Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Multivariate Analysis 93 (2005) 223–237
Journal of Multivariate Analysis
http://www.elsevier.com/locate/jmva

Similar tests for covariance structures in multivariate linear models

G. Forchini*

Department of Econometrics and Business Statistics, Faculty of Business and Economics, Monash University, Clayton VIC 3800, Australia

Received 12 March 2002

Available online 14 May 2004

Abstract

Nyblom (J. Multivariate Anal. 76 (2001) 294) has derived locally best invariant test for the covariance structure in a multivariate linear model. The class of invariant tests obtained by Nyblom [9] does not coincide with the class of similar tests for this testing set-up. This paper extends some of the results of Nyblom [9] by deriving the locally best similar tests for the covariance structure. Moreover, it develops a saddlepoint approximation to optimal weighted average power similar tests (i.e. tests which maximize a weighted average power).

AMS 2000 subject classifications: 62H15

Keywords: Multivariate linear model; Similar tests; Invariant tests; Covariance matrix; Locally best tests; Weighted average power tests

1. Introduction

In a multivariate linear regression model

$$Y = XB + E, \tag{1}$$

where Y and E are $n \times m$, X is $n \times p$ fixed and of rank p, and $n \geqslant p + m$, B is an $p \times m$ matrix of parameters, it is often assumed that the rows of E are independent. Violations of this assumption occur in plausible situations, in

*Fax: +44-01904-433759.
E-mail address: gf7@york.ac.uk.

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2004.04.001

particular in multivariate time series models (for references and examples see [4,9]). It is, thus, important to develop powerful tests capable of detecting these possible violations.

Nyblom [9] has considered the case where E has a multivariate normal distribution with mean matrix 0 and covariance matrix $\Omega(\theta)\otimes\Sigma$, where $\Omega(\theta)$ and Σ are, respectively, a known $n\times n$ matrix, and unknown $m\times m$ matrix, θ is an unknown scalar parameter, and $\Omega(\theta_0)=I_n$. Nyblom [9] analyses tests for $H_0:\theta=\theta_0$ against a suitable alternative of the form $H_1:\theta\in\Theta_1$, by focusing on tests invariant under the group of transformations

$$Y\rightarrow YP+XA$$

for an arbitrary $p\times m$ matrix A and a positive definite $m\times m$ matrix P. Nyblom's [9] main contribution is the derivation of locally most powerful tests in the class of tests invariant under the transformations above for one-tailed alternatives. Nyblom [9] also identifies situations in which the locally best test does not depend on the direction of the departure from the null hypothesis.

This paper extends the results of Nyblom [9] in two different ways. Firstly, we derive the locally best similar test of $H_0:\theta=\theta_0$ versus $H_1:\theta>\theta_0$ and $H_1:\theta<\theta_0$. The resulting test statistics have simple functional forms, and when the sample size n is large the locally best similar tests coincide with the locally best invariant tests of Nyblom [9]. Secondly, in order to allow for alternatives parameterized by a vector $\theta\in\mathbb{R}^q$ (rather than a scalar) and for two-tailed alternative, we discuss the construction of optimal weighted average power similar tests. The contribution of the paper here is in the use of the saddlepoint method to provide computable expressions for such tests in quite general situations.

The remaining part of the paper is organized as follows. Section 2 characterizes the class of similar test for $H_0:\theta=\theta_0$, and gives the density function of the matrix characterizing all similar tests. Sections 3 and 4 derive, respectively, the locally best similar test for one-tailed alternatives and an approximation for the optimal weighted average power test. Section 5 modifies the locally best invariant/similar test and the weighted average power test to make them invariant under permutations of the columns of Y. Section 6 discusses the calculation of the critical values for all tests. Possible applications are given in Section 7. Section 8 discusses Monte Carlo simulations comparing the power of the tests suggested in the paper with the test of Nyblom [9]. Finally the conclusions follow.

2. Characterization of similar tests

In this section, we assume that θ is a q-dimensional vector. The following theorem characterizes the class of similar tests for $H_0:\theta=\theta_0$ in (1).

Theorem 1. *The class of similar tests for* $H_0:\theta=\theta_0$ *against any alternative whatsoever is characterized by the vector* $V=CYT^{-1}$ *where* T *is an* $m\times m$

upper triangular matrix with positive diagonal elements such that $T'T = Y'M_XY$, $M_X = I_n - X(X'X)^{-1}X'$ *and* C *is an* $(n-k) \times n$ *matrix such that* $C'C = M_X$ *and* $CC' = I_{n-k}$.

Proof. Under the null hypothesis the statistics $\hat{B} = (X'X)^{-1}X'Y$ and $S = Y'M_XY$ are sufficient for the nuisance parameters (B, Σ) [8, Theorem 10.1.1]. Moreover, they are (boundedly) complete by Theorem 1, p. 144 of [7]. Let $W = CY$. Theorem A9.8 of [8] guarantees that W can be uniquely written as $W = VT$ where $V = WT^{-1}$ is an $n \times m$ matrix such that $V'V = I_m$ and T is an $m \times m$ upper triangular matrix with positive diagonal elements such that $T'T = S$. □

The class of similar tests coincide with the class of invariant tests under a suitable group of transformations as the following theorem makes clear.

Theorem 2. *Let* $U_{m\times m}$ *be the group of* $m \times m$ *upper triangular matrices with positive diagonal elements, and* $\mathcal{M}_{p\times m}$ *be the group of* $p \times m$ *matrices. The problem of testing* $H_0 : \theta = \theta_0$ *versus* $H_1 : \theta \in \Theta_1$ *is invariant under the group of transformations*

$$G = \{(A, D) : A \in U_{m\times m}, D \in \mathcal{M}_{p\times m}\},$$

with group operation $(A_1, D_1)\ (A_2, D_2) = (A_1A_2, D_1A_2 + D_2)$ *acting on* Y *by*

$$(A, D)Y = YA + XD.$$

Moreover, the maximal invariant under the action of G *is the matrix* V *defined in Theorem* 1.

Proof. The invariance under the action of G can be easily verified. The fact that V is a maximal invariant can be verified in two steps. First the action $Y \to Y + D$ has as a maximal invariant $W = CY$, where C is defined in the statement of the previous theorem. Then the maximal invariant under the action $W \to WA$ is easily found to be $V = WT^{-1}$. □

These results are analogous to those for the linear regression model derived by King and Hillier [6] and are also hinted at in [9]. Nyblom [9] does not investigate these tests further because they are deemed too complicated. However, we will show in Section 3 that the functional forms for the locally best tests are quite simple.

Under the null hypothesis the matrix V is uniformly distributed over the Stiefel manifold $V'V = I_m$. The distribution under the alternative hypothesis is given by the following theorem. Note that when $m = 1$, this reduces to Eq. (2) of King and Hillier [6].

Theorem 3. *Under the alternative hypothesis*

$$pdf(V;\theta) = \frac{\Gamma_m\left(\frac{n-p}{2}\right)}{(2\pi)^{\frac{(n-p+1)m}{2}}\pi^{\frac{(n-p)m}{2}}}|C\Omega(\theta)C'|^{-\frac{m}{2}} \times \left|V'(C\Omega(\theta)C')^{-1}V\right|^{-\frac{n-p}{2}+\frac{m+1}{2}}\prod_{i=1}^{m}|E_i'V'(C\Omega(\theta)C')^{-1}VE_i|^{-1},$$

where E_i is an $m \times i$ matrix such that $E_i = \binom{I_i}{0}$, $i = 1, 2, \ldots, m$.

Proof. Note that the $n-p \times m$ matrix $W = CY \sim N(0, [C\Omega(\theta)C'] \otimes \Sigma)$ so that

$$pdf(W) = (2\pi)^{-\frac{(n-p)m}{2}}|C\Omega(\theta)C'|^{-\frac{m}{2}}|\Sigma|^{-\frac{n-p}{2}} \times etr\left\{-\frac{1}{2}(C\Omega(\theta)C')^{-1}W\Sigma^{-1}W'\right\}$$

transform W to VT. The Jacobian is given by Theorem 2.1.13 in [8] as $(dW) = \prod_{i=1}^{m} t_{ii}^{n-p-i}(dT)(dH)$ where (dH) denotes the unnormalized invariant Haar measure on the Stiefel manifold $V'V = I_m$. Then

$$pdf(V) = (2\pi)^{-\frac{(n-p)m}{2}}|C\Omega(\theta)C'|^{-\frac{m}{2}}|\Sigma|^{-\frac{n-p}{2}} \times \int_{U_{m\times m}} etr\left\{-\frac{1}{2}V'(C\Omega(\theta)C')^{-1}VT\Sigma^{-1}T'\right\}\prod_{i=1}^{m} t_{ii}^{n-p-i}(dT),$$

where $U_{m\times m}$ is, as before, the space of all $m \times m$ upper triangular matrices with positive diagonal elements. This integral can be evaluated using Lemma A.1 in the appendix with $r = n-p$, $A = V'(C\Omega(\theta)C')^{-1}V$ and $D = \Sigma$. $\square$

Knowledge of the density of V under the alternative hypothesis allows us to construct optimal invariant/similar tests for $H_0 : \theta = \theta_0$ against a specific alternative of the form $H_1 : \theta = \theta_1$ where θ_1 is a fixed $q \times 1$ vector. In this case the optimal test has the form: reject H_0 if

$$|V'(C\Omega(\theta_1)C')^{-1}V|^{\frac{n-p}{2}-\frac{m+1}{2}}\prod_{i=1}^{m}|E_i'V'(C\Omega(\theta_1)C')^{-1}VE_i| < k_\alpha,$$

where k_α is a suitable constant such that the size of the test equals α. In the context of tests for on the covariance matrix for the linear regression model, King [5] has argued that with a careful selection of θ_1, this test can be powerful against more general alternatives. This choice, however, is problem dependent. In the rest of the paper, we will consider two classes of similar/invariant tests which are optimal according to two generally accepted criteria, and have general applicability.

3. Locally best invariant/similar tests

When $\theta \in \mathbb{R}$ and no uniformly most powerful test exists, one often appeals to a locally best invariant/similar test. This is a test which maximizes the slope of the power in a neighbourhood of $\theta = \theta_0$, and, for the testing set-up under consideration, it is characterized in the following theorem and corollary.

Theorem 4. *The locally best invariant test for $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$ has the form: reject H_0 if*

$$tr\{V'\Omega^{(1)}VD\} > k_\alpha,$$

where k_α is a suitable constant, D is a diagonal matrix with diagonal elements equal to $(n+m-p+1)/2 - i,\ i = 1, 2, \ldots, m,$ and $d\Omega(\theta) = \Omega^{(1)}(\theta)d\theta,\ \Omega^{(1)}(\theta_0) = \Omega^{(1)}$. The locally best invariant test for $H_0 : \theta = \theta_0$ versus $H_1 : \theta < \theta_0$ has the form: reject H_0 if

$$tr\{V'\Omega^{(1)}VD\} < k'_\alpha,$$

where k'_α is a suitable constant, and $\Omega^{(1)}$ and D are defined as above.

Proof. Since the logarithm is a strictly monotonic increasing function, the class of locally best invariant/similar test for $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$ has the form: reject H_0 if

$$\lim_{\theta \to 0^+} \left\{ \frac{\partial \ln(pdf(V))}{\partial \theta} \right\} > k_\alpha.$$

Let $\Omega^{(1)}$ be defined as above and $\hat{U} = M_X Y$. Then the derivative of $\ln(pdf(V))$ at $\theta = \theta_0$ is

$$\begin{aligned}
\lim_{\theta \to 0^+} \left(\frac{\partial \ln(pdf(V))}{\partial \theta} \right) &= -\frac{m}{2} tr\{C\Omega^{(1)}C'\} + \left(\frac{n-p}{2} - \frac{m+1}{2} \right) tr\{(V'\Omega^{(1)}V)\} \\
&\quad + \sum_{i=1}^{m} tr\{E_i'V'\Omega^{(1)}VE_i\} \\
&= -\frac{m}{2} tr\{C\Omega^{(1)}C'\} \\
&\quad + tr\left\{ V'\Omega^{(1)}V \left(\left(\frac{n-p}{2} - \frac{m+1}{2} \right) I_m + \sum_{i=1}^{m} E_iE_i' \right) \right\}.
\end{aligned}$$

The first part of the theorem follows from here. The second part can be obtained similarly by evaluating $\lim_{\theta \to 0^-} \left(\frac{\partial \ln(pdf(V))}{\partial \theta} \right)$. $\square$

Corollary 1. *For large n, the locally best invariant test $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$ is approximately: Reject H_0 if*

$$tr\{V'\Omega^{(1)}V\} > k_\alpha,$$

where k_α is a suitable constant. Analogously when $H_1 = \theta < \theta_0$ the locally best invariant test is approximately: Reject H_0 if

$$tr\{V'\Omega^{(1)}V\} < k'_\alpha,$$

where k'_α is a suitable constant.

Note that these correspond to the locally best invariant tests derived by Nyblom [9].

4. Optimal weighted average invariant/similar tests

Locally best invariant/similar tests can be constructed only if $\theta \in \mathbb{R}$. This restricts the class of problems to which this criterion can be successfully applied. Moreover, nothing guarantees the optimality of such tests for non-local alternatives. As an alternative to locally best invariant/similar test we consider weighted average invariant/similar tests of the form: Reject $H_0 : \theta = \theta_0$, against the alternative $H_1 : \theta \in \Theta_1$ if

$$T(\pi) = \int_{\Theta_1} \pi(\theta) pdf(V;\theta)\, dV > k_\alpha,$$

where, once more, k_α is a suitable constant, and $\pi(\theta)$ is a suitable weighting function. In the case under consideration $T(\pi)$ has the form

$$\int_{\Theta_1} \pi(\theta)|C\Omega(\theta)C'|^{-\frac{m}{2}}|V'(C\Omega(\theta)C')^{-1}V|^{-\frac{n-p}{2}+\frac{m+1}{2}} \times \prod_{i=1}^{m} |E_i'V'(C\Omega(\theta)C')^{-1}VE_i|^{-1} d\theta \tag{2}$$

and if n is large this integral can be approximated by using the saddlepoint method.

Theorem 5. *Suppose $|V'(C\Omega(\theta)C')^{-1}V|$ has an absolute minimum θ^* in the interior of Θ_1, i.e. θ^* solves*

$$d_{\theta^*} \ln|V'(C\Omega(\theta)C')^{-1}V| = tr\Big\{[V'(C\Omega(\theta)C')^{-1}V]^{-1} d_{\theta^*}[V'(C\Omega(\theta)C')^{-1}V]\Big\} = 0,$$

with

$$d^2_{\theta^*} \ln|V'(C\Omega(\theta)C')^{-1}V| > 0$$

for all $\theta \in \Theta_1$, where d_θ denotes the differential at $\theta = \theta$. Then

$$T(\pi) \sim \left(\frac{2\pi}{n}\right)^{\frac{q}{2}} \pi(\theta^*)|\Omega(\theta^*)|^{-\frac{m}{2}}|X'\Omega^{-1}(\theta^*)X|^{-\frac{m}{2}}|X'X|^{\frac{m}{2}}|V'(C\Omega(\theta^*)C')^{-1}V|^{-\frac{n-p}{2}+\frac{m+1}{2}} \times \prod_{i=1}^{m} |E_i'V'(C\Omega(\theta^*)C')^{-1}VE_i|^{-1}|\Delta(\theta^*)|^{-\frac{1}{2}} + O(n^{-\frac{q}{2}-1}), \tag{3}$$

where $\Delta(\theta^*)$ is the $n \times n$ Hessian matrix of $\ln|V'(C\Omega(\theta)C')^{-1}V|$ at $\theta = \theta^*$.

Proof. This result follows from Theorem 9.5.1 in [8]. The order of convergence is given by Estrada and Kanwal [3]. A simplification is achieved by noting that $|C\Omega(\theta)C'| = |\Omega(\theta)||X'\Omega^{-1}(\theta)X|/|X'X|$. $\square$

Extension to the case of a maximum on the boundary can be done but are much more involved numerically.

Often $\Omega^{-1}(\theta)$ has a simpler structure than $\Omega(\theta)$. This is the case, for example, if $\Omega^{-1}(\theta) = I_n + \sum_{i=2}^m \omega_i(\theta)\Omega_i$, where Ω_i are fixed $n \times n$ symmetric matrices, $\omega_i(\theta)$ are smooth functions of θ, m is a fixed number, and $\omega_i(0) = 0$. Also, the presence of the matrix C in the approximation makes the computations difficult. The following result simplify the calculations in many situations.

Lemma 1. *Let $\hat{E} = M_X Y$, then*

$$|V'(C\Omega(\theta)C')^{-1}V| = |\hat{E}'[\Omega(\theta)^{-1} - \Omega(\theta)^{-1}X(X'\Omega(\theta)^{-1}X)^{-1}X'\Omega(\theta)^{-1}]\hat{E}|/|\hat{E}'\hat{E}|.$$

Moreover, let $||A||_2$ be the spectral norm of the matrix A. Then if for every θ, matrix $||\hat{E}'\Omega^{-1}(\theta)\hat{E}||_2 = ||E'M_X\Omega^{-1}(\theta)M_XE||_2$ diverges for almost every E, as n goes to infinity, then

$$|V'(C\Omega(\theta)C')^{-1}V| \sim |\hat{E}'\Omega(\theta)^{-1}\hat{E}|/|\hat{E}'\hat{E}|$$

for almost every E.

Proof. Let $H = (C, \tilde{X})$, where $\tilde{X} = X(X'X)^{-1/2}$. Then the matrix H is orthogonal and $H'\Omega(\theta)^{-1}H = (H'\Omega(\theta)H)^{-1}$. From the inverse of a partition matrix it follows that

$$\begin{aligned}(C'\Omega(\theta)C)^{-1} &= C'\Omega(\theta)^{-1}C - C'\Omega(\theta)^{-1}\tilde{X}(\tilde{X}'\Omega(\theta)^{-1}\tilde{X})^{-1}\tilde{X}'\Omega(\theta)^{-1}C \\ &= C'[\Omega(\theta)^{-1} - \Omega(\theta)^{-1}X(X'\Omega(\theta)^{-1}X)^{-1}X'\Omega(\theta)^{-1}]C\end{aligned}$$

and the first statement of the lemma follows.

Let $\tilde{X} = \Omega(\theta)^{-1/2}X$ and $\tilde{E} = \Omega(\theta)^{-1/2}\hat{E}$. Then

$$|V'(C\Omega(\theta)C')^{-1}V| = (|\tilde{E}'\tilde{E}|/|\hat{E}'\hat{E}|)|I_m - \tilde{E}'\tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{E}[\tilde{E}'\tilde{E}]^{-1}|.$$

We need to show that $||\tilde{E}'\tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{E}[\tilde{E}'\tilde{E}]^{-1}||_2 \to 0$ when the conditions stated in the lemma hold. Because of the submultiplicative property and the fact that $\tilde{E}'\tilde{E}$ is a square matrix:

$$||\tilde{E}'\tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{E}[\tilde{E}'\tilde{E}]^{-1}||_2 \leqslant ||\tilde{E}'\tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{E}||_2/||\tilde{E}'\tilde{E}||_2.$$

Write $\tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}' = H\left(\begin{smallmatrix} I_k & 0 \\ 0 & 0\end{smallmatrix}\right)H'$ where H is an orthogonal matrix, then

$$||\tilde{E}'\tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{E}[\tilde{E}'\tilde{E}]^{-1}||_2 \leqslant \frac{||\tilde{E}'H\left(\begin{smallmatrix} I_k & 0 \\ 0 & 0\end{smallmatrix}\right)H'\tilde{E}||_2}{||\tilde{E}'HH'\tilde{E}||_2} = \frac{||\sum_{i=1}^k w_i w_i'||_2}{||\sum_{i=1}^n w_i w_i'||_2},$$

where w_i denotes row i of $H'\tilde{E}$. Note that

$$\left\| \sum_{i=1}^{k} w_i w_i' \right\|_2 \leqslant \sum_{i=1}^{k} \|w_i w_i'\|_2 = \sum_{i=1}^{k} w_i' w_i < \infty$$

because this is the norm of a finite sum of positive semidefinite matrices. Since the denominator diverges, the ratio tends to zero as the sample size increases and $\|\tilde{E}'\tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{E}[\tilde{E}'\tilde{E}]^{-1}\|_2 \to 0$. Since the determinant is a continuous function the second part of the lemma follows. $\square$

When the sample size is large, θ^* can often be approximated by the value of θ which minimizes

$$|\hat{E}'\Omega(\theta)^{-1}\hat{E}|/|\hat{E}'\hat{E}| = \left|\hat{E}'\hat{E} + \sum_{i=2}^{m} \omega_i(\theta)\hat{E}'\Omega_i\hat{E}\right|/|\hat{E}'\hat{E}| \tag{4}$$

and this is a simple optimization problem in most cases. Once θ^* has been found we can use it in (3) to calculate $|V'(C\Omega(\theta^*)C')^{-1}V|^{-\frac{n-p}{2}+\frac{m+1}{2}}$. Note that, in general, replacing $|V'(C\Omega(\theta^*)C')^{-1}V|$ with $|\hat{E}'\Omega(\theta)^{-1}\hat{E}|/|\hat{E}'\hat{E}|$ in this expression does not lead to a good approximation for $T(\pi)$ because small approximations errors are amplified by the fact that $|V'(C\Omega(\theta^*)C')^{-1}V|$ is raised to the power $-\frac{n-p}{2}+\frac{m+1}{2}$ in (2). However, (4) can often be used to approximate $|V'(C\Omega(\theta^*)C')^{-1}V|$ in $\Delta(\theta^*)$ since this term does not depend on n.

5. Invariance to the order of the columns of Y

Some people may feel uneasy with the fact that $pdf(V;\theta)$ in Theorem 3, and thus the locally best tests of Theorem 4, the weighted average power test in Eqs. (2) and (3) are not invariant under permutations of the columns of Y (and V) (see [1,9]). The tests proposed above can be easily modified to take this into account. Let $\mu = (\mu_1, \mu_2, \ldots, \mu_m)$ be a permutation of the integers $(1, 2, \ldots, m)$, and let P_μ be an $m \times m$ matrix in which all components are zero except the elements in position (μ_i, i), $i = 1, 2, \ldots, m$ which are 1. Then we can construct the weighted average power test by also averaging over all possible permutations of the columns of V,

$$Ti(\pi) = \sum_{\mu} \int_{\Theta_1} \pi(\theta) pdf(VP_\mu;\theta)/m! dV,$$

where the summation is over all permutations μ of $(1, 2, \ldots, m)$. We can also use the density of V averaged over all permutations of $(1, 2, \ldots, m)$ to derive the locally best tests. Then we have the following results analogous to Theorems 4 and 5, respectively.

Theorem 6. *If θ is a scalar, the locally best invariant tests for $H_0 : \theta = \theta_0$ versus one tailed alternative, which is also invariant to permutations of the columns of V coincide with the locally best tests of Nyblom [9] given in Corollary 1.*

Theorem 7. *Suppose that $\theta \in \mathbb{R}^q$, $|V'(C\Omega(\theta)C')^{-1}V|$ has an absolute minimum θ^* in the interior of Θ_1*

$$\begin{aligned} Ti(\pi) \sim {} & \left(\frac{2\pi}{n}\right)^{\frac{q}{2}} \pi(\theta^*)|\Omega(\theta)|^{-\frac{m}{2}}|X'\Omega^{-1}(\theta)X|^{-\frac{m}{2}}|X'X|^{\frac{m}{2}} \\ & \times |V'(C\Omega(\theta^*)C')^{-1}V|^{-\frac{n-p}{2}+\frac{m+1}{2}}|\Delta(\theta^*)|^{-\frac{1}{2}}(1/m!) \\ & \times \sum_{\mu} \prod_{i=1}^{m} |E_i'P_\mu' V'(C\Omega(\theta^*)C')^{-1}VP_\mu E_i|^{-1} + O\left(n^{-\frac{q}{2}-1}\right), \end{aligned}$$

where $\Delta(\theta^*)$ has been defined in Theorem 5.

6. Calculation of the critical values

The tests considered are similar so the critical values for the tests suggested can be easily calculated by a Monte Carlo simulation:

1. Generate Y by setting $B = 0$ and $E \sim N(0, I_n \otimes I_m)$ in (2).
2. Compute a value of the locally best similar test or the optimal weighted average similar test and save the result.
3. Repeat (1) and (2) M times.
4. Order the saved values of the locally best similar test or the optimal weighted average power similar test and, to estimate the critical value of size α, take the observation in position αM.

The accuracy of the method obviously increases as the number of repetitions M increases. Alternatively, the techniques of Monte Carlo tests can be used to calculate the p-values efficiently (see among others [2]).

7. Applications

The largest class of models to which we can apply the tests developed above consists of multivariate linear models for which there is separability between the n (time, say) and the m (individuals or space, say) dimensions. Harvey [4] call these models seemingly unrelated time series equations. In this case Ω represent the temporal covariance matrix. Nyblom [9] has considered the case where each rows of Y follows the same AR(1) or MA(1) process. The optimal weighted average power test can be easily constructed for more general cases in which the columns of Y follow common higher order AR, MA or ARMA models.

There are situations in which the tests developed above can be applied even though the covariance matrix of the multivariate linear model has a more complex structure than $\Omega(\theta) \otimes \Sigma$. In fact, it is sometimes possible to achieve this by first reducing the problem using further invariance arguments. For example, consider the mixed effects model

$$Y = ZD + VU + W, \tag{5}$$

where Z is a fixed $n \times p_1$ matrix, $W \sim N(0, \Omega(\theta) \otimes \Sigma)$, V is an $n \times p_2$ fixed matrix $(p_1 + p_2 = p)$, $U \sim N(G, I_{p_2} \otimes \Psi)$ and is independent of E, and Σ and Ψ are unknown $m \times m$ matrices and D and G are, respectively, a $p_1 \times m$ and $p_2 \times m$ matrix of unknown parameters. This is equivalent to (1) with $X = (Z, V)$, $B = (D', G')'$ and an error

$$E = VU + W \sim N(0, VV' \otimes \Psi + \Omega(\theta) \otimes \Sigma).$$

The testing problem is invariant under the group of transformations G defined in Theorem 2, so the maximal invariant is given by the matrix V defined in Theorem 1. By noting that the matrix C (defined in Theorem 1) is orthogonal to X, and thus, to V, it is easy to check that under the null hypothesis V is uniformly distributed over the Stiefel manifold, and under the alternative its density is given in Theorem 3.

In a random effects model G equals the zero matrix, and the matrix V does not appear among the regressors. The invariance considered in Theorem 3 is not enough to reduce the problem to the one analysed. However, we can note that the testing problem is also invariant to the transformation $Y \to Y + VU$ where U is any $N(G, I_p \otimes \Psi)$ random matrix, and the problem becomes analogous to the mixed effects model just considered.

8. A numerical example

In order to compare the performance of the locally best invariant/similar and the optimal weighted average power test with that of Nyblom's [9] test we consider a Monte Carlo experiment. The regressors are generated according to the model:

$$x_{ti} = c_{1i} + c_{2i} x_{t-1,i} + \varepsilon_{ti}, \quad i = 2, \ldots, p$$

and

$$x_{t1} = 1,$$

where c_{1i} and c_{2i} are uniformly distributed on the interval $(0, 1)$ and ε_{ti} is normally distributed with mean zero, variance generated from a chi-square distribution with p degrees of freedom, and ε_{0i} are independent χ^2_p. The random variables c_{1i} and c_{2i} are independent of each other and of ε_{ti}, ε_{0i} and the variance of ε_{ti}. After generating the regressors we construct model (1) and $E \sim N(0, \Omega(\theta) \otimes \Sigma)$. Because of the invariance arguments we can set $B = 0$ and $\Sigma = I_m$ without loss of generality.

We consider the case where the rows of Y form an AR(1) process with zero start-up value. The critical values are generated using 50,000 replications and the power is calculated using 10,000 replications for any value of θ under the alternative.

Two phenomena are always observed. Firstly, the power function of the locally best invariant/similar tests is indistinguishable from the power function of the Nyblom's [9] test. The power functions of the weighted average power tests $T(\pi)$ and $Ti(\pi)$ are basically equal.

Fig. 1 shows a typical result when testing for $H_0 : \theta = 0$ versus $\theta > 0$ (with $n = 0$, $p = 4$ and $m = 3$). The locally best invariant/similar test (LBI) and Nyblom's test have identical power functions. The power functions of $T(1)$ and $Ti(1)$ are also exactly alike. They are below the power functions of the locally best similar/invariant test and Nyblom's test (note that tests based on $T(1)$ and $Ti(1)$, with $\pi(\theta) = 1$ for all θ, are in fact two-tailed tests). By taking the weighting function as the indicator function $I_{\{\theta>0\}}$ which is equal to 1 for $\theta > 0$ and zero everywhere else, the resulting tests $T(I_{\{\theta>0\}})$ and $Ti(I_{\{\theta>0\}})$ are indistinguishable from the locally best tests (see Fig. 2) even ignoring the problem that the maximum occurs at the boundary in some cases. These conclusions hold for different values of n, m and p.

When testing for $H_0 : \theta = 1$ versus $\theta < 1$ the situation is very different. A typical result is reported in Fig. 3 ($n = 30$ and $p = 4$ and $m = 3$). In this case the weighted average power tests $T(1)$ and $Ti(1)$ seem to perform much better than the locally best tests even though the weighted average power tests are in fact two-tailed tests. Again, this conclusion holds for different values of n, m and p.

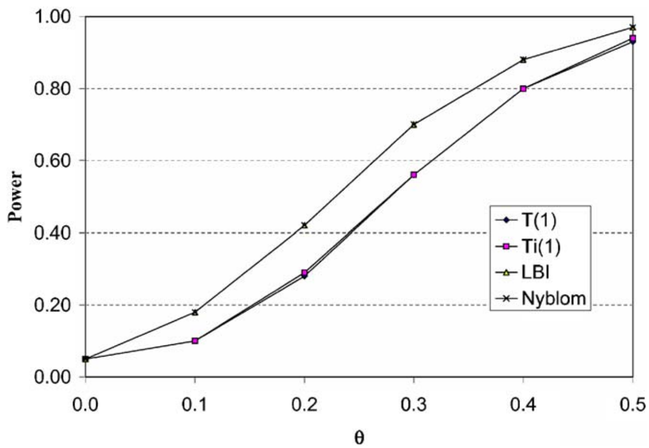

Fig. 1. Typical pattern for the power functions for the weighted average power ($T(1)$), the weighted average power test with invariance under permutations of the columns of V ($Ti(1)$), the locally best invariant/similar (LBI) and Nyblom's (Nyblom) tests for $H_0 : \theta = 0$ versus $H_1 : \theta > 0$ at the 5% significance level for $n = 30$, $m = 3$ and $p = 4$.

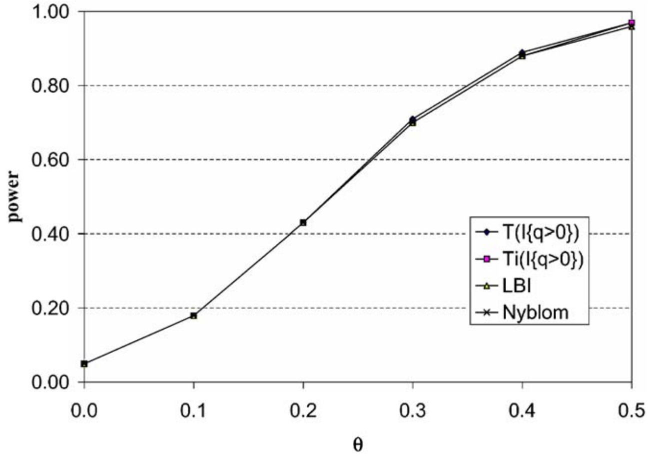

Fig. 2. Typical pattern for the power functions for the weighted average power $(T(I_{\{\theta>0\}}))$, the weighted average power test with invariance under permutations of the columns of V $(Ti(I_{\{\theta>0\}}))$, the locally best invariant/similar (LBI) and Nyblom's (Nyblom) tests for $H_0:\theta=0$ versus $H_1:\theta>0$ at the 5% significance level for $n=30$, $m=3$ and $p=4$.

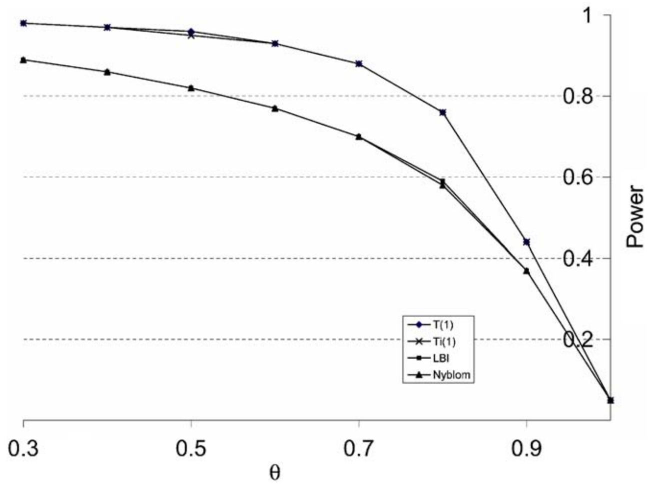

Fig. 3. Typical pattern for the power functions for the weighted average power $(T(1))$, the weighted average power test with invariance under permutations of the columns of V $(Ti(1))$, the locally best invariant/similar (LBI) and Nyblom's (Nyblom) tests for $H_0:\theta=1$ versus $H_1:\theta<1$ at the 5% significance level for $n=30$, $m=3$ and $p=4$.

9. Conclusions

When constructing tests for the covariance matrix of the multivariate linear regression model, it is natural to simplify the problem using invariance or similarity arguments, and to construct locally best invariant/similar tests. These can be used to test one-sided alternatives when the covariance matrix is parameterized in terms of a scalar parameter. For more complex models, optimal weighted average power similar/invariant tests can be employed, but often require the evaluation of complicated multivariate integrals, which limits their practical use. However, in the set-up under consideration, a saddlepoint technique can be used to approximate the test statistics on which optimal weighted average power similar/invariant tests are based. Simulations suggests that the weighted average power tests perform at least as well as locally best invariant/similar tests, and in some situations they perform significantly better.

Appendix A

Lemma A.1. *Let $U_{m\times m}$ be the space of all $m \times m$ upper triangular matrices with positive diagonal elements, A be $m \times m$, and D be $m \times m$, $m \leqslant r$. Then*

$$\int_{U_{m\times m}} etr\{-T'ATD^{-1}\} \prod_{i=1}^{m} t_{ii}^{r-i}(dT) = 2^{-m}\Gamma_m\left(\frac{r}{2}\right)|D|^{\frac{r}{2}}|A|^{-\frac{r}{2}+\frac{m+1}{2}} \prod_{i=1}^{m} |A_i|^{-1}$$

with

$$A_i = \begin{pmatrix} A_{i-1} & a_i \\ a_i' & a_{ii} \end{pmatrix}.$$

Proof. Note that

$$\int_{\mathbb{R}^{r\times m}} etr\{-Y'YD^{-1}\}(dY) = \frac{2^m \pi^{\frac{mr}{2}}}{\Gamma_m\left(\frac{r}{2}\right)} \int_{U_{m\times m}} etr\{-S'SD^{-1}\} \prod_{i=1}^{m} s_{ii}^{r-i}(dS)$$

and also that

$$\begin{aligned}\int_{\mathbb{R}^{r\times m}} etr\{-Y'YD^{-1}\}(dY) &= \frac{\pi^{\frac{mr}{2}}}{\Gamma_m\left(\frac{r}{2}\right)} \int_{S>0} etr\{-SD^{-1}\}|S|^{\frac{r-m-1}{2}}(dS) \\ &= \pi^{\frac{mr}{2}}|D|^{\frac{r}{2}}\end{aligned}$$

So

$$\int_{U_{m\times m}} etr\{-S'SD^{-1}\} \prod_{i=1}^{m} t_{ii}^{r-i}(dT) = 2^{-m}\Gamma_m\left(\tfrac{r}{2}\right)|D|^{\frac{r}{2}}$$

Let $S = BT$, and $B'B = A$, the Jacobian is $(dS) = \prod_{i=1}^{m} b_{ii}^{m+1-i}(dT)$

$$\int_{U_{m\times m}} etr\{-T'ATD^{-1}\} \prod_{i=1}^{m} (b_{ii}t_{ii})^{r-i} \prod_{i=1}^{m} b_{ii}^{m+1-i}(dT) = 2^{-m}\Gamma_m\left(\frac{r}{2}\right)|D|^{\frac{r}{2}}$$

$$\prod_{i=1}^{m} b_{ii}^{r+m+1-2i} \int_{U_{m\times m}} etr\{-T'ATD^{-1}\} \prod_{i=1}^{m} t_{ii}^{r-i}(dT) = 2^{-m}\Gamma_m\left(\frac{r}{2}\right)|D|^{\frac{r}{2}}$$

$$\int_{U_{m\times m}} etr\{-T'ATD^{-1}\} \prod_{i=1}^{m} t_{ii}^{r-i}(dT) = 2^{-m}\Gamma_m\left(\frac{r}{2}\right)|D|^{\frac{r}{2}} \prod_{i=1}^{m} b_{ii}^{-(r+m+1-2i)}.$$

Let

$$A_i = \begin{pmatrix} A_{i-1} & a_i \\ a_i' & a_{ii} \end{pmatrix}$$

If we write A_i as $A_i = B'_{(i)}B_{(i)}$ where $B_{(i)}$ is the $i \times i$ upper triangular with positive diagonal elements and $B_{(m)} = B$, then

$$\begin{aligned} b_{ii}^2 &= a_{ii} - a_i' A_{i-1}^{-1} a_i \\ &= |A_{i-1}|^{-1}|A_i|. \end{aligned}$$

So

$$\begin{aligned} \prod_{i=1}^{m} b_{ii}^{2i} &= |A_1|(|A_1|^{-1}|A_2|)^2(|A_2|^{-1}|A_3|)^3 \ldots (|A_{m-1}|^{-1}|A_m|)^m \\ &= |A_1|^{-1}|A_2|^{-1} \ldots |A_m|^{-1}|A_m|^m \\ &= |A_m|^{m+1} \prod_{i=1}^{m} |A_i|^{-1}. \end{aligned}$$

Note that $A_m = A$. Since we have

$$\begin{aligned} \prod_{i=1}^{m} b_{ii}^{-(r+m+1-2i)} &= \prod_{i=1}^{m} b_{ii}^{-(r+m+1)} \prod_{i=1}^{m} b_{ii}^{2i} \\ &= \left[\prod_{i=1}^{m} b_{ii}\right]^{-(r+m+1)} \prod_{i=1}^{m} b_{ii}^{2i} \\ &= |A|^{-\frac{r+m+1}{2}} \left[|A|^{m+1} \prod_{i=1}^{m} |A_i|^{-1}\right] \\ &= |A|^{-\frac{r}{2}+\frac{m+1}{2}} \prod_{i=1}^{m} |A_i|^{-1}. \end{aligned}$$

The lemma is proved. □

References

[1] R. Das, B.K. Sihna, Optimum invariant tests for random MANOVA models, Canad. J. Statist. 16 (1988) 193–199.
[2] J.-M. Dufour, L. Khalaf, Monte Carlo test methods in econometrics, in: B. Baltagi (Ed.), Companion to Theoretical Econometrics, Blackwell, Oxford, UK, 2001, pp. 494–519.
[3] R. Estrada, R.P. Kanwal, The asymptotic expansion of certain multi-dimensional generalized functions, J. Math. Anal. Appl. 163 (1992) 264–283.
[4] A. Harvey, Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press, Cambridge, UK, 1989.
[5] M. King, Towards a theory of point optimal testing. With comments and a reply by the author, Econometric Rev. 6 (1988) 169–255.
[6] M. King, G.H. Hillier, Locally best invariant tests of the error covariance matrix of the linear regression model, J. Roy. Statist. Soc. Ser. B 47 (1985) 98–102.
[7] E.L. Lehmann, Testing Statistical Hypotheses, Wiley, New York, 1986.
[8] R.J. Muirhead, Aspects of Multivariate Statistical Theory, Wiley, New York, 1982.
[9] J. Nyblom, Invariant tests for covariance structures in multivariate linear model, J. Multivariate Anal. 76 (2001) 294–315.