  $-Y_1, \dots, -Y_n, Y_1, \dots, Y_n$ . Il est donc naturel de définir le rang "signé" (centré) de  $\mathbf{Y}_i$  dans l'échantillon  $k$ -varié  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$  par

$$\mathbf{R}_i^{+(c)} := \frac{1}{2n} \sum_{j=1}^n [\mathbf{S}(\mathbf{Y}_i - \mathbf{Y}_j) + \mathbf{S}(\mathbf{Y}_i + \mathbf{Y}_j)].$$

Tout comme les signes spatiaux, ces rangs et rangs signés spatiaux sont équivariants sous transformations orthogonales, mais pas sous transformations affines, de sorte que les procédures qui en résultent sont invariantes par transformations orthogonales seulement (voir Möttönen et Oja (1995), et Möttönen et al. (1997)).

### 3.2.2 Les signes et les rangs affine-équivariants d'Oja.

La seconde généralisation due à Oja, par contre, débouche sur des procédures affine-invariantes. Les signes et rangs associés sont cependant plus délicats à définir.

Dans le cas univarié, le signe du réel  $y$  peut être vu comme la dérivée (par rapport à  $y$ ) du “volume”  $V(y) = |y|$  du simplexe déterminé par 0 et  $y$ , c'est-à-dire la longueur de l'intervalle  $[\min(0, y), \max(0, y)]$ . Dans le cas multivarié, l'origine de  $\mathbb{R}^k$  et un point  $\mathbf{y} \neq \mathbf{0}$  déterminent également un simplexe, le segment  $\{\lambda\mathbf{y}, \lambda \in [0, 1]\}$ . Le “volume” de ce simplexe étant  $\|\mathbf{y}\|$ , une généralisation multivariée du signe de  $\mathbf{y}$  est le gradient

$$\nabla\|\mathbf{y}\| = \frac{\mathbf{y}}{\|\mathbf{y}\|},$$

qui n'est rien d'autre que le signe spatial de  $\mathbf{y}$  défini précédemment.

L'idée menant au concept de signe affine-équivariant consiste à associer à  $\mathbf{y}$  un simplexe de  $\mathbb{R}^k$  d'intérieur non vide. Considérons pour ce faire un échantillon  $k$ -varié  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ . Soit  $\mathcal{Q} := \{\mathbf{q} = (i_1, i_2, \dots, i_{k-1}) \text{ où } 1 \leq i_1 < i_2 < \dots < i_{k-1} \leq n\}$ . Tout élément  $\mathbf{q}$  de  $\mathcal{Q}$  détermine un ensemble de  $k-1$  observations (celles qui sont indicées par les composantes de  $\mathbf{q}$ ). Ces  $k-1$  observations,  $\mathbf{y}$  et  $\mathbf{0} \in \mathbb{R}^k$  déterminent un simplexe (leur enveloppe convexe), de volume  $V_{\mathbf{q}}(\mathbf{y})$  (c'est en fait le paralléloétope associé, de volume  $k!V_{\mathbf{q}}(\mathbf{y})$ , que considère Oja). Le gradient de ce volume est un candidat pour l'extension du concept de signe (de  $\mathbf{y}$ ); mais, par souci de symétrie, on ne privilégiera aucun  $\mathbf{q} \in \mathcal{Q}$  en définissant le *signe affine-équivariant* de  $\mathbf{y}$  par

$$\tilde{\mathbf{S}}(\mathbf{y}) := \binom{n}{k-1}^{-1} (k!) \sum_{\mathbf{q} \in \mathcal{Q}} \nabla_{\mathbf{y}} V_{\mathbf{q}}(\mathbf{y}).$$

Le signe affine-équivariant de  $\mathbf{Y}_i$  est bien sûr défini par  $\tilde{\mathbf{S}}(\mathbf{Y}_i)$ .

Contrairement aux signes spatiaux, ces nouveaux signes ont de bonnes propriétés d'équivariance sous transformations affines. Plus précisément, on peut montrer que

$$\tilde{\mathbf{S}}_{\mathbf{M}}(\mathbf{M}\mathbf{Y}_i) = |\det \mathbf{M}| (\mathbf{M}^{-1})^T \tilde{\mathbf{S}}(\mathbf{Y}_i),$$

où  $\tilde{\mathbf{S}}_{\mathbf{M}}(\mathbf{M}\mathbf{Y}_i)$  désigne le signe affine-équivariant de  $\mathbf{M}\mathbf{Y}_i$  dans l'échantillon  $\mathbf{M}\mathbf{Y}_1, \dots, \mathbf{M}\mathbf{Y}_n$  (où  $\mathbf{M}$  est une matrice  $k \times k$  inversible).

De manière similaire, on peut définir le *rang affine-équivariant* de  $\mathbf{y}$  par

$$\tilde{\mathbf{R}}(\mathbf{y}) := \binom{n}{k}^{-1} (k!) \sum_{\mathbf{p} \in \mathcal{P}} \nabla_{\mathbf{y}} V_{\mathbf{p}}(\mathbf{y}),$$

où  $\mathcal{P} := \{\mathbf{p} = (i_1, i_2, \dots, i_k) \text{ où } 1 \leq i_1 < i_2 < \dots < i_k \leq n\}$  et  $V_{\mathbf{p}}(\mathbf{y})$  désigne le volume du simplexe déterminé par  $\mathbf{y}$  et les  $k$  observations indicées par  $\mathbf{p}$ , ainsi que le *rang signé affine-équivariant* de  $\mathbf{y}$  par  $\tilde{\mathbf{R}}^+(\mathbf{y}) := \tilde{\mathbf{R}}_{\text{sym}}(\mathbf{y})$ , où  $\tilde{\mathbf{R}}_{\text{sym}}(\mathbf{y})$  désigne le rang affine-équivariant de  $\mathbf{y}$  dans l'échantillon  $-\mathbf{Y}_1, \dots, -\mathbf{Y}_n, \mathbf{Y}_1, \dots, \mathbf{Y}_n$ . Le rang affine-équivariant (resp., le rang signé affine-équivariant) de  $\mathbf{Y}_i$ , défini par  $\tilde{\mathbf{R}}(\mathbf{Y}_i)$  (resp., par  $\tilde{\mathbf{R}}^+(\mathbf{Y}_i)$ ), satisfait à la relation d'équivariance affine  $\tilde{\mathbf{R}}_{\mathbf{M}}(\mathbf{M}\mathbf{Y}_i) = |\det \mathbf{M}| (\mathbf{M}^{-1})^T \tilde{\mathbf{R}}(\mathbf{Y}_i)$  (resp.,  $\tilde{\mathbf{R}}_{\mathbf{M}}^+(\mathbf{M}\mathbf{Y}_i) = |\det \mathbf{M}| (\mathbf{M}^{-1})^T \tilde{\mathbf{R}}^+(\mathbf{Y}_i)$ ).

L'équivariance affine de ces signes et rangs permet de définir des procédures de test (basées sur les signes ou les rangs signés) qui sont invariantes sous transformations affines, notamment pour le problème de position à un échantillon que nous avons considéré jusqu'ici (voir Hettmansperger et al. (1994)). Ces procédures affine-invariantes sont valides sous une hypothèse très générale de symétrie, à savoir une hypothèse de symétrie centrale (voir Hettmansperger et al. (1994) et (1997), Möttönen et al. (1998), et Oja (1999)). Il semble néanmoins difficile de construire des procédures optimales fondées sur ces signes et rangs multivariés. C'est la motivation principale qui nous a conduit à fonder nos procédures de rangs signés sur les signes et les rangs multivariés décrits dans les Sections 3.3 et 3.4.

### 3.3 Signes angulaires et rangs des distances à l'origine (I).

#### 3.3.1 Des signes angulaires.

Il est clair que la notion (absolue) de signe est bien souvent trop forte pour l'usage statistique qui en est fait habituellement. Pour illustrer ceci, remarquons que l'information contenue dans le vecteur des signes  $(S_1, \dots, S_n)$  d'un échantillon i.i.d. symétrique est équivalente à celle qui est contenue dans le vecteur  $(S_1, S_1 S_i, 1 < i \leq n)$ , ou encore,  $(S_1, S_i S_j, 1 < i < j \leq n)$ . Si le problème considéré est invariant par symétrie par rapport à l'origine, la valeur de  $S_1$  ne véhicule aucune information. Il en découle que, pour de tels problèmes, toute l'information sur  $\theta$  contenue dans les signes, réside en fait dans le vecteur

$$(S_i S_j, 1 < i < j \leq n).$$

Autrement dit, le signe absolu d'une observation peut toujours être ignoré; seul compte, pour toute paire d'observations, le fait que ces observations aient le même signe ou des signes opposés.

Il en va de même dans le cas multivarié. En effet, l'information contenue dans les signes

$$(\mathbf{U}_1(\boldsymbol{\Sigma}), \dots, \mathbf{U}_n(\boldsymbol{\Sigma})) \quad (9)$$

est équivalente (pour  $n$  suffisamment grand) à celle résidant dans

$$(\mathbf{U}_1(\boldsymbol{\Sigma}), \mathbf{U}'_i(\boldsymbol{\Sigma})\mathbf{U}_j(\boldsymbol{\Sigma}), 1 < i < j \leq n).$$

Et comme dans le cas univarié, le vecteur  $\mathbf{U}_1(\boldsymbol{\Sigma})$  peut, sans perte d'information, être omis, puisque le problème considéré (le problème (8)) est invariant par rotation (de manière équivalente, la racine carrée  $\boldsymbol{\Sigma}^{1/2}$ , qui n'est définie qu'à une transformation orthogonale près, peut toujours être choisie de sorte que  $\mathbf{U}_1(\boldsymbol{\Sigma})$  soit un vecteur fixé de  $\mathbb{R}^k$ ). En pratique cependant, et ceci est propre au cas multivarié, le vecteur

$$(\mathbf{U}'_i(\boldsymbol{\Sigma})\mathbf{U}_j(\boldsymbol{\Sigma}), 1 < i < j \leq n) \quad (10)$$

des cosinus des angles entre les  $\mathbf{U}_i(\boldsymbol{\Sigma})$  n'est pas une statistique, puisque la matrice de forme  $\boldsymbol{\Sigma}$  est un paramètre de nuisance dans le problème de test (8). Nous proposons donc de remplacer (10) par

$$(\mathbf{U}'_i(\hat{\boldsymbol{\Sigma}})\mathbf{U}_j(\hat{\boldsymbol{\Sigma}}), 1 < i < j \leq n), \quad (11)$$

pour un certain estimateur  $\hat{\boldsymbol{\Sigma}}$  de la forme  $\boldsymbol{\Sigma}$ . Le choix de cet estimateur est guidé par nos desiderata en termes d'invariance et de convergence en l'absence d'hypothèse sur les moments de la loi sous-jacente.

Un choix naturel est bien entendu la matrice de variance-covariance empirique

$$\hat{\boldsymbol{\Sigma}}_v := \frac{1}{n} \sum_{i=1}^n (\mathbf{Y}_i - \bar{\mathbf{Y}})(\mathbf{Y}_i - \bar{\mathbf{Y}})', \quad \text{où } \bar{\mathbf{Y}} := \frac{1}{n} \sum_{i=1}^n \mathbf{Y}_i. \quad (12)$$

Bien que les cosinus empiriques associés  $\mathbf{U}'_i(\hat{\boldsymbol{\Sigma}}_v)\mathbf{U}_j(\hat{\boldsymbol{\Sigma}}_v)$  soient invariants par transformations affines comme par transformations monotones radiales, il est peu judicieux de recourir à cet estimateur, puisqu'il requiert, pour estimer  $\boldsymbol{\Sigma}$  de façon convergente, que la loi sous-jacente ait des moments finis d'ordre deux. Puisque nous voulons développer des procédures qui soient valables sans aucune hypothèse de moments, nous proposons d'utiliser plutôt l'estimateur de forme de Tyler (1987). Celui-ci est défini comme suit : soit  $\hat{\boldsymbol{\Sigma}}_T^{-1/2} = \mathbf{C}_T$  l'unique (pour  $n > k(k-1)$ ) matrice  $k \times k$  triangulaire supérieure, de diagonale positive, dont l'élément supérieur gauche est "1", et qui satisfasse à la relation

$$\frac{1}{n} \sum_{i=1}^n \left( \frac{\mathbf{C}_T \mathbf{Y}_i}{\|\mathbf{C}_T \mathbf{Y}_i\|} \right) \left( \frac{\mathbf{C}_T \mathbf{Y}_i}{\|\mathbf{C}_T \mathbf{Y}_i\|} \right)' = \frac{1}{k} \mathbf{I}_k, \quad (13)$$

où  $\mathbf{I}_k$  désigne la matrice identité de dimension  $k$ . La matrice  $\mathbf{C}_T$  est donc choisie de sorte que la structure de covariance empirique du  $n$ -uplet

$$\left( \frac{\mathbf{C}_T \mathbf{Y}_1}{\|\mathbf{C}_T \mathbf{Y}_1\|}, \dots, \frac{\mathbf{C}_T \mathbf{Y}_n}{\|\mathbf{C}_T \mathbf{Y}_n\|} \right)$$

soit celle d'un échantillon uniformément distribué sur la sphère unité  $\mathcal{S}^{k-1} \subset \mathbb{R}^k$ . L'estimateur de Tyler est alors défini par  $\hat{\Sigma}_T := [(\mathbf{C}_T)' \mathbf{C}_T]^{-1}$ . Cet estimateur assure également l'invariance affine des cosinus estimés  $\mathbf{U}'_i(\hat{\Sigma}_T)\mathbf{U}_j(\hat{\Sigma}_T)$  (voir Randles 2000). Il est de plus clairement invariant sous des transformations radiales quelconques des observations, si bien que les propriétés d'invariance des cosinus estimés sont aussi bonnes si l'on utilise  $\hat{\Sigma}_T$  que si l'on utilise  $\hat{\Sigma}_v$ . Par contre, l'estimateur de Tyler estime  $\Sigma$ —à un facteur positif près, ce qui est clairement suffisant pour estimer les cosinus (10)—de manière convergente sans que soit requise aucune hypothèse sur la densité radiale de la loi elliptique sous-jacente. En particulier, cet estimateur tolère des moments d'ordre deux infinis (voir Tyler 1987).

Randles (2000) propose le test *de signe*  $\phi_S$  (de niveau asymptotique  $\alpha$ ) qui consiste à rejeter l'hypothèse nulle de (8) dès que

$$S := \frac{k}{n} \sum_{i,j=1}^n \mathbf{U}'_i(\hat{\Sigma}_T)\mathbf{U}_j(\hat{\Sigma}_T) > \chi_{k,1-\alpha}^2, \quad (14)$$

où  $\chi_{k,1-\alpha}^2$  désigne le quantile d'ordre  $1 - \alpha$  d'une variable chi-deux à  $k$  degré de liberté. Comme tout test de signe, cette procédure ne fait intervenir les observations qu'à travers leur direction depuis l'origine de  $\mathbb{R}^k$  (la distance à l'origine n'influence aucunement la valeur de la statistique de test ci-dessus). Le test de Randles est fondé sur des quantités invariantes sous les deux groupes de transformations considérés, et est donc libre sous la classe des lois elliptiques (sans conditions de moments). Ce test est de plus simple à mettre en oeuvre (en termes de temps de calcul), même pour  $k$  grand; voir Randles (2000) pour un algorithme itératif efficace permettant de calculer l'estimateur de Tyler.

Pour étudier les propriétés d'efficacité asymptotique de ce test, nous considérons l'extension multivariée du test de Student, à savoir le test de Hotelling. Dans sa version asymptotique, celui-ci consiste (au niveau  $\alpha$ ) à rejeter l'hypothèse nulle dès que

$$T^{2(n)} := n \bar{\mathbf{Y}}^{(n)'} (\mathbf{S}^{(n)})^{-1} \bar{\mathbf{Y}}^{(n)} > \chi_{k,1-\alpha}^2,$$

où  $\bar{\mathbf{Y}}^{(n)} := \frac{1}{n} \sum_{i=1}^n \mathbf{Y}_i$  et  $\mathbf{S}^{(n)} := \frac{1}{n-1} \sum_{i=1}^n (\mathbf{Y}_i - \bar{\mathbf{Y}}^{(n)})(\mathbf{Y}_i - \bar{\mathbf{Y}}^{(n)})'$  désignent respectivement la moyenne et la matrice de variance-covariance empiriques de l'échantillon. Ce test invariant par transformations linéaires est de niveau asymptotique  $\alpha$  sous la collection des distributions elliptiques de variance finie, et est optimal sous des densités gaussiennes.

Les efficacités asymptotiques relatives de ce test de signe par rapport au test de Hotelling  $T^2$ , sous les lois de Student  $k$ -variées à  $\nu$  degrés de liberté et normale (pour une quelconque valeur de  $\Sigma$ , puisque les deux procédures sont affine-invariantes), sont données par

$$\text{ARE}_{k,\nu}[\phi_S/T^2] = \frac{4}{k(\nu-2)} \left[ \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \right]^2 \quad \text{et} \quad \text{ARE}_{k,\mathcal{N}}[\phi_S/T^2] = \frac{2}{k} \left[ \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})} \right]^2, \quad (15)$$

respectivement, où  $\Gamma$  est la fonction Gamma d'Euler. La Table 1 reprend quelques valeurs numériques de ces efficacités relatives.

### 3.3.2 Des rangs des distances à l'origine.

Les cosinus  $\mathbf{U}'_i(\Sigma)\mathbf{U}_j(\Sigma)$  sont invariants sous l'action du groupe  $\mathcal{G}_\Sigma^{(n)}$  des transformations monotones radiales. Et une version empirique ad hoc de ceux-ci est invariante sous le groupe  $\mathcal{G}^{(n)}$  des transformations linéaires comme sous le groupe des transformations monotones radiales  $\mathcal{G}_\Sigma^{(n)}$ , pour tout  $\Sigma$ . En utilisant ces angles estimés, on peut donc développer, comme on vient de le voir, des procédures invariantes, donc libres.

Néanmoins, les cosinus  $\mathbf{U}'_i(\Sigma)\mathbf{U}_j(\Sigma)$  ne constituent pas un invariant maximal pour  $\mathcal{G}_\Sigma^{(n)}$ . Se restreindre aux angles se traduit donc par une perte d'information stricte pour le problème de test considéré. La question naturelle est donc : quel est l'invariant maximal?

Si le caractère radial des transformations de  $\mathcal{G}_\Sigma^{(n)}$  assure que les directions  $\mathbf{U}_i(\Sigma)$ —et, par conséquent, les cosinus  $\mathbf{U}'_i(\Sigma)\mathbf{U}_j(\Sigma)$ —sont invariants sous l'action de  $\mathcal{G}_\Sigma^{(n)}$ , la monotonie des transformations de ce

groupe implique, comme dans le cas univarié (voir la Section 1), que les rangs

$$R_i(\boldsymbol{\Sigma}) := \text{rank}(d_i(\boldsymbol{\Sigma}); d_1(\boldsymbol{\Sigma}), \dots, d_n(\boldsymbol{\Sigma}))$$

des distances à l'origine  $d_1(\boldsymbol{\Sigma}), \dots, d_n(\boldsymbol{\Sigma})$  sont également invariants. On vérifie aisément, par analogie avec le cas univarié, que

$$(\mathbf{U}_1(\boldsymbol{\Sigma}), \dots, \mathbf{U}_n(\boldsymbol{\Sigma}), R_1(\boldsymbol{\Sigma}), \dots, R_n(\boldsymbol{\Sigma})) \quad (16)$$

est un invariant maximal pour le groupe  $\mathcal{G}_{\boldsymbol{\Sigma}}^{(n)}$ . Dans le cas univarié, le vecteur (16) se réduit au vecteur (2) des signes et des rangs des valeurs absolues. Pour  $n$  suffisamment grand, le vecteur

$$(\mathbf{U}_1(\boldsymbol{\Sigma}), \mathbf{U}'_i(\boldsymbol{\Sigma})\mathbf{U}_j(\boldsymbol{\Sigma}), R_1(\boldsymbol{\Sigma}), \dots, R_n(\boldsymbol{\Sigma}), 1 < i < j \leq n)$$

est également un invariant maximal. Bien entendu, on omettra encore, pour les mêmes raisons que plus haut, le vecteur  $\mathbf{U}_1(\boldsymbol{\Sigma})$ . De manière similaire à ce qui a été fait pour les signes, nous surmonterons la difficulté constituée par le fait que, dans le cas multivarié, le vecteur aléatoire

$$(\mathbf{U}'_i(\boldsymbol{\Sigma})\mathbf{U}_j(\boldsymbol{\Sigma}), R_1(\boldsymbol{\Sigma}), \dots, R_n(\boldsymbol{\Sigma}), 1 < i < j \leq n) \quad (17)$$

n'est pas une statistique en remplaçant (17) par

$$(\mathbf{U}'_i(\hat{\boldsymbol{\Sigma}}_T)\mathbf{U}_j(\hat{\boldsymbol{\Sigma}}_T), R_1(\hat{\boldsymbol{\Sigma}}_T), \dots, R_n(\hat{\boldsymbol{\Sigma}}_T), 1 < i < j \leq n), \quad (18)$$

où le *pseudo-rang de Mahalanobis*  $R_i(\hat{\boldsymbol{\Sigma}}_T)$  désigne le rang de  $d_i(\hat{\boldsymbol{\Sigma}}_T)$  parmi les *pseudo-distances de Mahalanobis*  $d_1(\hat{\boldsymbol{\Sigma}}_T), \dots, d_n(\hat{\boldsymbol{\Sigma}}_T)$ . Les termes *pseudo-rang de Mahalanobis* et *pseudo-distance de Mahalanobis* sont utilisés ici pour insister sur le fait que l'estimateur  $\hat{\boldsymbol{\Sigma}}_T$  de Tyler est utilisé en lieu et place de la matrice de variance-covariance empirique  $\hat{\boldsymbol{\Sigma}}_v$  (si on utilisait  $\hat{\boldsymbol{\Sigma}}_v$ , on parlerait simplement de *rang de Mahalanobis* et de *distance de Mahalanobis*). Les pseudo-rangs de Mahalanobis—comme les rangs de Mahalanobis—sont invariants par rapport à  $\mathcal{G}^{(n)}$  et asymptotiquement invariants par rapport à  $\mathcal{G}_{\boldsymbol{\Sigma}}^{(n)}$ , dans la mesure où ils sont asymptotiquement équivalents aux vrais rangs  $R_i(\boldsymbol{\Sigma})$ , lesquels sont strictement invariants par rapport à  $\mathcal{G}_{\boldsymbol{\Sigma}}^{(n)}$  (cette équivalence asymptotique signifie plus précisément que les pseudo-rangs de Mahalanobis et les rangs de Mahalanobis sont tels que  $R_i(\hat{\boldsymbol{\Sigma}}_T) = R_i(\boldsymbol{\Sigma}) + o_P(n)$ , pour  $n \rightarrow \infty$ ).

Il est facile de fonder, sur ces signes et rangs multivariés, une extension multivariée du test de Wilcoxon (3); celle-ci consiste à rejeter l'hypothèse nulle de (8) si

$$\frac{3k}{n^3} \sum_{i,j=1}^n R_i(\hat{\boldsymbol{\Sigma}}_T) R_j(\hat{\boldsymbol{\Sigma}}_T) \mathbf{U}'_i(\hat{\boldsymbol{\Sigma}}_T) \mathbf{U}_j(\hat{\boldsymbol{\Sigma}}_T) > \chi_{k,1-\alpha}^2. \quad (19)$$

Ce test (de niveau asymptotique  $\alpha$ ) est strictement invariant sous transformations affines et asymptotiquement invariants sous les transformations monotones radiales, donc valable sous l'ensemble des lois elliptiques, sans aucune condition de moment. Des valeurs numériques de l'efficacité asymptotique relative de ce test par rapport au test de Hotelling sous des lois  $t$  et normales multivariées sont fournies dans la Table 1.

Comme nous l'avons vu, le principe d'invariance préconise de fonder les procédures de test pour le problème (8) sur des statistiques qui sont des fonctions des pseudo-rangs de Mahalanobis  $R_i(\hat{\boldsymbol{\Sigma}}_T)$  et des cosinus  $\mathbf{U}'_i(\hat{\boldsymbol{\Sigma}}_T)\mathbf{U}_j(\hat{\boldsymbol{\Sigma}}_T)$ . Le problème du choix de la statistique en question doit reposer sur des considérations de puissance, et sera abordé dans la Section 4. Nous présentons d'abord une dernière extension multivariée des concepts de signes et de rangs, qui peut avantageusement être substituée aux signes angulaires et aux rangs des distances à l'origine décrits ci-dessus.

### 3.4 Signes angulaires et rangs des distances à l'origine (II).

Dans cette section, nous expliquons comment récupérer asymptotiquement l'information contenue dans les cosinus  $\mathbf{U}'_i(\boldsymbol{\Sigma})\mathbf{U}_j(\boldsymbol{\Sigma})$  et les rangs des distances à l'origine  $R_i(\boldsymbol{\Sigma})$ , *sans estimer la matrice de forme  $\boldsymbol{\Sigma}$* . Ceci est réalisable au moyen des *interdirections de Randles* et des *lift-interdirections d'Oja*.

		Degrés de liberté de la loi $t$ sous-jacente							
$k$	test	3	4	6	8	10	15	20	$\infty$
1	$\phi_S$	1.621	1.125	0.879	0.798	0.757	0.710	0.690	0.637
	$\phi_W$	1.900	1.401	1.164	1.089	1.054	1.014	0.997	0.955
2	$\phi_S$	2.000	1.388	1.084	0.984	0.934	0.877	0.851	0.785
	$\phi_W$	1.748	1.317	1.123	1.066	1.041	1.015	1.005	0.985
3	$\phi_S$	2.162	1.500	1.172	1.063	1.009	0.947	0.920	0.849
	$\phi_W$	1.621	1.233	1.064	1.019	1.000	0.983	0.978	0.975
4	$\phi_S$	2.250	1.561	1.220	1.107	1.051	0.986	0.958	0.884
	$\phi_W$	1.533	1.171	1.018	0.979	0.964	0.954	0.952	0.961
6	$\phi_S$	2.344	1.626	1.271	1.153	1.094	1.027	0.997	0.920
	$\phi_W$	1.422	1.090	0.953	0.921	0.911	0.908	0.911	0.938
10	$\phi_S$	2.422	1.681	1.313	1.192	1.131	1.062	1.031	0.951
	$\phi_W$	1.315	1.007	0.882	0.855	0.848	0.851	0.857	0.907

Table 1: Efficacités asymptotiques relatives, par rapport au test de Hotelling, du test de signe  $\phi_S$  de Randles (2000), et de l'extension multivariée du test de Wilcoxon  $\phi_W$ , sous diverses distributions  $t$  et normale  $k$ -variées,  $k = 1, 2, 3, 4, 6, 10$ .

### 3.4.1 Une version “non paramétrique” des signes angulaires.

La contribution la plus ingénieuse de Randles, dans le domaine de la généralisation multivariée des signes et des rangs, est probablement l'introduction des *interdirections*. Les interdirections de Randles sont une extension des signes univariés, dans le sens où elles fournissent une autre approche dans l'estimation des cosinus  $\mathbf{U}'_i(\boldsymbol{\Sigma})\mathbf{U}_j(\boldsymbol{\Sigma})$ .

Considérons l'échantillon  $k$ -varié  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ . L'interdirection  $c_{ij} = c_{ij}^{(n)}$  associée à la paire  $(\mathbf{Y}_i, \mathbf{Y}_j)$  dans cet échantillon est, par définition, le nombre d'hyperplans passant par  $(k-1)$  des  $(n-2)$  points  $\mathbf{Y}_1, \dots, \mathbf{Y}_{i-1}, \mathbf{Y}_{i+1}, \dots, \mathbf{Y}_{j-1}, \mathbf{Y}_{j+1}, \dots, \mathbf{Y}_n$  qui séparent  $\mathbf{Y}_i$  et  $\mathbf{Y}_j$ , c'est-à-dire le nombre de tels hyperplans qui laissent  $\mathbf{Y}_i$  et  $\mathbf{Y}_j$  dans deux demi-espaces distincts (voir Figure 1).

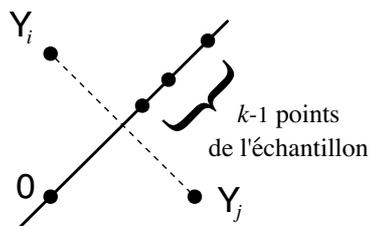


Figure 1: Définition des interdirections de Randles.

L'interdirection  $c_{ij}$  est donc un nombre entier compris entre 0 et  $\binom{n-2}{k-1}$ . A titre d'illustration, l'interdirection  $c_{ij}$  dans l'échantillon représenté à la Figure 2 vaut deux (deux hyperplans séparateurs sur un total de trois hyperplans à considérer).

Il est clair que la valeur des interdirections reste inchangée si on déplace les observations radialement le long des demi-droites issues de l'origine de  $\mathbb{R}^k$ . De même, on peut montrer que les interdirections sont invariantes par transformations linéaires (voir Randles 1989). Les interdirections sont donc invariantes sous  $\mathcal{G}^{(n)}$  et sous  $\mathcal{G}_{\boldsymbol{\Sigma}}^{(n)}$ , pour tout  $\boldsymbol{\Sigma}$ ; elles ont par conséquent exactement les mêmes propriétés d'invariance que les cosinus estimés  $\mathbf{U}'_i(\hat{\boldsymbol{\Sigma}}_T)\mathbf{U}_j(\hat{\boldsymbol{\Sigma}}_T)$ .

Voyons maintenant comment elles permettent d'estimer les cosinus vrais  $\mathbf{U}'_i(\boldsymbol{\Sigma})\mathbf{U}_j(\boldsymbol{\Sigma})$ . Si les observations sont indépendantes et uniformément distribuées sur la sphère unité  $\mathcal{S}^{k-1} \subset \mathbb{R}^k$ , il est clair que

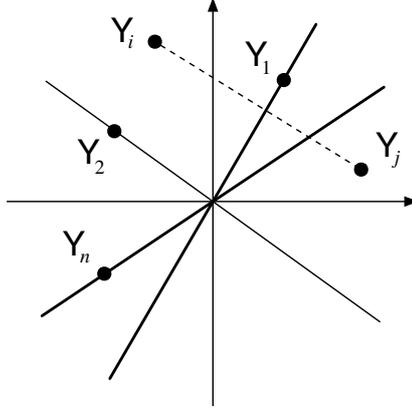


Figure 2: Illustration des interdirections de Randles.

la proportion d'hyperplans séparateurs pour la paire  $(\mathbf{Y}_i, \mathbf{Y}_j)$ , c'est-à-dire  $p_{ij}^{(n)} := c_{ij}^{(n)} / \binom{n-2}{k-1}$ , converge, quand la taille de l'échantillon tend vers l'infini, vers la proportion de l'angle compris entre  $\mathbf{Y}_i$  et  $\mathbf{Y}_j$  dans "l'angle maximal"  $\pi$ . Autrement dit, on a (voir Randles 1989)

$$p_{ij}^{(n)} = \frac{1}{\pi} \arccos \left( \frac{\mathbf{Y}_i' \mathbf{Y}_j}{\|\mathbf{Y}_i\| \|\mathbf{Y}_j\|} \right) + o_{\mathbb{P}}(1), \quad \text{pour } n \rightarrow \infty.$$

Que se passe-t-il sous  $\bigcup_f \mathcal{H}^{(n)}(\mathbf{0}, \Sigma, f)$ , c'est-à-dire si la loi sous-jacente est à symétrie elliptique de matrice de forme  $\Sigma$ ? Notons, pour un instant,  $p^{(n)}(\mathbf{Y}_i, \mathbf{Y}_j; \mathbf{Y}_1, \dots, \mathbf{Y}_n) := \binom{n-2}{k-1}^{-1} c^{(n)}(\mathbf{Y}_i, \mathbf{Y}_j; \mathbf{Y}_1, \dots, \mathbf{Y}_n)$ , où  $c^{(n)}(\mathbf{Y}_i, \mathbf{Y}_j; \mathbf{Y}_1, \dots, \mathbf{Y}_n)$  désigne l'interdirection associée à  $(\mathbf{Y}_i, \mathbf{Y}_j)$  dans l'échantillon  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ . Alors, en utilisant l'invariance des interdirections par transformations linéaires et par transformations radiales et le fait que, sous  $\bigcup_f \mathcal{H}^{(n)}(\mathbf{0}, \Sigma, f)$ , les vecteurs aléatoires  $\mathbf{U}_1(\Sigma), \mathbf{U}_2(\Sigma), \dots, \mathbf{U}_n(\Sigma)$  sont uniformément distribués sur  $\mathcal{S}^{k-1}$ , on obtient

$$\begin{aligned} p^{(n)}(\mathbf{Y}_i, \mathbf{Y}_j; \mathbf{Y}_1, \dots, \mathbf{Y}_n) &= p^{(n)}(\mathbf{U}_i(\Sigma), \mathbf{U}_j(\Sigma); \mathbf{U}_1(\Sigma), \dots, \mathbf{U}_n(\Sigma)) \\ &= \frac{1}{\pi} \arccos \left( \mathbf{U}_i'(\Sigma) \mathbf{U}_j(\Sigma) \right) + o_{\mathbb{P}}(1), \end{aligned}$$

pour  $n \rightarrow \infty$ , sous  $\bigcup_f \mathcal{H}^{(n)}(\mathbf{0}, \Sigma, f)$ . Par conséquent, si l'échantillon  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$  est à symétrie elliptique de matrice de forme  $\Sigma$ , la statistique  $\cos(\pi p_{ij}^{(n)})$  fournit un estimateur affine-invariant convergent du cosinus vrai  $\mathbf{U}_i'(\Sigma) \mathbf{U}_j(\Sigma)$ , et ce *sans qu'il soit nécessaire d'estimer la matrice de forme  $\Sigma$* .

L'équivalent fondé sur les interdirections du test de signe (14) consiste à rejeter  $\mathcal{H}_0$  si

$$\frac{k}{n} \sum_{i,j=1}^n \cos \pi p_{ij} > \chi_{k,1-\alpha}^2. \quad (20)$$

Ce test (de niveau asymptotique  $\alpha$ ) est asymptotiquement équivalent au test (14) fondé sur l'estimateur de Tyler; par conséquent, les efficacités asymptotiques relatives des deux tests par rapport au test de Hotelling coïncident (quelques valeurs numériques de celles-ci sous diverses lois  $t$  sont données dans la Table 1). La procédure fondée sur les interdirections, qui est plus lourde d'un point de vue calculatoire, a l'avantage d'être valide sous une classe de lois un peu plus large que celle des lois elliptiques, à savoir celle des *lois à directions elliptiques* (voir Randles 1989 pour plus détails).

### 3.4.2 Une version "non paramétrique" des rangs des distances à l'origine.

Il est remarquable que les interdirections, qui sont des statistiques affine-invariantes basées sur les hyperplans passant par les observations, permettent d'estimer les angles entre les observations sphéricisées,

sans qu'il soit nécessaire de connaître ni d'estimer la matrice de forme (qui permet de sphériciser les observations). De manière analogue, nous définissons un nouveau concept de distances (affine-invariantes) entre les observations et l'origine de  $\mathbb{R}^k$ , basées également sur les hyperplans passant par les observations.

Plus précisément, la *lift-interdirection*  $l_i = l_i^{(n)}$  associée à l'observation  $\mathbf{Y}_i$  dans l'échantillon  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$  est, par définition, le nombre d'hyperplans passant par  $k$  points parmi les  $(n-1)$  points  $\mathbf{Y}_1, \dots, \mathbf{Y}_{i-1}, \mathbf{Y}_{i+1}, \dots, \mathbf{Y}_n$ , et qui séparent  $\mathbf{Y}_i$  et  $-\mathbf{Y}_i$ .

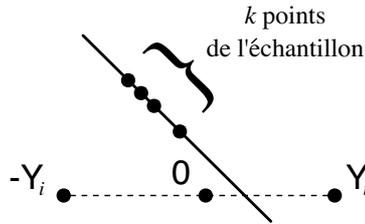


Figure 3: Définition des lift-interdirections.

La lift-interdirection  $l_i$  est donc un nombre entier compris entre 0 et  $\binom{n-1}{k}$ . A titre d'illustration, l'interdirection  $l_i$  dans l'échantillon représenté à la Figure 4 vaut deux (deux hyperplans séparateurs sur un total de trois hyperplans à considérer).

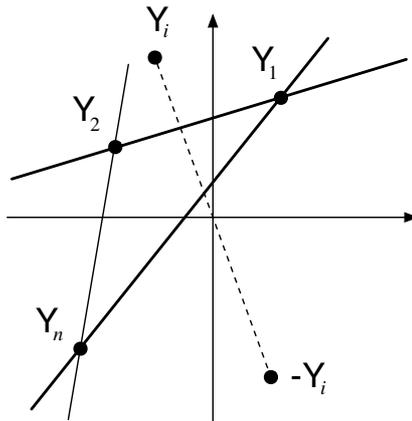


Figure 4: Illustration des lift-interdirections.

Dans le cas sphérique, la lift-interdirection associée à l'observation  $\mathbf{Y}_i$  converge vers une fonction monotone croissante de la distance  $d_i(\mathbf{I}_k) = \|\mathbf{Y}_i\|$  entre  $\mathbf{Y}_i$  et l'origine de  $\mathbb{R}^k$ . Comme d'autre part elles sont invariantes par transformations linéaires, ces lift-interdirections estiment, dans le cas elliptique, une fonction monotone croissante des distances entre les observations sphéricisées et l'origine, c'est-à-dire une fonction monotone croissante des  $d_i(\boldsymbol{\Sigma})$  (où  $\boldsymbol{\Sigma}$  désigne la "vraie" matrice de forme). Comme pour les interdirections, on parvient donc à estimer les quantités sphéricisées (dans ce cas précis, les distances à l'origine des observations sphéricisées) sans estimer la matrice de forme.

Puisque les lift-interdirections estiment une fonction monotone croissante des  $d_i(\boldsymbol{\Sigma})$ , on peut espérer que les rangs des lift-interdirections soient asymptotiquement équivalents aux pseudo-rangs de Mahalanobis, de sorte qu'on puisse remplacer ces derniers par les rangs des lift-interdirections, dans les procédures définies précédemment. Oja et Paindaveine (2003) prouvent que les deux types de rangs sont effectivement asymptotiquement équivalents. Néanmoins, pour fonder sur les rangs de ces distances

empiriques (et les signes interdirections) des procédures de rangs signés, on utilise plutôt, pour des raisons techniques, des *lift-interdirections symétrisées*, dans lesquelles on considère les hyperplans passant par les observations et par les symétriques des observations par rapport à l'origine de  $\mathbb{R}^k$ . Oja et Paindaveine (2003) vérifient que les deux types de lift-interdirections ont les mêmes propriétés d'invariance, sont asymptotiquement équivalentes, et que le caractère symétrique des lift-interdirections symétrisées permet de substituer avec succès leurs rangs aux pseudo-rangs de Mahalanobis dans les procédures définies précédemment.

Par exemple, le test (de niveau asymptotique  $\alpha$ ) de Wilcoxon complètement *fondés sur les hyperplans passant par les observations* consiste à rejeter l'hypothèse nulle du problème (8) si

$$\frac{3k}{n^3} \sum_{i,j=1}^n \hat{R}_i \hat{R}_j \cos(\pi p_{ij}) > \chi_{k,1-\alpha}^2, \quad (21)$$

où  $\hat{R}_i$  désigne le rang de la lift-interdirection symétrisée associée à l'observation  $\mathbf{Y}_i$  parmi les lift-interdirections symétrisées associées à  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ . Ce test jouit des mêmes propriétés d'invariance que le test (19) et est asymptotiquement équivalent à celui-ci sous la classe des lois elliptiques. Ces deux tests de Wilcoxon partagent donc les mêmes efficacités asymptotiques relatives par rapport au test de Hotelling (les valeurs correspondant au test  $\phi_W$  dans la Table 1 s'appliquent donc également ici).

Si ces procédures fondées sur les hyperplans sont particulièrement séduisantes d'un point de vue conceptuel (il est assez remarquable de pouvoir fonder des procédures—qui sont optimales, comme on le verra dans la Section 5—uniquement sur le fait que certains hyperplans séparent ou pas les paires d'observations), elles sont en revanche plus lourdes d'un point de vue calculatoire. Elles sont (de loin) les procédures les plus exigeantes en temps de calcul. Hallin et Paindaveine (2002d) ont effectué des simulations pour comparer les propriétés de robustesse des deux types de procédures (celles qui sont basées sur l'estimation de la matrice de forme et celles qui n'utilisent que les hyperplans séparateurs). Ces simulations indiquent que les deux types de tests se comportent essentiellement de la même façon si les distributions sous-jacentes sont elliptiques. Par contre, lorsque la distribution des observations n'est qu'approximativement elliptique (un mélange de deux lois elliptiques concentriques, ou de centres différents, etc), les procédures basées sur les hyperplans semblent un peu plus puissantes que les procédures basées sur l'estimation de la matrice de forme. De même, lorsqu'on étudie les qualités de robustesse de ces différentes procédures (en introduisant artificiellement un certain pourcentage d'observations aberrantes), les procédures fondées sur les hyperplans se montrent plus robustes.

## 4 Optimalité : la théorie de Hájek et Le Cam.

Nous avons déjà évoqué, dans la section précédente, nos exigences en termes d'invariance et d'optimalité. Nous avons vu plus haut comment remplir nos objectifs d'invariance en utilisant les rangs signés définis dans les Sections 3.3 et 3.4. Dans cette section, nous laissons de côté l'invariance, pour nous pencher sur la question de l'optimalité. Nous considérons une classe de modèles linéaires elliptiques, décrivons les divers types de problèmes que nous traiterons dans la suite, précisons de quels concepts d'optimalité il s'agira pour chacun de ces problèmes de test, et indiquons enfin comment cette optimalité pourra être atteinte dans chaque cas.

### 4.1 Une classe de modèles elliptiques.

Les modèles que nous considérerons dans la suite sont des cas particuliers du modèle linéaire général à erreurs VARMA. Celui-ci est caractérisé par l'équation

$$\mathbf{Y}^{(n)} = \mathbf{X}^{(n)} \boldsymbol{\beta} + \mathbf{U}^{(n)}, \quad (22)$$

où

$$\mathbf{Y}^{(n)} := \begin{pmatrix} Y_{1,1} & Y_{1,2} & \dots & Y_{1,k} \\ \vdots & \vdots & & \vdots \\ Y_{n,1} & Y_{n,2} & \dots & Y_{n,k} \end{pmatrix} =: \begin{pmatrix} \mathbf{Y}'_1 \\ \vdots \\ \mathbf{Y}'_n \end{pmatrix}$$

est le  $n$ -uple d'observations  $k$ -variées  $\mathbf{Y}_t$ ,  $t = 1, \dots, n$ ,

$$\mathbf{X}^{(n)} := \begin{pmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,m} \\ \vdots & \vdots & & \vdots \\ x_{n,1} & x_{n,2} & \dots & x_{n,m} \end{pmatrix} =: \begin{pmatrix} \mathbf{x}'_1 \\ \vdots \\ \mathbf{x}'_n \end{pmatrix}$$

est une matrice  $n \times m$  de régresseurs constants (*matrice de design*), et la matrice  $m \times k$

$$\boldsymbol{\beta} := \begin{pmatrix} \beta_{1,1} & \beta_{1,2} & \dots & \beta_{1,k} \\ \vdots & \vdots & & \vdots \\ \beta_{m,1} & \beta_{m,2} & \dots & \beta_{m,k} \end{pmatrix} =: \begin{pmatrix} \boldsymbol{\beta}'_1 \\ \vdots \\ \boldsymbol{\beta}'_m \end{pmatrix}$$

est le paramètre de régression. Plutôt que l'hypothèse traditionnelle suivant laquelle les erreurs

$$\mathbf{U}^{(n)} := \begin{pmatrix} U_{1,1} & U_{1,2} & \dots & U_{1,k} \\ \vdots & \vdots & & \vdots \\ U_{n,1} & U_{n,2} & \dots & U_{n,k} \end{pmatrix} =: \begin{pmatrix} \mathbf{U}'_1 \\ \vdots \\ \mathbf{U}'_n \end{pmatrix}$$

sont la réalisation d'un bruit blanc indépendant, nous supposons que ces vecteurs  $\mathbf{U}_t$ ,  $t = 1, \dots, n$ , constituent une réalisation finie (de longueur  $n$ ) d'une solution du modèle VARMA( $p, q$ )

$$\mathbf{A}(L) \mathbf{U}_t = \mathbf{B}(L) \boldsymbol{\varepsilon}_t, \quad t \in \mathbb{Z}, \quad (23)$$

où

$$\mathbf{A}(L) := \mathbf{I}_k - \sum_{i=1}^p \mathbf{A}_i L^i \quad \text{et} \quad \mathbf{B}(L) := \mathbf{I}_k + \sum_{i=1}^q \mathbf{B}_i L^i,$$

pour certaines matrices ( $k \times k$ )  $\mathbf{A}_1, \dots, \mathbf{A}_p, \mathbf{B}_1, \dots, \mathbf{B}_q$  à coefficients réels,  $L$  est l'opérateur *retard* (défini par  $L^i \boldsymbol{\varepsilon}_t := \boldsymbol{\varepsilon}_{t-i}$ ), et  $\{\boldsymbol{\varepsilon}_t | t \in \mathbb{Z}\}$  est un bruit blanc  $k$ -varié i.i.d. elliptique (de densité de probabilité (6)). Le paramètre d'intérêt sera constamment

$$\boldsymbol{\theta} := \left( (\text{vec } \boldsymbol{\beta}')', (\text{vec } \mathbf{A}_1)', \dots, (\text{vec } \mathbf{A}_p)', (\text{vec } \mathbf{B}_1)', \dots, (\text{vec } \mathbf{B}_q)' \right)' \in \mathbb{R}^K,$$

où  $\text{vec } \mathbf{C}$  désigne le vecteur obtenu en empilant les colonnes de  $\mathbf{C}$  les unes au-dessus des autres. Ce paramètre est donc de dimension  $K := km + k^2(p + q)$ . Les paramètres  $\boldsymbol{\Sigma}$  et  $f$  qui caractérisent la loi du bruit (voir (6)) sont généralement inconnus, et seront dès lors considérés comme des paramètres de nuisance. Dans un souci de cohérence avec la notation que nous avons utilisée dans le cadre du problème de position à un échantillon, nous désignerons par  $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$  l'hypothèse sous laquelle les observations  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$  ont été engendrées par (22) et (23), avec les valeurs  $\boldsymbol{\theta}$  du paramètre d'intérêt et  $(\boldsymbol{\Sigma}, f)$  des paramètres de nuisance.

Dans la suite, nous nous restreindrons aux deux classes de problèmes de test suivantes (la première étant un cas particulier de la seconde). Dans la première classe, il s'agira de tester l'hypothèse nulle sous laquelle  $\boldsymbol{\theta}$  prend une valeur fixée  $\boldsymbol{\theta}_0 \in \mathbb{R}^K$ . Le problème de test s'écrit donc

$$\begin{cases} \mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \\ \mathcal{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0. \end{cases} \quad (24)$$

Le problème de position à un échantillon qui a servi d'illustration jusqu'ici est bien entendu un problème de test de ce type. Dans la seconde classe de problèmes de test, la valeur du paramètre  $\boldsymbol{\theta}$  n'est pas spécifiée sous l'hypothèse nulle, mais satisfait seulement à un certain nombre de contraintes linéaires. De manière plus précise, pour une certaine matrice  $\mathbf{Y}$  de dimension  $(K \times r)$  et de rang maximal  $r$  ( $r \leq K$ ) et un  $K$ -vecteur  $\boldsymbol{\theta}_0$  fixés, le second type de problème s'écrit

$$\begin{cases} \mathcal{H}_0 : \boldsymbol{\theta} \in \boldsymbol{\theta}_0 + \mathcal{M}(\mathbf{Y}) \\ \mathcal{H}_1 : \boldsymbol{\theta} \notin \boldsymbol{\theta}_0 + \mathcal{M}(\mathbf{Y}), \end{cases} \quad (25)$$

où  $\mathcal{M}(\mathbf{Y})$  désigne l'espace vectoriel engendré par les colonnes de  $\mathbf{Y}$ . L'hypothèse nulle  $\mathcal{H}_0$  ci-dessus est souvent dite *linéaire*, car la contrainte  $\boldsymbol{\theta} \in \boldsymbol{\theta}_0 + \mathcal{M}(\mathbf{Y})$  signifie que le paramètre  $\boldsymbol{\theta}$  satisfait à une collection de  $K - r$  contraintes linéaires (ou plutôt affines). La plupart des problèmes de test d'intérêt pratique sont de ce type (problèmes d'analyse de la variance, test de l'ordre d'un VARMA, test de Durbin-Watson multivarié, ...).

## 4.2 Normalité locale asymptotique dans un modèle elliptique.

Comme l'ont montré Garel et Hallin (1995), le modèle linéaire général multivarié à erreurs VARMA décrit ci-dessus jouit de la propriété de *normalité locale asymptotique* (LAN), relativement au paramètre d'intérêt  $\boldsymbol{\theta}$ , les autres paramètres (associés au bruit blanc  $\boldsymbol{\varepsilon}_t$ ,  $t \in \mathbb{Z}$ ) étant fixés. Ceci signifie que, pour toute suite bornée  $(\boldsymbol{\tau}^{(n)})$  dans  $\mathbb{R}^K$  (déterminant une certaine perturbation locale du paramètre  $\boldsymbol{\theta}$ ), le logarithme du rapport de vraisemblance associé à l'hypothèse  $\mathcal{H}^{(n)}(\boldsymbol{\theta} + n^{-1/2}\boldsymbol{\tau}^{(n)}, \boldsymbol{\Sigma}, f)$  par rapport à l'hypothèse  $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$ , admet sous  $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$ , la décomposition asymptotique

$$L_{\boldsymbol{\theta} + n^{-1/2}\boldsymbol{\tau}^{(n)}/\boldsymbol{\theta}; \boldsymbol{\Sigma}, f}^{(n)}(\mathbf{Y}^{(n)}) = (\boldsymbol{\tau}^{(n)})' \boldsymbol{\Delta}_{\boldsymbol{\Sigma}, f}^{(n)}(\boldsymbol{\theta}) - \frac{1}{2} (\boldsymbol{\tau}^{(n)})' \boldsymbol{\Gamma}_{\boldsymbol{\Sigma}, f}(\boldsymbol{\theta}) \boldsymbol{\tau}^{(n)} + o_{\mathbb{P}}(1), \quad (26)$$

où la *suite centrale*  $\boldsymbol{\Delta}_{\boldsymbol{\Sigma}, f}^{(n)}(\boldsymbol{\theta})$ , toujours sous  $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$ , est asymptotiquement normale  $K$ -variée de moyenne nulle et de matrice de variance-covariance  $\boldsymbol{\Gamma}_{\boldsymbol{\Sigma}, f}(\boldsymbol{\theta})$ .

Pour interpréter ce résultat, considérons le modèle de position gaussien  $K$ -varié

$$\{ \mathcal{N}(\boldsymbol{\Gamma}_{\boldsymbol{\Sigma}, f}(\boldsymbol{\theta})\boldsymbol{\tau}, \boldsymbol{\Gamma}_{\boldsymbol{\Sigma}, f}(\boldsymbol{\theta})) \mid \boldsymbol{\tau} \in \mathbb{R}^K \}$$

caractérisant le comportement d'une observation unique, que nous noterons  $\boldsymbol{\Delta}$ . Il est facile de vérifier que le logarithme du rapport de vraisemblance associé à la loi gaussienne  $\mathcal{N}(\boldsymbol{\Gamma}_{\boldsymbol{\Sigma}, f}(\boldsymbol{\theta})\boldsymbol{\tau}, \boldsymbol{\Gamma}_{\boldsymbol{\Sigma}, f}(\boldsymbol{\theta}))$  par rapport à  $\mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\Sigma}, f}(\boldsymbol{\theta}))$  est précisément

$$\boldsymbol{\tau}' \boldsymbol{\Delta} - \frac{1}{2} \boldsymbol{\tau}' \boldsymbol{\Gamma}_{\boldsymbol{\Sigma}, f}(\boldsymbol{\theta}) \boldsymbol{\tau},$$

ce qui signifie (voir le second membre de (26)) que le logarithme du rapport de vraisemblance local en  $\boldsymbol{\theta}$  en (26) est asymptotiquement équivalent au logarithme du rapport de vraisemblance d'un modèle de position gaussien classique. Comme nous l'expliquons de façon plus formelle ci-dessous, ceci a d'importantes implications sur la construction de procédures *localement et asymptotiquement optimales* dans la suite d'*expériences locales* en question (ces termes sont définis plus loin).

## 4.3 Convergence des expériences statistiques et optimalité.

La normalité locale asymptotique entraîne, pour tout  $\boldsymbol{\theta}$ , la *convergence faible* de la suite d'*expériences locales* (localisées en  $\boldsymbol{\theta}$ )

$$\mathcal{E}^{(n)} := \left\{ \mathbb{R}^{nk}, \mathcal{B}^{nk}, \left\{ \mathbb{P}_{\boldsymbol{\theta} + n^{-1/2}\boldsymbol{\tau}, \boldsymbol{\Sigma}, f}^{(n)} \mid \boldsymbol{\tau} \in \mathbb{R}^K \right\} \right\}$$

vers le modèle de position gaussien  $K$ -varié

$$\mathcal{E} := \left\{ \mathbb{R}^K, \mathcal{B}^K, \left\{ \mathcal{N}(\boldsymbol{\Gamma}_{\boldsymbol{\Sigma}, f}(\boldsymbol{\theta})\boldsymbol{\tau}, \boldsymbol{\Gamma}_{\boldsymbol{\Sigma}, f}(\boldsymbol{\theta})) \mid \boldsymbol{\tau} \in \mathbb{R}^K \right\} \right\};$$

nous noterons  $\boldsymbol{\Delta} := \boldsymbol{\Delta}_{\boldsymbol{\Sigma}, f}$  l'unique observation associée à ce modèle-limite. Ce concept de convergence est basé sur une pseudo-distance, dite *distance de Le Cam*, entre les ensembles de fonctions de risque (de  $\mathbb{R}^K$  dans  $\mathbb{R}^+$ ) réalisables sous les expériences considérées, pour des fonctions de perte bornées. Dans ce contexte de tests d'hypothèses, ceci signifie en quelque sorte que, lorsque  $n \rightarrow \infty$ , toutes les courbes de puissance réalisables pour l'expérience  $\mathcal{E}^{(n)}$  convergent—point par point en  $\boldsymbol{\tau}$ , mais uniformément en l'ensemble de toutes les procédures de test possibles—vers les courbes de puissance associées au modèle-limite gaussien  $\mathcal{E}$ . A l'inverse, pour toute fonction de risque  $R$  réalisable dans le modèle-limite  $\mathcal{E}$ , il existe une suite de fonctions de risque associées à  $\mathcal{E}^{(n)}$  qui converge vers cette fonction de risque  $R$ . Il découle de ces considérations le principe général de construction de tests localement et asymptotiquement optimaux suivant :

Si un test  $\phi(\Delta)$  jouit d'une propriété d'optimalité (exacte) quelconque dans le modèle-limite gaussien  $\mathcal{E}$ , alors la suite de tests  $\phi(\Delta_{\Sigma,f}^{(n)}(\theta))$  hérite, localement et asymptotiquement, de la même propriété d'optimalité dans la suite locale d'expériences  $\mathcal{E}^{(n)}$  (pour une description plus précise des concepts d'optimalité locale et asymptotique considéré ici, on se reportera à Le Cam (1986, Section 11.9), Choi et al. (1996), ou Genot-Catalot et Picard (1993)).

Il suffit donc de connaître les tests optimaux dans le modèle-limite (la notion d'optimalité reste à préciser dans chaque cas) pour construire des tests localement et asymptotiquement optimaux dans le modèle d'origine. Nous décrivons brièvement ci-dessous quelques-uns de ces tests optimaux.

Dans le modèle-limite, l'hypothèse nulle du problème de test (24) s'écrit  $\mu := \Gamma_{\Sigma,f}(\theta)\tau = \mathbf{0} \in \mathbb{R}^K$ , ou de manière équivalente  $\tau = \mathbf{0}$ . Considérons alors le problème de test

$$\begin{cases} \mathcal{H}_0 : \tau = \mathbf{0} \\ \mathcal{H}_1^c : \tau' \Gamma_{\Sigma,f}(\theta) \tau > c, \quad c > 0, \end{cases} \quad (27)$$

où  $\tau' \Gamma_{\Sigma,f}(\theta) \tau = \mu' (\Gamma_{\Sigma,f}(\theta))^{-1} \mu$  (sous  $\mathcal{H}_1^c$ ,  $\mu$  est en-dehors de l'ellipsoïde de forme  $\Gamma_{\Sigma,f}(\theta)$  et de "rayon"  $c$ ). On peut montrer dans ce contexte que le test  $\phi^*$  rejetant  $\mathcal{H}_0$  dès que

$$\Delta' (\Gamma_{\Sigma,f}(\theta))^{-1} \Delta > \chi_{K,1-\alpha}^2 \quad (28)$$

est *maximin* pour le problème (27), dans la classe des tests de niveau  $\alpha$  et indépendamment de la valeur de  $c$ . Ceci signifie que  $\phi^*$  est de niveau  $\alpha$ , et que sa puissance satisfait à

$$\mathbb{E}_P[\phi^*] \geq \sup_{\phi: \mathbb{E}_{\mathcal{H}_0}[\phi] \leq \alpha} \inf_{P \in \mathcal{H}_1^c} \mathbb{E}_P[\phi], \quad \forall P \in \mathcal{H}_1^c.$$

Il est à noter que la statistique de test en (28) ne dépend pas de  $c$ , malgré le rôle joué par  $c$  dans la définition de la contre-hypothèse considérée.

Suivant le même schéma pour le problème de test (25), l'hypothèse nulle de (25) s'écrit, dans le modèle-limite  $\mathcal{E}$ ,  $\mathcal{H}_0 : \mu := \Gamma_{\Sigma,f}(\theta)\tau \in \mathcal{M}(\Gamma_{\Sigma,f}(\theta)\Upsilon)$ . Le test  $\phi^*$  rejetant  $\mathcal{H}_0$  dès que

$$\Delta' \left[ (\Gamma_{\Sigma,f}(\theta))^{-1} - \Upsilon(\Upsilon' \Gamma_{\Sigma,f}(\theta) \Upsilon)^{-1} \Upsilon' \right] \Delta > \chi_{K-r,1-\alpha}^2, \quad (29)$$

où  $r$  est le rang de  $\Upsilon$ , est le test *le plus stringent (most stringent)* pour ce problème de test, dans la classe des tests de niveau  $\alpha$ . Désignant par  $\mathcal{C}_\alpha$  la collection des tests de niveau  $\alpha$  pour le problème considéré, ceci signifie que  $\phi^* \in \mathcal{C}_\alpha$ , et que

$$\sup_{P \in \mathcal{H}_1} r_{\phi^*}(P) \leq \sup_{P \in \mathcal{H}_1} r_\phi(P), \quad \forall \phi \in \mathcal{C}_\alpha,$$

où le *regret*  $r_{\phi_0}(P)$  d'un test  $\phi_0$  en  $P \in \mathcal{H}_1$ , défini par

$$r_{\phi_0}(P) := \left[ \sup_{\phi \in \mathcal{C}_\alpha} \mathbb{E}_P[\phi] \right] - \mathbb{E}_P[\phi_0], \quad P \in \mathcal{H}_1,$$

est le déficit de puissance de  $\phi_0$  en  $P$  par rapport à la puissance la plus élevée qui peut être atteinte en  $P$  par les tests de la classe  $\mathcal{C}_\alpha$ .

Les statistiques en (28) et (29) requièrent bien entendu que la matrice de variance-covariance  $\Gamma_{\Sigma,f}(\theta)$  associée à  $\Delta$  dans le modèle-limite  $\mathcal{E}$  soit inversible. Ceci n'est pas toujours le cas (cf. le test des ordres d'un modèle VARMA); il faut alors substituer, aux statistiques en (28) et (29), des statistiques plus complexes faisant intervenir des inverses généralisés de  $\Gamma_{\Sigma,f}(\theta)$  (voir Hallin et Paindaveine (2003b) pour une discussion détaillée de ce point).

#### 4.4 Procédures localement et asymptotiquement optimales.

Il découle des deux sous-sections précédentes que la construction de procédures localement et asymptotiquement optimales pour les problèmes (24) et (25) peut être effectuée en remplaçant l'observation gaussienne  $\Delta$ , dans (28) et (29), par la suite centrale  $\Delta_{\Sigma, f}^{(n)}(\theta)$  associée à la décomposition LAN (26). Autrement dit, des procédures de test localement et asymptotiquement optimales peuvent être fondées sur les statistiques

$$(\Delta_{\Sigma, f}^{(n)}(\theta))' (\Gamma_{\Sigma, f}(\theta))^{-1} \Delta_{\Sigma, f}^{(n)}(\theta), \quad (30)$$

et

$$(\Delta_{\Sigma, f}^{(n)}(\theta))' \left[ (\Gamma_{\Sigma, f}(\theta))^{-1} - \Upsilon (\Upsilon' \Gamma_{\Sigma, f}(\theta) \Upsilon)^{-1} \Upsilon' \right] \Delta_{\Sigma, f}^{(n)}(\theta), \quad (31)$$

pour les problèmes (24) et (25), respectivement. Ainsi, en montrant que le modèle linéaire général multivarié à erreurs VARMA est localement et asymptotiquement normal, Garel et Hallin (1995) ont assuré l'existence de procédures localement et asymptotiquement optimales pour les problèmes considérés.

Les procédures décrites ci-dessus requièrent évidemment la connaissance de  $\Sigma$  et  $f$ , les paramètres du bruit blanc elliptique  $\varepsilon_t$ ,  $t \in \mathbb{Z}$ . Ces procédures sont donc hautement paramétriques; en particulier, elles ne seront valides que si les valeurs de  $\Sigma$  et  $f$  utilisées pour calculer la valeur de la statistique de test sont précisément les valeurs des paramètres du bruit blanc sous-jacent. Nous avons déjà indiqué que ces paramètres  $\Sigma$  et  $f$  sont généralement inconnus, et seraient donc considérés comme des paramètres de nuisance. Le principe d'invariance permet d'éliminer cette nuisance; celui-ci nous conduit à ne considérer que les procédures invariantes, qui, comme on l'a vu, sont de facto non paramétriques (valables sous une large classe de distributions elliptiques du bruit blanc). Il est remarquable que l'application du principe d'invariance soit possible, ici, sans aucun sacrifice sur le plan de l'efficacité.

### 5 Procédures optimales fondées sur les rangs multivariés.

Dans cette section, nous expliquons comment réconcilier nos objectifs d'invariance et nos objectifs d'optimalité. La stratégie que nous emploierons consiste à baser nos procédures sur des statistiques  $T_{\Sigma, f; \text{inv}}^{(n)}(\theta)$  qui sont

- (i) asymptotiquement équivalentes, sous  $\mathcal{H}^{(n)}(\theta, \Sigma, f)$ , à la statistique paramétrique  $T_{\Sigma, f}^{(n)}(\theta)$  localement et asymptotiquement optimale sous  $\mathcal{H}^{(n)}(\theta, \Sigma, f)$  (c'est-à-dire (30) ou (31), suivant le problème considéré), et
- (ii) fondées sur les signes et les rangs décrits dans les sections 3.3 ou 3.4, qui sont (asymptotiquement) invariants par rapport à des groupes de transformations engendrant la classe des lois elliptiques.

L'équivalence asymptotique en (i) implique que les procédures fondées sur  $T_{\Sigma, f; \text{inv}}^{(n)}(\theta)$  et  $T_{\Sigma, f}^{(n)}(\theta)$  auront les mêmes propriétés asymptotiques sous  $\mathcal{H}^{(n)}(\theta, \Sigma, f)$ ; en particulier, la procédure fondée sur  $T_{\Sigma, f; \text{inv}}^{(n)}(\theta)$  sera localement et asymptotiquement optimale sous  $\mathcal{H}^{(n)}(\theta, \Sigma, f)$ .

En outre, la statistique de test  $T_{\Sigma, f; \text{inv}}^{(n)}(\theta)$ , qui hérite de l'invariance (asymptotique) des signes et des rangs sur lesquels elle est fondée, sera (asymptotiquement) libre sous la classe des lois elliptiques. Par conséquent, il sera possible de baser sur  $T_{\Sigma, f; \text{inv}}^{(n)}(\theta)$  des tests qui ont le niveau nominal (asymptotique)  $\alpha$  sous la classe complète des lois elliptiques, c'est-à-dire sous  $\bigcup_{\Sigma} \bigcup_g \mathcal{H}^{(n)}(\theta, \Sigma, g)$  (contrairement à la procédure fondée sur  $T_{\Sigma, f}^{(n)}(\theta)$  qui n'est en général valide que sous  $\mathcal{H}^{(n)}(\theta, \Sigma, f)$ ). Il convient d'insister sur le fait que la densité  $f$ , dans ce qui précède, n'est pas nécessairement la "vraie" densité, mais plutôt la densité en laquelle, de façon arbitraire, les propriétés d'optimalité seront atteintes.

Nous illustrons ci-dessous la construction de ces procédures optimales de rangs signés multivariés sur quelques exemples concrets. Nous commençons par notre exemple générique : le problème de position à un échantillon.

## 5.1 Le problème de position à un échantillon.

Considérons de nouveau le modèle de position à un échantillon

$$\mathbf{Y}_i = \boldsymbol{\theta} + \boldsymbol{\varepsilon}_i, \quad i = 1, \dots, n,$$

où  $\boldsymbol{\varepsilon}_i$ ,  $i = 1, \dots, n$ , est un bruit blanc elliptique de matrice de forme  $\boldsymbol{\Sigma}$  et de densité radiale  $f$ . Nous notons  $\mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$  l'hypothèse associée. Nous désirons tester l'hypothèse  $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ , pour un certain  $k$ -vecteur  $\boldsymbol{\theta}_0$  fixé, sans spécifier ni la valeur de la matrice de forme  $\boldsymbol{\Sigma}$ , ni la densité radiale  $f$ . Sans perte de généralité, nous pouvons nous restreindre au cas pour lequel  $\boldsymbol{\theta}_0 := \mathbf{0}$ . Le problème de test considéré est donc le problème (8).

La normalité locale asymptotique, qui découle de Garel et Hallin (1995) sous l'hypothèse de différentiabilité en moyenne quadratique de la racine carrée  $\underline{f}^{1/2}$  de la densité de probabilité du bruit blanc, garantit l'existence de procédures paramétriques (à  $\boldsymbol{\Sigma}$  et  $f$  fixés) localement et asymptotiquement optimales (maximin, dans ce cas). Il est facile de vérifier que la procédure paramétrique localement et asymptotiquement optimale associée aux valeurs  $\boldsymbol{\Sigma}$  et  $f$  des paramètres de nuisance, consiste, au niveau asymptotique  $\alpha$ , à rejeter l'hypothèse nulle lorsque

$$T_{\boldsymbol{\Sigma}, f}^{(n)} := \boldsymbol{\Delta}_{\boldsymbol{\Sigma}, f}^{(n)'} \boldsymbol{\Gamma}_{\boldsymbol{\Sigma}, f}^{-1} \boldsymbol{\Delta}_{\boldsymbol{\Sigma}, f}^{(n)} = \frac{k}{n \mathcal{I}_{k, f}} \sum_{i, j=1}^n \varphi_f(d_i(\boldsymbol{\Sigma})) \varphi_f(d_j(\boldsymbol{\Sigma})) \mathbf{U}_i'(\boldsymbol{\Sigma}) \mathbf{U}_j(\boldsymbol{\Sigma}) > \chi_{k, 1-\alpha}^2, \quad (32)$$

où  $d_i(\boldsymbol{\Sigma}) := \|\boldsymbol{\Sigma}^{-1/2} \mathbf{Y}_i\|$ ,  $\mathbf{U}_i(\boldsymbol{\Sigma}) := \boldsymbol{\Sigma}^{-1/2} \mathbf{Y}_i / d_i(\boldsymbol{\Sigma})$ ,  $\varphi_f := -2(f^{1/2})' / f^{1/2}$  ( $(f^{1/2})'$  désigne la dérivée faible de  $f^{1/2}$ ), et où

$$\mathcal{I}_{k, f} := \left( \int_0^\infty f(r) r^{k-1} dr \right)^{-1} \left( \int_0^\infty \varphi_f^2(r) f(r) r^{k-1} dr \right)$$

est l'information de Fisher radiale associée à la densité  $\underline{f}$  (l'information de Fisher, sous les valeurs  $\boldsymbol{\Sigma}$  et  $f$  des paramètres de nuisance, est en effet  $(1/k) \mathcal{I}_{k, f} \boldsymbol{\Sigma}^{-1}$ ). La différentiabilité en moyenne quadratique de  $\underline{f}^{1/2}$  garantit que  $\mathcal{I}_{k, f} < \infty$  (voir Hallin et Paindaveine 2002a pour plus de détails).

La statistique de test invariante proposée, qui est fondée sur les signes et les rangs de la Section 3.4, et qui reconstruit la statistique paramétrique optimale (32) sous des densités radiales correctement spécifiées, est

$$T_{f; \text{inv}}^{(n)} := \frac{k}{n \mathcal{I}_{k, f}} \sum_{i, j=1}^n J_{k, f} \left( \frac{\hat{R}_i}{n+1} \right) J_{k, f} \left( \frac{\hat{R}_j}{n+1} \right) \cos(\pi p_{ij}), \quad (33)$$

où  $J_{k, f} := \varphi_f \circ \tilde{F}_k^{-1}$  ( $\tilde{F}_k$  désigne la fonction de répartition de  $d_i(\boldsymbol{\Sigma})$  sous  $\mathcal{H}^{(n)}(\mathbf{0}, \boldsymbol{\Sigma}, f)$ ). On peut montrer que la statistique non paramétrique (33) est asymptotiquement équivalente, sous  $\mathcal{H}^{(n)}(\mathbf{0}, \boldsymbol{\Sigma}, f)$ , à la statistique paramétrique (32). Il en découle que le test  $\phi_f = \phi_f^{(n)}$  qui rejette l'hypothèse dès que  $T_{f; \text{inv}}^{(n)} > \chi_{k, 1-\alpha}^2$  hérite, sous cette suite d'hypothèses, des propriétés d'optimalité de la procédure paramétrique. Le test  $\phi_f$  est donc localement et asymptotiquement optimal pour peu que la densité radiale de la loi elliptique sous-jacente soit  $f$ , et fait donc aussi bien que son équivalent paramétrique dans ce cas. Par ailleurs,  $\phi_f$  étant fondé sur une statistique asymptotiquement invariante, son risque de première espèce (asymptotiquement,  $\alpha$ ) reste le même sous l'hypothèse nulle  $\bigcup_{\boldsymbol{\Sigma}} \bigcup_g \mathcal{H}^{(n)}(\mathbf{0}, \boldsymbol{\Sigma}, g)$  tout entière. La procédure de rangs signés reste donc valide quelle que soit la loi elliptique sous-jacente (contrairement à la procédure paramétrique qui n'est en général valable que sous  $\mathcal{H}^{(n)}(\mathbf{0}, \boldsymbol{\Sigma}, f)$ ). Par conséquent,  $\phi_f$  est localement et asymptotiquement maximin, au niveau asymptotique  $\alpha$ , pour  $\bigcup_{\boldsymbol{\Sigma}} \bigcup_g \mathcal{H}^{(n)}(\mathbf{0}, \boldsymbol{\Sigma}, g)$  par rapport à des contre-hypothèses de la forme  $\bigcup_{\boldsymbol{\theta} \neq \mathbf{0}} \bigcup_{\boldsymbol{\Sigma}} \mathcal{H}^{(n)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}, f)$ .

Ceci montre que, quelle que soit la densité radiale choisie *a priori* ( $f$  est quelconque dans ce qui précède), cette méthodologie permet de construire des procédures non paramétriques qui sont localement et asymptotiquement optimales si la densité choisie  $f$  (qui détermine le choix des scores  $J_{k, f}$  dans la statistique de test) et la densité radiale de la loi elliptique sous-jacente coïncident. C'est donc toute une classe de procédures de test qui est proposée ici.

Bien entendu, toutes ces considérations n'apportent aucun renseignement sur l'efficacité de  $\phi_f$  sous les lois elliptiques dont la densité radiale est  $g$  ( $g \neq f$ ). C'est pour étudier ce point crucial que nous avons calculé les efficacités asymptotiques relatives de  $\phi_f$  par rapport au test de Hotelling, sous des densités radiales  $g$  quelconques (mais suffisamment régulières pour que le test de Hotelling soit valide). Il est à noter que le test de Hotelling constitue avec la procédure paramétrique gaussienne optimale; on peut en effet facilement vérifier que, pour tout  $\Sigma$  fixé, la statistique de Hotelling est asymptotiquement équivalente, sous  $\bigcup_g \mathcal{H}^{(n)}(\mathbf{0}, \Sigma, g)$ , à la version de la statistique (32) qui est associée à la densité radiale gaussienne  $f(r) := (2\pi)^{-1/2} \exp(-r^2/2)$ .

La Table 2 présente des valeurs numériques des efficacités asymptotiques relatives de quelques tests  $\phi_f$  par rapport au test de Hotelling, sous une classe de lois  $t$   $k$ -variées. Les scores utilisés sont les scores associés à diverses lois  $t$ , ainsi que les scores gaussiens. Cette étude indique que les procédures proposées ont un comportement uniformément satisfaisant (en termes d'efficacité asymptotique). En particulier, ces procédures—et ceci est un phénomène classique dans l'inférence fondée sur les rangs—sont particulièrement performantes sous les lois dont les queues (radiales) sont lourdes.

		Degrés de liberté de la distribution $t$ sous-jacente							
$k$	$f$	3	4	6	8	10	15	20	$\infty$
1	$t_3$	2.000	1.417	1.124	1.025	0.975	0.916	0.890	0.820
	$t_6$	1.926	1.414	1.167	1.087	1.049	1.005	0.985	0.936
	$t_{15}$	1.786	1.345	1.143	1.083	1.055	1.026	1.014	0.987
	$vdW$	1.639	1.257	1.093	1.048	1.030	1.013	1.007	1.000
2	$t_3$	2.143	1.488	1.157	1.045	0.987	0.920	0.889	0.807
	$t_6$	2.067	1.484	1.200	1.107	1.062	1.009	0.986	0.927
	$t_{15}$	1.910	1.407	1.173	1.102	1.068	1.032	1.018	0.984
	$vdW$	1.729	1.301	1.112	1.059	1.037	1.016	1.009	1.000
3	$t_3$	2.250	1.544	1.186	1.063	1.000	0.926	0.892	0.799
	$t_6$	2.174	1.540	1.227	1.124	1.073	1.014	0.988	0.919
	$t_{15}$	2.006	1.458	1.198	1.118	1.080	1.038	1.021	0.981
	$vdW$	1.798	1.336	1.128	1.069	1.043	1.019	1.011	1.000
4	$t_3$	2.333	1.589	1.210	1.079	1.012	0.932	0.896	0.794
	$t_6$	2.258	1.584	1.250	1.139	1.083	1.019	0.990	0.913
	$t_{15}$	2.084	1.499	1.220	1.132	1.090	1.044	1.025	0.979
	$vdW$	1.853	1.364	1.142	1.077	1.049	1.022	1.012	1.000
6	$t_3$	2.455	1.657	1.248	1.106	1.033	0.945	0.904	0.788
	$t_6$	2.382	1.652	1.286	1.163	1.101	1.028	0.995	0.905
	$t_{15}$	2.202	1.564	1.254	1.155	1.107	1.054	1.031	0.975
	$vdW$	1.935	1.408	1.164	1.092	1.059	1.027	1.016	1.000
10	$t_3$	2.600	1.741	1.299	1.145	1.065	0.968	0.922	0.785
	$t_6$	2.534	1.736	1.333	1.196	1.126	1.043	1.005	0.896
	$t_{15}$	2.355	1.649	1.302	1.188	1.132	1.068	1.040	0.969
	$vdW$	2.041	1.467	1.195	1.112	1.074	1.035	1.021	1.000

Table 2: Efficacités asymptotiques relatives de certains tests  $\phi_f$  par rapport au test de Hotelling, sous des lois de Student  $k$ -variées (à 3, 4, 6, 8, 10, 15, et 20 degrés de liberté) et sous la loi normale  $k$ -variée, pour différentes valeurs de  $k$ .

Ce bon comportement uniforme est confirmé par les deux résultats suivants. Le premier concerne la version gaussienne de nos tests, c'est-à-dire le test  $\phi_{vdW} := \phi_f$  (où  $f(r) := (2\pi)^{-1/2} \exp(-r^2/2)$ ), basé sur la statistique *de type van der Waerden*

$$\frac{1}{n} \sum_{i,j=1}^n \sqrt{\Psi_k^{-1}\left(\frac{\hat{R}_i}{n+1}\right)} \sqrt{\Psi_k^{-1}\left(\frac{\hat{R}_j}{n+1}\right)} \cos(\pi p_{ij}),$$

où  $\Psi_k$  désigne la fonction de répartition d’une variable chi-deux à  $k$  degrés de liberté. Le test  $\phi_{vdW}$  généralise la procédure de rangs signés univariée (4). Nous avons étendu (Hallin et Paindaveine 2002a) le résultat de Chernoff et Savage (1958), en prouvant que, *quelle que soit la dimension  $k$  des observations*, l’efficacité asymptotique relative de  $\phi_{vdW}$  par rapport au test de Hotelling est uniformément supérieure ou égale à 1, et que l’égalité ne se produit que sous des densités gaussiennes. Ce résultat est bien entendu le meilleur possible, puisque le test de Hotelling est optimal sous des densités gaussiennes, de sorte que le meilleur comportement relatif possible de  $\phi_{vdW}$  dans le cas normal est l’ex-aequo. Par contre, ceci montre que, quelle que soit la loi à symétrie elliptique non gaussienne—de variance finie (afin que le test de Hotelling soit valide)—sous-jacente, la procédure de van der Waerden proposée (qui, elle, ne requiert pas que ces variances soient finies) est strictement meilleure que le test de Hotelling (en termes d’efficacité asymptotique).

Le second résultat permettant d’appréhender la qualité uniforme de l’efficacité asymptotique des procédures proposées, concerne la procédure *de type Wilcoxon*  $\phi_W$ , associée à des fonctions scores *linéaires*. Le test  $\phi_W$ , qui rejette l’hypothèse nulle dès que

$$\frac{3k}{n(n+1)^2} \sum_{i,j=1}^n \hat{R}_i \hat{R}_j \cos(\pi p_{ij}) > \chi_{k,1-\alpha}^2,$$

généralise la procédure de Wilcoxon univariée (3). Ce test est une version *entièrement basée sur les hyperplans passant par les observations* du test de Wilcoxon multivarié proposé par Peters et Randles (1990), lequel est fondé sur les rangs de Mahalanobis  $R_i(\hat{\Sigma})$  et les interdirections. En utilisant les méthodes du calcul des variations, nous avons calculé (Hallin et Paindaveine 2002a) la borne inférieure (en la loi elliptique sous-jacente) des efficacités asymptotiques relatives de  $\phi_W$  par rapport au test de Hotelling. Cette borne est donnée, en fonction de la dimension  $k$  des observations, par

$$\frac{81}{500} \frac{(\sqrt{2k-1}+1)^5}{k^2(\sqrt{2k-1}+5)}. \quad (34)$$

Quelques valeurs numériques de cette borne sont reprises dans la Table 3 ci-dessous. La borne inférieure (34) est monotone décroissante pour  $k \geq 2$ , et tend vers  $81/125 = .648$ , lorsque la dimension des observations tend vers l’infini. On remarquera que les bornes inférieures associées aux cas bidimensionnel et tridimensionnel sont strictement supérieures au célèbre “.864”, obtenu dans le cas univarié par Hodges et Lehmann (1956). Nous renvoyons à Hallin et Paindaveine (2002a) pour une forme explicite des densités elliptiques en lesquelles ces bornes inférieures sont atteintes.

$k$	$\inf_g \text{ARE}_{k,g}(\phi_W/T^2)$
1	0.864
2	0.916
3	0.883
4	0.853
6	0.811
10	0.765
$+\infty$	0.648

Table 3: Quelques valeurs numériques de la borne inférieure (34) de l’efficacité asymptotique relative de la procédure de rangs signés généralisés de type Wilcoxon  $\phi_W$  par rapport au test de Hotelling  $T^2$  pour quelques valeurs de la dimension  $k$  des observations.

Cette remarquable efficacité asymptotique est loin d’être le seul avantage des procédures de rangs signés généralisés sur le test de Hotelling. Ces procédures sont en effet également valides sous une classe plus vaste de lois elliptiques (qui comprend, par exemple, les lois de Cauchy, qui n’ont pas de moyenne) et sont naturellement plus robustes, puisqu’elles sont fondées sur des rangs et sur ces fonctions constantes par morceaux que sont les interdirections et lift-interdirections.

## 5.2 Un problème sériel simple.

Dans cette section, nous quittons les problèmes de position pour considérer un problème sériel simple : le problème du test de l'hypothèse de bruit blanc indépendant. Plus précisément, nous désirons tester si le  $n$ -uple de  $k$ -vecteurs  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$  observés est la réalisation d'un processus indépendant et identiquement distribué (à symétrie elliptique), contre le fait que cet échantillon est la réalisation d'un processus VARMA.

Le modèle considéré est bien entendu un cas particulier purement sériel ( $\mathbf{X}^{(n)} = \mathbf{0}$ ) du modèle linéaire général multivarié à erreurs VARMA décrit dans la Section 4.1. Le paramètre d'intérêt est ici

$$\boldsymbol{\theta} := \left( (\text{vec } \mathbf{A}_1)', \dots, (\text{vec } \mathbf{A}_p)', (\text{vec } \mathbf{B}_1)', \dots, (\text{vec } \mathbf{B}_q)' \right)' \in \mathbb{R}^{k^2(p+q)},$$

et le problème de test de l'hypothèse de bruit blanc est le problème de test (24) avec  $\boldsymbol{\theta}_0 := \mathbf{0}$ , c'est-à-dire

$$\begin{cases} \mathcal{H}_0 : \boldsymbol{\theta} = \mathbf{0} \\ \mathcal{H}_1 : \boldsymbol{\theta} \neq \mathbf{0}. \end{cases}$$

Les mêmes conditions sur la densité radiale  $f$  que celles qui sont requises dans le cas du problème de position à un échantillon assurent la différentiabilité en moyenne quadratique de la racine carrée  $\underline{f}^{1/2}$  de la densité  $k$ -variée du bruit blanc, laquelle permet de prouver (Garel et Hallin 1995) que le modèle considéré est localement et asymptotiquement normal en  $\boldsymbol{\theta} = \mathbf{0}$  (à  $\boldsymbol{\Sigma}$  et  $f$  fixés).

La méthodologie utilisée plus haut peut donc être reprise dans ce cadre sériel. Des statistiques d'un type nouveau doivent cependant être introduites, car les statistiques de test (32) et (33) ne permettent pas de détecter le type de contre-hypothèses sérielles considérées ici (puissance locale nulle). Plus précisément, la procédure paramétrique localement asymptotiquement optimale (maximin, de nouveau) associée aux valeurs  $\boldsymbol{\Sigma}$  et  $f$  des paramètres du bruit blanc (donnée par (28)) consiste ici, au niveau asymptotique  $\alpha$ , à rejeter l'hypothèse nulle dès que

$$\begin{aligned} T_{\boldsymbol{\Sigma}, f}^{(n)} &:= \boldsymbol{\Delta}_{\boldsymbol{\Sigma}, f}^{(n)'} \boldsymbol{\Gamma}_{\boldsymbol{\Sigma}, f}^{-1} \boldsymbol{\Delta}_{\boldsymbol{\Sigma}, f}^{(n)} \\ &= \frac{k^2 \mu_{k-1;f}}{\mu_{k+1;f} \mathcal{I}_{k,f}} \sum_{i=1}^{\pi} (n-i)^{-1} \sum_{s,t=i+1}^n \varphi_f(d_s(\boldsymbol{\Sigma})) \varphi_f(d_t(\boldsymbol{\Sigma})) \\ &\quad d_{s-i}(\boldsymbol{\Sigma}) d_{t-i}(\boldsymbol{\Sigma}) \mathbf{U}'_s(\boldsymbol{\Sigma}) \mathbf{U}_t(\boldsymbol{\Sigma}) \mathbf{U}'_{s-i}(\boldsymbol{\Sigma}) \mathbf{U}_{t-i}(\boldsymbol{\Sigma}), \end{aligned} \quad (35)$$

où  $\mu_{l;f} := \int_0^\infty r^l f(r) dr$  et  $\pi := \max(p, q)$ , dépasse le quantile  $\chi_{k^2\pi, 1-\alpha}^2$  (la suite centrale  $\boldsymbol{\Delta}_{\boldsymbol{\Sigma}, f}^{(n)}$  intervenant dans la propriété LAN de ce problème est de dimension  $k^2\pi$ ). Ce nouveau type de statistique, forme quadratique en une mesure de dépendance sérielle, permet de détecter une éventuelle dépendance au sein des observations  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$  (on peut montrer que la version gaussienne de (35) est en fait une forme quadratique en les matrices de covariances croisées usuelles des  $\mathbf{Y}_i$ ).

Pour les mêmes raisons que dans le cas du problème de position à un échantillon, nous introduisons les procédures de rangs signés basées sur les statistiques

$$\begin{aligned} T_{K; \text{inv}}^{(n)} &:= \frac{k^2}{\mathbb{E}[K_1^2(U)] \mathbb{E}[K_2^2(U)]} \sum_{i=1}^{\pi} (n-i)^{-1} \sum_{s,t=i+1}^n K_1\left(\frac{\hat{R}_s}{n+1}\right) K_1\left(\frac{\hat{R}_t}{n+1}\right) \\ &\quad K_2\left(\frac{\hat{R}_{s-i}}{n+1}\right) K_2\left(\frac{\hat{R}_{t-i}}{n+1}\right) \cos(\pi p_{s,t}) \cos(\pi p_{s-i,t-i}), \end{aligned} \quad (36)$$

où  $U$  est uniformément distribuée sur  $]0, 1[$ ,  $K_1, K_2$  sont deux fonctions scores de carré sommable,  $\hat{R}_t$  désigne le rang de la lift-interdirection associée à l'observation  $\mathbf{Y}_t$ , et  $\cos(\pi p_{s,t})$  est l'estimateur de  $\mathbf{U}'_s(\boldsymbol{\Sigma}) \mathbf{U}_t(\boldsymbol{\Sigma})$  fondé sur les interdirections de Randles. La statistique de test (36) est strictement affine-invariante et est asymptotiquement invariante par rapport au groupe des transformations radiales monotones continues. Le test  $\phi_K^{(n)}$  qui consiste à rejeter l'hypothèse nulle dès que  $T_{K; \text{inv}}^{(n)} > \chi_{k^2\pi, 1-\alpha}^2$  est

asymptotiquement de niveau  $\alpha$ , et est valide sous une large classe de lois elliptiques (sans hypothèse de moments sur ces distributions sous-jacentes).

De plus, sous  $\mathcal{H}^{(n)}(\mathbf{0}, \Sigma, f)$  (c'est-à-dire, dans ce contexte, sous l'hypothèse nulle de bruit blanc indépendant, à symétrie elliptique de matrice de forme  $\Sigma$  et de densité radiale  $f$ ), la statistique  $T_{f;\text{inv}}^{(n)} := T_{K;\text{inv}}^{(n)}$  associée aux fonctions scores  $K_1 = J_{k,f} := \varphi_f \circ \tilde{F}_k^{-1}$  et  $K_2 = \tilde{F}_k^{-1}$  est asymptotiquement équivalente à la statistique paramétrique localement et asymptotiquement optimale  $T_{\Sigma,f}^{(n)}$ . Il en découle que le test associé, que nous noterons  $\phi_f = \phi_f^{(n)}$ , est localement et asymptotiquement maximin, dès que la densité radiale sous-jacente coïncide avec la densité utilisée pour calculer les scores  $K_1, K_2$ .

Comme pour le problème de position à un échantillon, la question de l'efficacité asymptotique de  $\phi_f$  sous des densités radiales  $g (\neq f)$  se pose avec force, puisque  $\phi_f$  est valide sous  $\bigcup_{\Sigma} \bigcup_g \mathcal{H}^{(n)}(\mathbf{0}, \Sigma, g)$ . Nous y répondons de manière analogue, en calculant les puissances locales et les efficacités asymptotiques relatives de  $\phi_f$  par rapport à la procédure gaussienne paramétrique, sous une loi elliptique de densité radiale  $g$  quelconque. Cette procédure gaussienne paramétrique  $\phi_{\mathcal{N}} = \phi_{\mathcal{N}}^{(n)}$ , fondée sur les matrices de covariances croisées classiques, est asymptotiquement équivalente, sous  $\bigcup_g \mathcal{H}^{(n)}(\mathbf{0}, \Sigma, g)$  (pour tout  $\Sigma$  fixé), à la version gaussienne ( $f(r) := (2\pi)^{-1/2} \exp(-r^2/2)$ ) de la procédure paramétrique fondée sur (35) (la procédure  $\phi_{\mathcal{N}}$  est donc valide sous  $\bigcup_{\Sigma} \bigcup_g \mathcal{H}^{(n)}(\mathbf{0}, \Sigma, g)$ ). Nous renvoyons à Hallin et Paindaveine (2002b) pour une expression—et des valeurs numériques sous des distributions  $t$   $k$ -variées—de ces efficacités asymptotiques relatives. Ces valeurs numériques indiquent que ces procédures sérielles de rangs signés généralisés se comportent de façon uniformément séduisante, même si, comme pour le problème de position qui a été considéré, c'est en présence de queues radiales lourdes que ces procédures sont particulièrement efficaces.

Pour confirmer ce bon comportement uniforme des procédures proposées, nous établissons les *équivalents sériels* des extensions multivariées des résultats (i) de Chernoff et Savage (1958), et (ii) de Hodges et Lehmann (1956), présentées dans le cadre du problème de position à un échantillon.

(i) Considérons la procédure non paramétrique de type van der Waerden que nous proposons pour le problème considéré, c'est-à-dire le test  $\phi_{vdW} := \phi_f$ , où  $f(r) := (2\pi)^{-1/2} \exp(-r^2/2)$ . Puisqu'il est associé à une densité radiale gaussienne, le test  $\phi_{vdW}$  est asymptotiquement optimal (comme son équivalent paramétrique  $\phi_{\mathcal{N}}$ ) sous des distributions gaussiennes, de sorte que l'efficacité asymptotique relative de  $\phi_{vdW}$  par rapport à  $\phi_{\mathcal{N}}$  est de 1 dans ce cas. Nous prouvons cependant que, comme dans le cas du modèle de position, cette efficacité asymptotique relative est strictement supérieure à 1 sous les lois elliptiques non gaussiennes (de variance finie, afin que le test paramétrique gaussien  $\phi_{\mathcal{N}}$  soit valable). Ceci généralise le résultat univarié de Hallin (1994).

(ii) Considérons le test  $\phi_{SP} = \phi_{SP}^{(n)} := \phi_K^{(n)}$  associé aux scores linéaires  $K_1(u) = K_2(u) := u$  pour tout  $u$ . L'efficacité asymptotique de cette procédure *de type Spearman* est elle aussi uniformément bonne comme le montre le résultat suivant. Nous avons calculé, comme dans le cadre du modèle de position, la borne inférieure de l'efficacité asymptotique relative de  $\phi_{SP}$  par rapport à  $\phi_{\mathcal{N}}$ , en fonction de la loi elliptique sous-jacente. Cette borne inférieure est donnée, en fonction de la dimension  $k$  des observations, par

$$\left[ 9 (2c_k^2 + k - 1)^4 \right] / 2^{10} k^2 c_k^4, \quad (37)$$

où, en désignant par  $J_r$  la fonction de Bessel de premier type d'ordre  $r$ , nous avons posé

$$c_k := \inf \left\{ x > 0 \mid \left( \sqrt{x} J_{\sqrt{2k-1}/2}(x) \right)' = 0 \right\}.$$

Ceci généralise le résultat univarié de Hallin et Tribel (2000). La Table 4 reprend quelques valeurs numériques de cette borne inférieure, en fonction de  $k$ . Un phénomène de "saut" similaire à celui qui a été observé dans le cas du modèle de position à un échantillon, est également présent dans cette situation sérielle : la borne inférieure associée au cas bivarié est en effet supérieure à la borne (.856) correspondant à la dimension 1; la suite des bornes inférieures (37) est monotone décroissante à partir de  $k = 2$  et tend vers  $9/16 = .5625$ , quand  $k$  tend vers l'infini. Les densités radiales qui réalisent ces bornes inférieures font également intervenir ces fonctions de Bessel (voir Hallin et Paindaveine (2002b)).

$k$	$\inf_g \text{ARE}_{k,g}(\phi_{SP}/\phi_{\mathcal{N}})$
1	0.856
2	0.913
3	0.878
4	0.845
6	0.797
10	0.742
$+\infty$	0.563

Table 4: Quelques valeurs numériques de la borne inférieure (37) de l'efficacité asymptotique relative de la procédure de rangs signés généralisés de type Spearman  $\phi_{SP}$  par rapport au test paramétrique gaussien  $\phi_{\mathcal{N}}$ , pour quelques valeurs de la dimension  $k$  des observations.

### 5.3 Restrictions linéaires.

La plupart des problèmes intéressants d'un point de vue pratique sont des problèmes de test d'hypothèses linéaires, c'est-à-dire du type (25). Pour de tels problèmes, la valeur du paramètre d'intérêt  $\theta$  n'est pas complètement spécifiée sous l'hypothèse nulle (comme c'est le cas pour les problèmes traités dans les Sections 5.1 et 5.2), mais satisfait seulement à un certain nombre de restrictions linéaires (nous renvoyons à la Section 5.3.2 ci-dessous pour des exemples).

#### 5.3.1 Procédures de rangs signés alignés.

Hallin et Paindaveine (2003d) construisent des procédures de rangs signés multivariés, qui sont fondées sur l'extension multivariée décrite dans la Section 3.3 des concepts de signes et de rangs. La stratégie employée est similaire à celle qui a été utilisée dans les Sections 5.1 et 5.2 : les tests sont basés sur une statistique qui, à la fois,

- (i) est asymptotiquement équivalente à la statistique paramétrique optimale (31) sous les valeurs correspondantes des paramètres de nuisance, et
- (ii) est fondée sur les rangs signés multivariés basés sur l'estimateur de Tyler.

Les procédures de rangs signés qui en résultent sont dès lors localement asymptotiquement optimales sous des densités radiales correctement spécifiées et sont valides sous la classe des lois elliptiques (de variance finie, dans ce cas précis, comme on le verra ci-dessous).

Néanmoins, un certain nombre de difficultés techniques propres aux problèmes de test d'hypothèses linéaires vient compliquer le développement de ces procédures. Plutôt que d'entrer dans les détails analytiques de leur construction, nous nous contentons d'expliquer en quoi ce type de problème est plus délicat à traiter, puis décrivons quelques exemples de problèmes de test d'hypothèses linéaires (ainsi que les procédures de rangs signés multivariés associées) particulièrement importants pour les applications.

La première difficulté vient du fait que le problème de test (25), dans le cadre du modèle linéaire général multivarié à erreurs VARMA, n'est pas toujours invariant par transformations linéaires. Si le problème de test n'est pas invariant par transformations linéaires, le principe d'invariance (linéaire), qui préconise de se restreindre à des statistiques de test affine-invariantes, est bien entendu sans objet. C'est donc le recours même aux rangs multivariés affine-invariants qui est remis en cause dans de telles situations. C'est la raison pour laquelle Hallin et Paindaveine (2003d) se restreignent à la classe des problèmes de test (25) qui sont invariants par transformations linéaires. Même si cette classe (nous renvoyons à Hallin et Paindaveine (2003a) pour sa caractérisation) est relativement réduite, elle contient néanmoins la plupart des problèmes d'intérêt pratique. Les procédures de rangs signés proposées dans Hallin et Paindaveine (2003d) satisfont au principe d'invariance puisqu'elles sont invariantes par transformations affines (et asymptotiquement invariantes par transformations radiales monotones continues).

Si la méthodologie employée s'appuie sur celle qui a été développée dans les sections précédentes, une difficulté supplémentaire se présente toutefois : la valeur du paramètre  $\theta$  n'est pas fixée sous l'hypothèse

nulle. Ce problème est connu, dans le contexte des méthodes de rangs, sous le nom de “*problème d’alignement*”. Les procédures de rangs signés développées dans ce cas sont des fonctions de signes et de rangs calculés en utilisant un estimateur préliminaire  $\hat{\boldsymbol{\theta}}$  de  $\boldsymbol{\theta}$  (qui est contraint sous l’hypothèse nulle de (25), c’est-à-dire tel que  $\hat{\boldsymbol{\theta}} \in \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$ ). Le résultat-clé permettant de substituer ces *rangs signés alignés* aux “vrais” rangs signés (ceux basés sur la vraie valeur de  $\boldsymbol{\theta}$ ) est un résultat de linéarité asymptotique, qui est prouvé dans Hallin et Paindaveine (2003c). L’estimateur préliminaire doit bien entendu satisfaire à un certain nombre d’hypothèses d’équivariance (pour assurer l’invariance des procédures de rangs signés alignés) et de convergence (afin que les rangs signés alignés reconstruisent asymptotiquement les vrais rangs signés). C’est cette hypothèse de convergence qui limite le domaine de validité des procédures de rangs signés alignés à la classe des lois elliptiques *de variance finie*; typiquement, la  $\sqrt{n}$ -convergence des estimateurs classiques de  $\boldsymbol{\theta}$  requiert en effet des moments finis d’ordre deux (estimateur des moindres carrés pour la composante non sérielle de  $\boldsymbol{\theta}$ , estimateurs de type Yule-Walker pour la partie sérielle de  $\boldsymbol{\theta}$ , etc).

Comme nous l’avons mentionné ci-dessus, nos procédures de rangs signés sont, dans ce type de problèmes, fondées sur les rangs signés basés sur l’estimation de la matrice de forme  $\boldsymbol{\Sigma}$ . Dans la plupart des cas, la statistique paramétrique localement et asymptotiquement optimale (31) ne dépend pas des “signes paramétriques”  $\mathbf{U}_i(\boldsymbol{\Sigma})$  uniquement à travers les cosinus  $\mathbf{U}'_i(\boldsymbol{\Sigma})\mathbf{U}_j(\boldsymbol{\Sigma})$ , mais fait intervenir des fonctions plus compliquées des  $\mathbf{U}_i(\boldsymbol{\Sigma})$ . Dans une telle situation, les interdirections de Randles (les signes de la Section 3.4), qui permettent d’estimer ces cosinus—mais pas les  $\mathbf{U}_i(\boldsymbol{\Sigma})$  eux-mêmes, sont inutiles. Par contre, les signes multivariés  $\mathbf{U}_i(\hat{\boldsymbol{\Sigma}})$  fondés sur un estimateur de la matrice de forme permettent d’estimer une fonction quelconque des  $\mathbf{U}_i(\boldsymbol{\Sigma})$ . Dans Hallin et Paindaveine (2003d), nous proposons néanmoins un concept d’*interdirections absolues* qui permet de reconstruire individuellement les signes  $\mathbf{U}_i(\boldsymbol{\Sigma})$ .

Avant de considérer quelques exemples de procédures de rangs signés permettant de tester des restrictions linéaires sur  $\boldsymbol{\theta}$ , mentionnons que, dans les problèmes mixtes (à la fois de position et sériel), les efficacités asymptotiques relatives des procédures de rangs signés par rapport aux procédures gaussiennes paramétriques sont des combinaisons linéaires convexes de celles obtenues dans les cas du problème de position à un échantillon (voir Section 5.1) et du problème purement sériel considéré dans la Section 5.2.

### 5.3.2 Quelques exemples d’intérêt pratique.

La classe des problèmes de test (25) qui sont invariants par transformations linéaires comprend de nombreux problèmes d’un grand intérêt pratique. Nous donnons ci-dessous trois exemples de tels problèmes, et pour chacun d’eux, décrivons la procédure de rangs signés multivariés associée.

#### (a) Test de Durbin-Watson multivarié.

Le problème de Durbin-Watson multivarié est la version du problème (25) associée à  $\boldsymbol{\theta}_0 = \mathbf{0}$  et

$$\boldsymbol{\Upsilon} = \begin{pmatrix} \mathbf{I}_{km} \\ \mathbf{0}_{k^2(p+q) \times km} \end{pmatrix},$$

où  $\mathbf{I}_r$  désigne la matrice identité  $r \times r$ . Il permet de tester l’indépendance sérielle des termes d’erreur du modèle linéaire considéré contre des hypothèses sous lesquelles le processus des erreurs est un VARMA d’ordres au plus  $p$  et  $q$  (le paramètre de régression  $\boldsymbol{\beta}$  joue ici le rôle d’une nuisance).

La procédure de rangs signés multivariés associée est basée sur la statistique de test

$$T_{K;\text{inv}}^{(n)} = \frac{k^2}{\mathbb{E}[K_1^2(U)]\mathbb{E}[K_2^2(U)]} \sum_{i=1}^{\max(p,q)} (n-i)^{-1} \sum_{s,t=i+1}^n K_1\left(\frac{R_s(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T)}{n+1}\right) K_1\left(\frac{R_t(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T)}{n+1}\right) \times K_2\left(\frac{R_{s-i}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T)}{n+1}\right) K_2\left(\frac{R_{t-i}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T)}{n+1}\right) \mathbf{U}'_{s-i}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T) \mathbf{U}_{t-i}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T) \mathbf{U}'_s(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T) \mathbf{U}_t(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T), \quad (38)$$

où, si on désigne par  $d_t(\boldsymbol{\beta}, \boldsymbol{\Sigma}) := \|\boldsymbol{\Sigma}^{-1/2} \mathbf{Z}_t(\boldsymbol{\beta})\|$  la norme de Mahalanobis des résidus  $\mathbf{Z}_t(\boldsymbol{\beta}) := \mathbf{Y}_t - \boldsymbol{\beta}' \mathbf{x}_t$ , on a posé  $R_t(\boldsymbol{\beta}, \boldsymbol{\Sigma}) = \text{rank}(d_t(\boldsymbol{\beta}, \boldsymbol{\Sigma}); d_1(\boldsymbol{\beta}, \boldsymbol{\Sigma}), \dots, d_n(\boldsymbol{\beta}, \boldsymbol{\Sigma}))$  et  $\mathbf{U}_t(\boldsymbol{\beta}, \boldsymbol{\Sigma}) = \boldsymbol{\Sigma}^{-1/2} \mathbf{Z}_t(\boldsymbol{\beta}) / d_t(\boldsymbol{\beta}, \boldsymbol{\Sigma})$ . Il est à

noter que, dans la statistique de test (38), l'estimateur de Tyler  $\hat{\Sigma}_T$  est calculé dans l'échantillon des résidus estimés  $\mathbf{Z}_t(\hat{\boldsymbol{\beta}})$ ,  $t = 1, \dots, n$ . (si le modèle ne contient pas de tendance linéaire, la statistique de test (38) coïncide avec la version de la statistique de test (36) qui est fondée sur les rangs signés de la Section 3.3 pour le problème de test de bruit blanc indépendant).

Le test (au niveau asymptotique  $\alpha$ ) de Durbin-Watson généralisé que nous proposons consiste à rejeter l'hypothèse nulle dès que  $T_{K;\text{inv}}^{(n)}$  est supérieure au quantile d'ordre  $1 - \alpha$  d'une variable chi-deux à  $k^2 \max(p, q)$  degrés de liberté. Dans cet exemple, on pourrait également utiliser les rangs signés de la Section 3.4 afin d'obtenir des procédures fondées sur les hyperplans passant par les résidus estimés (puisque la statistique optimale ne dépend des directions des résidus qu'à travers les cosinus des angles compris entre les résidus sphéricisés).

### (b) Test de l'ordre d'un modèle VARMA.

Dans ce second exemple, nous considérons le problème qui consiste à tester un modèle VARMA( $p_0, q_0$ ) contre un modèle VARMA( $p, q$ ) d'ordre plus élevé, c'est-à-dire  $p \geq p_0$ ,  $q \geq q_0$  et  $\max(p - p_0, q - q_0) > 0$  (nous supposons, pour simplifier, que le modèle n'inclut pas de tendance linéaire). Ceci est à nouveau un cas particulier du problème de test (25), puisqu'il est obtenu en posant  $\boldsymbol{\theta}_0 = \mathbf{0}$  et

$$\boldsymbol{\Upsilon} = \begin{pmatrix} \mathbf{I}_{k^2 p_0} & \mathbf{0}_{k^2 p_0 \times k^2 q_0} \\ \mathbf{0}_{k^2(p-p_0) \times k^2 p_0} & \mathbf{0}_{k^2(p-p_0) \times k^2 q_0} \\ \mathbf{0}_{k^2 q_0 \times k^2 p_0} & \mathbf{I}_{k^2 q_0} \\ \mathbf{0}_{k^2(q-q_0) \times k^2 p_0} & \mathbf{0}_{k^2(q-q_0) \times k^2 q_0} \end{pmatrix}.$$

Le cas particulier  $p - p_0 = q - q_0 = 1$  joue un rôle important dans certaines procédures d'identification de modèles (voir, par exemple, Pötscher 1983, 1985, ou Garel et Hallin 1999 pour le cas univarié). Pour ne pas compliquer les notations, nous ne considérerons ci-dessous que le cas  $p - p_0 = 1$ ,  $q = q_0 = 0$ .

Les procédures de rangs signés multivariées proposées dans ce cas précis sont fondées sur les matrices de covariances croisées généralisées

$$\boldsymbol{\Gamma}_{i;K}^{(n)} := \hat{\Sigma}_T'^{-1/2} \left( \frac{1}{n-i} \sum_{t=i+1}^n K_1 \left( \frac{R_t(\hat{\boldsymbol{\theta}}, \hat{\Sigma}_T)}{n+1} \right) K_2 \left( \frac{R_{t-i}(\hat{\boldsymbol{\theta}}, \hat{\Sigma}_T)}{n+1} \right) \mathbf{U}_t(\hat{\boldsymbol{\theta}}, \hat{\Sigma}_T) \mathbf{U}'_{t-i}(\hat{\boldsymbol{\theta}}, \hat{\Sigma}_T) \right) \hat{\Sigma}_T^{(n)1/2}, \quad (39)$$

où les distances de Mahalanobis  $d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}) := \|\boldsymbol{\Sigma}^{-1/2} \mathbf{Z}_t(\boldsymbol{\theta})\|$ , leurs rangs  $R_t(\boldsymbol{\theta}, \boldsymbol{\Sigma})$  et les signes  $\mathbf{U}_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}) = \boldsymbol{\Sigma}^{-1/2} \mathbf{Z}_t(\boldsymbol{\theta}) / d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma})$  sont calculés dans la collection des résidus  $\mathbf{Z}_t(\boldsymbol{\theta}) := \mathbf{Y}_t - \sum_{i=1}^{p_0} \mathbf{A}_i \mathbf{Y}_{t-i}$ .

Désignons par  $\mathbf{G}_u$ ,  $u \in \mathbb{N}$ , les matrices de Green associées à l'opérateur autorégressif  $\mathbf{A}(L)$  (rappelons que celles-ci sont définies par les récursions linéaires  $\mathbf{G}_u - \sum_{i=1}^{\min(p_0, u)} \mathbf{A}_i \mathbf{G}_{u-i} = \delta_{u0} \mathbf{I}_k$ , où  $\delta_{u0} = 1$  si  $u = 0$ , et  $\delta_{u0} = 0$  si  $u \neq 0$ ). Avec ces notations, la procédure (de niveau asymptotique  $\alpha$ ) proposée pour tester l'hypothèse de dépendance AR( $p_0$ ) contre l'hypothèse de dépendance AR( $p_0 + 1$ ) consiste à rejeter l'hypothèse nulle dès que

$$T_{K;\text{inv}}^{(n)} = \left( \mathbf{S}_K^{(n)} \right)' \mathbf{J}_K^{(n)} \mathbf{S}_K^{(n)} > \chi_{k^2, 1-\alpha}^2, \quad (40)$$

où, si  $\mathbf{v}^{(n)}$  désigne le  $p_0$ -vecteur dont la  $i$ ème composante est  $\sum_{t=\max(i, 2)}^{n-1} (n-t)^{1/2} \mathbf{G}_{t-i}(\hat{\boldsymbol{\theta}}) (\text{vec } \boldsymbol{\Gamma}_{t;K}^{(n)}(\hat{\boldsymbol{\theta}}))$ ,  $\mathbf{S}_K^{(n)} := ((n-1)^{1/2} (\text{vec } \boldsymbol{\Gamma}_{1;K}^{(n)})', \mathbf{v}^{(n)})'$  et

$$\mathbf{J}_K^{(n)} := \frac{k^2}{\mathbb{E}[K_1^2(U)] \mathbb{E}[K_2^2(U)]} \left[ \left( \begin{pmatrix} \hat{\Sigma}_T \otimes \hat{\Sigma}_T^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{w}^{2(n)} \end{pmatrix}^{-1} - \begin{pmatrix} \mathbf{I}_{k^2} & \mathbf{0} \\ \mathbf{I}_{k^2 p_0} & \mathbf{0} \end{pmatrix} (\mathbf{W}^{2(n)})^{-1} \begin{pmatrix} \mathbf{I}_{k^2} & \mathbf{0} \\ \mathbf{I}_{k^2 p_0} & \mathbf{0} \end{pmatrix}' \right) \right].$$

Ci-dessus,  $\mathbf{w}^{2(n)}$  et  $\mathbf{W}^{2(n)}$  désignent les matrices  $k^2 p_0 \times k^2 p_0$  dont les blocs en position  $(i, j)$  ( $i, j = 1, \dots, p_0$ ) sont  $\sum_{t=\max(i, j, 2)}^{n-1} \mathbf{G}_{t-i}(\hat{\boldsymbol{\theta}}) \hat{\Sigma}_T \mathbf{G}'_{t-j}(\hat{\boldsymbol{\theta}}) \otimes \hat{\Sigma}_T^{-1}$  et  $\sum_{t=\max(i, j)}^{n-1} \mathbf{G}_{t-i}(\hat{\boldsymbol{\theta}}) \hat{\Sigma}_T \mathbf{G}'_{t-j}(\hat{\boldsymbol{\theta}}) \otimes \hat{\Sigma}_T^{-1}$ , respectivement.

La statistique de test (40) a la même structure algébrique que celle qui est obtenue dans le cas univarié (voir Hallin et Puri (1994), ou Garel et Hallin (1999)). La statistique permettant de tester l'ordre d'un MA (typiquement, MA( $q_0$ ) contre MA( $q_0 + 1$ )), qui fait intervenir des combinaisons linéaires compliquées des matrices de covariances croisées  $\mathbf{\Gamma}_{i;K}^{(n)}$ , est par contre plus complexe dans le cas multivarié.

**(c) Détection de différentes phases dans la tendance.**

Soit  $(t_1^{(n)}, \dots, t_{m-1}^{(n)})$ ,  $n \in \mathbb{N}$ , une suite de  $(m-1)$ -uples telle que  $t_0^{(n)} := 0 < t_1^{(n)} < \dots < t_{m-1}^{(n)} < t_m^{(n)} := n$  pour tout  $n$ . Considérons le modèle (22), avec la suite de constantes de régression

$$\mathbf{x}_t^{(n)} = \mathbf{e}_i^{(m)}, \quad \text{pour } t_{i-1}^{(n)} < t \leq t_i^{(n)},$$

où  $\mathbf{e}_i^{(m)}$  désigne le  $i$ ème vecteur de la base canonique de  $\mathbb{R}^m$ . Le modèle associé caractérise un processus VARMA( $p, q$ ), avec une tendance qui dépend du temps (plus précisément, un processus de moyenne  $\beta_i$  dans la période  $\mathcal{C}_i = \mathcal{C}_i^{(n)} := \{t_{i-1}^{(n)} + 1, \dots, t_i^{(n)}\}$ ). Dans ce contexte, l'hypothèse nulle du problème de test (25) associé au choix  $\boldsymbol{\theta}_0 = \mathbf{0}$  et

$$\mathbf{\Upsilon} = \begin{pmatrix} \mathbf{I}_k & \mathbf{0}_{k \times k^2(p+q)} \\ \mathbf{I}_k & \mathbf{0}_{k \times k^2(p+q)} \\ \vdots & \vdots \\ \mathbf{I}_k & \mathbf{0}_{k \times k^2(p+q)} \\ \mathbf{0}_{k^2(p+q) \times k} & \mathbf{I}_{k^2(p+q)} \end{pmatrix}$$

correspond à l'hypothèse d'absence de différents régimes, c'est-à-dire à l'hypothèse  $\mathcal{H}_0 : \beta_1 = \dots = \beta_m$ . Les coefficients des opérateurs AR et MA du bruit sont ici des paramètres de nuisance. En l'absence de composante sérielle dans le modèle (c'est-à-dire quand  $p = q = 0$ ), le problème considéré se réduit au problème de position à  $m$  échantillons, qui est le problème classique de l'analyse de la variance multivariée à un facteur.

Nous décrivons les statistiques de rangs signés multivariés proposées dans le cas où les erreurs  $\mathbf{U}_t$  du modèle linéaire considéré sont engendrées par le processus VAR(1) d'équation  $\mathbf{U}_t = \mathbf{A}\mathbf{U}_{t-1} + \boldsymbol{\varepsilon}_t$ . Posons  $\boldsymbol{\lambda}^{(n)} := ((\lambda_1^{(n)})^{-1/2}, \dots, (\lambda_m^{(n)})^{-1/2})'$ , où  $\lambda_j^{(n)} := n_j/n := (t_j^{(n)} - t_{j-1}^{(n)})/n$ . La statistique de test proposée est alors

$$T_{K;\text{inv}}^{(n)} = \left( \mathbf{S}_K^{(n)} \right)' \mathbf{J}_K^{(n)} \mathbf{S}_K^{(n)},$$

avec

$$\mathbf{S}_K^{(n)} := \left[ \mathbf{I}_{km} + \sqrt{\frac{n}{n-1}} (\mathbf{I}_m \otimes \hat{\mathbf{A}}') L^{-1} \right] \begin{pmatrix} \frac{1}{\sqrt{n_1}} \sum_{t \in \mathcal{C}_1} K \left( \frac{R_1(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}}_T)}{n+1} \right) \hat{\boldsymbol{\Sigma}}_T^{-1/2'} \mathbf{U}_1(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}}_T) \\ \frac{1}{\sqrt{n_2}} \sum_{t \in \mathcal{C}_2} K \left( \frac{R_2(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}}_T)}{n+1} \right) \hat{\boldsymbol{\Sigma}}_T^{-1/2'} \mathbf{U}_2(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}}_T) \\ \vdots \\ \frac{1}{\sqrt{n_m}} \sum_{t \in \mathcal{C}_m} K \left( \frac{R_m(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}}_T)}{n+1} \right) \hat{\boldsymbol{\Sigma}}_T^{-1/2'} \mathbf{U}_m(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}}_T) \end{pmatrix},$$

et

$$\mathbf{J}_K^{(n)} := \frac{k}{\mathbb{E}[K^2(U)]} \left[ (\mathbf{I}_m - \boldsymbol{\lambda}^{(n)} \boldsymbol{\lambda}^{(n)'}) \otimes [\hat{\boldsymbol{\Sigma}}_T^{-1} + \hat{\mathbf{A}}' \hat{\boldsymbol{\Sigma}}_T^{-1} \hat{\mathbf{A}}]^{-1} \right],$$

où  $K$  est une fonction score de carré sommable et les distances de Mahalanobis  $d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}) := \|\boldsymbol{\Sigma}^{-1/2} \mathbf{Z}_t(\boldsymbol{\theta})\|$ , leurs rangs  $R_t(\boldsymbol{\theta}, \boldsymbol{\Sigma})$  et les signes  $\mathbf{U}_t(\boldsymbol{\theta}, \boldsymbol{\Sigma}) := \boldsymbol{\Sigma}^{-1/2} \mathbf{Z}_t(\boldsymbol{\theta}) / d_t(\boldsymbol{\theta}, \boldsymbol{\Sigma})$  sont calculés dans la collection des

résidus  $\mathbf{Z}_t(\boldsymbol{\theta}) := (\mathbf{Y}_t - \boldsymbol{\beta}' \mathbf{x}_t^{(n)}) - \mathbf{A}(\mathbf{Y}_{t-1} - \boldsymbol{\beta}' \mathbf{x}_{t-1}^{(n)})$ . Le test (de niveau asymptotique  $\alpha$ ) associé rejette l'hypothèse nulle dès que  $T_{K;\text{inv}}^{(n)} > \chi_{k(m-1), 1-\alpha}^2$ .

Dans le cas du problème de position à  $m$  échantillons (absence de composante sérielle), la statistique de test prend la forme (posons simplement  $\mathbf{A} = \mathbf{0}$ )

$$T_{K;\text{inv}}^{(n)} = \frac{k}{\mathbb{E}[K^2(U)]} \left[ \sum_{j=1}^m \frac{1}{n_j} \sum_{i, \bar{i} \in \mathcal{C}_j} K\left(\frac{R_i(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T)}{n+1}\right) K\left(\frac{R_{\bar{i}}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T)}{n+1}\right) \mathbf{U}'_i(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T) \mathbf{U}_{\bar{i}}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T) \right. \\ \left. - \sum_{j, \bar{j}=1}^m \frac{n}{\sqrt{n_j} \sqrt{n_{\bar{j}}}} \sum_{i, \bar{i} \in \mathcal{C}_j} K\left(\frac{R_i(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T)}{n+1}\right) K\left(\frac{R_{\bar{i}}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T)}{n+1}\right) \mathbf{U}'_i(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T) \mathbf{U}_{\bar{i}}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_T) \right],$$

laquelle est une version, basée sur les rangs signés de la Section 3.3, de la statistique de test de Randles et Um (1998) (qui est fondée sur les rangs de Mahalanobis (Section 3.3) et les signes de la Section 3.4, c'est-à-dire les interdirections de Randles). Bien entendu, dans ce cas précis, il est possible de construire une version de la statistique de test qui soit entièrement fondée sur les hyperplans passant par les résidus estimés  $\mathbf{Z}_t(\hat{\boldsymbol{\beta}}) := \mathbf{Y}_t - \hat{\boldsymbol{\beta}}' \mathbf{x}_t^{(n)}$  (comme c'était le cas pour le test de Durbin-Watson décrit plus haut), puisque les  $\mathbf{U}_i(\boldsymbol{\Sigma})$  n'interviennent dans la statistique de test optimale qu'à travers les cosinus  $\mathbf{U}'_i(\boldsymbol{\Sigma}) \mathbf{U}_j(\boldsymbol{\Sigma})$ .

## 6 Une simulation.

Dans cette dernière section, nous présentons les résultats d'une expérience de Monte-Carlo visant à étudier le comportement des tests de rangs signés pour des tailles d'échantillon finies. Nous considérons le test de Durbin-Watson décrit dans la Section 5.3.2.

Nous avons engendré  $N = 2500$  échantillons indépendants  $(\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_{400})$  de taille  $n = 400$  en utilisant diverses lois bivariées centrées réduites (la loi normale et les lois  $t$  à 3, 8 et 15 degrés de liberté). Les variables de Student  $\boldsymbol{\varepsilon}_t$  à  $\nu$  degrés de liberté ont été engendrées en utilisant le fait que  $\boldsymbol{\varepsilon}_t =_d \mathbf{Z}_t / \sqrt{Y_t/\nu}$  ( $=_d$  désigne l'égalité en loi), où le vecteur  $\mathbf{Z}_t$  de loi normale bivariée centrée réduite et la variable  $Y_t$  de loi chi-deux à  $\nu$  degrés de liberté sont indépendants. Pour chaque échantillon, nous avons calculé les séries d'"observations"  $\mathbf{Y}_1, \dots, \mathbf{Y}_{400}$  associées aux quatre ( $m = 0, 1, 2, 3$ ) modèles linéaires

$$\mathbf{Y}_t = \boldsymbol{\beta}_1 I_{[1 \leq t \leq 200]} + \boldsymbol{\beta}_2 I_{[201 \leq t \leq 400]} + \mathbf{U}_t,$$

$$\mathbf{U}_t - (m\mathbf{A}) \mathbf{U}_{t-1} = \boldsymbol{\varepsilon}_t, \quad \mathbf{U}_0 = \mathbf{0}, \quad (41)$$

où

$$\boldsymbol{\beta}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\beta}_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{et} \quad \mathbf{A} = \begin{pmatrix} 0.05 & 0.02 \\ -0.01 & 0.04 \end{pmatrix}.$$

Pour chaque série, nous avons effectué les 7 tests suivants (au niveau  $\alpha = 5\%$ ) :

- le test paramétrique gaussien  $\phi_{\mathcal{N}}$ , qui est la version gaussienne de la statistique paramétrique (31) (c'est-à-dire celle qui s'obtient en prenant  $f(r) := (2\pi)^{-1/2} \exp(-r^2/2)$  et en remplaçant  $\boldsymbol{\Sigma}$  par la matrice de variance-covariance empirique des résidus  $\mathbf{Z}_t(\hat{\boldsymbol{\beta}})$ ,
- le test de rangs signés de type van der Waerden  $\phi_{vDW}$  (fondé sur la version de (38) dont les fonctions scores sont  $K_1(u) = K_2(u) = (\Psi_2^{-1}(u))^{1/2}$ , où  $\Psi_2$  désigne la fonction de répartition d'une variable chi-deux à deux degrés de liberté),
- les tests de rangs signés  $\phi_{f_5}, \phi_{f_8}, \phi_{f_{15}}$ , qui sont les tests fondés sur les versions de (38) associées aux fonctions scores  $K_1(u) = \varphi_{f_\nu} \circ \tilde{F}_{\nu 2}^{-1}(u)$ ,  $K_2 = \tilde{F}_{\nu 2}^{-1}(u)$  ( $\nu = 5, 8, 15$ ), où  $f_\nu$  désigne la densité radiale de la loi de Student bivariée à  $\nu$  degrés de liberté,
- le test de signes  $\phi_S$  et le test de rangs signés de type Spearman  $\phi_{SP}$ , qui sont fondés sur les versions de (38) de fonctions scores  $K_1(u) = K_2(u) = 1$  et  $K_1(u) = K_2(u) = u$ , respectivement.

Pour chaque test, l'estimateur de  $\beta' = (\beta_1, \beta_2)$  qui a été utilisé est l'estimateur des moindres carrés ordinaires

$$\hat{\beta}' = \left( \frac{1}{100} \sum_{t=1}^{100} \mathbf{Y}_t, \frac{1}{100} \sum_{t=101}^{200} \mathbf{Y}_t \right),$$

et l'estimateur de Tyler  $\hat{\Sigma}_T$  a été calculé en utilisant l'algorithme itératif décrit dans Randles (2000). Les itérations ont été stoppées dès que la norme de Frobenius de la différence entre les deux membres de (13) tombe en dessous de  $10^{-6}$ .

Les fréquences de rejet sont données dans la Table 5. Les écart-type correspondants sont (pour  $N = 2500$  répétitions) de .0044, .0080 et .0100, pour les fréquences d'ordre  $p = .05$  ( $p = .95$ ),  $p = .20$  ( $p = .80$ ) et  $p = .50$ , respectivement. La Table 5 contient aussi les efficacités asymptotiques relatives des différents tests par rapport au test paramétrique gaussien  $\phi_{\mathcal{N}}$  sous les diverses lois considérées.

test	loi	Matrice d'autorégression $m\mathbf{A}$					ARE	loi	Matrice d'autorégression $m\mathbf{A}$					ARE
		$\mathbf{0}$	$\mathbf{A}$	$2\mathbf{A}$	$3\mathbf{A}$				$\mathbf{0}$	$\mathbf{A}$	$2\mathbf{A}$	$3\mathbf{A}$		
$\phi_{\mathcal{N}}$	$\mathcal{N}$	.0556	.1420	.5036	.8888	1.000	$t_8$	.0572	.1408	.4956	.8912	1.000		
$\phi_{v dW}$		.0536	.1408	.5064	.8944	1.000		.0596	.1340	.5148	.9144	1.047		
$\phi_{f_{15}}$		.0532	.1416	.4912	.8900	0.981		.0548	.1412	.5404	.9272	1.098		
$\phi_{f_8}$		.0544	.1356	.4708	.8744	0.942		.0552	.1436	.5376	.9244	1.111		
$\phi_{f_5}$		.0496	.1148	.4252	.8340	0.868		.0468	.1392	.5128	.9152	1.090		
$\phi_S$		.0532	.0976	.3216	.6832	0.617		.0532	.1096	.3684	.7616	0.711		
$\phi_{SP}$		.0564	.1456	.4960	.8904	0.970		.0556	.1380	.5112	.9072	1.022		
$\phi_{\mathcal{N}}$	$t_{15}$	.0556	.1388	.5056	.8832	1.000	$t_3$	.0468	.1528	.4840	.8492	1.000		
$\phi_{v dW}$		.0584	.1380	.5120	.8968	1.013		.0528	.1872	.6716	.9620	1.400		
$\phi_{f_{15}}$		.0536	.1432	.5196	.9020	1.032		.0516	.2124	.7320	.9800	1.614		
$\phi_{f_8}$		.0512	.1404	.5128	.8952	1.020		.0456	.2192	.7612	.9876	1.774		
$\phi_{f_5}$		.0440	.1264	.4760	.8740	0.973		.0424	.2256	.7832	.9908	1.953		
$\phi_S$		.0536	.0992	.3460	.7212	0.662		.0532	.1500	.5312	.8916	1.000		
$\phi_{SP}$		.0588	.1456	.5100	.8956	0.990		.0536	.1792	.6492	.9584	1.305		

Table 5: Fréquences de rejet, sous différentes valeurs de la matrice d'autorégression  $m\mathbf{A}$  (cf. (41)) et sous différentes lois  $t$  et la loi normale bivariée, du test paramétrique gaussien  $\phi_{\mathcal{N}}$ , des tests de rangs signés  $\phi_{f_5}$ ,  $\phi_{f_8}$ ,  $\phi_{f_{15}}$ ,  $\phi_{v dW}$ ,  $\phi_{SP}$ , ainsi que du test de signe  $\phi_S$ ; ces simulations sont basées sur 2500 échantillons de taille 400.

Tous les tests satisfont apparemment à la contrainte de niveau de 5% (la demi-longueur d'un intervalle de confiance à 95% est d'environ .01). Le classement des puissances estimées des diverses procédures est essentiellement compatible avec les efficacités asymptotiques relatives. Ainsi, sous des densités gaussiennes, les puissances estimées des tests  $\phi_{f_\nu}$  augmentent avec  $\nu$ , comme attendu. De même, l'optimalité asymptotique du test  $\phi_{f_\nu}$  sous la loi de Student à  $\nu$  degrés de liberté est confirmée.

## References

- [1] Brockwell, P. J., et R. A. Davis (1987). *Time Series: Theory and Methods*, Springer, New York.
- [2] Chernoff, H., et I. R. Savage (1958). Asymptotic normality and efficiency of certain nonparametric tests, *Ann. Math. Statist.* **29**, 972-994.
- [3] Choi, S., W. J. Hall, et A. Schick (1996). Asymptotically uniformly most powerful tests in parametric and semiparametric models, *Ann. Statist.* **24**, 841-861.
- [4] Gareil, B., et M. Hallin (1995). Local asymptotic normality of multivariate ARMA processes with a linear trend, *Ann. Inst. Statist. Math.* **47**, 551-579.
- [5] Genot-Catalot, V., et D. Picard (1993). *Éléments de Statistique asymptotique*. Springer-Verlag, Paris.

- [6] Hallin, M. (1986). Non-stationary  $q$ -dependent processes and time-varying moving-average models : invertibility properties and the forecasting problem, *Adv. Appl. Prob.* **18**, 170-210.
- [7] Hallin, M. (1994). On the Pitman-nonadmissibility of correlogram-based methods, *Jour. Time Series Anal.* **15**, 607-612.
- [8] Hallin, M., J.-Fr. Ingenbleek, et M. L. Puri (1985). Linear serial rank tests for randomness against ARMA alternatives, *Ann. Statist.* **13**, 1156-1181.
- [9] Hallin, M., J.-Fr. Ingenbleek, et M. L. Puri (1989). Asymptotically most powerful rank tests for multivariate randomness against serial dependence, *J. Multivariate Anal.* **30**, 34-71.
- [10] Hallin, M., et G. Mélarde (1988). Rank-based tests for randomness against first-order serial dependence, *J. Amer. Statist. Assoc.* **83**, 1117-1128.
- [11] Hallin, M., et D. Paindaveine (2002a). Optimal tests for multivariate location based on interdirections and pseudo-Mahalanobis ranks, *Ann. Statist.* **30**, 1103-1133.
- [12] Hallin, M., et D. Paindaveine (2002b). Optimal procedures based on interdirections and pseudo-Mahalanobis ranks for testing multivariate elliptic white noise against ARMA dependence, *Bernoulli* **8**, 787-816.
- [13] Hallin, M., et D. Paindaveine (2002c). Multivariate signed ranks : Randles' interdirections or Tyler's angles? In Y. Dodge, Ed., *Statistical Data Analysis Based on the  $L_1$  Norm and Related Procedures*, Birkhäuser, Basel, 271-282.
- [14] Hallin, M., et D. Paindaveine (2003a). Affine invariant linear hypotheses for the multivariate general linear model with VARMA error terms. In Chr. Léger and M. Moore, Eds, *Mathematical Statistics and Applications, Festschrift for Constance van Eeden*, I.M.S. Lecture Notes-Monograph Series, I.M.S., Hayward, California, in press.
- [15] Hallin, M., et D. Paindaveine (2003b). Rank-based optimal tests of the adequacy of an elliptic VARMA model, submitted.
- [16] Hallin, M., et D. Paindaveine (2003c). Asymptotic linearity of serial and nonserial multivariate signed rank statistics, submitted.
- [17] Hallin, M., et D. Paindaveine (2003d). Affine invariant aligned rank tests for the multivariate general linear model with ARMA errors, submitted.
- [18] Hallin, M., et D. Paindaveine (2003e). Rank-based methods for univariate and multivariate time series models. *Statistical Science*, to appear.
- [19] Hallin, M., et M.L. Puri (1988). Optimal rank-based procedures for time-series analysis : testing an ARMA model against other ARMA models, *Ann. Statist.* **16**, 402-432.
- [20] Hallin, M., et M.L. Puri (1991). Time-series analysis via rank-order theory : signed-rank tests for ARMA models. *J. Multivariate Anal.* **39**, 1-29.
- [21] Hallin, M., et M.L. Puri (1994). Aligned rank tests for linear models with autocorrelated error terms. *J. Multivariate Anal.* **50**, 175-237.
- [22] Hallin, M., et M.L. Puri (1994). A multivariate Wald-Wolfowitz rank test against serial dependence. *Canadian J. Statist.* **23**, 55-65.
- [23] Hallin, M., et O. Tribel (2000). The efficiency of some nonparametric competitors to correlogram-based methods. In F.T. Bruss and L. Le Cam, Eds, *Game Theory, Optimal Stopping, Probability, and Statistics, Papers in honor of T.S. Ferguson on the occasion of his 70th birthday*, I.M.S. Lecture Notes-Monograph Series, I.M.S., Hayward, California, 249-262.
- [24] Hallin, M., et B. J. M. Werker (1999). Optimal testing for semi-parametric AR models : from Gaussian Lagrange multipliers to autoregression rank scores and adaptive tests. In S. Ghosh, Ed., *Asymptotics, Nonparametrics, and Time Series*, M. Dekker, New York, 295-350.
- [25] Hallin, M., et B. J. M. Werker (2003). Semiparametric efficiency, distribution-freeness, and invariance, *Bernoulli* **9**, 55-65.

- [26] Hannan, E. J. (1970). *Multiple Time series*. J. Wiley, New York.
- [27] Hettmansperger, T. P., J. Nyblom, et H. Oja (1994). Affine invariant multivariate one-sample sign tests, *J. Roy. Statist. Soc. Ser. B* **56**, 221-234.
- [28] Hettmansperger, T. P., J. Möttönen, et H. Oja (1997). Affine invariant multivariate one-sample signed-rank tests, *J. Amer. Statist. Assoc.* **92**, 1591-1600.
- [29] Hodges, J. L., et E. L. Lehmann (1956). The efficiency of some nonparametric competitors of the  $t$ -test, *Ann. Math. Statist.* **27**, 324-335.
- [30] Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer-Verlag, New York.
- [31] Möttönen, J., et H. Oja (1995). Multivariate spatial sign and rank methods, *J. Nonparam. Statist.* **5**, 201-213.
- [32] Möttönen, J., H. Oja, et J. Tienari (1997). On the efficiency of multivariate spatial sign and rank methods, *Ann. Statist.* **25**, 542-552.
- [33] Möttönen, J., T. P. Hettmansperger, H. Oja, et J. Tienari (1998). On the efficiency of the multivariate affine invariant rank methods, *J. Multivariate Anal.* **66**, 118-132.
- [34] Oja, H. (1999). Affine invariant multivariate sign and rank tests and corresponding estimates : a review, *Scand. J. Statist.* **26**, 319-343.
- [35] Peters, D., et R. H. Randles (1990). A multivariate signed-rank test for the one-sample location problem, *J. Amer. Statist. Assoc.* **85**, 552-557.
- [36] Puri, M. L., et P. K. Sen (1971). *Nonparametric Methods in Multivariate Analysis*. J. Wiley, New York.
- [37] Randles, R. H. (1989). A distribution-free multivariate sign test based on interdirections, *J. Amer. Statist. Assoc.* **84**, 1045-1050.
- [38] Randles, R. H. (2000). A simpler, affine-invariant, multivariate, distribution-free sign test, *J. Amer. Statist. Assoc.* **95**, 1263-1268.
- [39] Randles, R. H. et Y. Um (1998). Nonparametric tests for the multivariate multi-sample location problem, *Statistica Sinica* **8**, 801-812.
- [40] Tyler, D. E. (1987). A distribution-free M-estimator of multivariate scatter, *Ann. Statist.* **15**, 234-251.

Davy PAINDAVEINE  
 Département de Mathématique,  
 I.S.R.O., et E.C.A.R.E.S.  
 Université Libre de Bruxelles  
 Campus de la Plaine CP 210  
 B-1050 Bruxelles BELGIUM  
 dpaindav@ulb.ac.be