On estimating the rank of semidefinite matrix*

Stephen G. Donald† University of Texas at Austin

Natércia Fortuna‡ CEF.UP, Universidade do Porto

Vladas Pipiras§ University of North Carolina at Chapel Hill

October 5, 2014

Abstract

This paper concerns the problem of rank estimation in situations where the object of interest is a semidefinite matrix using an estimator that is either indefinite or semidefinite and satisfies an asymptotic normality condition. Several rank tests are examined, based on either available approaches or basic new results. A number of related issues are discussed such as the choice of matrix estimators and rank tests based on finer assumptions than those of the asymptotic normality of matrix estimators. Several examples where rank estimation for a semidefinite matrix is of interest are studied and serve as a guide.

Keywords: matrix rank, symmetric matrix, indefinite and semidefinite estimators, eigenvalues, matrix decompositions, estimation, asymptotic normality.

JEL classification: C12, C13.

1 Introduction

Inference regarding the rank $\mathrm{rk}\{M\}$ of an unknown, real-valued matrix M is an important and well-studied problem in Econometrics and Statistics. A number of rank tests have been proposed including the LDU (Lower-Diagonal-Upper triangular decomposition) test of Gill and Lewbel (1992) and Cragg and Donald (1996), the Minimum Chi-Squared (MINCHI2) test of Cragg and Donald (1997), the SVD (Singular Value Decomposition) tests in Ratsimalahelo (2002) and Kleibergen and Paap (2006), and the characteristic root test of Robin and Smith (2000). The problem of rank estimation is reviewed in Camba-Mendez and Kapetanios (2009) where several applications (for example, IV modeling, demand systems, cointegration) are also discussed.

*The paper expands, revises and corrects the authors' working paper "On rank estimation in semidefinite matrices", CEF.UP Working Paper 2 (2010), Faculdade de Economia do Porto, Porto, Portugal.

†Department of Economics, University of Texas at Austin, 1 University Station, Austin, TX 78712, USA, *donald@eco.utexas.edu*

‡Faculdade de Economia, Universidade do Porto, Rua Dr. Roberto Frias, 4200-464 Porto, Portugal, *nfortuna@fep.up.pt*, CEF.UP - Centre for Economics and Finance at University of Porto. This research has been financed by Portuguese Public Funds through FCT (Fundação para a Ciência e a Tecnologia) in the framework of the project PEst-OE/EGE/UI4105/2014. Financial support from FEDER and FCT (research grant PTDC/EGE-ECO/122820/2010) is also gratefully acknowledged.

§Dept. of Statistics and Operations Research, UNC at Chapel Hill, CB#3260, Hanes Hall, Chapel Hill, NC 27599, USA, *pipiras@email.unc.edu*. The research work is supported in part by NSA Grant H98230-13-1-0220.

A standard assumption in all these rank tests is that one has an estimator $\widehat{M}$ for M such that

$$\sqrt{N}(\mathrm{vec}(\widehat{M}) - \mathrm{vec}(M)) \xrightarrow{d} \mathcal{N}(0, W), \tag{1.1}$$

where N is the sample size or any other relevant parameter such that $N \to \infty$, where $\xrightarrow{d}$ denotes convergence in distribution (convergence in probability will be denoted by $\xrightarrow{p}$) and vec is a standard matrix operation. It is standard to test the null hypothesis $H_0 : \mathrm{rk}\{M\} \le k$ (or $H_0 : \mathrm{rk}\{M\} = k$) against the alternative $H_1 : \mathrm{rk}\{M\} > k$. Using tests of these hypotheses various methods, such as sequential testing, have been proposed to estimate $\mathrm{rk}\{M\}$. The limiting covariance matrix W in (1.1) is often assumed to be nonsingular, though some departures from this assumption have been considered in Robin and Smith (2000), and Camba-Mendez and Kapetanios (2009).

In this paper, we are interested in tests for the rank of a $p \times p$ symmetric semidefinite matrix M. Without loss of generality, we focus on the case where M is *positive* semidefinite. Since M is symmetric, we shall focus on symmetric estimators $\widehat{M}$. In this case, it is convenient to write (1.1) as

$$\sqrt{N}(\mathrm{vech}(\widehat{M}) - \mathrm{vech}(M)) \xrightarrow{d} \mathcal{N}(0, W_0), \tag{1.2}$$

where vech is a standard operation on symmetric matrices, vectorizing the upper-triangular part of a symmetric matrix. Two different cases need to be distinguished regarding (1.2), namely those of:

$$\begin{aligned} \text{Case 1} &: \ \textit{indefinite} \text{ matrix estimators } \widehat{M}, \\ \text{Case 2} &: \ \textit{semidefinite} \text{ matrix estimators } \widehat{M}. \end{aligned} \tag{1.3}$$

In Case 1, one may still often assume that (1.2) holds with *nonsingular* matrix W_0. This is the case considered in Donald, Fortuna and Pipiras (2007). In Case 2, under rank deficiency for M, the matrix W_0 in (1.2) is necessarily *singular* (see Proposition 2.1 in Donald et al. (2007)).

In this paper, we focus on Case 2 in (1.3). After preliminary notation and assumptions are presented in Section 2, we discuss several examples in Section 3. In some examples, we show that for certain rank hypotheses one can determine exactly from the data the rank and hence rank testing is not really relevant for such cases. In other cases, this is not possible and one must adapt existing tests or derive new tests. Also noted in the examples is the fact that there are some cases where the estimator itself is positive semidefinite and other cases where one can use an indefinite matrix estimator. Our discussion of testing first considers the case of using a semidefinite matrix estimator. For this case we consider, in Section 4, the possibility of adapting existing tests and show that while some can be adjusted to deliver proper tests, for others there are difficulties that we discuss. In Section 5, we consider the possibility of using (scaled) eigenvalues of the estimated matrix directly and show that while they have the potential to distinguish between the hypotheses, it does not seem to be possible to properly control the size of the test. In Section 6, we add an additional assumption that opens up the possibility of properly controlling the size of the eigenvalue based test, and discuss this in the context of examples. Section 7 considers tests based on an indefinite estimator as considered earlier in Donald et al. (2007). Section 8 contains a simulation study that sheds light on properties of various tests in small samples while concluding remarks are contained in Section 9. Proofs for the results are in the Appendix.

2 Notation and other preliminaries

The following notation will be used throughout the paper. As in Section 1, M is an unknown, $p \times p$, symmetric, semidefinite matrix with real-valued entries. Suppose without loss of generality that

M is *positive* semidefinite. Its estimator $\widehat{M} = \widehat{M}(N)$, where $N \to \infty$ is the sample size or other parameter, is symmetric. We shall focus on the case where $\widehat{M}$ is positive semidefinite. $\widehat{M}$ satisfies either (1.1) or (1.2), that is,

$$\sqrt{N}(\mathrm{vec}(\widehat{M}) - \mathrm{vec}(M)) \xrightarrow{d} \mathcal{N}(0, W), \tag{2.1}$$

$$\sqrt{N}(\mathrm{vech}(\widehat{M}) - \mathrm{vech}(M)) \xrightarrow{d} \mathcal{N}(0, W_0). \tag{2.2}$$

The relation between W and W_0 is $W = D_p W_0 D_p'$, where D_p is the $p^2 \times p(p+1)/2$ duplication matrix (see, e.g., Magnus and Neudecker (1999), pp. 48–53). We shall also write (2.1) or (2.2) as

$$\sqrt{N}(\widehat{M} - M) \xrightarrow{d} \mathcal{Y}, \tag{2.3}$$

where $\mathcal{Y}$ is a normal (Gaussian) matrix. The rank of any matrix A is denoted by $\mathrm{rk}\{A\}$, its trace by $\mathrm{tr}(A)$ (if A is quadratic) and its transpose by A'.

To accommodate the example of spectral density matrices, we shall also consider separately the case of Hermitian semidefinite matrices M with complex-valued entries. In this case, (2.1) is replaced by

$$\sqrt{N}(\mathrm{vec}(\widehat{M}) - \mathrm{vec}(M)) \xrightarrow{d} \mathcal{N}^c(0, W^c), \tag{2.4}$$

where $\mathcal{N}^c$ indicates complex normal. By definition, (2.4) is equivalent to

$$\sqrt{N} \begin{pmatrix} \mathrm{vec}(\Re\widehat{M}) - \mathrm{vec}(\Re M) \\ \mathrm{vec}(\Im\widehat{M}) - \mathrm{vec}(\Im M) \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{2} \begin{pmatrix} \Re W^c & \Im W^c \\ \Im W^c & \Re W^c \end{pmatrix}\right), \tag{2.5}$$

where $\Re$ and $\Im$ stand for the real and imaginary parts, respectively. The notation A^* will stand for the Hermitian transpose of a matrix A with complex-valued entries. For later reference, we make the distinction between the real and complex cases explicitly as

$$\begin{aligned} &\text{Case R} \quad : \quad \text{entries of } M \text{ are real-valued,} \\ &\text{Case C} \quad : \quad \text{entries of } M \text{ are complex-valued.} \end{aligned} \tag{2.6}$$

Finally, for later use, let $Q = (Q_1\ Q_2)$ be an orthogonal matrix in Case R of (2.6) (and unitary matrix in Case C of (2.6)) such that

$$Q^* M Q = \begin{pmatrix} Q_1^* \\ Q_2^* \end{pmatrix} M (Q_1\ Q_2) = \mathrm{diag}\{\upsilon_1, \upsilon_2, \dots, \upsilon_p\}, \tag{2.7}$$

where

$$0 = \upsilon_1 = \dots = \upsilon_{p-r} < \upsilon_{p-r+1} \le \dots \le \upsilon_p \tag{2.8}$$

are the ordered eigenvalues of M, and $\mathrm{rk}\{M\} = r$. The submatrix Q_1 in (2.7) is $p \times (p-r)$, and Q_2 is $p \times r$. Note, in particular, that $Q_1^* M Q_1 = 0$.

3 Examples concerning ranks of semidefinite matrices

In this section, we gather a number of examples where rank estimation in symmetric, semidefinite matrices is of interest. We also make a few remarks on the situations where rank estimation is not appropriate, even if semidefinite matrices and their estimators are involved.

Example 3.1 (Linear regression with heteroscedastic error terms.) Consider a linear regression model

$$y_i = \beta x_i + \epsilon_i, \quad i = 1, \dots, n, \tag{3.1}$$

where y_i is $p \times 1$, β is $p \times q$ and unknown, (x_i, ϵ_i) are i.i.d. vectors with x_i being $q \times 1$ and ϵ_i being $p \times 1$. Suppose that $E(\epsilon_i | x_i) = 0$ and

$$E(\epsilon_i \epsilon_i' | x_i) = p(x_i)^{-1} \Sigma(x_i), \tag{3.2}$$

where $\Sigma(x)$ is a $p \times p$, conditional covariance matrix depending on x and $p(x) > 0$ is a density of x_i.[1] The matrix $\Sigma(x)$ is symmetric, positive semidefinite, and could be estimated through the symmetric, *semidefinite* matrix estimator

$$\widehat{\Sigma}(x) = \frac{1}{n} \sum_{k=1}^{n} (y_k - \widehat{\beta} x_k)(y_k - \widehat{\beta} x_k)' K_h(x - x_k), \tag{3.3}$$

where $\widehat{\beta}$ is the least squares estimator of β and $K_h(x) = K(x/h)/h^q$ is a scaled kernel function, where $h > 0$ is a bandwidth. It is necessary to assume that $\Sigma(x)$, $p(x)$ are sufficiently smooth at x (see, for example, Pagan and Ullah (1999)).

The following basic result shows that, under suitable assumptions, the estimator $\widehat{\Sigma}(x)$ is asymptotically normal for $\Sigma(x)$. It is proved in Appendix A. Here are some assumptions and notation used. Set $x^j = x_1^{j_1} \dots x_q^{j_q}$ and $|j| = j_1 + \dots + j_q$ for $j = (j_1, \dots, j_q) \in (\mathbb{N} \cup \{0\})^q$ and $x = (x_1, \dots, x_q)' \in \mathbb{R}^q$. We suppose that K is a kernel of order $r \geq 2$ in the sense that $\int K(x)dx = 1$, $\int x^j K(x) dx = 0$, for $1 \leq |j| < r$, K is symmetric and has a bounded support. Suppose (x_i, ϵ_i) are i.i.d. vectors supported on $\mathcal{H} := \mathcal{H}_x \times \mathcal{H}_\epsilon := (a_x, b_x) \times (a_\epsilon, b_\epsilon)$ and having density $p(x, \epsilon) > 0$ on $\mathcal{H}$. Let $C^m(\mathcal{H}_x)$ be the set of functions on $\mathcal{H}_x$ whose mth derivative is continuous on $\mathcal{H}_x$, and $\|K\|_2^2 = \int K(x)^2 dx$.

Proposition 3.1 *Consider the model (3.1)-(3.2) and suppose the assumptions above. Suppose also that* $\Sigma(x) \in C^r(\mathcal{H}_x)$ *and* $p(x)$, $E(\epsilon_k \epsilon_k' \otimes \epsilon_k \epsilon_k' | x_k = x)$, $E(|\epsilon_k \epsilon_k'|^{2+\delta} | x_k = x) \in C^0(\mathcal{H}_x)$ *with* $\delta > 0$ *and that* $E x_i x_i'$ *is invertible. Then, as* $n \to \infty$, $nh^q \to \infty$, $nh^{q+2r} \to 0$, $h \to 0$, *for* $x \in \mathcal{H}_x$,

$$\sqrt{nh^q}(\mathrm{vec}(\widehat{\Sigma}(x)) - \mathrm{vec}(\Sigma(x))) \xrightarrow{d} \mathcal{N}(0, W(x)), \tag{3.4}$$

where

$$W(x) = \|K\|_2^2 p(x) E(\epsilon_k \epsilon_k' \otimes \epsilon_k \epsilon_k' | x_k = x). \tag{3.5}$$

Under the assumptions of Proposition 3.1, when $\Sigma(x) = 0$ and in the case $p = 1$, the limiting covariance matrix is necessarily $W(x) = 0$. This follows directly from the fact that $\widehat{\Sigma}(x) \geq 0$ or also since $\Sigma(x) = 0$ implies $\epsilon_k = 0$ given $x_k = x$. In the general case $p \geq 1$ and when $\mathrm{rk}\{\Sigma(x)\} = r < p$, there is orthogonal Q such that $(Q\epsilon_k)_l = 0$ given $x_k = x$, where $(z)_l$ indicates the lth component of a vector z, $l = 1, \dots, p - r$. The matrix $W(x)$ has the same rank as

$$(Q \otimes Q) W(x) (Q' \otimes Q') = \|K\|_2^2 p(x) E(Q\epsilon_k (Q\epsilon_k)' \otimes Q\epsilon_k (Q\epsilon_k)' | x_k = x).$$

This yields

$$\mathrm{rk}\{W(x)\} \leq r^2 = \mathrm{rk}\{\Sigma(x)\}^2, \tag{3.6}$$

by using the facts that $Q\epsilon_k (Q\epsilon_k)'$ given $x_k = x$, has $p - r$ zero rows and hence rank $\leq r$, and that $\mathrm{rk}\{A \otimes B\} = \mathrm{rk}\{A\}\mathrm{rk}\{B\}$.

[1] The notation $p(x)$ or $\widehat{p}(x)$ should not be confused with p in the dimension $p \times p$ of the matrix M.

Remark 3.1 In Example 3.1, it was essential to assume the dependence of the matrix $\Sigma(x)$ on x and then estimating its rank for fixed x. In the case where ϵ_i's are independent of x_i's and hence $\Sigma(x)p(x)^{-1} = E\epsilon_i\epsilon_i'$ does not depend on x, estimating the rank of the constant matrix $E\epsilon_i\epsilon_i'$ is not appropriate. To see why, note that if Σ has less than full rank, then there is an orthogonal matrix Q such that

$$EQ\epsilon_i(Q\epsilon_i)' = \mathrm{diag}\{0,\dots,0,\alpha_{p-r+1},\dots,\alpha_p\}, \tag{3.7}$$

where $\alpha_k > 0$ and $r = \mathrm{rk}\{\Sigma\}$. But this means that, for some Q and β,

$$(Q(y_i - \beta x_i))_k = 0 \quad \text{a.s.}, \quad k = 1,\dots,p-r, \tag{3.8}$$

where $(z)_k$ indicates the kth component of a vector z. The exact linear relationship between y and x in (3.8) could, in principle, be first checked with data and if found, eliminated by reducing p. Similar observations can also be found, for example, in connection to principal components (where it is not appropriate to consider principal components with zero variance). See, for example, p. 27 in Jolliffe (2002).

Example 3.2 (Spectral density matrices.) Let

$$\Sigma(w) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} Ex_kx_0'e^{-iwk}, \quad w \in (-\pi,\pi], \tag{3.9}$$

be the spectral density matrix of a stationary, say zero mean, p-variate time series $\{x_k\}_{k\in\mathbb{Z}}$. Under mild assumptions (e.g. $\sum_k |Ex_kx_0'| < \infty$), the spectral density matrix is Hermitian, positive semidefinite and with complex-valued entries in general (see, for example, Brillinger (1975), Hannan (1970)). One interesting problem, considered by Camba-Mendez and Kapetanios (2005), is to test for

$$\mathrm{rk}\{\Sigma(w)\} \tag{3.10}$$

for *fixed* w. An important application of this is to cointegration. Recall that $\mathrm{rk}\{\Sigma(w)\}$ at $w = 0$ is p minus the cointegration rank of the p-variate series $\{y_k\}_{k\in\mathbb{Z}}$, where $\Sigma(w)$ is defined from $x_k = y_k - y_{k-1}$ (see, for example, Hayashi (2000), Maddala and Kim (1998)).

The spectral density matrix $\Sigma(w)$ could be estimated through a smoothed periodogram (and the *semidefinite* matrix estimator)

$$\widehat{\Sigma}(w) = \frac{1}{2m+1}\sum_{k=-m}^{m}\bar{\Sigma}\left(w + \frac{2\pi k}{n}\right) \tag{3.11}$$

where n is the sample size in $\{x_1,\dots,x_n\}$, $m = m(n)$, and as usual,

$$\bar{\Sigma}(w) = \frac{1}{2\pi}\sum_{k=-(n-1)}^{n-1}\widehat{\Gamma}_k e^{-ikw}, \quad \widehat{\Gamma}_k = \frac{1}{n}\sum_{t=1}^{n-|k|} x_t x_{t+k}'. \tag{3.12}$$

Here is the basic asymptotic normality result of the type (2.4), shown in Theorem 11 of Hannan (1970), p. 289. Suppose $\{x_k\}_{k\in\mathbb{Z}}$ is a stationary, zero mean time series such that

$$x_k = \sum_{j=-\infty}^{\infty} A_j\epsilon_{k-j}, \quad k \in \mathbb{Z}, \tag{3.13}$$

where ϵ_j are i.i.d. vectors with $E\epsilon_j = 0$, $E|\epsilon_j|^4 < \infty$, $\sum_{j=-\infty}^{\infty} |A_j| < \infty$, and

$$\lim_{n\to\infty} \sup_w m^{1/2} |\Sigma(w) - E\widehat{\Sigma}(w)| = 0. \tag{3.14}$$

Proposition 3.2 *With the above notation and assumptions, as $n, m \to \infty$, $m/n \to 0$,*

$$(2m)^{1/2}(\mathrm{vec}(\widehat{\Sigma}(w)) - \mathrm{vec}(\Sigma(w))) \xrightarrow{d} \mathcal{N}^c(0, W^c(w)), \tag{3.15}$$

where the asymptotic covariance between $\widehat{\Sigma}(w)_{ij}$ and $\widehat{\Sigma}(w)_{uv}$ is given by

$$W^c(w) = \begin{cases} 2\Sigma(w)_{iu}\Sigma(w)_{vj}, & \text{if} \quad w \neq 0, \pm\pi, \\ 2(\Sigma(w)_{iu}\Sigma(w)_{vj} + \Sigma(w)_{iv}\Sigma(w)_{uj}), & \text{if} \quad w = 0, \pm\pi. \end{cases} \tag{3.16}$$

For example, when $p = 2$, (3.16) becomes

$$W^c(w) = 2 \begin{pmatrix} \Sigma(w)_{11}\Sigma(w)_{11} & \Sigma(w)_{11}\Sigma(w)_{21} & \Sigma(w)_{12}\Sigma(w)_{11} & \Sigma(w)_{12}\Sigma(w)_{21} \\ \Sigma(w)_{21}\Sigma(w)_{11} & \Sigma(w)_{21}\Sigma(w)_{21} & \Sigma(w)_{22}\Sigma(w)_{11} & \Sigma(w)_{22}\Sigma(w)_{21} \\ \Sigma(w)_{11}\Sigma(w)_{12} & \Sigma(w)_{11}\Sigma(w)_{22} & \Sigma(w)_{12}\Sigma(w)_{12} & \Sigma(w)_{12}\Sigma(w)_{22} \\ \Sigma(w)_{21}\Sigma(w)_{12} & \Sigma(w)_{21}\Sigma(w)_{22} & \Sigma(w)_{22}\Sigma(w)_{12} & \Sigma(w)_{22}\Sigma(w)_{22} \end{pmatrix} =: 2W_1^c(w), \tag{3.17}$$

when $w = 0, \pm\pi$, and

$$W^c(w) = 2(W_1^c(w) + \Sigma(w) \otimes \Sigma(w)), \tag{3.18}$$

when $w \neq 0, \pm\pi$. It can be seen from these relations that $\mathrm{rk}\{W^c(w)\}$ depends on $\Sigma(w)$.

Remark 3.2 Similarly to Remark 3.1, it would not be appropriate to test in Example 3.2 for

$$r = \sup_w \mathrm{rk}\{\Sigma(w)\} < p. \tag{3.19}$$

Indeed, if (3.19) holds, then there is a unitary matrix Q such that $Q\Sigma(w)Q^* = \mathrm{diag}\{0, \dots, 0, \alpha_{p-r+1}, \dots, \alpha_p\}$ for all w, where $\alpha_k > 0$. Then, $\sum_{k=-\infty}^{\infty} E(Qx_0)(Qx_k)^* e^{-iwk} = \mathrm{diag}\{0, \dots, 0, \alpha_{p-r+1}, \dots, \alpha_p\}$, which yields, in particular, that

$$\sum_{k=-\infty}^{\infty} E(Qx_0)_l (Qx_k)_l^* e^{-iwk} = 0, \quad l = 1, \dots, p - r, \tag{3.20}$$

(as before, $(z)_l$ indicates the lth component of z) or that

$$(Qx_k)_l = 0 \quad \text{a.s.,} \quad \text{all} \quad k \in \mathbb{Z}, l = 1, \dots, p - r. \tag{3.21}$$

This type of exact linear relationship can, in principle, be seen from the data and eliminated without loss of generality.

Example 3.3 (Semiparametric multiple-index model.) Donkers and Schafgans (2008) consider the semiparametric multiple-index model

$$g(x) := E(y|x) = H(x'\beta_1, \dots, x'\beta_p), \tag{3.22}$$

where dependent variable $y \in \mathbb{R}$ (the more general case of $y \in \mathbb{R}^s$ could also be considered) and explanatory variables $x \in \mathbb{R}^l$. The function H is unknown but sufficiently smooth, $\beta_1, \dots, \beta_p$ are unknown parameters and p is an unknown number of indices. Let n be the sample size of the data $y_1, \dots, y_n$ and $x_1, \dots, x_n$.

Of interest in our context is the $l \times l$ matrix

$$M = E\left(w(x)\frac{\partial g(x)}{\partial x}\frac{\partial g(x)'}{\partial x}\right), \tag{3.23}$$

where $w(x)$ is a suitable trimming function. It can be estimated through

$$\widehat{M}_S = \frac{1}{n}\sum_{i=1}^{n} w(x_i)\left(\frac{\widehat{G}^{(1)}(x_i)}{\widehat{p}(x_i)} - \frac{\widehat{G}(x_i)\widehat{p}^{(1)}(x_i)}{\widehat{p}(x_i)^2}\right)\left(\frac{\widehat{G}^{(1)}(x_i)}{\widehat{p}(x_i)} - \frac{\widehat{G}(x_i)\widehat{p}^{(1)}(x_i)}{\widehat{p}(x_i)^2}\right)', \tag{3.24}$$

where, with a kernel function K, and for $k = 0, 1$,

$$\widehat{p}^{(k)}(x_i) = \frac{1}{(n-1)h^{l+k}}\sum_{j=1, j\neq i}^{n} K^{(k)}\left(\frac{x_i - x_j}{h}\right), \tag{3.25}$$

$$\widehat{G}^{(k)}(x_i) = \frac{1}{(n-1)h^{l+k}}\sum_{j=1, j\neq i}^{n} y_j K^{(k)}\left(\frac{x_i - x_j}{h}\right) \tag{3.26}$$

with $f^{(0)}(x) = f(x)$ and $f^{(1)}(x) = (\partial f/\partial x)(x)$ for any function f. Note that both $\widehat{M}_S$ and M are symmetric, positive *semidefinite*.

Donkers and Schafgans (2008) suggest estimating p in (3.22) as the rank $\mathrm{rk}\{M\}$. (In practice, larger p is fixed, and the true p is estimated as $\mathrm{rk}\{M\}$.) Using the results of Samarov (1993), these authors state the asymptotic normality result along the lines of (2.2).

Proposition 3.3 *(Donkers and Schafgans (2008), Samarov (1993)) Under the assumptions stated in Donkers and Schafgans (2008), Samarov (1993),*

$$\sqrt{n}(\mathrm{vec}(\widehat{M}_S) - \mathrm{vec}(M)) \xrightarrow{d} \mathcal{N}(0, W), \tag{3.27}$$

where

$$W = \mathrm{Var}(\mathrm{vec}(R(x_i, y_i))) \tag{3.28}$$

with

$$R(x,y) = w(x)\left(\frac{\partial g(x)}{\partial x}\frac{\partial g(x)'}{\partial x} - (y - g(x))\left(\frac{\frac{\partial p(x)}{\partial x}\frac{\partial g(x)'}{\partial x}}{p(x)} + \frac{\frac{\partial g(x)}{\partial x}\frac{\partial p(x)'}{\partial x}}{p(x)} + 2\frac{\partial^2 g(x)}{\partial x \partial x'}\right)\right). \tag{3.29}$$

As in Example 3.1, one can show easily that $\mathrm{rk}\{W_0\}$ in (3.28) is constrained by $r = \mathrm{rk}\{M\}$. In fact, a stronger statement can be made. As shown in Lemma 1 of Donkers and Schafgans (2008), $\mathrm{rk}\{W\} = lr - r(r-1)/2$.

Example 3.4 (Number of factors in nonparametric relationship.) Donald (1997) considers a nonparametric model

$$y_i = F(x_i) + \epsilon_i, \quad i = 1, \dots, n, \tag{3.30}$$

where dependent variables $y_i \in \mathbb{R}^p$, explanatory variables $x_i \in \mathbb{R}^q$, and error terms ϵ_i have zero mean and nonsingular covariance matrix $\Sigma = E\epsilon_i\epsilon_i'$. The function F is unknown but supposed sufficiently smooth. A "local" version of (3.30) and related problems are considered in Fortuna (2008).

Of interest in our context is the semidefinite matrix

$$M = Ep(x_i)F(x_i)F(x_i)', \tag{3.31}$$

where $p(x)$ is the density of x_i. Its rank $r = \mathrm{rk}\{M\}$ is the number of factors in the nonparametric relationship (3.30) in the sense that $F(x) = AH(x)$ for some $p \times r$ matrix A and $r \times 1$ function $H(x)$.

To test for $\mathrm{rk}\{M\}$, Donald (1997) estimates the matrix M through the *indefinite* matrix estimator

$$\widehat{M}_I = \frac{1}{n(n-1)} \sum_{i \neq j} y_i y_j' K_h(x_i - x_j), \tag{3.32}$$

where K is a kernel function, $K_h(x) = K(x/h)/h^q$ and $h > 0$ is a bandwidth. It has the following assymptotics.

Proposition 3.4 *(Donald (1997), Donald et al. (2007)) Under the assumptions of Donald (1997),* $nh^{q/2}(\mathrm{vec}(\widehat{M}_I) - \mathrm{vec}(M))$ *is asymptotically normal. Moreover, one has* $\widehat{M}_I = \widehat{M}_1 + \widehat{M}_2$ *where (i)* $u'Mu = 0$ *for a vector* u *implies* $u'\widehat{M}_1 u = 0$*, (ii)* $\widehat{M}_1 - M = O_p((nh^{q/2})^{-1})$*, and*

$$nh^{q/2}\mathrm{vech}(\widehat{M}_2) \xrightarrow{d} \mathcal{N}(0, W_0), \tag{3.33}$$

where

$$W_0 = V^{-1} D_p^+ (\Sigma\Sigma \otimes \Sigma\Sigma) D_p^{+\prime} \tag{3.34}$$

with $V = (2\|K\|_2^2 Ep(x_i))^{-1/2}$*, density* $p(x)$ *of* x_i *and the Moore-Penrose inverse* D_p^+ *of the duplication matrix* D_p*.*

Alternatively, the matrix M can be estimated by the *semidefinite* matrix estimator

$$\widehat{M}_S = \frac{1}{n} \sum_{i=1}^{n} \widehat{p}(x_i)^{-1} \widehat{G}(x_i) \widehat{G}(x_i)', \tag{3.35}$$

where similarly to Example 3.3, for example,

$$\widehat{G}(x_i) = \frac{1}{(n-1)h^q} \sum_{j \neq i} y_j K\left(\frac{x_i - x_j}{h}\right). \tag{3.36}$$

As in Example 3.3, using the results of Samarov (1993), one can establish the following result. The assumptions and a short proof are in Appendix A.

Proposition 3.5 *(Samarov (1993)) Under the assumptions stated in Appendix A,*

$$\sqrt{n}(\mathrm{vec}(\widehat{M}_S) - \mathrm{vec}(M)) \xrightarrow{d} \mathcal{N}(0, W), \tag{3.37}$$

where

$$W = \mathrm{Var}\left(p(x_i)(F(x_i) \otimes I_p + I_p \otimes F(x_i)) y_i\right). \tag{3.38}$$

Note that, under the model (3.30), the limiting covariance matrix W in (3.38) is given by

$$W = \mathrm{Var}(p(x_i)(F(x_i) \otimes I_p + I_p \otimes F(x_i))F(x_i)) + \mathrm{Var}(p(x_i)(F(x_i) \otimes I_p + I_p \otimes F(x_i))\epsilon_i).$$

The two estimators $\widehat{M}_I$ in (3.32) and $\widehat{M}_S$ in (3.35) are related, with an obvious informal way to go from $\widehat{M}_I$ to $\widehat{M}_S$. It is therefore quite surprising that the normalizations used in Propositions 3.4 and 3.5 ($nh^{q/2}$ and $n^{1/2}$, respectively) are different. However, to see clearly that the two normalizations are needed, the reader is encouraged to compute the asymptotic variances of $\widehat{M}_I$ and $\widehat{M}_S$ when $p = 1$ and $F(x) \equiv 0$.

Let us also note that Proposition 3.4 follows from Donald (1997), Donald et al. (2007) assuming, in particular, that $nh^{3q/2} \to \infty$. This assumption implies that $nh^{q/2}/n^{1/2} = n^{1/2}h^{q/2} \geq (nh^{3q/2})^{1/2} \to \infty$ and hence that the normalization $nh^{q/2}$ is larger than $n^{1/2}$.

4 On extensions of available rank tests

As seen from Proposition 2.1 in Donald et al. (2007) and examples considered in Section 3, under rank deficiency for M, the limiting covariance matrix W_0 in (2.2) or W in (2.1) is singular. Several extensions of available rank tests were proposed for the case of singular limiting covariance matrices. We examine here a number of such proposals in our context. The focus is on the case R in (2.6). But the case C is also considered briefly by reexamining a test suggested by Camba-Méndez and Kapetanios (2005, 2009).

4.1 Tests involving generalized inverses

Available rank tests for nonsingular limiting covariance matrices W in (2.1) often involve inverses of these matrices. When they are singular, rank tests can be extended naturally by using generalized inverses. For example, one such typical extension is for the LDU rank test of Gill and Lewbel (1992), and Cragg and Donald (1996). This LDU test is based on Gaussian elimination with complete pivoting. Ignoring the issue of ties for simplicity, suppose without loss of generality that the matrices are permuted beforehand and hence that permutations are not necessary in the Gaussian elimination procedure. If the matrix M is partitioned as

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}, \tag{4.1}$$

where M_{11} is $r \times r$, then r steps of the Gaussian elimination procedure lead to the Schur complement

$$\Lambda_r = M_{22} - M_{21}M_{11}^{-1}M_{12} \tag{4.2}$$

(the inverse M_{11}^{-1} exists only for $r \leq \mathrm{rk}\{M\}$). Let also

$$\Phi = (-M_{21}M_{11}^{-1} \quad I_{p-r}), \quad \Gamma = \Phi \otimes \Phi, \tag{4.3}$$

and introduce analogous notation $\widehat{M}_{11}$, $\widehat{M}_{12}$, $\widehat{M}_{21}$, $\widehat{M}_{22}$, $\widehat{\Phi}$ and $\widehat{\Gamma}$. Let A^+ denote the Moore-Penrose inverse of a matrix A, and χ^2_m stand for the χ^2-distribution with m degrees of freedom.

Proposition 4.1 *(Camba-Méndez and Kapetanios (2001, 2009)) Suppose that (2.1) holds, the rank of W is known and there is $\widehat{W} \xrightarrow{p} W$ such that $\mathrm{rk}\{\widehat{W}\} = \mathrm{rk}\{W\}$. Then, under $\mathrm{rk}\{M\} = r$,*

$$\widehat{\xi}_{\mathrm{ldu}}(r) := N\mathrm{vec}(\widehat{\Lambda}_r)'(\widehat{\Gamma}\widehat{W}\widehat{\Gamma}')^+\mathrm{vec}(\widehat{\Lambda}_r) \xrightarrow{d} \chi^2_m, \tag{4.4}$$

where $m = \min\{(p-r)^2, \mathrm{rk}\{W\}\}$.

The basic reason for assuming $\text{rk}\{\widehat{W}\} = \text{rk}\{W\}$ is that $\widehat{W} \xrightarrow{p} W$ does not imply in general that their generalized inverses converge. See the references above, as well as Lütkepohl and Burda (1997), Andrews (1987).

A similar result could be obtained, for example, using vech instead of vec operation in (4.4) (and, in addition, using the so-called symmetric pivoting as in Donald et al. (2007)) or for other rank tests such as the MINCHI2 test of Cragg and Donald (1996). However, these tests are not appropriate in our context because $\text{rk}\{W\}$ is unknown. As the examples considered in Section 3 and the next general result show, the rank of W is constrained by the rank of M itself, which is unknown. See Appendix A for a proof.

Proposition 4.2 *Suppose that $\widehat{M}$ is a semidefinite matrix estimator for M, and that (2.1) holds with covariance matrix W. Then, with $r = \text{rk}\{M\}$,*

$$\text{rk}\{W\} \leq r(2p - r). \tag{4.5}$$

4.2 Test of Robin and Smith

Robin and Smith (2000) suggested another interesting test allowing for singular covariance matrix W in (2.1). A simplified version of their characteristic root test adapted to our context is the following. In (2.1), suppose, in addition, that

$$0 < \text{rk}\{W\} = s \leq p^2. \tag{4.6}$$

Let $C = (c_1, \dots, c_p)$ consist of eigenvectors of $MM' = M^2$ such that $CC' = I_p$. Partition the matrix C as $C = (C_{p-r}\ C_r)$ where $r = \text{rk}\{M\}$ and C_{p-r} is $p \times (p - r)$. Assume also that

$$t := \text{rk}\{(C'_{p-r} \otimes C'_{p-r})W(C_{p-r} \otimes C_{p-r})\} > 0. \tag{4.7}$$

Let now $\widehat{\lambda}_1 \leq \dots \leq \widehat{\lambda}_p$ be the ordered eigenvalues of $\widehat{M}^2$, and consider the test statistic

$$\widehat{\xi}_{\text{cr}}(r) = N \sum_{j=1}^{p-r} \widehat{\lambda}_j. \tag{4.8}$$

Under $\text{rk}\{M\} = r$, the test statistic $\widehat{\xi}_{\text{cr}}(r)$ has the limiting distribution described by

$$\xi_{\text{cr}}(r) = \sum_{j=1}^{t} \lambda_j^r Z_j^2 = \sum_{j=1}^{(p-r)^2} \lambda_j^r Z_j^2, \tag{4.9}$$

where $\{Z_j\}$ are independent $\mathcal{N}(0, 1)$ random variables, and $\lambda_1^r \geq \lambda_2^r \geq \dots \geq \lambda_t^r > 0 = \lambda_{t+1}^r = \dots = \lambda_{(p-r)^2}^r$ are the eigenvalues of $(C'_{p-r} \otimes C'_{p-r})W(C_{p-r} \otimes C_{p-r})$. In practice, the limiting distribution is approximated by

$$\sum_{j=1}^{(p-r)^2} \widehat{\lambda}_j^r Z_j^2, \tag{4.10}$$

where $\widehat{\lambda}_j^r$ are the ordered eigenvalues of $(\widehat{C}'_{p-r} \otimes \widehat{C}'_{p-r})\widehat{W}(\widehat{C}_{p-r} \otimes \widehat{C}_{p-r})$.

Assumption (4.7) of Robin and Smith (2000) is not satisfied in the semidefinite context. In fact, as shown in Section 5 below (see the discussion following Corollary 5.1), $t = 0$ and hence

$\lambda_j^r = 0$ for all j. To see why this is expected, consider Example 3.4 and the matrix estimator $\widehat{M}_S$ in (3.35). The matrix C above also consists of eigenvectors of M, and we can suppose that $C'MC = \mathrm{diag}\{0,\dots,0,\upsilon_{p-r+1},\dots,\upsilon_p\}$ where $0 < \upsilon_{p-r+1} \le \cdots \le \upsilon_p$ and $r = \mathrm{rk}\{M\}$. Since $C'MC = Ef(x_i)C'F(x_i)F(x_i)'C$ in that example, it follows that $C'_{p-r}F(x)F(x)'C_{p-r} = 0$ and hence

$$C'_{p-r}F(x) = 0. \tag{4.11}$$

In view of (3.38), (4.11) implies that $(C'_{p-r}\otimes C'_{p-r})W(C_{p-r}\otimes C_{p-r}) = 0$. The same conclusion for the estimator $\widehat{M}_S$ in (3.24) of Example 3.3 is reached in Lemma 2 of Donkers and Schafgans (2008).

4.3 Test of Donkers and Schafgans

Following the relation (4) in Donkers and Schafgans (2008), another possibility is to consider the test statistic (of a minimum discrepancy type)

$$\widehat{\xi}_{\mathrm{md}}(r) = N \min_{\mathrm{rk}\{\mathcal{M}\}=r,\mathcal{M}=\mathcal{M}'} (\mathrm{vech}(\widehat{M}) - \mathrm{vech}(\mathcal{M}))'(\mathrm{vech}(\widehat{M}) - \mathrm{vech}(\mathcal{M})). \tag{4.12}$$

Under $\mathrm{rk}\{M\} = r$, Donkers and Schafgans (2008) seem to suggest that $\widehat{\xi}_{\mathrm{md}}(r)$ has the limiting distribution $\mathcal{N}(0,W_0)'\mathcal{N}(0,W_0)$. Critical values for the latter distribution can be generated by drawing from $\mathcal{N}(0,\widehat{W}_0)$.

In fact, we do not expect the limiting distribution of $\widehat{\xi}_{\mathrm{md}}(r)$ to be as described above. As in the proof of Theorem 4.2 in Donald et al. (2007) (see also Cragg and Donald (1996, 1997)), the limit of $\widehat{\xi}_{\mathrm{md}}(r)$ is expected to be $\mathcal{N}(0,W_0)'(I-A)\mathcal{N}(0,W_0)$ for a suitable idempotent matrix A. Moreover, if the operations vech are replaced by vec and the minimum is replaced by that over $\mathrm{rk}\{\mathcal{M}\}\le r$ in (4.12), then $\widehat{\xi}_{\mathrm{md}}(r) = N\sum_{j=1}^{p-r}\widehat{\lambda}_j$, where $\widehat{\lambda}_j$ are the ordered eigenvalues of $\widehat{M^2}$, and one is back to the test of Robin and Smith considered above.

4.4 Test of Camba-Méndez and Kapetanios

Another related test for the rank of Hermitian positive semidefinite matrix was suggested by Camba-Méndez and Kapetanios (2005, 2009). Consider the more general Case C of (2.6). As for the LDU test, the matrix M is partitioned as (4.1) and it is supposed for simplicity that M_{11} has full rank (which can be supposed after permutations involved in the Gaussian elimination procedure). The focus is again on the Schur complement Λ_r given in (4.2). Then, $\mathrm{rk}\{M\} = r$ is equivalent to $\Lambda_r = 0$ which can be shown to be equivalent to $\mathrm{diag}\{\Lambda_r\} = 0$ (where $\mathrm{diag}\{A\}$ denotes the diagonal elements of a vector A). Since the alternative to $\mathrm{diag}\{\Lambda_r\} = 0$ is that at least one of the diagonal elements of Λ_r is strictly positive, Camba-Méndez and Kapetanios suggest using the multivariate one-sided tests of Kudo (1963), Kudo and Choi (1975). More precisely, consider the test statistic

$$\widehat{\xi}_{\mathrm{kudo}}(r) = N\mathrm{diag}\{\widehat{\Lambda}_r\}'\Psi^{-1}\mathrm{diag}\{\widehat{\Lambda}_r\}, \tag{4.13}$$

where Ψ (supposed to be known for simplicity) appears in the asymptotic normality result for $\widehat{\Lambda}_r$, namely,

$$\sqrt{N}\mathrm{diag}\{\widehat{\Lambda}_r\} \xrightarrow{d} \mathcal{N}(0,\Psi). \tag{4.14}$$

An application of the result of Kudo then suggests that the asymptotic distribution $\xi_{\mathrm{kudo}}(r)$ of (4.13) is given by

$$P(\xi_{\mathrm{kudo}}(r) > x) = \sum_{q=0}^{p-r} w_q P(\chi_q^2 \ge x), \tag{4.15}$$

where w_q are suitable weights.

The application of the results of Kudo above, however, is not justified. The Schur complement is also positive semidefinite (see, for example, Zhang (2005)), and hence the elements of $\mathrm{diag}\{\widehat{\Lambda}_r\}$ are real and nonnegative. But then (4.14) is possible only with $\Psi \equiv 0$ and the definition (4.13) and its limit in (4.15) are not valid.

Remark 4.1 Though application of the results of Kudo are not appropriate in the context above, they can be used in Case 1 of (1.3) considered by Donald et al. (2007).

4.5 Approach of Lütkepohl and Burda

Finally, following Lütkepohl and Burda (1997), another possibility is to consider a matrix estimator

$$\widehat{M}_u = \widehat{M} + \frac{u}{\sqrt{N}},$$

where u is a symmetric random matrix, independent of $\widehat{M}$ and satisfying $\mathrm{vech}(u) \stackrel{d}{=} \mathcal{N}(0, I)$. Then,

$$\sqrt{N}(\mathrm{vech}(\widehat{M}_u) - \mathrm{vech}(M)) \xrightarrow{d} \mathcal{N}(0, W_0 + I) \tag{4.16}$$

and $W_0 + I$ is now positive definite (since $a'(W_0 + I)a = a'W_0a + a'a = 0$ if and only if $a = 0$). The rank of the matrix M can then be tested by using any of the matrix rank tests found in Donald et al. (2007).

5 Asymptotics of eigenvalues, and consistency of rank tests

Further light can be shed on several extensions of the available rank tests (see Section 4) by examining the asymptotic behavior of the eigenvalues of the matrix estimator $\widehat{M}$. Thus, let $0 \leq \widehat{v}_1 \leq \widehat{v}_2 \leq \cdots \leq \widehat{v}_p$ be the ordered eigenvalues of $\widehat{M}$.

Theorem 5.1 *Under the assumption (2.1), and with* $\mathrm{rk}\{M\} = r$,

$$\sqrt{N}\widehat{v}_j \xrightarrow{p} 0, \quad j = 1, \ldots, p - r, \tag{5.1}$$

$$\sqrt{N}\widehat{v}_j \xrightarrow{p} +\infty, \quad j = p - r + 1, \ldots, p. \tag{5.2}$$

The theorem is proved in Appendix A. Letting

$$\widehat{\xi}_{\mathrm{eig}}(k) = \sqrt{N^{\beta}} \sum_{j=1}^{p-k} \widehat{v}_j^{\beta}, \tag{5.3}$$

with some fixed $\beta > 0$, we have the following immediate corollary of the the result above.

Corollary 5.1 *Under the assumption (2.1),*

$$\text{under} \quad \mathrm{rk}\{M\} \leq k, \qquad \widehat{\xi}_{\mathrm{eig}}(k) \xrightarrow{p} 0, \tag{5.4}$$

$$\text{under} \quad \mathrm{rk}\{M\} > k, \qquad \widehat{\xi}_{\mathrm{eig}}(k) \xrightarrow{p} +\infty. \tag{5.5}$$

Since $\widehat{v}_k^2$ are the eigenvalues of $\widehat{M}^2$, the test statistic $\widehat{\xi}_{\mathrm{eig}}(r)$ in (5.3) when $\beta = 2$ coincides with the characteristic root test statistic $\widehat{\xi}_{\mathrm{cr}}(r)$ in (4.8). Corollary 5.1 then shows that the limit $\xi_{\mathrm{cr}}(r)$ in (4.9) is identical to zero and hence that $\lambda_j^r = 0$ for all j in (4.9) and that $t = 0$ in (4.7).

Since $\widehat{\xi}_{\mathrm{cr}}(r) = \widehat{\xi}_{\mathrm{eig}}(r)$, Corollary 5.1 also seems to suggest that the test of Robin and Smith (Section 4.2) is consistent. Note, however, that the critical values in the test are based on (4.10), and hence converge to 0 (since $\xi_{\mathrm{cr}}(r) = 0$). Corollary 5.1 then does *not* imply consistency. On the other hand, if the critical values are based on $\mathcal{N}(0, \widehat{W}_0)'\mathcal{N}(0, \widehat{W}_0)$ following Donkers and Schafgans (Section 4.3), then the test is consistent by Corollary 5.1.

Though the two hypotheses $\mathrm{rk}\{M\} \leq k$ and $\mathrm{rk}\{M\} > k$ can be distinguished in the asymptotic sense by Corollary 5.1, the use of the test in finite samples and controlling the size of the test is not obvious given these results. This is the conclusion reached in Section 8 based on simulations in one of the examples considered.

6 Towards rank tests based on more refined asymptotics

In view of the results of Section 5, it is natural to expect that a faster rate than $\sqrt{N}$ in (5.1)–(5.2) may yield a convergence in distribution result with a nondegenerate limit in (5.1) (and consequently in (5.4)). We examine this and related questions in this section.

The following general assumption will be useful. Let $Q = (Q_1\ Q_2)$ be as in (2.7) so that $Q_1'MQ_1 = 0$. Then, $\sqrt{N}Q_1'(\widehat{M} - M)Q_1 = \sqrt{N}Q_1'\widehat{M}Q_1 \xrightarrow{d} Q_1'\mathcal{Y}Q_1$ in view of (2.3). Since $\sqrt{N}Q_1'\widehat{M}Q_1$ is semidefinite, Lemma A.1 implies that $\sqrt{N}Q_1'\widehat{M}Q_1 \xrightarrow{p} 0$. Therefore, in some cases, it may be natural to expect that

$$a_N Q_1'\widehat{M}Q_1 \xrightarrow{d} \mathcal{A}, \tag{6.1}$$

where $a_N \to \infty$ and a_N grows faster than $\sqrt{N}$ (that is, $a_N/\sqrt{N} \to \infty$). In fact, in several examples considered below, $\mathcal{A}$ is deterministic and the convergence in (6.1) is in probability.

Theorem 6.1 *Under the assumption (2.1) and (6.1), with the notation of Section 5 and when* $\mathrm{rk}\{M\} = r$,

$$a_N\widehat{v}_j \xrightarrow{p} \alpha_j, \quad j = 1, \dots, p - r, \tag{6.2}$$

$$a_N\widehat{v}_j \xrightarrow{p} +\infty, \quad j = p - r + 1, \dots, p, \tag{6.3}$$

where $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_{p-r}$ *are the ordered eigenvalues of* $\mathcal{A}$. *Moreover,*

$$\text{under} \quad \mathrm{rk}\{M\} \leq k, \qquad \widehat{\xi}_{\mathrm{eig}}(k) \xrightarrow{d} \sum_{j=1}^{p-k} \alpha_j^{\beta}, \tag{6.4}$$

$$\text{under} \quad \mathrm{rk}\{M\} > k, \qquad \widehat{\xi}_{\mathrm{eig}}(k) \xrightarrow{p} +\infty, \tag{6.5}$$

where $\widehat{\xi}_{\mathrm{eig}}(k)$ *is defined by (5.3) but using the normalization* a_N^{β} *instead of* $N^{\beta/2}$.

The proof of Theorem 6.1 is analogous to that of Theorem 5.1 and is omitted. We also note that, strictly speaking, the eigenvalues α_j above are those of $Q_1'MQ_1$, where Q_1 is $p \times (p - k)$ (not $p \times (p - r)$, since r is unknown), and thus depend on k.

In all examples considered below, the convergence in (6.1) is in probability to a deterministic limit $\mathcal{A}$, corresponding to the law of large numbers asymptotics determined by the limit of

$a_N E Q_1' \widehat{M} Q_1$. Supposing there is a consistent estimator $\widehat{\mathcal{A}} \xrightarrow{p} \mathcal{A}$, Theorem 6.1 then leads to a consistent rank test: a critical value for a test of $H_0 : \mathrm{rk}\{M\} = k$ could be defined as

$$(1+c)\sum_{j=1}^{p-k} \widehat{\alpha}_j^{\beta} =: \widehat{\zeta}_{1+c}(k) \tag{6.6}$$

for some (arbitrary) $c > 0$, where $\widehat{\alpha}_j$ are the ordered eigenvalues of $\widehat{\mathcal{A}}$. The resulting test is not fully satisfactory but could lead to a more useful rank test in the future (see a discussion at the end of this section, and Section 8 below).

Example 6.1 (Linear regression with heteroscedastic error terms.) Consider the model (3.1)–(3.2) in Example 3.1. To show that the estimator (3.3) satisfies the assumption (6.1), we make the following additional assumptions. Suppose that the error terms ϵ_i can be expressed as

$$\epsilon_i = (p(x_i))^{-1/2} f(x_i)\eta_i, \tag{6.7}$$

where $f(x)$ is a $p \times p$ matrix depending on x and η_i are $p \times 1$ i.i.d. vectors with i.i.d. entries such that $E\eta_i = 0$, $E\eta_i\eta_i' = I$ and $E(\eta_i)_l^4 =: E\eta^4 < \infty$. Note that, under the assumption (6.7), $\Sigma(x) = f(x)f(x)'$. Moreover, note that $Q_1'\Sigma(x)Q_1 = 0$ with Q_1 as in (2.7) if and only if

$$Q_1' f(x) = 0. \tag{6.8}$$

The next result is proved in Appendix A. It shows that, under additional assumptions, the estimator $\widehat{\Sigma}(x)$ in (3.3) satisfies the assumption (6.1).

Proposition 6.1 *With the above notation and that of Example 3.1, suppose in addition that $K(u) = K_0(u_1)\dots K_0(u_q)$, $u = (u_1,\dots,u_q)$, with a kernel K_0 of order 2, that x_k has a density $p(x)$ on $\mathcal{H}_x := (a_x, b_x)$ with $p(x) > 0$ and $p \in C^1(\mathcal{H}_x)$, and that $f \in C^1(\mathcal{H}_x)$. Then, as $n \to \infty$, $h \to 0$, and $nh^{q/2+2} \to \infty$, for $x \in \mathcal{H}_x$,*

$$h^{-2}\mathrm{vec}(Q_1'\widehat{\Sigma}(x)Q_1) \xrightarrow{p} \mathrm{vec}(Q_1'\Pi(x)Q_1), \tag{6.9}$$

where $\Pi(x) = (\Pi(x)_{ij})_{i,j=1,\dots,p}$ with

$$\Pi(x)_{ij} = c_0 \mathrm{tr}\Big(\sum_{k=1}^{p} \frac{\partial f_{ik}}{\partial x}(x)\frac{\partial f_{kj}}{\partial x}(x)'\Big) \tag{6.10}$$

and $c_0 = \int v^2 K_0(v)dv$.

The convergence (6.9) corresponds to the law of large numbers asymptotics. This can be seen from the proof of Proposition 6.1 and also from the following argument. Observe first that $\Sigma(x)_{ij} = \sum_{k=1}^{p} f_{ik}(x)f_{kj}(x)$ and hence that

$$\Sigma^{(2)}(x)_{ij} = \sum_{k=1}^{p}\Big(\frac{\partial^2 f_{ik}(x)}{\partial x^2} f_{kj}(x) + 2\frac{\partial f_{ik}}{\partial x}(x)\frac{\partial f_{kj}}{\partial x}(x)' + f_{ik}(x)\frac{\partial f_{kj}}{\partial x}(x)\Big). \tag{6.11}$$

Then, in view of (6.8), $\Pi(x)_{ij}$ in (6.10) can be replaced by $c_0\mathrm{tr}(\Sigma^{(2)}(x)_{ij})/2$. The latter quantity is the law of large numbers limit. Indeed, note that $h^{-2}Q_1'E\widehat{\Sigma}(x)Q_1$ is expected to behave as $h^{-2}Q_1' E\epsilon_k\epsilon_k' K_h(x-x_k)Q_1 = h^{-2}Q_1' \int \Sigma(x_k)K_h(x-x_k)dx_kQ_1 =: I_h$. Since $Q_1'\Sigma(x)Q_1 = 0$, an argument based on Taylor expansion shows that I_h behaves as $Q_1'\Upsilon(x)Q_1$, where $\Upsilon(x)_{ij} = c_0\mathrm{tr}(\Sigma^{(2)}(x)_{ij})/2$. Thus, $\Pi(x)_{ij} = \Upsilon(x)_{ij}$.

Note also that the discussion above suggests a natural estimator for $\Pi(x)_{ij}$ as $c_0\mathrm{tr}(\widehat{\Sigma}^{(2)}(x)_{ij})/2$, where

$$\widehat{\Sigma}^{(2)}(x)_{ij} = \frac{1}{nh^{q+2}}\sum_{k=1}^{n}\Big((y_k-\widehat{\beta}x_k)(y_k-\widehat{\beta}x_k)'\Big)_{ij}K^{(2)}(\frac{x-x_k}{h}). \tag{6.12}$$

Finally, observe the following relationship between the normalizations $\sqrt{nh^q}$ and h^{-2} used in (3.4) and (6.9), respectively. As expected in the framework of this section (see (6.1)), $h^{-2}/\sqrt{nh^q} = 1/\sqrt{nh^{q+4}} \to \infty$ since $nh^{q+4} = nh^{q+2r} \to 0$ for a kernel K of order $r=2$ as in Proposition 3.1.

Example 6.2 (Number of factors in nonparametric relationship.) Consider the semidefinite matrix estimator $\widehat{M}_S$ given in (3.35). Since $Q_1'F(x) = 0$, observe that

$$Q_1'\widehat{M}_SQ_1 = Q_1'\frac{1}{n}\sum_{i=1}^{n}\widehat{p}(x_i)^{-1}\widehat{\epsilon}(x_i)\widehat{\epsilon}(x_i)'Q_1, \tag{6.13}$$

where

$$\widehat{\epsilon}(x_i) = \frac{1}{n-1}\sum_{j\neq i}\epsilon_jK_h(x_i-x_j). \tag{6.14}$$

As in Example 6.1, the law of large numbers asymptotics is expected wherein the limit of $Q_1'\widehat{M}_SQ_1$ is determined by $EQ_1'\widehat{M}_SQ_1$. One expects $EQ_1'\widehat{M}_SQ_1$ to behave asymptotically as

$$Q_1'Ep(x_i)^{-1}\widehat{\epsilon}(x_i)\widehat{\epsilon}(x_i)'Q_1 = Q_1'\frac{1}{(n-1)^2}Ep(x_i)^{-1}\sum_{j\neq i}\epsilon_j\epsilon_j'K_h(x_i-x_j)^2Q_1$$

$$= \frac{\|K\|_2^2}{(n-1)h^q}Q_1'\Sigma Q_1Ep(x_i)^{-1}K_{2,h}(x_i-x_j) \sim \frac{\|K\|_2^2}{nh^q}Q_1'\Sigma Q_1, \tag{6.15}$$

where $K_2(x) = K(x)^2/\|K\|_2^2$, and hence that

$$nh^qQ_1'\widehat{M}_SQ_1 \xrightarrow{p} \|K\|_2^2Q_1'\Sigma Q_1. \tag{6.16}$$

We shall not prove (6.16) rigorously here. We shall nevertheless use this convergence in the simulations considered in Section 8.

Remark 6.1 Example 3.3 concerning semiparametric multiple-index model could be dealt with as Example 6.2 above. Dealing with the spectral density matrix estimator of Example 3.2 is more challenging and left for future work.

Examples 6.1 and 6.2 above suggest that $\mathcal{A}$ in (6.1) is a constant matrix, and that (6.1) corresponds to the law of large numbers asymptotics. It is natural to go beyond (6.1) by postulating the asymptotics of central limit theorem as

$$b_N(a_NQ_1'\widehat{M}Q_1 - \mathcal{A}) \xrightarrow{d} \mathcal{B}, \tag{6.17}$$

where $b_N \to \infty$ and $\mathcal{B}$ is a random matrix. The assumption (6.17) is related to the asymptotic behavior of

$$b_N(a_N\widehat{v}_j - \alpha_j), \quad j = 1, \dots, p-r \tag{6.18}$$

(see, for example, Eaton and Taylor (1991)).

When testing for $\text{rk}\{M\} \leq k$, it would then be natural to consider the test statistic $b_N \sum_{j=1}^{k}(a_N\widehat{v}_j - \widehat{\alpha}_j)$. There are several open problems in this approach. First, the rate of convergence of $\widehat{\alpha}_j$ to α_j needs to be understood and compared to b_N. Second, consider, for example, Example 6.1. One expects (see the proof of Proposition 6.1) that

$$b_N(h^{-2}\text{vech}(Q_1'\widehat{\Sigma}(x)Q_1) - \text{vech}(Q_1'\Pi(x)Q_1)) \xrightarrow{d} \mathcal{N}(0, \widetilde{W}_0), \tag{6.19}$$

for suitable $b_N \to \infty$. The limiting covariance matrix $\widetilde{W}_0$, however, is not positive definite unless $Q_1'\Pi(x)Q_1$ is of full rank. The latter is now a rank-like assumption on $\Pi(x)$, involving the second derivative of $\Sigma(x)$. Thus, the rank problem about $\Sigma(x)$ leads to another such problem involving the second derivative of $\Sigma(x)$. Despite these open problems, the preliminary simulations reported in Section 8 suggest that the approach is promising.

7 Use of indefinite matrix estimators

The previous section concerned the possibility of obtaining rank tests with good properties under the assumption (2.1) or (2.2) where matrix estimator $\widehat{M}$ is semidefinite. Another natural possibility is to search for an indefinite matrix estimator which may allow one to assume (2.2) with nonsingular W_0 and hence to use the well-developed framework of Donald et al. (2007).

The choice of indefinite matrix estimator $\widehat{M}$ depends on the problem at hand and may not be evident. For example, in Example 3.4, we already specified two matrix estimators: the indefinite $\widehat{M}_I$ in (3.32) and the semidefinite $\widehat{M}_S$ in (3.35). The rank tests for $\widehat{M}_I$ can be carried out in the framework of Donald et al. (2007). In the next example, we show that an indefinite estimator can be introduced naturally in Example 3.3 as well.

Example 7.1 (Semiparametric multiple-index model.) Suppose that the model (3.22) of Example 3.3 can be written as

$$y_i = g(x_i) + \epsilon_i, \quad i = 1, \dots, n, \tag{7.1}$$

where ϵ_i are independent of x_i, $E\epsilon_i = 0$, $E\epsilon_i^2 = \sigma^2 > 0$, and $g(x)$ is given by (3.22). A natural indefinite estimator of the matrix M in (3.23) (with suitable weight $w(x)$) is

$$\widehat{M}_I = \frac{1}{n(n-1)} \sum_{i \neq j} \widehat{p}(x_i)^{-1}\widehat{p}(x_j)^{-1} y_i y_j \frac{1}{h^{l+2}} K^{(2)}\left(\frac{x_i - x_j}{h}\right). \tag{7.2}$$

The basic idea behind (7.2) is that, asymptotically, $E\widehat{M}_I$ behaves as (with $i \neq j$)

$$Ep(x_i)^{-1}p(x_j)^{-1}y_i y_j \frac{1}{h^{l+2}} K^{(2)}\left(\frac{x_i - x_j}{h}\right) = Ep(x_i)^{-1}p(x_j)^{-1}g(x_i)g(x_j)\frac{1}{h^{l+2}}K^{(2)}\left(\frac{x_i - x_j}{h}\right)$$

$$= Ep(x_i)^{-1}p(x_j)^{-1}\frac{\partial g(x_i)}{\partial x}\frac{\partial g(x_j)'}{\partial x}\frac{1}{h^l}K\left(\frac{x_i - x_j}{h}\right) \to Ep(x_i)^{-1}\frac{\partial g(x_i)}{\partial x}\frac{\partial g(x_i)'}{\partial x} = M. \tag{7.3}$$

As in Example 3.4, write $\widehat{M}_I = \widehat{M}_1 + \widehat{M}_2$, where $\widehat{M}_1 = \widehat{M}_{1,1} + 2\widehat{M}_{1,2}$ with

$$\widehat{M}_{1,1} = \frac{1}{n(n-1)} \sum_{i \neq j} \widehat{p}(x_i)^{-1} \widehat{p}(x_j)^{-1} \frac{\partial g}{\partial x}(x_i) \frac{\partial g}{\partial x}(x_j)' K_h(x_i - x_j),$$

$$\widehat{M}_{1,2} = \frac{1}{n(n-1)} \sum_{i \neq j} \widehat{p}(x_i)^{-1} \widehat{p}(x_j)^{-1} \frac{\partial g}{\partial x}(x_i) \epsilon_j \frac{1}{h^{l+1}} K^{(1)} \left(\frac{x_i - x_j}{h} \right)$$

and $\widehat{M}_2 = \widehat{M}_{2,1} + 2\widehat{M}_{2,2} + \widehat{M}_{2,3}$ with

$$\widehat{M}_{2,1} = \frac{1}{n(n-1)} \sum_{i \neq j} \widehat{p}(x_i)^{-1} \widehat{p}(x_j)^{-1} \left(g(x_i) g(x_j) \frac{1}{h^{l+2}} K^{(2)} \left(\frac{x_i - x_j}{h} \right) - \frac{\partial g}{\partial x}(x_i) \frac{\partial g}{\partial x}(x_j)' K_h(x_i - x_j) \right),$$

$$\widehat{M}_{2,2} = \frac{1}{n(n-1)} \sum_{i \neq j} \widehat{p}(x_i)^{-1} \widehat{p}(x_j)^{-1} \left(g(x_i) \epsilon_j \frac{1}{h^{l+2}} K^{(2)} \left(\frac{x_i - x_j}{h} \right) - \frac{\partial g}{\partial x}(x_i) \epsilon_j \frac{1}{h^{l+1}} K^{(1)} \left(\frac{x_i - x_j}{h} \right) \right),$$

$$\widehat{M}_{2,3} = \frac{1}{n(n-1)} \sum_{i \neq j} \widehat{p}(x_i)^{-1} \widehat{p}(x_j)^{-1} \epsilon_i \epsilon_j \frac{1}{h^{l+2}} K^{(2)} \left(\frac{x_i - x_j}{h} \right).$$

Note that $u'Mu = 0$ implies $u'\partial g/\partial x = 0$ and hence $u'\widehat{M}_1 u = 0$. One could then expect that the framework of Donald et al. (2007) applies and that the asymptotics of $\widehat{M}_2$ is driven by $\widehat{M}_{2,3}$. In the latter regard, note that (with $i \neq j$)

$$\begin{aligned}
& E\mathrm{vech}(\widehat{M}_{2,3})\mathrm{vech}(\widehat{M}_{2,3})' \\
& \approx \frac{2\sigma^4}{n^2} E p(x_i)^{-2} p(x_j)^{-2} \frac{1}{h^{2l+4}} \mathrm{vech}\left(K^{(2)} \left(\frac{x_i - x_j}{h} \right) \right) \mathrm{vech}\left(K^{(2)} \left(\frac{x_i - x_j}{h} \right) \right)' \\
& \sim \frac{2\sigma^4}{n^2 h^{l+4}} E p(x_i)^{-2} p(x_j)^{-2} \frac{1}{h^l} \mathrm{diag}\left(\mathrm{vech}\left(K^{(2)} \left(\frac{x_i - x_j}{h} \right) \right)^2 \right) \\
& \sim \frac{2\sigma^4}{n^2 h^{l+4}} E p(x_i)^{-3} \mathrm{diag}\left(\mathrm{vech}\left(\|K^{(2)}(u)\|_2^2 \right) \right) =: \frac{1}{n^2 h^{l+4}} V,
\end{aligned} \tag{7.4}$$

where, for a vector $x = (x_1, \ldots, x_m)$, $x^2 = (x_1^2, \ldots, x_m^2)$ and $\|K^{(2)}(u)\|_2^2$ is the matrix consisting of entries $\int |\partial^2 K(u)/\partial u_i \partial u_j|^2 du$. In particular, one expects that

$$n h^{\frac{l}{2}+2} \mathrm{vech}(\widehat{M}_{2,3}) \xrightarrow{d} \mathcal{N}(0, V), \tag{7.5}$$

where V is nonsingular, as required in the framework of Donald et al. (2007).

Remark 7.1 Whether indefinite matrix estimators can be introduced naturally in Examples 3.1 and 3.2 is an open question.

8 Numerical study

We present here a small numerical study examining some of the tests discussed above. We consider the setting of the number of factors in a nonparametric relationship, as in Examples 3.4 and 6.2. We are interested in estimating the rank of the matrix M in (3.31) by using the indefinite and semidefinite matrix estimator $\widehat{M}_I$ and $\widehat{M}_S$ in (3.32) and (3.35), respectively,

The specific nonparametric model is that appearing in Donald (1997), namely,

$$y_i = \delta A H(x_i) + \epsilon_i, \quad i = 1, \dots, n, \tag{8.1}$$

where x_i are i.i.d. $U(0,1)$ (uniform on $(0,1)$), ϵ_i are i.i.d. $\mathcal{N}(0, I_3)$, $\delta = 1$ or $1/2$ controls the signal-to-noise ratio,

$$H(x) = D \begin{pmatrix} x^2 e^{-x^2/0.2531} \\ 1 \\ \cos((x+1)^{2.5}) \end{pmatrix}$$

with D such that $EH(x_i)H(x_i)' = I$, and

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}$$

with the choices $(a_{22}, a_{23}, a_{33}) = (0,0,0)$, $(a_{22}, a_{23}, a_{33}) = (1/2, 1/2, 0)$ and $(a_{22}, a_{23}, a_{33}) = (1/2, 1/2, 1/2)$ corresponding to $\mathrm{rk}\{M\} = 1$, 2 and 3, respectively.

We shall examine here the performance of the following tests. When based on the indefinite estimator $\widehat{M}_I$, the tests are denoted EIG_id, EIG2_id and LDU_id, and are, respectively, the EIG test (based on the sum of eigenvalues), the MINCHI2 test (based on the sum of squared eigenvalues) and the LDU test (based on the LDU decomposition) considered in Donald et al. (2007).

When based on the semidefinite estimator $\widehat{M}_S$, the tests are denoted LB_LDU_sd, LB_SVD_sd, EIG_sd and EIGf_sd. LB_LDU_sd and LB_SVD_sd refer to the Lütkepohl and Burda approach discussed in Section 4.5, with the only difference that we also explore the possibility of multiplying the perturbation term u by a constant c_0 (in which case the limiting covariance matrix is $W_0 + c_0^2 I$). The specific tests are, respectively, the LDU and the SVD tests of Donald et al. (2007) applied to the perturbed indefinite estimator of $\widehat{M}_S$. (The EIG and MINCHI2 tests of Donald et al. (2007) are not used here since they require the limiting covariance matrix to have a special form.) EIG_sd and EIGf_sd refer to the approaches based on eigenvalues discussed in Sections 5 and 6, respectively. Since these tests do not involve a limiting (non-degenerate) distribution, they will be treated differently.

Figure 1 presents several PP-plots for the tests EIG_id, EIG2_id and LDU_id. These plots have probability $p \in (0,1)$ on the vertical axis against $\alpha_k(p) = P(\widehat{\xi}(k) > c_k(p))$ on the horizontal axis, where $\widehat{\xi}(k)$ is a test statistic and $c_k(p)$ is the nominal critical value such that $P(\xi(k) > c_k(p)) = p$ with a corresponding limiting distribution $\xi(k)$ (normal for EIG_id, and chi-squared for EIG2_id and LDU_id). The probability $\alpha_k(p)$ is computed based on 1000 Monte Carlo replications.

In all the simulations reported here, we consider the model (8.1) with $\mathrm{rk}\{M\} = r = 2$ and $\delta = 1/2$, the sample size $n = 250$ or $n = 500$, and the bandwidth $h = 0.2$. The PP-plots in Figure 1 correspond to testing for the rank $k = 1$ (top) and $k = 2$ (bottom), for the sample sizes $n = 250$ (left) and $n = 500$ (right). The plots for $k = 2$ (bottom) and small values of p show that the considered tests are undersized, that is, $p > \alpha_k(p) = P(\widehat{\xi}(k) > c_k(p))$, with the EIG2_id test having the smallest size distortion. The plots for $k = 1$ (top) give idea about the power of the tests, with EIG2_id being the most powerful (that is, $\alpha_k(p)$ closest to 1 for small p).

Figure 2 presents analogous PP-plots for LB_LDU_sd and LB_SVD_sd tests. As indicated above, we vary a constant c_0 in the perturbation term $c_0 u$ by considering $c_0 = 1$, 0.1 and 0.01. The plots suggest that the approach of Lütkepohl and Burda is sensitive to the choice of c_0. For larger c_0 (e.g. $c_0 = 1$ as originally suggested), the power of the test is too small. For smaller c_0, the power is larger (e.g. $c_0 = 0.01$ leads to the power close to 100%) but the size of the test gets too distorted. Based on these observations, the tests are not recommended, until a suitable method of selecting c_0 is found.

Figure 3 concerns the EIG_sd test. We provide the boxplots of the corresponding test statistic $\widehat{\xi}(k)$ $(= \widehat{\xi}_{\mathrm{eig}}(k))$ for $n = 250$ (left) and $n = 500$ (right). Note the different vertical scales for $k = 0$ and $k = 1, 2$. While $\widehat{\xi}(1) \stackrel{p}{\to} \infty$ and $\widehat{\xi}(2) \stackrel{p}{\to} 0$ in theory, note that the distinction between $k = 1$ and $k = 2$ is impossible in practice without a more detailed analysis of the behavior of $\widehat{\xi}(k)$.

The final Figure 4 concerns the EIGf_sd test. We provide the boxplots of the ratios $\widehat{\xi}(k)/\widehat{\zeta}_1(k)$, where $\widehat{\xi}(k)$ is the test statistic and $\widehat{\zeta}_1(k)$ is its estimated limit (6.6) with $c = 0$. For example, the ratio below 2 would correspond to $\widehat{\xi}(k) \leq 2\widehat{\zeta}_1(k) = \widehat{\zeta}_2(k)$. Note from the plots that, as expected, the ratios concentrate around 1 for $k = 2$, and are larger than 1 for $k = 0$ and 1. Though not completely formalized yet, the approach appears promising. Even an ad hoc choice of $c = 2$ in the test (ratio smaller than 3) leads to the powers of 49% ($n = 250$) and 85% ($n = 500$) when testing for $k = 1$, and the sizes of 4% ($n = 250$) and 5% ($n = 500$) when testing for $k = 2$.

9 Conclusions

In this paper, we considered inference on the rank of a symmetric, semidefinite matrix having an asymptotically normal symmetric matrix estimator. We made an important distinction between indefinite and semidefinite estimators. A number of examples of interest served as a guide throughout.

For indefinite matrix estimators, a well-developed framework of Donald et al. (2010) is expected to apply. How to construct indefinite estimators, however, may not be evident in all the situations of interest: we produced such an estimator in the example of multiple index models but not in several other examples (spectral density matrix, and covariance matrix in heteroscedastic regression).

In contrast, the case of semidefinite matrix estimators seems to be much more delicate. We argued that the tests suggested in the literature are not suitable (sometimes contrary to what has been claimed). We also explored a new test based on additional asymptotic assumptions. The approach appears promising but further work is needed to deliver a fully satisfactory test.

A Technical proofs

Proof of Proposition 3.1: Observe that

$$\widehat{\Sigma}(x) - \Sigma(x) = \widehat{\Sigma}_1(x) + \widehat{\Sigma}_2(x) + \widehat{\Sigma}_3(x) + \widehat{\Sigma}_3(x)' + \widehat{\Sigma}_4(x),$$

where

$$
\begin{aligned}
\widehat{\Sigma}_1(x) &= \frac{1}{n}\sum_{k=1}^{n}(\epsilon_k\epsilon_k' K_h(x-x_k) - E\epsilon_k\epsilon_k' K_h(x-x_k)),\\
\widehat{\Sigma}_2(x) &= E\epsilon_k\epsilon_k' K_h(x-x_k) - \Sigma(x),\\
\widehat{\Sigma}_3(x) &= \frac{1}{n}\sum_{k=1}^{n}\epsilon_k x_k' K_h(x-x_k)\left(\frac{1}{n}\sum_{k=1}^{n}x_k x_k'\right)^{-1}\frac{1}{n}\sum_{k=1}^{n}x_k\epsilon_k',\\
\widehat{\Sigma}_4(x) &= \frac{1}{n}\sum_{k=1}^{n}\epsilon_k x_k'\left(\frac{1}{n}\sum_{k=1}^{n}x_k x_k'\right)^{-1}\left(\frac{1}{n}\sum_{k=1}^{n}x_k x_k' K_h(x-x_k)\right)\left(\frac{1}{n}\sum_{k=1}^{n}x_k x_k'\right)^{-1}\frac{1}{n}\sum_{k=1}^{n}x_k\epsilon_k'.
\end{aligned}
$$

By using the assumptions of the proposition, note that $\widehat{\Sigma}_3(x) = O_p((1/\sqrt{nh^q})(1/\sqrt{n})) = O_p(1/nh^{q/2})$, $\widehat{\Sigma}_4(x) = O_p((1/\sqrt{n})(1/\sqrt{n})) = O_p(1/n)$. Using the properties of the kernel function K and $\Sigma(x)$, we also have $\widehat{\Sigma}_2(x) = O(h^r)$. It is then enough to show that, with $W(x)$ in (3.5),

$$\sqrt{nh^q}\mathrm{vec}(\widehat{\Sigma}_1(x)) \xrightarrow{d} \mathcal{N}(0, W(x)),$$

or, equivalently, that $\sqrt{nh^q}a'\mathrm{vec}(\widehat{\Sigma}_1(x)) \xrightarrow{d} \mathcal{N}(0, a'W(x)a)$ for any $a \in \mathbb{R}^{p^2}$.

Note that

$$a'\mathrm{vec}(\widehat{\Sigma}_1(x)) = \frac{1}{n}\sum_{k=1}^{n}(a'(\epsilon_k\otimes\epsilon_k)K_h(x-x_k) - Ea'(\epsilon_k\otimes\epsilon_k)K_h(x-x_k))$$

and

$$
\begin{aligned}
&nh^q E(a'\mathrm{vec}(\widehat{\Sigma}_1(x)))^2 = h^q E(a'(\epsilon_k\otimes\epsilon_k)K_h(x-x_k) - Ea'(\epsilon_k\otimes\epsilon_k)K_h(x-x_k))^2\\
&= \|K\|_2^2 E(a'(\epsilon_k\otimes\epsilon_k))^2 K_{2,h}(x-x_k) - h^q(Ea'(\epsilon_k\otimes\epsilon_k)K_h(x-x_k))^2 \to a'W(x)a,
\end{aligned}
$$

where $K_{2,h}(x) = K(x/h)^2/h^q\|K\|_2^2$. We may suppose without loss of generality that $a'W(x)a > 0$ (in case of $a'W(x)a = 0$, the result is trivial). By Lindeberg-Feller central limit theorem and its Lyapunov condition, it is enough to show that

$$\frac{(nh^q)^{\frac{2+\delta}{2}}}{n^{1+\delta}}E|a'(\epsilon_k\otimes\epsilon_k)K_h(x-x_k)|^{2+\delta} \to 0.$$

Note that this is equivalent to

$$\frac{1}{(nh^q)^{\delta/2}}E|a'(\epsilon_k\otimes\epsilon_k)|^{2+\delta}K_{2+\delta,h}(x-x_k) \to 0,$$

where $K_{2+\delta,h}(x) = |K(x/h)|^{2+\delta}/h^q\|K\|_{2+\delta}^{2+\delta}$, and the latter convergence holds by the assumptions of the proposition. □

The following are basic assumptions for Proposition 3.5. Let $G(x) = F(x)p(x)$, where $p(x)$ is the density of x_i. Suppose $p(x)$ is bounded away from zero on a convex, open, bounded set U of $\mathbb{R}^q$. Let also $H_{\nu+1}(U)$ consist of functions which have partial derivatives up to order ν satisfying the global Lipschitz condition in the sense of Samarov (1993), p. 839. Suppose ν is an integer such that $\nu \geq q+4$.

(C1) Partial derivatives of p and G up to the order $\nu+1$ are bounded and $p, G \in H_{\nu+2}(U)$.

(C2) $n^{1/2}h^{\nu+1} \to 0$, $n^{1/2}h^{q+4} \to \infty$, as $n \to \infty$.

(C3) Suppose kernel function $K(x)$ is a bounded continuous function with support in the unit cube $\{\|x\| \leq 1\}$ and such that $K(x) = K(-x)$, $\int K(x)dx = 1$, and $\int K(x)x^l dx = 0$, for $l = 1, \dots, \nu$, where $x^l = x_1^{l_1} \dots x_q^{l_q}$ for $x = (x_1, \dots, x_q)' \in \mathbb{R}^q$ and $l_1 + \cdots + l_q = l$ with nonnegative integers l_i.

(C4) W in (3.38) is finite.

Proof of Proposition 3.5: We briefly outline the proof of Theorem 1 in Samarov (1993), p. 845, as adapted to the problem here. Note that

$$\widehat{T}_n := \text{vec}(\widehat{M_S}) = \frac{1}{n}\sum_{i=1}^n \widehat{p}(x_i)^{-1}(\widehat{G}(x_i) \otimes \widehat{G}(x_i)).$$

It is shown first using Taylor expansion that $\widehat{T}_n - V_1 - V_2 = o_p(n^{-1/2})$, where

$$\begin{aligned}
V_1 &= \frac{1}{n}\sum_{i=1}^n \bar{p}(x_i)^{-1}(\bar{G}(x_i) \otimes \bar{G}(x_i)), \\
V_2 &= -\frac{1}{n}\sum_{i=1}^n \bar{p}(x_i)^{-2}(\bar{G}(x_i) \otimes \bar{G}(x_i))(\widehat{p}(x_i) - \bar{p}(x_i)) \\
&\quad + \frac{1}{n}\sum_{i=1}^n \bar{p}(x_i)^{-1}((\bar{G}(x_i) \otimes I_p) + (I_p \otimes \bar{G}(x_i)))(\widehat{G}(x_i) - \bar{G}(x_i))
\end{aligned}$$

and $\bar{p}(x) = EK_h(x - x_i)$, $\bar{G}(x) = E(y_i K_h(x - x_i))$. Furthermore, $V_1 - V_3 = o_p(n^{-1/2})$ and $V_2 - V_4 = o_p(n^{-1/2})$, where

$$\begin{aligned}
V_3 &= \frac{1}{n}\sum_{i=1}^n p(x_i)^{-1}(G(x_i) \otimes G(x_i)), \\
V_4 &= -\frac{1}{n}\sum_{i=1}^n p(x_i)^{-2}(G(x_i) \otimes G(x_i))(\widehat{p}(x_i) - \bar{p}(x_i)) \\
&\quad + \frac{1}{n}\sum_{i=1}^n p(x_i)^{-1}((G(x_i) \otimes I_p) + (I_p \otimes G(x_i)))(\widehat{G}(x_i) - \bar{G}(x_i)).
\end{aligned}$$

The next step is to approximate the U-statistic V_4 through its projection V_5 defined by

$$V_5 = \frac{1}{n}\sum_{j=1}^n (t_1(x_j, y_j) - t_2(x_j)),$$

where

$$\begin{aligned}
t_1(x, y) &= Ep(x_i)^{-1}(G(x_i) \otimes I_p) + (I_p \otimes G(x_i))(yK_h(x - x_i) - \bar{G}(x_i)), \\
t_2(x) &= Ep(x_i)^{-2}(G(x_i) \otimes G(x_i))(K_h(x - x_i) - \bar{p}(x_i)).
\end{aligned}$$

Furthermore, one can show that $V_5 - \widetilde{V}_5 = o_p(n^{-1/2})$, where

$$\widetilde{V}_5 = \frac{1}{n}\sum_{j=1}^{n}(\widetilde{t}_1(x_j, y_j) - \widetilde{t}_2(x_j))$$

and

$$\begin{aligned} \widetilde{t}_1(x,y) &= (G(x)\otimes I_p + I_p \otimes G(x))y - E(G(x_i)\otimes I_p + I_p\otimes G(x_i))y_i, \\ \widetilde{t}_2(x) &= p(x)^{-1}(G(x)\otimes G(x)) - Ep(x_i)^{-1}(G(x_i)\otimes G(x_i)). \end{aligned}$$

Gathering all these observations and noting that $V_3 - \sum_{j=1}^{n}\widetilde{t}_2(x_j) = \text{vec}(M)$, one concludes that

$$\text{vec}(M_S) - \text{vec}(M) = \frac{1}{n}\sum_{i=1}^{n}\widetilde{t}_1(x_i, y_i) + o_p(n^{-1/2}).$$

The asymptotic normality result now follows from the central limit theorem. $\square$

PROOF OF PROPOSITION 4.2: Let Q be an orthogonal matrix as in (2.7) and $D = \text{diag}\{\upsilon_1, \ldots, \upsilon_p\}$. Then,

$$\sqrt{N}(\text{vec}(Q'\widehat{M}Q) - \text{vec}(D)) \xrightarrow{d} \mathcal{N}(0, \widetilde{W}),$$

where $\widetilde{W} = (Q\otimes Q)W(Q'\otimes Q')$ has the same rank as W. If $Q = (Q_1\ Q_2)$ as in (2.7), then $\sqrt{N}\text{vec}(Q_1'\widehat{M}Q_1)$ is also asymptotically normal. By Lemma A.1 below, the only way this can happen is when $\sqrt{N}\text{vec}(Q_1'\widehat{M}Q_1) \xrightarrow{d} 0$. Since $Q_1'\widehat{M}Q_1$ is of dimension $(p-r)^2$, this means that $(p-r)^2$ elements of $\sqrt{N}\text{vec}(Q'\widehat{M}Q)$ are asymptotically 0. Hence, $\text{rk}\{\widetilde{W}\} \le p^2 - (p-r)^2 = r(2p-r)$. $\square$

The following elementary lemma was used in the proof above.

Lemma A.1 *If X_N is symmetric, semidefinite matrix such that $\sqrt{N}\text{vec}(X_N) \xrightarrow{d} \mathcal{N}(0,Z)$, then $Z \equiv 0$.*

PROOF: The assumption can be written as $\sqrt{N}X_N \xrightarrow{d} \mathcal{X}$, where $\mathcal{X}$ is a normal (Gaussian) matrix. Since X_N is, say, positive semidefinite, it follows that $a'\mathcal{X}a \ge 0$ a.s. for any vector a. The result now follows from another elementary lemma next. $\square$

Lemma A.2 *Let $\mathcal{X}$ be a symmetric, normal (Gaussian) matrix. If all eigenvalues of $\mathcal{X}$ are nonnegative, then $\mathcal{X} = 0$.*

PROOF: If $\mathcal{X}$ is symmetric and its eigenvalues are nonnegative, then it is positive semidefinite. In particular,

$$\sum_{i,j=1}^{m} a_i x_{ij} a_j \ge 0 \tag{A.1}$$

for all a_i, where $\mathcal{X} = (x_{ij})_{i,j=1,\ldots,m}$. Taking $a_i = 1$, $a_j = 0$, $j \ne i$, leads to $x_{ii} \ge 0$ and $a_i = 1$, $a_j = 1$, $a_k = 0$, $k \ne i, j$, leads to $x_{ij} + x_{ji} = 2x_{ij} \ge 0$. Since x_{ij} are all normal, this can happen only when $x_{ij} \equiv 0$, or $\mathcal{X} = 0_{m\times m}$. $\square$

Proof of Theorem 5.1: Arguing as in the proofs of Theorems 4.3 and 4.5 in Donald et al. (2007) (see also Eaton and Tyler (1991)), $\sqrt{N}\widehat{\upsilon}_k$, $k = 1, \dots, p - r$, are asymptotically the ordered eigenvalues of $Q_2'\sqrt{N}(\widehat{M} - M)Q_2$ or $Q_2'\mathcal{Y}Q_2$, where $Q = (Q_1\ Q_2)$ is the matrix in the proof of Proposition 4.2 above and $\mathcal{Y}$ appears in (2.3). Since $\sqrt{N}\widehat{\upsilon}_k \geq 0$, by using Lemma A.1, this can happen only when $Q_2'\mathcal{Y}Q_2 = 0$. This leads to (5.1). The relation (5.2) follows from $\widehat{\upsilon}_k \xrightarrow{p} \upsilon_k > 0$, $k = p - r + 1, \dots, p$. □

Proof of Proposition 6.1: As in the proof of Proposition 3.1, $h^{-2}Q_1'\widehat{\Sigma}(x)Q_1 = h^{-2}Q_1'(\widehat{\Sigma}_1(x) + \widehat{\Sigma}_2(x))Q_1 + O_p(h^{-2}/nh^{q/2})$. We have $h^{-2}/nh^{q/2} = 1/nh^{q/2+2} \to 0$ by the assumption that $nh^{q/2+2} \to \infty$. For the first term, on the other hand, note that

$$I_0 := \mathrm{vec}(h^{-2}Q_1'(\widehat{\Sigma}_1(x) + \widehat{\Sigma}_2(x))Q_1)$$

$$= \frac{h^{-2}}{n}\sum_{k=1}^{n}\mathrm{vec}(Q_1'\epsilon_k\epsilon_k'Q_1)K_h(x - x_k) = \frac{h^{-2}}{n}\sum_{k=1}^{n}\mathrm{vec}(Q_1'\bar{\epsilon}_k\bar{\epsilon}_k'Q_1)K_h(x - x_k),$$

where $\bar{\epsilon}_k = \bar{f}(x_k)\eta_k$ and $\bar{f}(x_k) = f(x_k) - f(x)$. Then,

$$I_0 = (Q_1' \otimes Q_1)\frac{h^{-2}}{n}\sum_{k=1}^{n}\mathrm{vec}(\bar{\epsilon}_k\bar{\epsilon}_k')K_h(x - x_k)$$

$$= (Q_1' \otimes Q_1)\frac{h^{-2}}{n}\sum_{k=1}^{n}(\mathrm{vec}(\bar{\epsilon}_k\bar{\epsilon}_k')K_h(x - x_k) - E\mathrm{vec}(\bar{\epsilon}_k\bar{\epsilon}_k')K_h(x - x_k))$$

$$+(Q_1' \otimes Q_1)h^{-2}E\mathrm{vec}(\bar{\epsilon}_k\bar{\epsilon}_k')K_h(x - x_k) =: (Q_1' \otimes Q_1)I_1 + (Q_1' \otimes Q_1)I_2.$$

It is then enough to show that $I_1 = o_p(1)$ and that $I_2 \xrightarrow{p} \Pi(x)$.

For $I_1 = o_p(1)$, it is enough to show that $EI_1'I_1 = o(1)$ or, since $I_1'I_1 = \mathrm{trace}(I_1I_1')$, that $EI_1I_1' = o(1)$. Under the assumptions of the proposition, note that $\mathrm{vec}(\bar{\epsilon}_k\bar{\epsilon}_k') = p(x_k)^{-1}\mathrm{vec}(\bar{f}(x_k)\eta_k\eta_k'\bar{f}(x_k)') = p(x_k)^{-1}(\bar{f}(x_k) \otimes \bar{f}(x_k))\mathrm{vec}(\eta_k\eta_k')$ and $E\mathrm{vec}(\eta_k\eta_k')\mathrm{vec}(\eta_k\eta_k')' = E\eta^4 \cdot I$. This yields

$$EI_1I_1' = \frac{h^{-4}}{n}\Big(Ep(x_k)^{-2}(\bar{f}(x_k) \otimes \bar{f}(x_k))(\bar{f}(x_k)' \otimes \bar{f}(x_k)')K_h^2(x - x_k)\ E\eta^4 - (EI_2)^2\Big)$$

$$= \frac{h^{-4}\|K\|_2^2E\eta^4}{nh^q}Ep(x_k)^{-2}(\bar{f}(x_k)\bar{f}(x_k)' \otimes \bar{f}(x_k)\bar{f}(x_k)')K_{2,h}(x - x_k) - \frac{h^{-4}}{n}(EI_2)^2,$$

where $K_2(u) = K^2(u)/\|K\|_2^2$. As will be shown below $h^{-2}EI_2 = O(1)$, so that the second term above is asymptotically negligible. For the first term, its entry consists of the sum of terms of the general form

$$\frac{h^{-4}}{nh^q}Ep(x_k)^{-2}\bar{f}_{i_1i_2}(x_k)\bar{f}_{i_3i_4}(x_k)\bar{f}_{j_1j_2}(x_k)\bar{f}_{j_3j_4}(x_k)K_{2,h}(x - x_k)$$

$$= \frac{h^{-4}}{nh^q}\int_{\mathcal{H}_x}p(y)^{-1}\bar{f}_{i_1i_2}(y)\bar{f}_{i_3i_4}(y)\bar{f}_{j_1j_2}(y)\bar{f}_{j_3j_4}(y)K_{2,h}(x - y)dy$$

$$= \frac{1}{nh^q}\int_{\mathcal{H}_x}p(y)^{-1}\frac{\partial f_{i_1i_2}}{\partial y}(y^*_{i_1i_2})(\frac{y - x}{h})\frac{\partial f_{i_3i_4}}{\partial y}(y^*_{i_3i_4})(\frac{y - x}{h})\times$$

$$\times\frac{\partial f_{j_1j_2}}{\partial y}(y^*_{j_1j_2})(\frac{y - x}{h})\frac{\partial f_{j_3j_4}}{\partial y}(y^*_{j_3j_4})(\frac{y - x}{h})K_{2,h}(x - y)dy$$

$$= \frac{1}{nh^q} \int_{\mathbb{R}^q} p(x)^{-1} \frac{\partial f_{i_1 i_2}}{\partial y}(x) z \frac{\partial f_{i_3 i_4}}{\partial y}(x) z \frac{\partial f_{j_1 j_2}}{\partial y}(x) z \frac{\partial f_{j_3 j_4}}{\partial y}(x) z K_2(z) dz + o(\frac{1}{nh^q}),$$

by the mean value theorem, where y^* is a point in $[x,y]$ (or $[y,x]$). Hence, $EI_1 I_1' = o(1)$ since $nh^q \to \infty$ by assumption.

It remains to show that $I_2 \xrightarrow{p} \Pi(x)$. By arguing as above, note that

$$I_2 = h^{-2} E \text{vec}(\bar{\epsilon}_k \bar{\epsilon}_k) K_h(x - x_k) = h^{-2} E(\bar{f}(x_k) \otimes \bar{f}(x_k)) \text{vec}(\eta_k \eta_k') p(x_k)^{-1} K_h(x - x_k)$$

$$= \text{vec}(h^{-2} E \bar{f}(x_k) \bar{f}(x_k)' p(x_k)^{-1} K_h(x - x_k))$$

$$= \text{vec}\Big(h^{-2} E \Big(\sum_{m=1}^{p} \bar{f}_{im}(x_k) \bar{f}_{mj}(x_k) p(x_k)^{-1} K_h(x - x_k) \Big)_{i,j=1,\dots,p} \Big)$$

$$\to \text{vec}\Big(\Big(\int_{\mathbb{R}^q} \sum_{m=1}^{p} \frac{\partial f_{im}}{\partial x}(x) z \frac{\partial f_{mj}}{\partial x}(x) z \, K(z) dz \Big)_{i,j=1,\dots,p} \Big)$$

$$= c_0 \text{vec}\Big(\text{tr}\Big(\sum_{m=1}^{p} \frac{\partial f_{im}}{\partial x}(x) \frac{\partial f_{mj}}{\partial x}(x)' \Big)_{i,j=1,\dots,p} \Big) = \text{vec}\Big((\Pi(x)_{ij})_{i,j=1,\dots,p} \Big).$$

□

References

Andrews, D. W. K. (1987), 'Asymptotic results for generalized Wald tests', *Econometric Theory* **3**(3), 348–358.

Brillinger, D. R. (1975), *Time Series*, Holt, Rinehart and Winston, Inc., New York. Data analysis and theory, International Series in Decision Processes.

Camba-Mendez, G. & Kapetanios, G. (2001), 'Testing the rank of the Hankel covariance matrix: a statistical approach', *IEEE Transactions on Automatic Control* **46**(2), 331–336.

Camba-Mendez, G. & Kapetanios, G. (2005), 'Estimating the rank of the spectral density matrix', *Journal of Time Series Analysis* **26**(1), 37–48.

Camba-Mendez, G. & Kapetanios, G. (2009), 'Statistical tests and estimators of the rank of a matrix and their applications in econometric modelling', *Econometric Reviews* **28**(6), 581–611.

Cragg, J. G. & Donald, S. G. (1996), 'On the asymptotic properties of LDU-based tests of the rank of a matrix', *Journal of the American Statistical Association* **91**(435), 1301–1309.

Cragg, J. G. & Donald, S. G. (1997), 'Inferring the rank of a matrix', *Journal of Econometrics* **76**(1-2), 223–250.

Donald, S. G. (1997), 'Inference concerning the number of factors in a multivariate nonparametric relationship', *Econometrica* **65**(1), 103–131.

Donald, S. G., Fortuna, N. & Pipiras, V. (2007), 'On rank estimation in symmetric matrices: the case of indefinite matrix estimators', *Econometric Theory* **23**(6), 1217–1232.

Donald, S. G., Fortuna, N. & Pipiras, V. (2010), On rank estimation in semidefinite matrices, CEF.UP Working Paper 2, Faculdade de Economia do Porto, Porto, Portugal.

Donkers, B. & Schafgans, M. (2008), 'Specification and estimation of semiparametric multiple-index models', *Econometric Theory* **24**(6), 1584–1606.

Eaton, M. L. & Tyler, D. E. (1991), 'On Wielandt's inequality and its application to the asymptotic distribution of the eigenvalues of a random symmetric matrix', *The Annals of Statistics* **19**(1), 260–271.

Fortuna, N. (2008), 'Local rank tests in a multivariate nonparametric relationship', *Journal of Econometrics* **142**(1), 162–182.

Gill, L. & Lewbel, A. (1992), 'Testing the rank and definiteness of estimated matrices with applications to factor, state-space and ARMA models', *Journal of the American Statistical Association* **87**(419), 766–776.

Hannan, E. J. (1970), *Multiple Time Series*, John Wiley and Sons, Inc., New York-London-Sydney.

Hayashi, F. (2000), *Econometrics*, Princeton University Press, Princeton, NJ.

Jolliffe, I. T. (2002), *Principal Component Analysis*, Springer Series in Statistics, second edn, Springer-Verlag, New York.

Kleibergen, F. & Paap, R. (2006), 'Generalized reduced rank tests using the singular value decomposition', *Journal of Econometrics* **133**(1), 97–126.

Kudô, A. (1963), 'A multivariate analogue of the one-sided test', *Biometrika* **50**, 403–418.

Kudô, A. & Choi, J. R. (1975), 'A generalized multivariate analogue of the one sided test', *Memoirs of the Faculty of Science. Kyushu University. Series A. Mathematics* **29**(2), 303–328.

Lütkepohl, H. & Burda, M. M. (1997), 'Modified Wald tests under nonregular conditions', *Journal of Econometrics* **78**(2), 315–332.

Maddala, G. S. & Kim, I.-M. (1998), *Unit Roots, Cointegration and Structural Change*, Themes in Modern Econometrics, Cambridge University Press, Cambridge.

Magnus, J. R. & Neudecker, H. (1999), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons Ltd., Chichester. Revised reprint of the 1988 original.

Pagan, A. & Ullah, A. (1999), *Nonparametric Econometrics*, Themes in Modern Econometrics, Cambridge University Press, Cambridge.

Ratsimalahelo, Z. (2002), Rank test based on matrix perturbation theory, Preprint.

Robin, J.-M. & Smith, R. J. (2000), 'Tests of rank', *Econometric Theory* **16**(2), 151–175.

Samarov, A. M. (1993), 'Exploring regression structure using nonparametric functional estimation', *Journal of the American Statistical Association* **88**(423), 836–847.

Zhang, F., ed. (2005), *The Schur Complement and its Applications*, Vol. 4 of *Numerical Methods and Algorithms*, Springer-Verlag, New York.

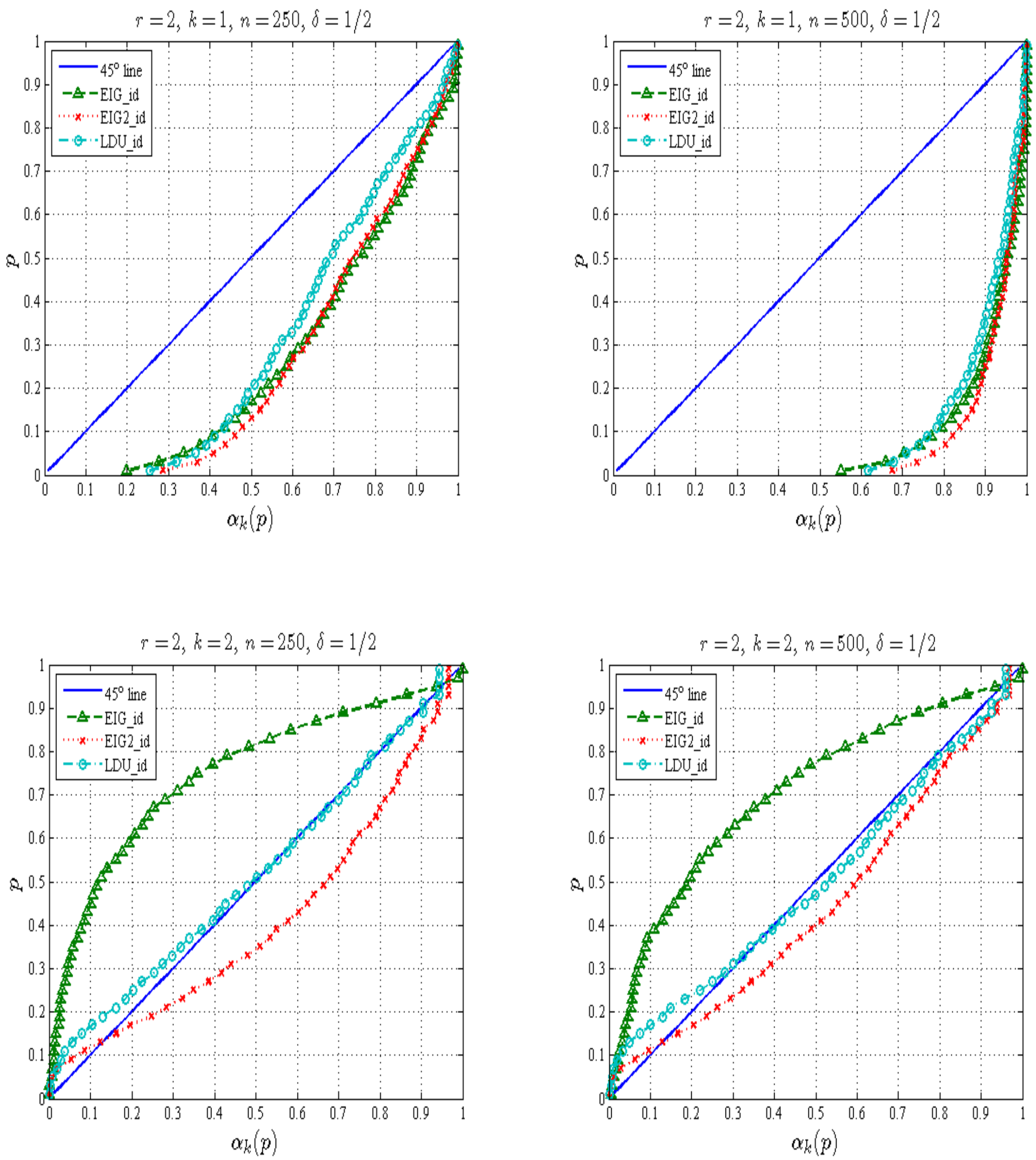

Figure 1: PP-plots for EIG_id, EIG2_id and LDU_id tests.

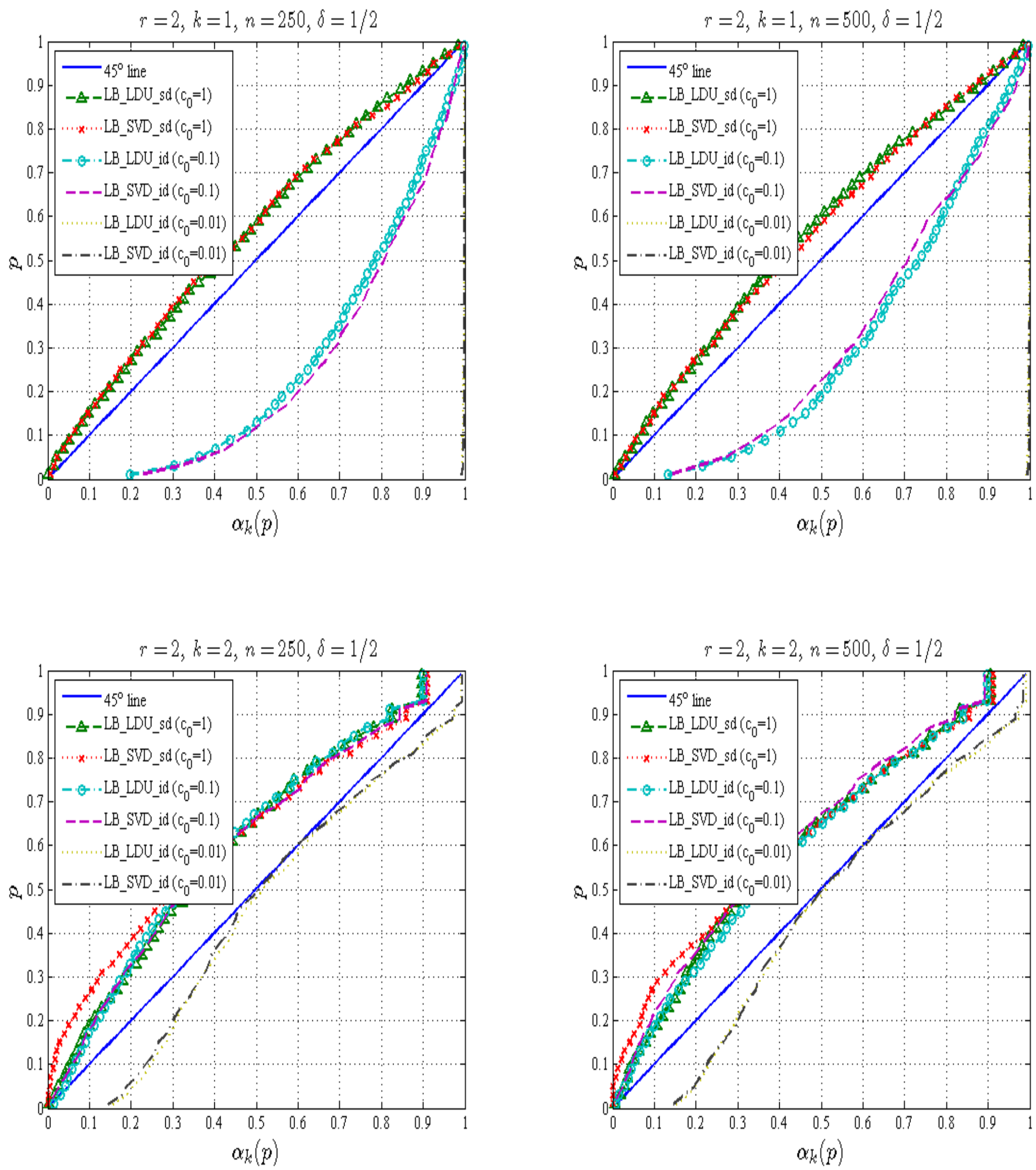

Figure 2: PP-plots for LB_LDU_sd and LB_SVD_sd tests.

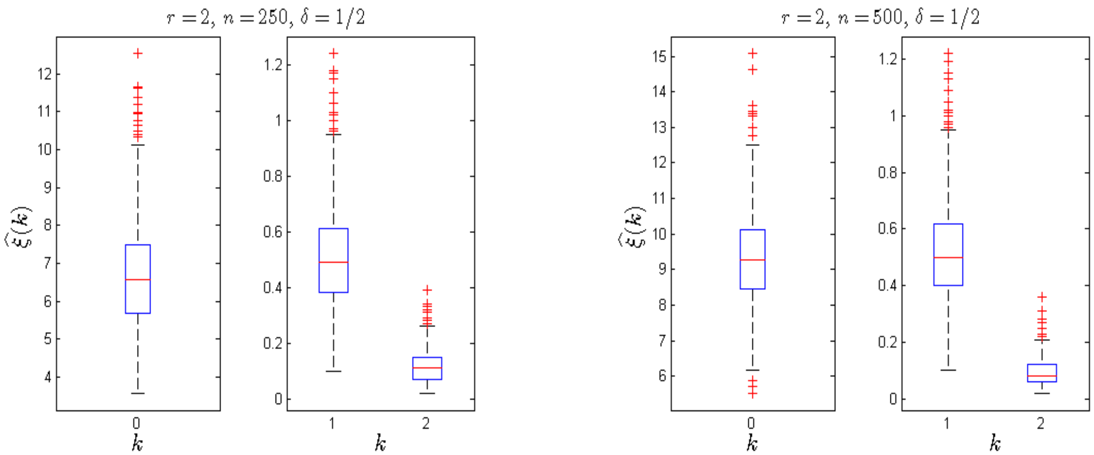

Figure 3: Boxplots of the test statistic $\widehat{\xi}(k)$ in the EIG_sd test.

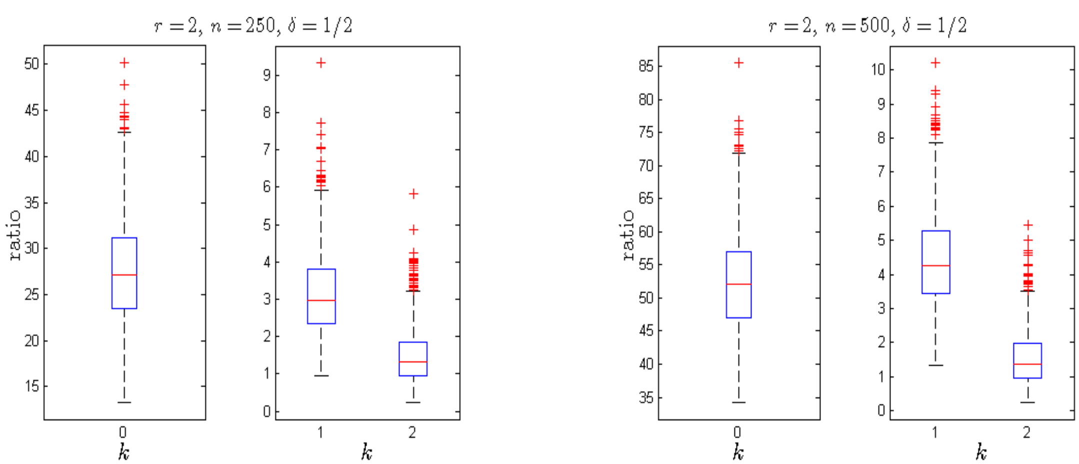

Figure 4: Boxplots of the ratios $\widehat{\xi}(k)/\widehat{\zeta}(k)$ in the EIGf_sd test.