# AFFINE INVARIANT LINEAR HYPOTHESES FOR THE MULTIVARIATE GENERAL LINEAR MODEL WITH VARMA ERROR TERMS

MARC HALLIN AND DAVY PAINDAVEINE
*Université Libre de Bruxelles*

Affine invariance is often considered a *natural* requirement when testing hypotheses in a multivariate context. This invariance issue is considered here in the problem of testing linear constraints on the parameters of a multivariate linear model with VARMA error terms. We give a characterization of the collection of null hypotheses that are invariant under the group of affine transformations, hence compatible with a requirement of affine invariant testing. We comment the results and discuss some examples.

## 1. Introduction

Affine invariance/equivariance often is considered a natural requirement in multivariate statistical inference. The rationale for such a requirement is that the data at hand, or the noise underlying the model, should be treated as intrinsically multivariate objects, irrespective of any particular choice of a coordinate system. This requirement plays a fundamental role in most recent developments in the area of robust multivariate analysis, where the concepts of spatial quantiles, spatial signs, spatial ranks, location or regression depth and contours, ..., all refer to either rotational or affine invariance/equivariance (see for instance Oja (1999) for a recent review). In such a context, reasonable testing procedures should be invariant—as soon, of course, as the null hypothesis itself is invariant.

Robust multivariate inference so far has been developed essentially for independent observations (location and regression models, MANOVA, principal components, ...). However, testing methods based on multivariate signs and ranks (more precisely, *interdirections* and the so-called *pseudo-Mahalanobis ranks*) recently have been extended (Hallin and Paindaveine, 2002a–c) to time-series problems. More specifically, these papers are treating the problem of testing linear hypotheses in the multivariate general linear model with VARMA error terms (equivalently, a VARMA model with linear trend) described below. As test statistics based on interdirections and pseudo-Mahalanobis ranks are automatically invariant under linear transformations, a preliminary question naturally arises: are affine invariance properties in this setting still meaningful? And, in case they are, which are the invariant null hypotheses?

This question, which is of a purely algebraic nature, is addressed here in full generality, and the class of invariant linear hypotheses, hence the class of testing problems that qualify for being treated by means of interdirections and pseudo-Mahalanobis rank test statistics, is characterized. It appears (Propositions 2.1 and 2.2) that the answer relies on the commuting properties of the matrices characterizing the hypothesis to be tested.

The model under study is the multivariate linear model

$$\mathbf{Y}^{(n)} = \mathbf{X}^{(n)}\boldsymbol{\beta} + \mathbf{U}^{(n)}, \tag{1.1}$$

where

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

is an $n$-tuple of $k$-variate observations $\mathbf{Y}_t$, $t = 1, \dots, n$,

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

is an $n \times m$ matrix of constant regressors (the design matrix), and

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

is the $m \times k$ regression parameter. We do not make the traditional assumption that the error term

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

is white noise, but rather assume $\mathbf{U}_t$, $t = 1, \dots, n$, to be a finite realization (of length $n$) of a solution of the multivariate linear stochastic difference equation (a VARMA$(p, q)$ model)

$$\mathbf{A}(L)\mathbf{U}_t = \mathbf{B}(L)\boldsymbol{\varepsilon}_t, \quad t \in \mathbb{Z}, \tag{1.2}$$

where

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

for some $k \times k$ real matrices $\mathbf{A}_1, \dots, \mathbf{A}_p, \mathbf{B}_1, \dots, \mathbf{B}_q$, $\{\boldsymbol{\varepsilon}_t \mid t \in \mathbb{Z}\}$ is a $k$-dimensional white-noise process, and $L$ stands for the lag operator. Note that we do not make any assumption that (1.2) be causal and invertible.

The same model also can be written as

$$Y_{t,i} = \sum_{j=1}^{m} \beta_{j,i} x_{t,j} + U_{t,i}, \quad t = 1, \ldots, n, \; i = 1, \ldots, k,$$

and is sometimes referred to as a "multivariate ARMA model with a linear trend."

Denote by

$$\boldsymbol{\theta} := \big((\mathrm{vec}\,\boldsymbol{\beta}')', (\mathrm{vec}\,\mathbf{A}_1)', \ldots, (\mathrm{vec}\,\mathbf{A}_p)', (\mathrm{vec}\,\mathbf{B}_1)', \ldots, (\mathrm{vec}\,\mathbf{B}_q)'\big)' \in \mathbb{R}^K$$

($\mathrm{vec}\,\mathbf{C}$, as usual, stands for the vector resulting from stacking the columns of a matrix $\mathbf{C}$ on top of each other) the parameter of the model, with dimension $K := km + k^2(p+q)$.

Writing $\mathcal{M}(\boldsymbol{\Upsilon})$ for the linear subspace of $\mathbb{R}^K$ spanned by the columns of a full-rank $K \times r$ matrix $\boldsymbol{\Upsilon}$ $(r < K)$, we consider the problem of testing the null hypothesis $\mathcal{H}_0$ under which $(\boldsymbol{\theta} - \boldsymbol{\theta}_0) \in \mathcal{M}(\boldsymbol{\Upsilon})$ against an alternative of the form $\mathcal{H}_1 : (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \notin \mathcal{M}(\boldsymbol{\Upsilon})$, where $\boldsymbol{\theta}_0 \in \mathbb{R}^k$. Such null hypotheses are usually referred to as *linear hypotheses*, since belonging to some $r$-dimensional affine subspace in $\mathbb{R}^K$ is equivalent to satisfying some set of $(K - r)$ linearly independent linear constraints.

## 2. Affine invariant testing problems

### 2.1. A characterization of affine invariant hypotheses

The goal of this paper is to determine the class of affine subspaces $\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$ determining affine-invariant linear hypotheses. More precisely, denote by $\mathrm{GL}(k, \mathbb{R})$ the group of real $(k \times k)$ invertible matrices. Then the affine transformation $\boldsymbol{\varepsilon}_t \mapsto \mathbf{M}\boldsymbol{\varepsilon}_t$, $\mathbf{M} \in \mathrm{GL}(k, \mathbb{R})$ of the noise induces the transformation

$$\begin{aligned}&(\boldsymbol{\beta}, \mathbf{A}_1, \ldots, \mathbf{A}_p, \mathbf{B}_1, \ldots, \mathbf{B}_q) \\ &\qquad \mapsto (\boldsymbol{\beta}\mathbf{M}', \mathbf{M}\mathbf{A}_1\mathbf{M}^{-1}, \ldots, \mathbf{M}\mathbf{A}_p\mathbf{M}^{-1}, \mathbf{M}\mathbf{B}_1\mathbf{M}^{-1}, \ldots, \mathbf{M}\mathbf{B}_q\mathbf{M}^{-1})\end{aligned}$$

of the parameter. In terms of $\boldsymbol{\theta}$, this induced transformation is $\boldsymbol{\theta} \mapsto \mathbf{g}_{\mathbf{M}}^{(m,p+q)}\boldsymbol{\theta}$, where

$$\mathbf{g}_{\mathbf{M}}^{(r_1,r_2)} := \begin{pmatrix} \mathbf{I}_{r_1} \otimes \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{r_2} \otimes (\mathbf{M}'^{-1} \otimes \mathbf{M}) \end{pmatrix}$$

(recall indeed that $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\,\mathrm{vec}\,\mathbf{B}$). Here and in the sequel, we use the same notation for a group of linear transformations (acting on some $d$-dimensional real space), and the corresponding collection of $d \times d$ matrices. Letting $\mathcal{G}_{r_2}^{r_1}(k) := \{\mathbf{g}_{\mathbf{M}}^{(r_1,r_2)}, \mathbf{M} \in \mathrm{GL}(k, \mathbb{R})\} \subset \mathrm{GL}(kr_1 + k^2 r_2, \mathbb{R})$, thus, we are investigating under which conditions on $\boldsymbol{\theta}_0$ and $\boldsymbol{\Upsilon}$

$$\mathbf{g}_{\mathbf{M}}^{(m,p+q)}(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})) = \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon}),$$

for all $\mathbf{g}_{\mathbf{M}}^{(m,p+q)} \in \mathcal{G}_{p+q}^{m}(k)$.

As we shall see in the proof of Proposition 2.2, the main task consists in characterizing the class of vector spaces $\mathcal{M}(\mathbf{\Upsilon})$ that are invariant under $\mathcal{G}_{p+q}^{m}(k)$, meaning that $\mathbf{g}_{\mathbf{M}}^{(m,p+q)}\mathcal{M}(\mathbf{\Upsilon}) = \mathcal{M}(\mathbf{\Upsilon})$ or, equivalently, that $\mathcal{M}(\mathbf{g}_{\mathbf{M}}^{(m,p+q)}\mathbf{\Upsilon}) = \mathcal{M}(\mathbf{\Upsilon})$, for all $\mathbf{g}_{\mathbf{M}}^{(m,p+q)} \in \mathcal{G}_{p+q}^{m}(k)$.

Write $M_{m,n}(\mathbb{R})$ for the set of all real $m \times n$ matrices, and let $M_n(\mathbb{R}) := M_{n,n}(\mathbb{R})$. Also let $\widetilde{M}_{m,n}(\mathbb{R})$ stand for the set of full-rank $m \times n$ matrices in $M_{m,n}(\mathbb{R})$. Finally, denote by $\mathbf{e}_i \in \mathbb{R}^k$ the $i$th vector in the canonical basis of $\mathbb{R}^k$, that is, the $i$th column of the $k \times k$ identity matrix $\mathbf{I}_k := (\mathbf{e}_1, \ldots, \mathbf{e}_k)$. The following lemma gives some technical results that will be used later on.

**Lemma 2.1.** *Let* $\mathbf{L}_k := (1/k)\sum_{i,j=1}^{k}(\mathbf{e}_i\mathbf{e}_j' \otimes \mathbf{e}_i\mathbf{e}_j') \in M_{k^2}(\mathbb{R})$, $\mathbf{l}_k := \operatorname{vec}\mathbf{I}_k$ *and* $\mathbf{Q}_k := \mathbf{I}_{k^2} - \mathbf{L}_k$. *Then,*

(i) $\mathbf{L}_k = (1/k)\mathbf{l}_k\mathbf{l}_k'$ *and* $\mathbf{l}_k'\mathbf{l}_k = k$;

(ii) $\mathbf{L}_k$ *and* $\mathbf{Q}_k$ *are symmetric and idempotent;*

(iii) $\mathbf{L}_k(\operatorname{vec}\mathbf{V}) = (1/k)(\operatorname{tr}\mathbf{V})\mathbf{l}_k$ *for all* $\mathbf{V} \in M_k(\mathbb{R})$;

(iv) $\mathbf{L}_k(\mathbf{M}'^{-1} \otimes \mathbf{M}) = (\mathbf{M}'^{-1} \otimes \mathbf{M})\mathbf{L}_k = \mathbf{L}_k$ *for all* $\mathbf{M} \in \mathrm{GL}(k, \mathbb{R})$;

(v) $\mathbf{Q}_k(\mathbf{M}'^{-1} \otimes \mathbf{M}) = (\mathbf{M}'^{-1} \otimes \mathbf{M})\mathbf{Q}_k$ *for all* $\mathbf{M} \in \mathrm{GL}(k, \mathbb{R})$.

The matrix $\mathbf{L}_k$ appearing in this lemma is quite similar to the classical *commutation matrix*

$$\mathbf{K}_k := \frac{1}{k}\sum_{i,j=1}^{k}(\mathbf{e}_i\mathbf{e}_j' \otimes \mathbf{e}_j\mathbf{e}_i'),$$

satisfying the commutation property

$$\mathbf{K}_k(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{B} \otimes \mathbf{A})\mathbf{K}_k \quad \text{for all } \mathbf{A}, \mathbf{B} \in M_k(\mathbb{R}), \tag{2.1}$$

whereas the commutation property of $\mathbf{L}_k$ is given in part (iv) of Lemma 2.1. Note however that (2.1) is not, strictly speaking, a commutation property, as $\mathbf{B} \otimes \mathbf{A}$ stands in the right-hand side, and not $\mathbf{A} \otimes \mathbf{B}$. On the contrary, part (iv) of Lemma 2.1 is a strict commutation property. But of course, the *improper* commutation property in (2.1) is to hold for all $\mathbf{A}, \mathbf{B} \in M_k(\mathbb{R})$, whereas the *strict* commutation property in part (iv) of Lemma 2.1 is only required for $\mathbf{A}, \mathbf{B}$ of the form $\mathbf{M}, \mathbf{M}'^{-1}$, with $\mathbf{M} \in \mathrm{GL}(k, \mathbb{R})$.

It follows from Lemma 2.1 that $\mathbf{Q}_k$ and $\mathbf{L}_k$ are mutually orthogonal projection matrices in $\mathbb{R}^{k^2}$. We will denote by $\mathbf{P}_k := \mathbf{Q}_k(\mathbf{I}_{k^2-1} \vdots \mathbf{0}_{k^2-1\times 1})'$ the matrix whose columns are the $(k^2-1)$ first columns of $\mathbf{Q}_k$. Then the projection $\mathbf{Q}_k$ (resp. $\mathbf{L}_k$) maps $\mathbb{R}^{k^2}$ onto the hyperplane $\mathcal{M}(\mathbf{P}_k)$ (resp. onto the line $\mathcal{M}(\mathbf{l}_k)$). Still from Lemma 2.1, we learn that both $\mathcal{M}(\mathbf{P}_k)$ and $\mathcal{M}(\mathbf{l}_k)$ are invariant under $\mathcal{G}_1^0(k)$. In view of this, the following result is not really surprising.

**Proposition 2.1.** *The vector space $\mathcal{M}(\boldsymbol{\Upsilon})$ is invariant under the group of transformations $\mathcal{G}_{p+q}^m(k)$ if and only if*

$$\boldsymbol{\Upsilon} = \begin{pmatrix} \mathbf{Z}_m \otimes \mathbf{I}_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{p+q} \otimes \mathbf{P}_k & \mathbf{W}_{p+q} \otimes \mathbf{l}_k \end{pmatrix} \mathbf{G}, \tag{2.2}$$

*where $\mathbf{Z}_m, \mathbf{V}_{p+q}$, and $\mathbf{W}_{p+q}$ are (possibly void when either $r_Z$, $r_V$, or $r_W$ are zero) full-rank matrices with dimensions $m \times r_Z$, $(p+q) \times r_V$, and $(p+q) \times r_W$, respectively, and $\mathbf{G} \in \mathrm{GL}(r, \mathbb{R})$, with $r = r_Z k + r_V(k^2 - 1) + r_W < K$.*

Proposition 2.1 characterizes the class of null hypotheses of the form $\mathcal{H}_0$: $\boldsymbol{\theta} \in \mathcal{M}(\boldsymbol{\Upsilon})$ that are invariant under affine transformations. Proposition 2.2 now extends this characterization to the more general case of hypotheses of the form $\mathcal{H}_0$: $\boldsymbol{\theta} \in \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$.

**Proposition 2.2.** *The affine subspace $\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$ is invariant under the group of transformations $\mathcal{G}_{p+q}^m(k)$ if and only if $\boldsymbol{\Upsilon}$ is as in (2.2) and*

$$\boldsymbol{\theta}_0 = \begin{pmatrix} \mathbf{0} \\ \mathbf{w}_{p+q} \otimes \mathbf{l}_k \end{pmatrix} + \boldsymbol{\Upsilon} \boldsymbol{\omega}_r, \tag{2.3}$$

*where $\mathbf{w}_{p+q}$ and $\boldsymbol{\omega}_r$ denote arbitrary vectors with dimensions $p + q$ and $r$, respectively.*

Of course, we can assume, without loss of generality, that $\boldsymbol{\omega}_r = \mathbf{0}$ in Proposition 2.2. Condition (2.3) then takes the form

$$\boldsymbol{\theta}_0 := \big((\mathrm{vec}\, \boldsymbol{\beta}_0')', (\mathrm{vec}\, \mathbf{A}_{1,0})', \ldots, (\mathrm{vec}\, \mathbf{A}_{p,0})', (\mathrm{vec}\, \mathbf{B}_{1,0})', \ldots, (\mathrm{vec}\, \mathbf{B}_{q,0})'\big)' \in \mathbb{R}^K,$$

where $\boldsymbol{\beta}_0 = \mathbf{0}$, and each $\mathbf{A}_{i,0}$ and each $\mathbf{B}_{j,0}$ is a multiple of the $k \times k$ identity matrix.

## 2.2. Some examples

In order to shed some light on the statistical meaning of the somewhat abstract results of Propositions 2.1 and 2.2, we now discuss and illustrate some of their consequences. Without loss of generality, we henceforth assume that $\mathbf{G} = \mathbf{I}_r$.

Furthermore, note that the block-diagonal structure of the array $\boldsymbol{\Upsilon}$ in (2.2) implies that no affine invariant linear hypothesis can mix the "trend parameters" with the VARMA or serial ones. Therefore, in the following examples, we treat separately the constraints dealing with the two types of parameters. The constraint matrices $\boldsymbol{\Upsilon}$ accordingly have a block-row of zeros (associated with the unconstrained parameters) which, for notational simplicity, we omit in the sequel. For instance, in Examples 2.1 and 2.2, we write

$$\boldsymbol{\Upsilon} = (\mathbf{Z}_m \otimes \mathbf{I}_k) \quad \text{instead of} \quad \boldsymbol{\Upsilon} = \begin{pmatrix} \mathbf{Z}_m \otimes \mathbf{I}_k \\ \mathbf{0} \end{pmatrix}.$$

### 2.2.1. Affine invariant hypotheses involving the trend parameters

The situation for the linear model part of (1.1) is rather simple—much simpler than for the VARMA part. A linear constraint of the form $(\boldsymbol{\theta} - \boldsymbol{\theta}_0) \in \mathcal{M}(\boldsymbol{\Upsilon})$, with $\boldsymbol{\Upsilon}$ of the form $\boldsymbol{\Upsilon} := (\mathbf{Z}_m \otimes \mathbf{I}_k)$ and $\boldsymbol{\beta}_0 := \mathbf{0}$ indeed yields a straightforward generalization of the usual univariate linear constraints ($k = 1$); these constraints now simply involve the $m$ $k$-variate regression parameters (i.e., the lines of $\boldsymbol{\beta}$). Of course, in the model under study, the error term is the realization of an unspecified VARMA process, the characteristics of which are nuisance parameters.

**Example 2.1.** Letting $\mathbf{Z}_m := (\mathbf{I}_{m-1} \vdots \mathbf{0}_{m-1\times 1})'$, we obtain the important particular case of testing the significance of the last $k$-dimensional regressor.

**Example 2.2.** As another example, $\mathbf{Z}_m := (1, \ldots, 1)' \in \mathbb{R}^m$ characterizes the null hypothesis under which the $m$ lines of $\boldsymbol{\beta}$ are equal to each other. Such a hypothesis appears, for instance, in the $k$-variate $m$-sample location problem (MANOVA), still in the presence of possibly intercorrelated and serially dependent error terms.

These two important problems thus are affine-invariant.

In the more traditional situation where the errors are i.i.d. normal, the invariance properties of MANOVA models are well documented; the usual Gaussian likelihood ratio tests for the hypotheses in Examples 2.1 and 2.2 then also are affine-invariant (see, e.g., Bilodeau and Brenner, 1999, pp. 158–159).

### 2.2.2. Affine invariant hypotheses involving the serial parameters

**Example 2.3.** Note that the matrix $\boldsymbol{\Upsilon}_1 := (\mathbf{V}_{p+q} \otimes \mathbf{P}_k \vdots \mathbf{V}_{p+q} \otimes \mathbf{l}_k)$ (with $\mathrm{rank}(\mathbf{V}_{p+q}) < p + q$) defines the same null hypothesis as the matrix $\boldsymbol{\Upsilon}_2 := \mathbf{V}_{p+q} \otimes \mathbf{I}_{k^2}$. This hypothesis is the multivariate version of the more traditional univariate linear constraints on the parameters of an ARMA model; the linear model structure of the trend here plays the role of the nuisance.

Of particular interest is the special case characterized by a matrix $\boldsymbol{\Upsilon}$ of the form

$$\boldsymbol{\Upsilon} = \mathbf{V}_{p+q} \otimes \mathbf{I}_{k^2} := \begin{pmatrix} \mathbf{I}_{p_0} & \mathbf{0}_{p_0 \times q_0} \\ \mathbf{0}_{\pi \times p_0} & \mathbf{0}_{\pi \times q_0} \\ \mathbf{0}_{q_0 \times p_0} & \mathbf{I}_{q_0} \\ \mathbf{0}_{\pi \times p_0} & \mathbf{0}_{\pi \times q_0} \end{pmatrix} \otimes \mathbf{I}_{k^2}.$$

This matrix characterizes the problem of testing VARMA$(p_0, q_0)$ against VARMA$(p_0 + \pi, q_0 + \pi)$ dependence $(p = p_0 + \pi,\ q = q_0 + \pi,\ \pi > 0)$. The particular case where $\pi = 1$ plays an important role in several model identification procedures (see, e.g., Pötscher, 1983, 1985, or Garel and Hallin, 1999 for the univariate case).

**Example 2.4.** A matrix of the form $\mathbf{\Upsilon} = \mathbf{V}_{p+q} \otimes \mathbf{P}_k$ yields the same linear constraints as above, but further requires that $\operatorname{tr} \mathbf{A}_1 = \cdots = \operatorname{tr} \mathbf{A}_p = \operatorname{tr} \mathbf{B}_1 = \cdots = \operatorname{tr} \mathbf{B}_q = 0$. Such "trace constraints" can be limited to specific lags by considering $\mathbf{\Upsilon}$ matrices of the form

$$\mathbf{\Upsilon} = \big((\mathbf{V}_{p+q} \vdots \widetilde{\mathbf{V}}_{p+q}) \otimes \mathbf{P}_k \vdots \widetilde{\mathbf{V}}_{p+q} \otimes \mathbf{l}_k\big)$$

or, equivalently,

$$\mathbf{\Upsilon} = (\mathbf{V}_{p+q} \otimes \mathbf{P}_k \vdots \widetilde{\mathbf{V}}_{p+q} \otimes \mathbf{I}_{k^2}).$$

For instance, $\mathbf{V}_{3+0} = (1, 0, 2)'$ and $\widetilde{\mathbf{V}}_{3+0} = (0, 1, 1)'$ yield the null hypothesis $\mathbf{A}_3 = 2\mathbf{A}_1 + \mathbf{A}_2$, $\operatorname{tr} \mathbf{A}_1 = 0$.

**Example 2.5.** A matrix $\mathbf{\Upsilon}$ of the form $\mathbf{\Upsilon} = \mathbf{W}_{p+q} \otimes \mathbf{l}_k$ characterizes the same hypothesis as in Example 2.3, but further specifies that all the $\mathbf{A}_i$'s and $\mathbf{B}_j$'s are of the form $a_i \mathbf{I}_k$ and $b_j \mathbf{I}_k$, respectively. These constraints can also be restricted to a few specific lags, via $\mathbf{\Upsilon}$ matrices of the form

$$\mathbf{\Upsilon} = (\widetilde{\mathbf{W}}_{p+q} \otimes \mathbf{P}_k \vdots (\mathbf{W}_{p+q} \vdots \widetilde{\mathbf{W}}_{p+q}) \otimes \mathbf{l}_k)$$

or, equivalently,

$$\mathbf{\Upsilon} = (\widetilde{\mathbf{W}}_{p+q} \otimes \mathbf{I}_{k^2} \vdots \mathbf{W}_{p+q} \otimes \mathbf{l}_k).$$

For instance, $\widetilde{\mathbf{W}}_{2+1} = (1, 0, 2)'$ and $\mathbf{W}_{2+1} = (0, 1, -1)'$ yield the null hypothesis $\mathbf{B}_1 = 2\mathbf{A}_1 - \mathbf{A}_2$, with $\mathbf{A}_2 = a_2 \mathbf{I}_k$ for some $a_2 \in \mathbb{R}$.

**Example 2.6.** It is also possible to consider linear constraints on the traces of the $\mathbf{A}_i$'s and $\mathbf{B}_j$'s. This can be obtained by letting $\mathbf{\Upsilon} = (\mathbf{I}_{p+q} \otimes \mathbf{P}_k \vdots \mathbf{W}_{p+q} \otimes \mathbf{l}_k)$, with $\operatorname{rank}(\mathbf{W}_{p+q}) < p + q$. For instance, the hypothesis that all traces are equal corresponds to the special case of $\mathbf{W}_{p+q} = (1, 1, \ldots, 1)' \in \mathbb{R}^{p+q}$.

**Example 2.7.** Examples 2.3 through 2.6 are dealing with the case $\boldsymbol{\theta}_0 = \mathbf{0}$ covered by Proposition 2.1. Here is an application of Proposition 2.2, with $\boldsymbol{\theta}_0 \neq \mathbf{0}$. Letting for simplicity $\mathbf{\Upsilon}$ be the void matrix (so that $\mathcal{M}(\mathbf{\Upsilon}) = \{\mathbf{0}\}$), the null hypotheses under which the VARMA coefficients $\mathbf{A}_{i,0}$ and $\mathbf{B}_{j,0}$ are of the form $a_{i,0}\mathbf{I}_k$ and $b_{j,0}\mathbf{I}_k$, respectively, for some specified constants $a_{i,0}$ and $b_{j,0}$, satisfy the conditions of Proposition 2.2, and thus are affine-invariant. Of special interest is the particular case $a_{i,0} = 0$, $b_{j,0} = 0$ for all $i$ and $j$, which corresponds to testing randomness against serial VARMA dependence.

One should not have the feeling, however, that all hypotheses of practical interest are invariant under affine transformations. Here are two examples of non-invariant hypotheses.

**Counterexample 2.1.** Turning back to the purely non serial case (i.i.d. errors), the most usual problem in analysis of covariance models—that of comparing $k$ univariate regression equations, each of them involving $m$ regressors—is not affine-invariant. Indeed, the corresponding null hypothesis $\mathcal{H}_0$: $\boldsymbol{\theta} \in \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$, with

$$\boldsymbol{\theta}_0 = \mathbf{0}, \quad \text{and} \quad \boldsymbol{\Upsilon} = \mathbf{I}_m \otimes \mathbf{1}_k, \text{ with } \mathbf{1}_k := (1, \ldots, 1)' \in \mathbb{R}^k$$

(stipulating that the $k$ columns of $\boldsymbol{\beta}$—not the $m$ lines—are equal to each other), involves a matrix $\boldsymbol{\Upsilon}$ that clearly fails to satisfy (2.2).

**Counterexample 2.2.** Considering the purely serial model, i.e., the VARMA$(p, q)$ model (1.2), assume that $k = ds$ with $d \in \mathbb{N}$, $d \geq 2$, and partition $\mathbf{U}'_t = (U_{t,1}, \ldots, U_{t,k})$ into $d$ $s$-variate subvectors: $\mathbf{U}'_t = (\mathbf{U}'_{t,1}, \ldots, \mathbf{U}'_{t,d})$. Denote by $\mathbf{J}_l^{(s,d)}$ the $ds \times s$ matrix

$$\begin{pmatrix} \mathbf{0}_{(l-1)s\times s} \\ \mathbf{I}_s \\ \mathbf{0}_{(d-l)s\times s} \end{pmatrix}.$$

Then the matrix $\boldsymbol{\Upsilon}$ characterizing the hypothesis under which, after adequate partitioning into blocks of dimension $s \times s$, the matrices $\mathbf{A}_1, \ldots, \mathbf{A}_p$ and $\mathbf{B}_1, \ldots, \mathbf{B}_q$ are all block-diagonal (with unspecified diagonal blocks) is

$$\mathbf{I}_{p+q} \otimes \begin{pmatrix} \mathbf{I}_s \otimes \mathbf{J}_1^{(s,d)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_s \otimes \mathbf{J}_2^{(s,d)} & & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_s \otimes \mathbf{J}_d^{(s,d)} \end{pmatrix}.$$

Proposition 2.2 implies that this hypothesis is *not* affine-invariant.

**Counterexample 2.3.** Consider again the same VARMA$(p, q)$ model as in Counterexample 2.2, with the same partition of $\mathbf{U}_t$, but assume that (1.2) is obtained by stacking $d$ independent $s$-variate models ($k = ds$). The corresponding $\mathbf{A}_i$'s and $\mathbf{B}_j$'s then, after adequate partitioning, are naturally block-diagonal. Denoting by $\mathbf{A}_i^{(l)}$ the $l$th diagonal block in $\mathbf{A}_i$, by $\mathbf{B}_j^{(l)}$ the $l$th diagonal block in $\mathbf{B}_j$, consider the problem of comparing the blocks $\mathbf{A}_i^{(l)}$ and $\mathbf{B}_j^{(l)}$ for various values of $l$. For instance, consider the hypothesis under which $\mathbf{A}_i^{(l)} = \mathbf{A}_i^{(1)}$ and $\mathbf{B}_j^{(l)} = \mathbf{B}_j^{(1)}$, for all $l = 1, \ldots, d$, $i = 1, \ldots, p$ and

$j = 1, \ldots, q$ (comparison of $d$ $s$-dimensional VARMA$(p, q)$ models). Again, Proposition 2.2 implies that this hypothesis, though perfectly relevant from a practical point of view, is not affine-invariant.

The reason for this lack of invariance with respect to the group of affine transformations $\mathrm{GL}(k, \mathbb{R})$ is that this group in fact is not adapted to the testing problems at hand in each of these counterexamples. Mixing the components of $\boldsymbol{\varepsilon}_t$ by transforming $\boldsymbol{\varepsilon}_t$ into $\mathbf{M}\boldsymbol{\varepsilon}_t$ for arbitrary $\mathbf{M} \in \mathrm{GL}(ds, \mathbb{R})$ makes little sense, since this creates cross-correlations between the $k$ regression submodels or the $d$ time series under study. And, a *fully* affine invariant procedure in this case would do a poor job. Full affine invariance in such situations should be weakened into a lesser requirement of invariance with respect to some appropriate subgroup of $\mathrm{GL}(ds, \mathbb{R})$. In Counterexample 2.2, the adequate subgroup (making the hypothesis invariant) would be

$$\left\{\mathbf{M} \in \mathrm{GL}(ds, \mathbb{R}) : \mathbf{M} = \sum_{l=1}^{d} \mathbf{e}_l^{(d)} \mathbf{e}_l^{(d)\prime} \otimes \mathbf{M}_l, \mathbf{M}_l \in \mathrm{GL}(s, \mathbb{R})\right\}, \tag{2.4}$$

where $\mathbf{e}_l^{(d)} \in \mathbb{R}^d$ denotes the $l$th vector of the canonical basis of $\mathbb{R}^d$. This subgroup is of course isomorphic to $GL(s, \mathbb{R}) \times \cdots \times GL(s, \mathbb{R})$ ($d$ times), where "$\times$" denotes the direct product of groups. In Counterexample 2.3, the right subgroup is even smaller, taking the form

$$\{\mathbf{M} \in \mathrm{GL}(ds, \mathbb{R}) : \mathbf{M} = \mathbf{I}_d \otimes \mathbf{M}_1, \mathbf{M}_1 \in \mathrm{GL}(s, \mathbb{R})\}, \tag{2.5}$$

clearly, a subgroup of group (2.4).

## 3. Conclusions

Affine invariance plays a fundamental role in a variety of robust inference methods for multivariate observations. Such methods recently have been considered (Hallin and Paindaveine 2002a-c) for hypothesis testing in the context of multivariate linear models with VARMA errors. However, not all linear hypotheses qualify for such methods, which only make sense for affine-invariant hypotheses. Therefore, an algebraic characterization of the class of affine-invariant hypotheses is proposed in Propositions 2.1 and 2.2, and examples of invariant and noninvariant problems are discussed.

**Acknowledgements.** The research of M.H. was supported by an A.R.C. contract of the Communauté française de Belgique and a P.A.I. from the Belgian federal government.

APPENDIX

## A. Proofs

We start with the proof of Lemma 2.1.

*Proof of Lemma* 2.1. (i) Only the first statement needs some proof. From the definition of $\mathbf{L}_k$, we have

$$\mathbf{L}_k := \frac{1}{k}\sum_{i,j=1}^{k}(\mathbf{e}_i\mathbf{e}_j' \otimes \mathbf{e}_i\mathbf{e}_j') = \frac{1}{k}\sum_{i,j=1}^{k}(\operatorname{vec}\mathbf{e}_i\mathbf{e}_i')(\operatorname{vec}\mathbf{e}_j\mathbf{e}_j')'$$
$$= \frac{1}{k}\Big(\sum_{i=1}^{k}\operatorname{vec}\mathbf{e}_i\mathbf{e}_i'\Big)\Big(\sum_{j=1}^{k}\operatorname{vec}\mathbf{e}_j\mathbf{e}_j'\Big)' = \frac{1}{k}\mathbf{l}_k\mathbf{l}_k'.$$

(ii) It follows from (i) that $\mathbf{L}_k$ is symmetric. It is also idempotent, since

$$(\mathbf{L}_k)^2 = \frac{1}{k^2}\mathbf{l}_k\mathbf{l}_k'\mathbf{l}_k\mathbf{l}_k' = \frac{1}{k}\mathbf{l}_k\mathbf{l}_k' = \mathbf{L}_k.$$

(iii) The identity $(\mathbf{A}\otimes\mathbf{B})(\operatorname{vec}\mathbf{V}) = \operatorname{vec}(\mathbf{B}\mathbf{V}\mathbf{A}')$ yields

$$\mathbf{L}_k(\operatorname{vec}\mathbf{V}) = \frac{1}{k}\sum_{i,j}\operatorname{vec}(\mathbf{e}_i\mathbf{e}_j'\mathbf{V}\mathbf{e}_j\mathbf{e}_i') \tag{A.1}$$
$$= \frac{1}{k}\sum_{i,j}\operatorname{vec}(V_{jj}\mathbf{e}_i\mathbf{e}_i') = \frac{1}{k}(\operatorname{tr}\mathbf{V})(\operatorname{vec}\mathbf{I}_k).$$

(iv) From (A.1), we have

$$\mathbf{L}_k(\mathbf{M}'^{-1}\otimes\mathbf{M})(\operatorname{vec}\mathbf{V}) = \mathbf{L}_k\operatorname{vec}(\mathbf{M}\mathbf{V}\mathbf{M}^{-1}) = \frac{1}{k}\operatorname{tr}(\mathbf{M}\mathbf{V}\mathbf{M}^{-1})\mathbf{l}_k$$
$$= \frac{1}{k}(\operatorname{tr}\mathbf{V})\mathbf{l}_k = \mathbf{L}_k(\operatorname{vec}\mathbf{V})$$

for all $\mathbf{V} \in M_k(\mathbb{R})$, which implies $\mathbf{L}_k(\mathbf{M}'^{-1}\otimes\mathbf{M}) = \mathbf{L}_k$, for all $\mathbf{M} \in GL(k,\mathbb{R})$. Taking transposes, one obtains $(\mathbf{M}^{-1}\otimes\mathbf{M}')\mathbf{L}_k = \mathbf{L}_k$ for all $\mathbf{M}$. As for (v), it trivially follows from the previous result. □

As already stressed in Section 2.1, Lemma 2.1 shows that both $\mathcal{M}(\mathbf{P}_k)$ and $\mathcal{M}(\mathbf{l}_k)$ are invariant under $\mathcal{G}_1^0(k)$. Before turning to the proof of Proposition 2.1, we first establish a couple of further lemmas, the general purpose of which is to make sure that no proper subspace of $\mathcal{M}(\mathbf{P}_k)$ is invariant under $\mathcal{G}_1^0(k)$. We first recall the concept of *centralizer* in group theory. Let $\mathcal{H}$ be some subgroup of a group $\mathcal{G}$. Then, the centralizer $C_{\mathcal{G}}(\mathcal{H})$ of $\mathcal{H}$ in $\mathcal{G}$ is defined as the collection of all elements in $\mathcal{G}$ that commute with every element in $\mathcal{H}$.

**Lemma A.1.** *For any* $\mathbf{V} \in \widetilde{M}_{K,r}(\mathbb{R})$ $(r \leq K)$, *denote by* $\mathbf{\Pi}(\mathbf{V}) := \mathbf{V}(\mathbf{V}'\mathbf{V})^{-1}\mathbf{V}'$ *the matrix of the orthogonal projection in* $\mathbb{R}^K$ *on the subspace* $\mathcal{M}(\mathbf{V})$. *Then* $\mathcal{M}(\mathbf{V})$ *is invariant under* $\mathcal{G}^m_{p+q}(k)$ *if and only if* $\mathbf{\Pi}(\mathbf{V}) \in C_{M_K(\mathbb{R})}(\mathcal{G}^m_{p+q}(k))$.

*Proof.* Note that $\mathcal{M}(\mathbf{V})$ is invariant under $\mathcal{G}^m_{p+q}(k)$ if and only if $\mathbf{V}$ is such that $\mathbf{\Pi}(\mathbf{V})\mathbf{g}_{\mathbf{M}}^{(m,p+q)}\mathbf{V} = \mathbf{g}_{\mathbf{M}}^{(m,p+q)}\mathbf{V}$ for all $\mathbf{g}_{\mathbf{M}}^{(m,p+q)} \in \mathcal{G}^m_{p+q}(k)$, i.e., iff

$$\mathbf{\Pi}(\mathbf{V})\mathbf{g}_{\mathbf{M}}^{(m,p+q)}\mathbf{V} = \mathbf{g}_{\mathbf{M}}^{(m,p+q)}\mathbf{\Pi}(\mathbf{V})\mathbf{V} \quad \text{for all } \mathbf{g}_{\mathbf{M}}^{(m,p+q)} \in \mathcal{G}^m_{p+q}(k). \tag{A.2}$$

This proves the sufficient condition. To prove the necessary one, assume that (A.2) holds. Now, for any $\mathbf{w} \in \big(\mathcal{M}(\mathbf{V})\big)^{\perp}$, where $\big(\mathcal{M}(\mathbf{V})\big)^{\perp}$ denotes the Euclidean orthogonal complement of $\mathcal{M}(\mathbf{V})$ in $\mathbb{R}^K$,

$$\begin{aligned}\mathbf{V}'[\mathbf{g}_{\mathbf{M}}^{(m,p+q)}\mathbf{w}] = [\mathbf{g}_{\mathbf{M}'}^{(m,p+q)}\mathbf{V}]'\mathbf{w} &= [\mathbf{\Pi}(\mathbf{V})\mathbf{g}_{\mathbf{M}'}^{(m,p+q)}\mathbf{V}]'\mathbf{w} \\ &= \mathbf{V}'\mathbf{g}_{\mathbf{M}}^{(m,p+q)}\,\mathbf{\Pi}(\mathbf{V})\mathbf{w} = \mathbf{0},\end{aligned}$$

for all $\mathbf{g}_{\mathbf{M}}^{(m,p+q)} \in \mathcal{G}^m_{p+q}(k)$. It follows that, for some matrix $\mathbf{V}_{\perp}$ whose columns form a basis of $\big(\mathcal{M}(\mathbf{V})\big)^{\perp}$, we have, for all $\mathbf{g}_{\mathbf{M}}^{(m,p+q)} \in \mathcal{G}^m_{p+q}(k)$, $\mathbf{\Pi}(\mathbf{V})\mathbf{g}_{\mathbf{M}}^{(m,p+q)}\mathbf{V}_{\perp} = \mathbf{0} = \mathbf{g}_{\mathbf{M}}^{(m,p+q)}\mathbf{\Pi}(\mathbf{V})\mathbf{V}_{\perp}$. Piecing this together with (A.2) yields $\mathbf{\Pi}(\mathbf{V})\mathbf{g}_{\mathbf{M}}^{(m,p+q)} = \mathbf{g}_{\mathbf{M}}^{(m,p+q)}\mathbf{\Pi}(\mathbf{V})$, for all $\mathbf{g}_{\mathbf{M}}^{(m,p+q)} \in \mathcal{G}^m_{p+q}(k)$. $\square$

For any couple $\mathbf{A}, \mathbf{B}$, where $\mathbf{A}$ is a $(m_1 \times n_1)$ matrix and $\mathbf{B}$ a $(m_2 \times n_2)$ one, let

$$\mathbf{J}(\mathbf{A};\mathbf{B}) := \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}.$$

**Lemma A.2.** *Denote by* $\mathcal{P}_m$ *the set of* $m \times m$ *symmetric and idempotent matrices. Then, for all* $r_1, r_2 \in \mathbb{N}$,

(i) *the centralizer of* $\mathcal{G}^{r_1}_{r_2}(k)$ *in* $M_{kr_1+k^2r_2}(\mathbb{R})$ *is*

$$\begin{aligned} &C_{M_{kr_1+k^2r_2}(\mathbb{R})}\big(\mathcal{G}^{r_1}_{r_2}(k)\big) \\ &\quad = \{\mathbf{J}(\mathbf{C}\otimes\mathbf{I}_k; \mathbf{A}\otimes\mathbf{I}_{k^2} + \mathbf{B}\otimes\mathbf{L}_k), \mathbf{C} \in M_{r_1}(\mathbb{R}), \mathbf{A}, \mathbf{B} \in M_{r_2}(\mathbb{R})\} \\ &\quad = \{\mathbf{J}(\mathbf{C}\otimes\mathbf{I}_k; \mathbf{A}\otimes\mathbf{Q}_k + \mathbf{B}\otimes\mathbf{L}_k), \mathbf{C} \in M_{r_1}(\mathbb{R}), \mathbf{A}, \mathbf{B} \in M_{r_2}(\mathbb{R})\};\end{aligned}$$

(ii) $C_{M_{kr_1+k^2r_2}(\mathbb{R})}(\mathcal{G}^{r_1}_{r_2}(k)) \cap \mathcal{P}_{kr_1+k^2r_2}$ *is given by*

$$\begin{aligned}&\{\mathbf{J}(\mathbf{\Pi}(\mathbf{Z}_{r_1})\otimes\mathbf{I}_k; \mathbf{\Pi}(\mathbf{V}_{r_2})\otimes\mathbf{Q}_k + \mathbf{\Pi}(\mathbf{W}_{r_2})\otimes\mathbf{L}_k), \\ &\qquad \mathbf{Z}_{r_1} \in \widetilde{M}_{r_1,r_Z}(\mathbb{R}), \mathbf{V}_{r_2} \in \widetilde{M}_{r_2,r_V}(\mathbb{R}), \mathbf{W}_{r_2} \in \widetilde{M}_{r_2,r_W}(\mathbb{R})\},\end{aligned}$$

*with* $0 \leq r_Z \leq r_1$ *and* $0 \leq r_V, r_W \leq r_2$.

*Proof.* (i) It is clearly sufficient to prove that the centralizer of $\mathcal{G}_0^{r_1}(k)$ in $M_{kr_1}(\mathbb{R})$ is

$$\{\mathbf{C}\otimes\mathbf{I}_k, \mathbf{C}\in M_{r_1}(\mathbb{R})\}, \tag{A.3}$$

and that the centralizer of $\mathcal{G}_{r_2}^0(k)$ in $M_{k^2r_2}(\mathbb{R})$ is

$$\{\mathbf{A}\otimes\mathbf{I}_{k^2}+\mathbf{B}\otimes\mathbf{L}_k,\ \mathbf{A},\mathbf{B}\in M_{r_2}(\mathbb{R})\}. \tag{A.4}$$

Let us first prove that $C_{M_{kr_1}(\mathbb{R})}\big(\mathcal{G}_0^{r_1}(k)\big)$ is given by (A.3). Let $\mathbf{D}$ belong to the centralizer of $\mathcal{G}_0^{r_1}(k)$ in $M_{kr_1}(\mathbb{R})$, with $k\times k$ block $\mathbf{D}_{mn}$ in position $(m,n)$. Consider the arrays $\mathbf{E}_{ij}:=\mathbf{e}_i\mathbf{e}_j'$ and $\mathbf{M}_{ij}:=\mathbf{I}_k+\mathbf{E}_{ij}\in \mathrm{GL}(k,\mathbb{R})$, for $i,j=1,\dots,k$. Then, $\mathbf{g}_{\mathbf{M}_{ij}}^{(r_1,0)}\mathbf{D}=\mathbf{D}\mathbf{g}_{\mathbf{M}_{ij}}^{(r_1,0)}$ if and only if $\mathbf{E}_{ij}\mathbf{D}_{mn}=\mathbf{D}_{mn}\mathbf{E}_{ij}$ for all $m,n=1,\dots,r_1$ and all $i,j=1,\dots,k$. This yields $\mathbf{D}_{mn}=c_{mn}\,\mathbf{I}_k$, i.e., $\mathbf{D}=\mathbf{C}\otimes\mathbf{I}_k$ (with $\mathbf{C}=(c_{mn})$). This shows that the centralizer $C_{M_{kr_1}(\mathbb{R})}(\mathcal{G}_0^{r_1}(k))$ is a subset of (A.3). The result follows, since each element in (A.3) clearly commutes with every element in $\mathcal{G}_0^{r_1}(k)$.

Secondly, we prove that the centralizer of $C_{M_{k^2r_2}(\mathbb{R})}\big(\mathcal{G}_{r_2}^0(k)\big)$ is given by (A.4). Let us first establish this result for $r_2=1$. Using Lemma 2.1, every element of (A.4) (with $r_2=1$) is seen to belong to $C_{M_{k^2}(\mathbb{R})}(\mathcal{G}_1^0(k))$, so that it is sufficient to show that $C_{M_{k^2}(\mathbb{R})}(\mathcal{G}_1^0(k))$ is a subset of (A.4).

To achieve this, let $\mathbf{D}$ be in the centralizer of $C_{M_{k^2}(\mathbb{R})}\big(\mathcal{G}_1^0(k)\big)$, and denote by $\mathbf{D}_{mn}$ the $k\times k$ block in position $(m,n)$ in $\mathbf{D}$. Considering the arrays $\mathbf{M}_{ij}$, $i,j=1,\dots,k$ again, note that $\mathbf{M}_{ij}^{-1}:=\mathbf{I}_k-\mathbf{E}_{ij}$ for $i\neq j$, and $\mathbf{M}_{ii}^{-1}:=\mathbf{I}_k-\frac{1}{2}\mathbf{E}_{ii}$. Now, $\mathbf{D}$ does commute with $\mathbf{g}_{\mathbf{M}_{ij}}^{(0,1)}$ if and only if it does with $\mathbf{F}_{ij}$, where $\mathbf{F}_{ij}:=(\mathbf{I}_k\otimes\mathbf{E}_{ij})-\mathbf{E}_{ji}\otimes(\mathbf{I}_k+\mathbf{E}_{ij})$ for $i\neq j$, and $\mathbf{F}_{ii}:=(\mathbf{I}_k\otimes\mathbf{E}_{ii})-\mathbf{E}_{ii}\otimes[\frac{1}{2}(\mathbf{I}_k+\mathbf{E}_{ii})]$.

If $\mathbf{D}$ is in $C_{M_{k^2}(\mathbb{R})}\big(\mathcal{G}_1^0(k)\big)$, the $k\times k$ block in position $(m,m)$ in $\mathbf{F}_{ii}\mathbf{D}-\mathbf{D}\mathbf{F}_{ii}$, namely, $\mathbf{E}_{ii}\mathbf{D}_{mm}-\mathbf{D}_{mm}\mathbf{E}_{ii}$, is equal to $\mathbf{0}$ for all $m,i=1,\dots,k$, which implies that the matrices $\mathbf{D}_{mm}$ are diagonal: write

$$\mathbf{D}_{mm}=\begin{pmatrix}\lambda_1^{(m)} & & 0\\ & \ddots & \\ 0 & & \lambda_k^{(m)}\end{pmatrix}.$$

The block in position $(m,n)$, $m\neq n$, in $\mathbf{F}_{ij}\mathbf{D}-\mathbf{D}\mathbf{F}_{ij}$ is $\mathbf{E}_{ij}\mathbf{D}_{mn}-\mathbf{D}_{mn}\mathbf{E}_{ij}$ for all $i\neq n$ and $j\neq m$. All these blocks are $\mathbf{0}$, so that $\mathbf{D}_{mn}$ is of the form

$$\mathbf{D}_{mn}=\lambda_{mn}\,\mathbf{I}_k+\mu_{mn}\,\mathbf{E}_{mn},\quad \text{all } m\neq n,$$

for some $\lambda_{mn},\mu_{mn}\in\mathbb{R}$.

Again, considering the block in position $(j,i)$, $i\neq j$, in $\mathbf{F}_{ij}\mathbf{D}-\mathbf{D}\mathbf{F}_{ij}=\mathbf{0}$ yields

$$\mathbf{E}_{ij}\mathbf{D}_{ji}-\mathbf{D}_{ji}\mathbf{E}_{ij}-(\mathbf{I}_k+\mathbf{E}_{ij})\mathbf{D}_{ii}+\mathbf{D}_{jj}(\mathbf{I}_k+\mathbf{E}_{ij})=\mathbf{0}$$

for all $i \neq j$. This implies $\lambda_i^{(i)} = \lambda_1^{(1)}$ for all $i = 1, \ldots, k$, $\lambda_j^{(i)} = \lambda_2^{(1)}$ for all $i \neq j$, and $\mu_{ij} = \lambda_1^{(1)} - \lambda_2^{(1)}$ for all $i \neq j$. Finally, the block in position $(m, i)$, $i \neq j$, $m \neq j$, in $\mathbf{F}_{ij}\mathbf{D} - \mathbf{D}\mathbf{F}_{ij} = \mathbf{0}$ yields

$$\mathbf{E}_{ij}\mathbf{D}_{mi} - \mathbf{D}_{mi}\mathbf{E}_{ij} + \mathbf{D}_{mj}(\mathbf{I}_k + \mathbf{E}_{ij}) = \mathbf{0}$$

for all $i \neq j$. This implies $\lambda_{mj} = 0$ for all $m \neq j$.

Collecting all these results shows that $\mathbf{D} = \lambda_2^{(1)}\mathbf{I}_{k^2} + (\lambda_1^{(1)} - \lambda_2^{(1)})k\,\mathbf{L}_k$, for some $\lambda_1^{(1)}, \lambda_2^{(1)} \in \mathbb{R}$, which proves that $C_{M_{k^2}(\mathbb{R})}(\mathcal{G}_1^0(k))$ is included in (A.4) (with $r_2 = 1$).

The result for $r_2 > 1$ then follows. Indeed, denote by $\mathbf{G}_{ij}$ the $k^2 \times k^2$ block in position $(i, j)$ in the $k^2 r_2 \times k^2 r_2$ matrix $\mathbf{G}$. Then $\mathbf{G} \in C_{M_{k^2 r_2}(\mathbb{R})}(\mathcal{G}_{r_2}^0(k))$ if and only if $\mathbf{G}_{ij} \in C_{M_{k^2}(\mathbb{R})}(\mathcal{G}_1^0(k))$ for all $i, j = 1, \ldots, r_2$, i.e., iff every $\mathbf{G}_{ij}$ is of the form $a_{ij}\mathbf{I}_{k^2} + b_{ij}\mathbf{L}_k$, which yields the desired result.

(ii) Let $\mathbf{J}(\mathbf{C}\otimes\mathbf{I}_k; \mathbf{A}\otimes\mathbf{Q}_k + \mathbf{B}\otimes\mathbf{L}_k)$ in $C_{M_{kr_1+k^2r_2}(\mathbb{R})}(\mathcal{G}_{r_2}^{r_1}(k))$ be symmetric and idempotent. Then $\mathbf{C}^2 = \mathbf{C} = \mathbf{C}'$ and

$$\mathbf{A} \otimes \mathbf{Q}_k + \mathbf{B} \otimes \mathbf{L}_k = \mathbf{A}^2 \otimes \mathbf{Q}_k + \mathbf{B}^2 \otimes \mathbf{L}_k = \mathbf{A}' \otimes \mathbf{Q}_k + \mathbf{B}' \otimes \mathbf{L}_k,$$

which yields

$$(\mathbf{A}^2 - \mathbf{A}) \otimes \mathbf{Q}_k + (\mathbf{B}^2 - \mathbf{B}) \otimes \mathbf{L}_k = (\mathbf{A}' - \mathbf{A}) \otimes \mathbf{Q}_k + (\mathbf{B}' - \mathbf{B}) \otimes \mathbf{L}_k = \mathbf{0}.$$

Linear independence between $\mathbf{Q}_k$ and $\mathbf{L}_k$ in $M_{k^2}(\mathbb{R})$ implies that $\mathbf{A}^2 = \mathbf{A} = \mathbf{A}'$ and $\mathbf{B}^2 = \mathbf{B} = \mathbf{B}'$. This shows that $\mathbf{C}$, $\mathbf{A}$, and $\mathbf{B}$ are orthogonal projection matrices, on $\mathcal{M}(\mathbf{Z}_{r_1})$, $\mathcal{M}(\mathbf{V}_{r_2})$ and $\mathcal{M}(\mathbf{W}_{r_2})$, say, respectively. □

We are now ready to prove Proposition 2.1.

*Proof of Proposition* 2.1. Let us first prove the sufficiency part of the proposition. Using the fact that $\mathbf{P}_k'\mathbf{l}_k = (\mathbf{I}_{k^2-1} \vdots \mathbf{0}_{k^2-1\times 1})\mathbf{Q}_k\mathbf{l}_k = \mathbf{0}$, one obtains that

$$\text{(A.5)}\quad \boldsymbol{\Upsilon} = \mathbf{J}(\mathbf{Z}_m \otimes \mathbf{I}_k; (\mathbf{V}_{p+q} \otimes \mathbf{P}_k \vdots \mathbf{W}_{p+q} \otimes \mathbf{l}_k))\mathbf{G} \qquad \text{for some } \mathbf{G} \in \mathrm{GL}(r, \mathbb{R}),$$

implies

$$\text{(A.6)}\quad \begin{aligned}\boldsymbol{\Pi}(\boldsymbol{\Upsilon}) &= \mathbf{J}(\boldsymbol{\Pi}(\mathbf{Z}_m) \otimes \mathbf{I}_k; \boldsymbol{\Pi}(\mathbf{V}_{p+q}) \otimes \boldsymbol{\Pi}(\mathbf{P}_k) + \boldsymbol{\Pi}(\mathbf{W}_{p+q}) \otimes \boldsymbol{\Pi}(\mathbf{l}_k))\\ &= \mathbf{J}(\boldsymbol{\Pi}(\mathbf{Z}_m) \otimes \mathbf{I}_k; \boldsymbol{\Pi}(\mathbf{V}_{p+q}) \otimes \mathbf{Q}_k + \boldsymbol{\Pi}(\mathbf{W}_{p+q}) \otimes \mathbf{L}_k),\end{aligned}$$

which, in view of Lemma A.2, is equivalent to

$$\boldsymbol{\Pi}(\boldsymbol{\Upsilon}) \in C_{M_K(\mathbb{R})}(\mathcal{G}_{p+q}^m(k)),$$

hence also, by Lemma A.1, to the statement that $\mathcal{M}(\boldsymbol{\Upsilon})$ is invariant under the group of transformations $\mathcal{G}^m_{p+q}(k)$.

The necessity part as usual is more intricate. But we already worked a lot. It remains to show that the implication from (A.5) to (A.6) can be reversed. Note first that

$$\begin{aligned}\boldsymbol{\Pi}(\boldsymbol{\Upsilon})\mathbf{J}\big(\mathbf{Z}_m \otimes \mathbf{I}_k; (\mathbf{V}_{p+q} \otimes \mathbf{P}_k \vdots \mathbf{W}_{p+q} \otimes \mathbf{l}_k)\big) \\ = \mathbf{J}\big(\mathbf{Z}_m \otimes \mathbf{I}_k; (\mathbf{V}_{p+q} \otimes \mathbf{P}_k \vdots \mathbf{W}_{p+q} \otimes \mathbf{l}_k)\big),\end{aligned}$$

since there exist full-rank arrays $\mathbf{Z}_m$, $\mathbf{V}_{p+q}$ and $\mathbf{W}_{p+q}$ such that $\boldsymbol{\Pi}(\boldsymbol{\Upsilon})$ is given by (A.6). This yields

$$\text{(A.7)} \qquad \mathcal{M}\Big(\mathbf{J}\big(\mathbf{Z}_m \otimes \mathbf{I}_k; (\mathbf{V}_{p+q} \otimes \mathbf{P}_k \vdots \mathbf{W}_{p+q} \otimes \mathbf{l}_k)\big)\Big) \subset \mathcal{M}(\boldsymbol{\Upsilon}).$$

Now,

$$\begin{aligned}\text{(A.8)} \quad & \dim \mathcal{M}(\boldsymbol{\Upsilon}) \\ & = \operatorname{rank}\big(\boldsymbol{\Pi}(\boldsymbol{\Upsilon})\big) \\ & = \operatorname{rank}(\boldsymbol{\Pi}(\mathbf{Z}_m)\otimes\mathbf{I}_k) + \operatorname{rank}\big(\boldsymbol{\Pi}(\mathbf{V}_{p+q})\otimes\boldsymbol{\Pi}(\mathbf{P}_k) + \boldsymbol{\Pi}(\mathbf{W}_{p+q})\otimes\boldsymbol{\Pi}(\mathbf{l}_k)\big) \\ & \leq \operatorname{rank}\big(\boldsymbol{\Pi}(\mathbf{Z}_m\otimes\mathbf{I}_k)\big) + \dim[\operatorname{Im}\boldsymbol{\Pi}(\mathbf{V}_{p+q}\otimes\mathbf{P}_k) \oplus \operatorname{Im}\boldsymbol{\Pi}(\mathbf{W}_{p+q}\otimes\mathbf{l}_k)],\end{aligned}$$

where "$\oplus$" denotes the direct sum in $\mathbb{R}^{k^2(p+q)}$ of vector subspaces. Noting that

$$\mathcal{M}\big((\mathbf{V}_{p+q} \otimes \mathbf{P}_k \vdots \mathbf{W}_{p+q} \otimes \mathbf{l}_k)\big) = \mathcal{M}(\mathbf{V}_{p+q} \otimes \mathbf{P}_k) \oplus \mathcal{M}(\mathbf{W}_{p+q} \otimes \mathbf{l}_k),$$

(A.8) reduces to

$$\begin{aligned}\dim\big(\mathcal{M}(\mathbf{Z}_m \otimes \mathbf{I}_k)\big) + \dim\big(\mathcal{M}(\mathbf{V}_{p+q} \otimes \mathbf{P}_k) \oplus \mathcal{M}(\mathbf{W}_{p+q} \otimes \mathbf{l}_k)\big) \\ = \dim \mathcal{M}\Big(\mathbf{J}\big(\mathbf{Z}_m \otimes \mathbf{I}_k; (\mathbf{V}_{p+q} \otimes \mathbf{P}_k \vdots \mathbf{W}_{p+q} \otimes \mathbf{l}_k)\big)\Big) \\ \geq \dim \mathcal{M}(\boldsymbol{\Upsilon}).\end{aligned}$$

Comparing with (A.7), we deduce that

$$\mathcal{M}\Big(\mathbf{J}\big(\mathbf{Z}_m \otimes \mathbf{I}_k; (\mathbf{V}_{p+q} \otimes \mathbf{P}_k \vdots \mathbf{W}_{p+q} \otimes \mathbf{l}_k)\big)\Big) = \mathcal{M}(\boldsymbol{\Upsilon}),$$

so that (A.5) and (A.6) actually are equivalent. □

Finally, we prove Proposition 2.2.

*Proof of Proposition* 2.2. The subspace $\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$ is invariant under the group $\mathcal{G}^m_{p+q}(k)$ iff $\mathbf{g}_{\mathbf{M}}^{(m,p+q)}(\boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})) = \mathbf{g}_{\mathbf{M}}^{(m,p+q)}\boldsymbol{\theta}_0 + \mathcal{M}(\mathbf{g}_{\mathbf{M}}^{(m,p+q)}\boldsymbol{\Upsilon}) = \boldsymbol{\theta}_0 + \mathcal{M}(\boldsymbol{\Upsilon})$ for all $\mathbf{g}_{\mathbf{M}}^{(m,p+q)} \in \mathcal{G}^m_{p+q}(k)$, i.e., iff

$$\mathcal{M}(\mathbf{g}_{\mathbf{M}}^{(m,p+q)}\boldsymbol{\Upsilon}) = \mathcal{M}(\boldsymbol{\Upsilon}) \tag{A.9}$$

and

$$\mathbf{g}_{\mathbf{M}}^{(m,p+q)}\boldsymbol{\theta}_0 - \boldsymbol{\theta}_0 \in \mathcal{M}(\boldsymbol{\Upsilon}), \tag{A.10}$$

for all $\mathbf{g}_{\mathbf{M}}^{(m,p+q)} \in \mathcal{G}^m_{p+q}(k)$. Condition (A.9) means that $\mathcal{M}(\boldsymbol{\Upsilon})$ is invariant under the group $\mathcal{G}^m_{p+q}(k)$, which, in view of Proposition 2.1, holds iff $\boldsymbol{\Upsilon}$ satisfies (2.2). Now, (A.10) is equivalent to $[\mathbf{I}_K - \Pi(\boldsymbol{\Upsilon})](\mathbf{g}_{\mathbf{M}}^{(m,p+q)}\boldsymbol{\theta}_0 - \boldsymbol{\theta}_0) = \mathbf{0}$, i.e., using the fact that $\Pi(\boldsymbol{\Upsilon})$ then necessarily commutes with all $\mathbf{g}_{\mathbf{M}}^{(m,p+q)}$, to

$$\mathbf{g}_{\mathbf{M}}^{(m,p+q)}[\mathbf{I}_K - \Pi(\boldsymbol{\Upsilon})]\boldsymbol{\theta}_0 = [\mathbf{I}_K - \Pi(\boldsymbol{\Upsilon})]\boldsymbol{\theta}_0, \tag{A.11}$$

for all $\mathbf{g}_{\mathbf{M}}^{(m,p+q)} \in \mathcal{G}^m_{p+q}(k)$. Clearly, for any $(p+q)$-vector $\mathbf{w}_{p+q}$, $[\mathbf{I}_K - \Pi(\boldsymbol{\Upsilon})]\boldsymbol{\theta}_0 = (\mathbf{0}', (\mathbf{w}_{p+q} \otimes \mathbf{l}_k)')'$ satisfies (A.11). Conversely, if condition (A.11) holds, $\mathcal{M}([\mathbf{I}_K - \Pi(\boldsymbol{\Upsilon})]\boldsymbol{\theta}_0)$ is invariant under $\mathcal{G}^m_{p+q}(k)$, which, still in view of Proposition 2.1, implies that $[\mathbf{I}_K - \Pi(\boldsymbol{\Upsilon})]\boldsymbol{\theta}_0 = (\mathbf{0}', (\mathbf{w}_{p+q} \otimes \mathbf{l}_k)')'$ for some $(p+q)$-vector $\mathbf{w}_{p+q}$. Condition (2.3) and the proposition follow. □

## REFERENCES


Bilodeau, M., and Brenner, D.(1999). *Theory of Multivariate Statistics.* Springer-Verlag, New York.

Garel, B. and Hallin, H. (1999). Rank-based AR order identification, *J. Amer. Statist. Assoc.* 94, 1357–1371.

Hallin, M., and Paindaveine, D. (2002a). Optimal procedures based on interdirections and pseudo-Mahalanobis ranks for testing multivariate elliptic white noise against ARMA dependence. *Bernoulli* 8, 787–815.

Hallin, M., and Paindaveine, D. (2002b). Rank-based optimal tests of the adequacy of an elliptic VARMA model. Submitted.

Hallin, M., and Paindaveine, D. (2002c). Affine-invariant aligned rank tests for the multivariate general linear model with ARMA errors. Submitted.

Hallin, M., and Puri, M.L. (1994). Aligned rank tests for linear models with autocorrelated error terms, *J. Multivariate Anal.* 50, 175–237.

Oja, H. (1999). Affine invariant multivariate sign and rank tests and corresponding estimates: a review, *Scand. J. Statist.* 26, 319–343.

Pötscher, B.M. (1983). Order estimation in ARMA models by Lagrangian multiplier tests, *Ann. Statist.* 11, 872–885.

Pötscher, B.M. (1985). The behaviour of the Lagrange multiplier test in testing the orders of an ARMA model, *Metrika* 32, 129–150.

Marc Hallin
Département de Mathématique,
I.S.R.O. and E.C.A.R.E.S.
Université Libre de Bruxelles
Campus de la Plaine CP 210
1050 Bruxelles
Belgium
mhallin@ulb.ac.be

Davy Paindaveine
Département de Mathématique,
I.S.R.O. and E.C.A.R.E.S.
Université Libre de Bruxelles
Campus de la Plaine CP 210
1050 Bruxelles
Belgium
dpaindav@ulb.ac.be