cef.up
Centro de Economia e Finanças da UPorto

CEF.UP Working Paper
2010-02

On Rank Estimation in Semidefinite Matrices

Stephen G. Donald
Natércia Fortuna
Vladas Pipiras

Centro de Economia e Finanças da UP
www.fep.up.pt

On rank estimation in semidefinite matrices

Stephen G. Donald*
University of Texas at Austin

Natércia Fortuna†
CEF.UP, Universidade do Porto

Vladas Pipiras‡
University of North Carolina at Chapel Hill

April 23, 2010

Abstract

This work concerns the problem of rank estimation in semidefinite matrices, having either indefinite or semidefinite matrix estimator satisfying a typical asymptotic normality condition. Several rank tests are examined, based on either available rank tests 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 asymptotic normality of matrix estimators. Several examples where rank estimation in semidefinite matrices is of interest are studied and serve as guide throughout the work.

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

JEL classification: C12, C13.

1 Introduction

Testing for and estimation of the rank $\text{rk}\{M\}$ 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, 2003) and Kleibergen and Paap (2006), and the characteristic root test of Robin and Smith (2000). The problem of rank estimation is reviewed in Camba-Méndez and Kapetanios (2005*b*, 2008) where several applications (for example, IV modeling, demand systems, cointegration) are also discussed.

A starting assumption in all tests above is having an asymptotically normal estimator $\widehat{M}$ of M in the sense that

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

*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* (corresponding author). CEF.UP - Centre for Economics and Finance at University of Porto - is supported by the Fundação para a Ciência e a Tecnologia (FCT), Portugal. Fortuna is grateful to FCT and Fundo Social Europeu (FSE), in the scope of Quadro Comunitário de Apoio III, for a financial support through the grant SFRH/BPD/26439/2006.

‡Department of Statistics and Operations Research, UNC at Chapel Hill, CB#3260, Hanes Hall, Chapel Hill, NC 27599, USA, *pipiras@email.unc.edu*

where N is the sample size or any other relevant parameter such that $N \to \infty$, $\to_d$ denotes convergence in distribution and vec is a standard matrix operation. The actual testing is for $H_0 : \mathrm{rk}\{M\} \le r$ (or $H_0 : \mathrm{rk}\{M\} = r$) against $H_1 : \mathrm{rk}\{M\} \ge r$, and $\mathrm{rk}\{M\}$ itself is estimated by a number of available methods, for example, sequential testing. The limiting covariance matrix W in (1.1) is often assumed to be nonsingular, though some departures from this assumption are also present in, for example, Robin and Smith (2000), Camba-Méndez and Kapetanios (2005*b*, 2008).

In this work, we are interested in testing for the rank of a $p \times p$ matrix M when the matrix M is (for example, 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 matrix operation on symmetric matrices. Two different cases need to be distinguished regarding (1.2), namely those of:

$$\begin{aligned} &\text{Case 1} \quad : \quad \text{indefinite matrix estimators } \widehat{M}, \\ &\text{Case 2} \quad : \quad \text{(semi)definite matrix estimators } \widehat{M}. \end{aligned} \tag{1.3}$$

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)). We focus in this paper on Case 2. Though it seems separate from Case 1, the two cases are, in fact, related, for example, in flexibility of choosing indefinite or semidefinite matrix estimator (Sections 8).

In addition to general interest in the problem, we were also motivated by the following. We were interested in understanding several cases of rank estimation in semidefinite matrices previously considered in the literature (Section 4), and several tests already proposed for the problem (Section 5).

Before proceeding further, it is important to make the following simple observation regarding (1.1) or (1.2) in Case 2. It sheds light on where semidefiniteness comes into play in (1.1) or (1.2). Consider, for example, the case $p = 1$, that is, $\dim\{M\} = 1$. Semidefiniteness of $\widehat{M}$ now means that $\widehat{M} \ge 0$, and rank deficiency of M translates into $\mathrm{rk}\{M\} = 0$ or $M = 0$. Under $\mathrm{rk}\{M\} = 0$, (1.1) or (1.2) becomes $\sqrt{N}\widehat{M} \to_d \mathcal{N}(0, \sigma^2)$ and the only way this can happen is when

$$\sqrt{N}\widehat{M} \xrightarrow{d} 0 \tag{1.4}$$

($\to_d$ can be replaced by $\to_p$). Thus, $\widehat{M}$ is estimated "too efficiently", borrowing the term used in Lütkepohl and Burda (1997).

The structure of the paper is the following. Some basic general notation is introduced in Section 2. Several examples where rank in semidefinite matrices might be of interest, are given in Sections 3 and 4. These examples will be used for guidance throughout the paper. When matrix W or W_0 is singular in (1.1) or (1.2), respectively, one may and has adapted rank tests previously available in the literature for the case of nonsingular W or W_0. For example, inverses involving W or W_0 in available rank tests could be tried to be replaced by generalized inverses. This and other situations are discussed in Section 5. Under the assumption (1.1) or (1.2), asymptotics of eigenvalues of $\widehat{M}$ are established in Section 6. This leads to consistent though not satisfactory rank tests from a practical perspective. In Section 7, finer rank tests are discussed for specific examples and involve finer asymptotics than those in (1.1) or (1.2). For example, in 1-dimension discussed around (1.4), one may ask whether a rate faster than $\sqrt{N}$ leads to nondegenerate limit. Alternatively, one may attempt to start with indefinite matrix estimator of M in the first place. This is discussed in Section 8.

2 Some notation and other preliminaries

Here is some basic general notation that will be used throughout the paper. As in Section 1, M is an unknown, $p \times p$, 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. Above and throughout, $\to_d$ and $\to_p$ stand for convergence in distribution and probability, respectively. The rank of a matrix A is denoted by $\mathrm{rk}\{A\}$.

To incorporate the example of spectral density matrices, we will 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 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

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

Finally, for later use, let $Q = (Q_1\ Q_2)$ be an orthogonal (unitary) matrix such that

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

where

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

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

3 Several uninteresting but instructive examples

Here are several examples showing that not every case of rank estimation in semidefinite matrices is of interest. The examples are instructive for cases of interest considered in the next section.

Example 3.1 (Rank of covariance matrix of error terms in linear regression.) 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 an unknown $p \times q$ matrix, x_i are $q \times 1$, i.i.d. variables and ϵ_i are zero mean, i.i.d. error terms, independent of x_i, with

$$E\epsilon_i \epsilon_i' = \Sigma. \tag{3.2}$$

One problem to consider might appear the estimation of rank $\mathrm{rk}\{\Sigma\}$, having a positive definite matrix estimator

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

where $\widehat{\beta}$ is, for example, the least squares estimator of β. This problem, however, is obviously not that meaningful. If Σ has lower than full rank, then there is an orthogonal matrix Q such that

$$EQ\epsilon_i (Q\epsilon_i)' = \mathrm{diag}\{0, \dots, 0, \alpha_{p-r+1}, \dots, \alpha_p\}, \tag{3.4}$$

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

$$(Q(y_i - \beta x_i))_k = 0 \quad \text{a.s.}, \quad k = 1, \dots, p - r, \tag{3.5}$$

where $(z)_k$ indicates the k-th component of a vector z. The exact linear relationship between y and x in (3.5) could, in principle, be first checked with data and if found, eliminated by reducing p. Thus, one can suppose without loss of generality that Σ is of full rank.

Remark 3.1 A similar but simpler situation is to consider $y_i = \epsilon_i$, $i = 1, \dots, n$, with $E\epsilon_i \epsilon_i' = \Sigma$ and $\widehat{\Sigma} = n^{-1} \sum_{k=1}^{n} y_k y_k'$. Supposing $\mathrm{rk}\{\Sigma\} < p$ is not meaningful for the same reasons as in Example 3.1. It should also be noted that similar discussion can be found, for example, in connection to principal components (where it is not meaningful to consider principal components with zero variance). See, for example, p. 27 in Jolliffe (2002).

Example 3.2 (Some cases of reduced-rank spectral density matrices.) Though Example 3.1 is elementary, its variations appear in the literature. For example, Camba-Méndez and Kapetanios (2005*a*), p. 38, consider an example where it might be of interest to say that

$$r = \sup_{w} \mathrm{rk}\{\Sigma(w)\} < p, \tag{3.6}$$

where

$$\Sigma(w) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} Ex_k x_0' e^{-iwk}, \quad w \in (-\pi, \pi], \tag{3.7}$$

is the spectral density matrix of a stationary, say zero mean, p-multivariate time series $\{x_k\}_{k\in\mathbb{Z}}$. Under mild assumptions (e.g. $\sum_k |Ex_k x_0'| < \infty$), the spectral density matrix is Hermitian, positive semidefinite and with complex-valued entries in general (see, for example, Brillinger (1975), Hannan (1970)). If (3.6) holds, then, for all w and some unitary matrix Q,

$$Q\Sigma(w)Q^* = \mathrm{diag}\{0,\ldots,0,\alpha_{p-r+1},\ldots,\alpha_p\}, \tag{3.8}$$

where $\alpha_k \geq 0$. Then

$$\sum_{k=-\infty}^{\infty} E(Qx_0)(Qx_k)^* e^{-iwk} = \mathrm{diag}\{0,\ldots,0,\alpha_{p-r+1},\ldots,\alpha_p\}, \tag{3.9}$$

which yields, in particular, that

$$\sum_{k=-\infty}^{\infty} E(Qx_0)_l (Qx_k)_l^* e^{-iwk} = 0, \quad l = 1,\ldots,p-r, \tag{3.10}$$

(as in Example 3.1, $(z)_k$ indicates the k-th component of z) or that

$$(Qx_k)_l = 0 \quad \text{a.s.,} \quad \text{all} \quad k \in \mathbb{Z}, l = 1,\ldots,p-r. \tag{3.11}$$

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

Remark 3.2 In fairness to Camba-Méndez and Kapetanios (2005*a*), p. 38, these authors also discuss their example in light of time series factor model. This model cannot be dealt with the arguments of Example 3.2 above.

4 A number of interesting cases

In this section, we gather a number of examples where rank estimation in semidefinite matrices is of interest.

Example 4.1 (Linear regression with heteroscedastic error terms.) Consider again a linear regression model

$$y_i = \beta x_i + \epsilon_i, \quad i = 1,\ldots,n, \tag{4.1}$$

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 $p \times 1$. But suppose now that

$$E(\epsilon_i \epsilon_i' | x_i) = \Sigma(x_i) p(x_i)^{-1}, \tag{4.2}$$

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

$$\widehat{\Sigma}(x) = \frac{1}{n}\sum_{k=1}^{n}(y_k - \widehat{\beta}x_k)(y_k - \widehat{\beta}x_k)' K_h(x - x_k), \tag{4.3}$$

where $\widehat{\beta}$ is the least squares estimator of β and $K_h(x) = (1/h)K(x/h^q)$ 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 basic assumptions. We suppose that K is the kernel of order r in the sense that $\int K(x)dx = 1$, $\int x^j K(x)dx = 0$, $j = 1, \dots, r-1$, $\int x^r K(x)dx < \infty$, K is symmetric, and with bounded support. Suppose (x_i, ϵ_i) are i.i.d. vectors supported on $\mathcal{H} := \mathcal{H}_x \times \mathcal{H}_\epsilon := (a_x, b_x) \times (a_\epsilon, b_\epsilon)$ (possibly $a = \infty$, $b = -\infty$) and having density $p(x, \epsilon) > 0$ on $\mathcal{H}$. Let $C^m(\mathcal{H}_x)$ denote the functions on $\mathcal{H}_x$ whose m-th derivative is continuous on $\mathcal{H}_x$, and $\|K\|_2^2 = \int K(x)^2 dx$.

Proposition 4.1 *Consider the model (4.1)-(4.2) and suppose the assumptions above. Suppose also that* $\Sigma(x) \in C^r(\mathcal{H}_r)$, $p(x)$, $E(\epsilon_k\epsilon_k' \otimes \epsilon_k\epsilon_k' | x_k = x)$, $E(|\epsilon_k|^{4+2\delta} | x_k = x) \in C^0(\mathcal{H}_x)$ *with* $\delta > 0$ *and* $Ex_i x_i'$ *is invertible. Then, as* $n \to \infty$, $nh^q \to \infty$, $nh^{q+2r} \to 0$, $h \to 0$, *for* $x \in \mathcal{H}_x$,

$$\sqrt{Nh^q}(\mathrm{vec}(\widehat{\Sigma}(x)) - \mathrm{vec}(\Sigma(x))) \xrightarrow{d} \mathcal{N}(0, W(x)), \tag{4.4}$$

where

$$W(x) = \|K\|_2^2 p(x) E(\epsilon_k\epsilon_k' \otimes \epsilon_k\epsilon_k' | x_k = x). \tag{4.5}$$

Under the assumptions of Proposition 4.1, 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 $x_h = x$. 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

$$\mathrm{rk}\{W(x)\} \leq r^2 = \mathrm{rk}\{\Sigma(x)\}^2, \tag{4.6}$$

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 $\mathrm{rk}\{A \otimes B\} = \mathrm{rk}\{A\}\mathrm{rk}\{B\}$.

Example 4.2 (Spectral density matrices.) With the notation of Example 3.2 one may be interested in testing for

$$\mathrm{rk}\{\Sigma(w)\} \tag{4.7}$$

for *fixed* w. This is the problem considered in Camba-Méndez and Kapetanios (2005*a*). An important potential application of this is to cointegration. Recall that $\mathrm{rk}\{\Sigma(w)\}$ at $w = 0$ is p minus the cointegration rank of the p-multivariate 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

$$\widehat{\Sigma}(w) = \frac{1}{2m+1} \sum_{k=-m}^{m} \bar{\Sigma}\left(w + \frac{2\pi k}{n}\right) \tag{4.8}$$

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

$$\bar{\Sigma}(w)=\frac{1}{2\pi}\sum_{k=-(n-1)}^{n-1}\widehat{\Gamma}_k e^{-ikw},\quad \widehat{\Gamma}_k=\frac{1}{n}\sum_{t=1}^{n-|k|}x_t x'_{t+k}. \tag{4.9}$$

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

$$x_k=\sum_{j=-\infty}^{\infty}A_j\epsilon_{k-j},\quad k\in\mathbb{Z}, \tag{4.10}$$

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

$$\lim_{n\to\infty}\sup_w m^{1/2}|\Sigma(w)-E\widehat{\Sigma}(w)|=0. \tag{4.11}$$

Proposition 4.2 *With the above notation and assumptions, as $n,m\to\infty$, $m/n\to 0$,*

$$(2m)^{1/2}(\mathrm{vec}(\widehat{\Sigma}(w))-\mathrm{vec}(\Sigma(w)))\xrightarrow{d}\mathcal{N}^c(0,W^c(w)), \tag{4.12}$$

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

$$W^c(w)=\begin{cases}2\Sigma(w)_{iu}\Sigma(w)_{vj}, & \text{if}\quad w\neq 0,\pm\pi,\\ 2(\Sigma(w)_{iu}\Sigma(w)_{vj}+\Sigma(w)_{iv}\Sigma(w)_{uj}), & \text{if}\quad w=0,\pm\pi.\end{cases} \tag{4.13}$$

For example, when $p=2$, (4.13) becomes

$$W^c(w)=2\begin{pmatrix}\Sigma(w)_{11}\Sigma(w)_{11} & \Sigma(w)_{11}\Sigma(w)_{21} & \Sigma(w)_{12}\Sigma(w)_{11} & \Sigma(w)_{12}\Sigma(w)_{21}\\ \Sigma(w)_{21}\Sigma(w)_{11} & \Sigma(w)_{21}\Sigma(w)_{21} & \Sigma(w)_{22}\Sigma(w)_{11} & \Sigma(w)_{22}\Sigma(w)_{21}\\ \Sigma(w)_{11}\Sigma(w)_{12} & \Sigma(w)_{11}\Sigma(w)_{22} & \Sigma(w)_{12}\Sigma(w)_{12} & \Sigma(w)_{12}\Sigma(w)_{22}\\ \Sigma(w)_{21}\Sigma(w)_{12} & \Sigma(w)_{21}\Sigma(w)_{22} & \Sigma(w)_{22}\Sigma(w)_{12} & \Sigma(w)_{22}\Sigma(w)_{22}\end{pmatrix}=:2W_0^c(w), \tag{4.14}$$

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

$$W^c(w)=2(W_0^c(w)+\Sigma(w)\otimes\Sigma(w)), \tag{4.15}$$

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

Example 4.3 (Multiple index mean regression model.) Donkers and Schafgans (2003, 2005) consider the multiple index mean regression model

$$g(x):=E(y|x)=H(x'\beta_1,\dots,x'\beta_p), \tag{4.16}$$

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, and $\beta_1,\dots,\beta_p$ are unknown parameters. Let n be the sample size in $\{y_1,\dots,y_n,x_1,\dots,x_n\}$.

Of interest to our context is the matrix

$$M=E\left(w(x)\frac{\partial g(x)}{\partial x}\frac{\partial g(x)'}{\partial x}\right), \tag{4.17}$$

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

$$\widehat{M}_S = \frac{1}{n}\sum_{i=1}^{n} w(x_i)\left(\frac{\widehat{G}'(x_i)}{\widehat{f}(x_i)} - \frac{\widehat{G}(x_i)\widehat{f}'(x_i)}{\widehat{f}(x_i)^2}\right)\left(\frac{\widehat{G}'(x_i)}{\widehat{f}(x_i)} - \frac{\widehat{G}(x_i)\widehat{f}'(x_i)}{\widehat{f}(x_i)^2}\right)', \tag{4.18}$$

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

$$\widehat{f}^{(k)}(x_i) = \frac{1}{(n-1)h^{l+k}}\sum_{j=1, j\neq i}^{n} K^{(k)}\left(\frac{x_i - x_j}{h}\right), \tag{4.19}$$

$$\widehat{G}^{(k)}(x_i) = \frac{1}{(n-1)h^{l+k}}\sum_{j=1, j\neq i}^{n} K^{(k)}\left(\frac{x_i - x_j}{h}\right) y_j \tag{4.20}$$

(one obviously sets here $\widehat{f}^{(0)} = \widehat{f}$ and $\widehat{f}^{(1)} = \widehat{f}'$). Note that $\widehat{M}_S$ is positive semidefinite, and so is M.

Donkers and Schafgans (2003, 2005) suggest to estimate p in (4.16) 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 show the asymptotic normality result (2.2).

Proposition 4.3 *(Donkers and Schafgans (2003, 2005), Samarov (1993)) Under the assumptions stated in Donkers and Schafgans (2003, 2005),*

$$\sqrt{n}(\mathrm{vech}(\widehat{M}_S) - \mathrm{vech}(M)) \xrightarrow{d} \mathcal{N}(0, W_0), \tag{4.21}$$

where

$$W_0 = \mathrm{Var}(\mathrm{vec}(R(x_i, y_i))) \tag{4.22}$$

with

$$R(x,y) = w(x)\left(\frac{\partial g(x)}{\partial x}\frac{\partial g(x)}{\partial x}' - (y - g(x))\left(\frac{\frac{\partial f(x)}{\partial x}\frac{\partial g(x)}{\partial x}'}{f(x)} + \frac{\frac{\partial g(x)}{\partial x}\frac{\partial f(x)}{\partial x}'}{f(x)} + 2\frac{\partial^2 g(x)}{\partial x \partial x'}\right)\right). \tag{4.23}$$

As in Example 4.1, one can easily show that $\mathrm{rk}\{W_0\}$ in (4.22) is constrained by $\mathrm{rk}\{M\}$.

Example 4.4 (Number of factors in nonparametric relationship.) Donald (1997) considers a nonparametric model

$$y_i = F(x_i) + \epsilon_i, \quad i = 1, \ldots, n, \tag{4.24}$$

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 (4.24) and related problems are considered in Fortuna (2008).

Of interest to our context is a semidefinite matrix

$$M = Ef(x_i)F(x_i)F(x_i)', \tag{4.25}$$

where $f(x)$ is the density of x_i. Its rank $r = \mathrm{rk}\{M\}$ is the number of factors in nonparametric relationship (4.24) 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 *indefinite* matrix estimator

$$\widehat{M}_I = \frac{1}{n(n-1)}\sum_{i\neq j} y_i y_j' K_h(x_i - x_j), \tag{4.26}$$

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 4.4 *(Donald (1997), Donald et al. (2007)) Under the assumptions of Donald (1997), $nh^{q/2}(\mathrm{vec}(\widehat{M})-\mathrm{vec}(M))$ is asymptotically normal. Of interest to rank testing, one has $\widehat{M}_I=\widehat{M}_1+\widehat{M}_2$ where (i) $u'Mu=0$ for a vector u implies $u'\widehat{M}_1u=0$, (ii) $\widehat{M}_1-M=O_p((nh^{q/2})^{-1})$, and*

$$nh^{q/2}\mathrm{vech}(\widehat{M}_2)\xrightarrow{d}\mathcal{N}(0,W), \tag{4.27}$$

where

$$W=V^{-1}D_p^+(\Sigma\Sigma\otimes\Sigma\Sigma)D_p^{+'} \tag{4.28}$$

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

Alternatively, the matrix M can be sought to be estimated through *semidefinite* matrix estimator

$$\widehat{M}_S=\frac{1}{n}\sum_{i=1}^{n}\widehat{f}(x_i)^{-1}\widehat{G}(x_i)\widehat{G}(x_i)', \tag{4.29}$$

where similarly to Example 4.1, for example,

$$\widehat{G}(x_i)=\frac{1}{(n-1)h^q}\sum_{j\neq i}y_jK\left(\frac{x_i-x_j}{h}\right). \tag{4.30}$$

As in Example 4.1, using the results of Samarov (1993), one can establish the following result. The assumptions and a short proof are moved to Appendix A.

Proposition 4.5 *(Samarov (1993)) Under the assumptions stated in Appendix A,*

$$\sqrt{n}(\mathrm{vec}(\widehat{M}_S)-\mathrm{vec}(M))\xrightarrow{d}\mathcal{N}(0,W), \tag{4.31}$$

where

$$W=\mathrm{Var}\left(f(x_i)(F(x_i)\otimes I_p+I_p\otimes F(x_i))y_i\right). \tag{4.32}$$

Note that, under model (4.24), the limiting covariance matrix W in (4.32) is given by

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

The two estimators $\widehat{M}_I$ in (4.26) and $\widehat{M}_S$ in (4.29) 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 4.4 and 4.5 are different (one is $nh^{q/2}$ and the other is $n^{1/2}$). However, to see clearly that 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 $F(x)\equiv 0$.

5 On extensions of available rank tests

Recall from Proposition 2.1 in Donald et al. (2007) and as seen from examples considered in Section 4, 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 case R in (2.6). But the case C is also considered to the end by reexamining a test recently suggested by Camba-Méndez and Kapetanios.

5.1 Tests involving generalized inverses

Available rank tests for nonsingular limiting covariance matrices W in (2.1) typically involve inverses of these matrices. When they are singular, rank tests can be extended naturally by using generalized inverses. For example, here is one such typical extension for the LDU rank test of Gill and Lewbel (1992), and Cragg and Donald (1996).

The LDU test is based on Gaussian elimination with complete pivoting. Ignoring the issue of ties for simplicity, one may 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

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}, \tag{5.1}$$

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

$$\Lambda_r = M_{22} - M_{21} M_{11}^{-1} M_{12} \tag{5.2}$$

(because of M_{11}^{-1}, this is meaningful only for $r \leq \mathrm{rk}\{M\}$). Let also

$$\Phi = (-M_{21}M_{11}^{-1} \quad I_{p-r}), \quad \Gamma = \Phi \otimes \Phi, \tag{5.3}$$

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 5.1 *(Camba-Méndez and Kapetanios (2001, 2008)) Suppose that (2.1) holds, the rank of W is known and there is $\widehat{W} \to_p W$ such that $\mathrm{rk}\{\widehat{W}\} = \mathrm{rk}\{W\}$. Then, under $\mathrm{rk}\{M\} = r$,*

$$\widehat{\xi}_{\mathrm{ldu}}(r) = N \mathrm{vec}(\widehat{\Lambda}_r)'(\widehat{\Gamma}\widehat{W}\widehat{\Gamma}')^+ \mathrm{vec}(\widehat{\Lambda}_r) \xrightarrow{d} \chi^2_m, \tag{5.4}$$

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

Remark 5.1 The basic reason for assuming $\mathrm{rk}\{\widehat{W}\} = \mathrm{rk}\{W\}$ is that $\widehat{W} \to_p W$ does not imply in general that their generalized inverses converge. See 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 (5.4) (in addition, using the so-called symmetric pivoting) or for other rank tests such as the MINCHI2 test of Cragg and Donald (1996). However, those tests are not useful in our context because $\mathrm{rk}\{W\}$ is unknown. This can be seen from the examples considered in Section 4. Furthermore, as those examples 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 5.2 *Suppose that $\widehat{M}$ is a semidefinite matrix estimator for M, and that (2.1) holds with covariance matrix W. Then, with $r = \mathrm{rk}\{M\}$,*

$$\mathrm{rk}\{W\} \leq r(2p - r). \tag{5.5}$$

Even though $\mathrm{rk}\{W\}$ is unknown, several authors (Lütkepohl and Burda (1997), Camba-Méndez and Kapetanios (2005*b*)) suggest getting its rough estimate (for example, using asymptotic results for eigenvalues of $\widehat{W}$) and then substituting a reduced rank estimate $\widehat{W}$ into (5.4). However, with this procedure, the problem of testing for $\mathrm{rk}\{M\}$ is essentially replaced by that of estimation of another rank, namely, $\mathrm{rk}\{W\}$. Furthermore, using the same argument, $\mathrm{rk}\{M\}$ itself could roughly be estimated directly.

5.2 Tests of Robin and Smith

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

$$0 < \mathrm{rk}\{W\} = s \leq p^2. \tag{5.6}$$

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

$$t = \mathrm{rk}\{(C'_{p-r} \otimes C'_{p-r})W(C_{p-r} \otimes C_{p-r})\} > 0. \tag{5.7}$$

One can easily see that $t \leq \min\{s, (p-r)^2\}$.

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

$$\widehat{\xi}_{\mathrm{cr}}(r) = N \sum_{j=1}^{p-r} \widehat{\lambda}_j. \tag{5.8}$$

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

$$\xi_{\mathrm{cr}}(r) = \sum_{j=1}^{t} \lambda_j^r Z_j^2 = \sum_{j=1}^{(p-r)^2} \lambda_j^r Z_j^2, \tag{5.9}$$

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

$$\sum_{j=1}^{(p-r)^2} \widehat{\lambda}_j^r Z_j^2, \tag{5.10}$$

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

The above test of Robin and Smith (2000) cannot be applied to our context under their assumptions (5.6) and (5.7). To see why this is expected, consider Example 4.4 and the matrix estimators $\widehat{M}_S$ in (4.29). 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 $0 < \upsilon_{p-r+1} \leq \dots \leq \upsilon_p$ 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

$$C'_{p-r}F(x) = 0. \tag{5.11}$$

In view of (4.32), (5.11) implies that

$$(C'_{p-r} \otimes C'_{p-r})W(C_{p-r} \otimes C_{p-r}) = 0. \tag{5.12}$$

Hence, the assumption (5.7) of Robin and Smith (2000) stated above is not satisfied. On the other hand, as shown in the next section, this result should not be surprising at all (see remark in that section).

5.3 Tests of Camba-Méndez and Kapetanios

Finally, another related test was recently suggested by Camba-Méndez and Kapetanios (2001, 2005*b*, 2008). Proposed as a test for the rank of Hermitian positive semidefinite matrix, it works as follows. Consider the more general Case C of (2.6). As for the LDU test, the matrix M is partitioned as (5.1) and it is supposed for simplicity that M_{11} has full rank (which can be supposed after permutations involved in Gaussian elimination procedure). The focus is again on the Schur complement Λ_r given in (5.2). 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, Camba-Méndez and Kapetanios (2008), for example, suggest to consider the test statistic

$$\widehat{\xi}_{\mathrm{kudo}}(r) = N\mathrm{diag}\{\widehat{\Lambda}_r\}'\Psi^{-1}\mathrm{diag}\{\widehat{\Lambda}_r\}, \tag{5.13}$$

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

$$\sqrt{N}\mathrm{diag}\{\widehat{\Lambda}_r\} \xrightarrow{d} \mathcal{N}(0, \Psi). \tag{5.14}$$

These authors say that application of the result of Kudo yields the asymptotic distribution $\widehat{\xi}_{\mathrm{kudo}}(r)$ of (5.13) given by

$$P(\widehat{\xi}_{\mathrm{kudo}}(r) > x) = \sum_{q=0}^{p-r} w_q P(\chi_q^2 \geq x), \tag{5.15}$$

where w_q are suitable weights.

Applications of the results of Kudo, however, is not justified in the above context. The Schur complement is also positive semidefinite (see, for example, Zhang (2005)), and hence the elements of $\mathrm{diag}\{\widehat{\Lambda}_r\}$ are real, nonnegative. But then (5.14) is possible only with $\Psi \equiv 0$ and relation (5.13) and its limit in (5.15) are not meaningful.

Remark 5.2 Though application of the results of Kudo are not meaningful in the context above, they can be used in Case 1 of (1.3) considered by Donald et al. (2007).

6 Asymptotics of eigenvalues, and consistent rank tests

Though available rank tests discussed in Section 5 do not appear helpful, consistent rank tests can, in fact, be easily established under assumption (2.1). Let $0 \leq \widehat{v}_1 \leq \widehat{v}_2 \leq \cdots \leq \widehat{v}_p$ be the ordered eigenvalues of the matrix estimator $\widehat{M}$. The following basic result concerns asymptotics of these eigenvalues.

Theorem 6.1 *Under assumption (2.1), and with* $\mathrm{rk}\{M\} = r$*,*

$$\sqrt{N}\widehat{v}_j \xrightarrow{p} 0, \quad j = 1, \ldots, p-r, \tag{6.1}$$

$$\sqrt{N}\widehat{v}_j \xrightarrow{p} +\infty, \quad j = p-r+1, \ldots, p. \tag{6.2}$$

The theorem is proved in Appendix A. Letting

$$\widehat{\xi}_{\mathrm{eig}}(k) = \sqrt{N^{\beta}} \sum_{j=1}^{p-k} \widehat{v}_j^{\beta}, \tag{6.3}$$

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

Corollary 6.1 *Under assumption (2.1),*

$$\textit{under} \quad \mathrm{rk}\{M\} \le k, \qquad \widehat{\xi}_{\mathrm{eig}}(k) \xrightarrow{p} 0, \tag{6.4}$$

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

Corollary 6.1 shows that the rank test based on (6.3) is consistent. This rank test, however, is not satisfactory from a practical perspective. Even with a fixed (arbitrary) critical value, the conclusion of the test would depend on a multiplication of $\widehat{M}$ by a constant.

Remark 6.1 Since $\widehat{v}_k^2$ are the eigenvalues of $\widehat{M}^2$, Corollary 6.1 shows that the characteristic root test based on (5.8) of Robin and Smith (2000) (see Section 5 above) is also consistent. The same corollary also shows that $\lambda_j^r = 0$ for all j in (5.9) or hence that $t = 0$ in (5.7).

7 Finer rank tests

Though Theorem 6.1 and its Corollary 6.1 provide a consistent test under (2.1), the test is not satisfactory from a practical perspective. For example, it is natural to expect that a faster rate than $\sqrt{N}$ in (6.1) may yield a finer result (with nondegenerate limit in (6.1)). We examine this and related questions in this section.

The following general assumptions are relevant. Let $Q = (Q_1\, Q_2)$ be as in (2.7) so that $Q_2' M Q_2 = 0$. Under assumption (2.1), $\sqrt{N} Q_2' \widehat{M} Q_2 \to_p 0$. Therefore, in some cases, it may be natural to expect that

$$a_N Q_2' \widehat{M} Q_2 \xrightarrow{d} \mathcal{A}, \tag{7.1}$$

where $a_N \to \infty$ and a_N grows faster than $\sqrt{N}$ (that is, $a_N/\sqrt{N} \to \infty$).

Theorem 7.1 *Under assumption (2.1) and (7.1), with notation of Section 6 and when* $\mathrm{rk}\{M\} = r$,

$$a_N \widehat{v}_j \xrightarrow{p} \alpha_j, \quad j = 1, \ldots, p-r, \tag{7.2}$$

$$a_N \widehat{v}_j \xrightarrow{p} +\infty, \quad j = p-r+1, \ldots, p, \tag{7.3}$$

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

$$\textit{under} \quad \mathrm{rk}\{M\} \le k, \qquad \widehat{\xi}_{\mathrm{eig}}(k) \xrightarrow{p} \sum_{j=1}^{p-k} \alpha_j^{\beta}, \tag{7.4}$$

$$\textit{under} \quad \mathrm{rk}\{M\} > k, \qquad \widehat{\xi}_{\mathrm{eig}}(k) \xrightarrow{p} +\infty, \tag{7.5}$$

where $\widehat{\xi}_{\mathrm{eig}}(k)$ *is defined by (6.3) but using normalization* a_N^{β} *instead of* $N^{\beta/2}$.

The proof of Theorem 7.1 is analogous to that of Theorem 6.1 and is omitted. Theorem 7.1 gives consistent rank tests which could be implemented in practice. Supposing there is a consistent estimator $\widehat{\mathcal{A}} \to_p \mathcal{A}$, a critical value for a test of $H_0 : \mathrm{rk}\{M\} = k$ could be defined as

$$(1+c)\sum_{j=1}^{p-k} \widehat{\alpha}_j^{\beta} \tag{7.6}$$

for some (arbitrary) $c > 0$, where $\widehat{\alpha}_j$ are the ordered eigenvalues of $\widehat{\mathcal{A}}$. Note that the corresponding test is invariant to a multiplication of $\widehat{M}$ by a constant.

Example 7.1 (Linear regression with heteroscedastic error terms.) The next result shows that, under additional assumptions, the estimator $\widehat{\Sigma}(x)$ in (4.3) satisfies the assumption (7.1). The result is proved in Appendix A.

Proposition 7.1 *With the above notation and that of Example 4.1, suppose that* $\Sigma(x) \in C^2(\mathcal{H}_x)$, $p(x)$, $E(|\epsilon_k|^4)|x_k = x) \in C^0(\mathcal{H}_x)$ *and* $Ex_i x_i'$ *is invertible. Then, as* $n \to \infty$, $nh^{5q} \to \infty$, $h \to 0$, *for* $x \in \mathcal{H}_x$,

$$h^{-2q} Q_2' \widehat{\Sigma}(x) Q_2 \xrightarrow{p} \frac{1}{6} Q_2' \Sigma(x)'' Q_2. \tag{7.7}$$

Example 7.2 (Number of factors in nonparametric relationship.) Consider a semidefinite matrix estimator $\widehat{M}_S$ given in (4.29). With this estimator and the above notation, observe that

$$Q_2' \widehat{M}_S Q_2 = Q_2' \frac{1}{n} \sum_{i=1}^{n} \widehat{f}(x_i)^{-1} \widehat{\epsilon}(x_i) \widehat{\epsilon}(x_i)' Q_2, \tag{7.8}$$

where

$$\widehat{\epsilon}(x_i) = \frac{1}{n-1} \sum_{j \neq i} \epsilon_j K_h(x_i - x_j). \tag{7.9}$$

Then, one expects that $EQ_2' \widehat{M}_S Q_2$ is asymptotically

$$Q_2' E f(x_i)^{-1} \widehat{\epsilon}(x_i) \widehat{\epsilon}(x_i)' Q_2 = Q_2' \frac{1}{(n-1)^2} E f(x_i)^{-1} \sum_{j \neq i} \epsilon_j \epsilon_j' K_h(x_i - x_j)^2 Q_2$$

$$= \frac{\|K\|_2^2}{(n-1)h^q} Q_2' \Sigma Q_2 E f(x_i)^{-1} K_{2,h}(x_i - x_j) \sim \frac{\|K\|_2^2}{nh^q} Q_2' \Sigma Q_2, \tag{7.10}$$

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

$$nh^q Q_2' \widehat{M}_S Q_2 \xrightarrow{p} \|K\|_2^2 Q_2' \Sigma Q_2. \tag{7.11}$$

Remark 7.1 Example 4.3 could be dealt with as Example 7.2 above. Dealing with the matrix estimator of Example 4.2 is more challenging and left for a future work.

Examples 7.1 and 7.2 above suggest that $\mathcal{A}$ is a constant matrix, and that (7.1) corresponds to Law of Large Numbers asymptotics. It is natural to go beyond (7.1) by postulating the asymptotics of Central Limit Theorem as

$$b_N(a_N Q_2' \widehat{M} Q_2 - \mathcal{A}) \xrightarrow{d} \mathcal{B}, \tag{7.12}$$

where $b_N \to \infty$ and $\mathcal{B}$ is a random matrix.

There are several potential difficulties with (7.12). First, the assumption (7.12) is related to the asymptotic behavior of

$$b_N(a_N \widehat{v}_j - \alpha_j), \quad j = 1, \dots, p - r \tag{7.13}$$

(see, for example, Eaton and Taylor (1991)). Note, however, that α_k is unknown here and the rate of convergence of $\widehat{\alpha}_k$ to α_k is not immediate. Second, another difficulty with (7.12) is the following. For example, in Example 7.1 above, one expects (and this is not too difficult to show) that, in the case $p = 1$ for simplicity,

$$b_N(h^{-2q}\widehat{\Sigma}(x) - \frac{1}{6}\Sigma(x)'') \xrightarrow{d} \mathcal{N}(0, \sigma^2), \tag{7.14}$$

for suitable $b_N \to \infty$. The limiting variance σ^2, however, is strictly positive only when $\Sigma(x)'' \neq 0$. The latter is now an assumption on the rank of $\Sigma(x)''$. Testing for its rank, in fact, will lead to another problem of matrix rank estimation, and this can continue an arbitrary number of times.

8 Use of indefinite matrix estimators

The previous section concerns finer rank tests under assumption (2.1) or (2.2) where matrix estimator $\widehat{M}$ is semidefinite. Another obvious possibility is to search for an indefinite matrix estimator which may allow 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 estimator $\widehat{M}$ depends on the problem at hand, and may not be immediately obvious. For example, in Example 4.4, we already specified two matrix estimators: indefinite $\widehat{M}_I$ in (4.26) and semidefinite $\widehat{M}_S$ in (4.29). 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 4.3 as well. Our discussion will not be completely rigorous.

Example 8.1 (Multiple index mean regression model.) Suppose for simplicity that the model (4.16) in Example 4.3 is given by

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

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 (4.16). A natural indefinite estimator of the matrix M in (4.16) (with suitable weight $w(x)$) is

$$\widehat{M}_I = \frac{1}{n(n-1)} \sum_{i \neq j} \widehat{f}(x_i)^{-1} \widehat{f}(x_j)^{-1} y_i y_j \frac{1}{h^{l+2}} K^{(2)}\left(\frac{x_i - x_j}{h}\right). \tag{8.2}$$

The basic idea behind (8.2) is that, asymptotically, $E\widehat{M}_I$ behaves as (with $i \neq j$)

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

$$= Ef(x_i)^{-1}f(x_j)^{-1}\frac{\partial g(x_i)}{\partial x}\frac{\partial g(x_j)}{\partial x}'\frac{1}{h^l}K\left(\frac{x_i-x_j}{h}\right) \to Ef(x_i)^{-1}\frac{\partial g(x_i)}{\partial x}\frac{\partial g(x_i)}{\partial x}' = M. \tag{8.3}$$

As in Example 4.4, write $\widehat{M}_I = \widehat{M}_1 + \widehat{M}_2$, where

$$\widehat{M}_2 = \frac{1}{n(n-1)}\sum_{i\neq j} f(x_i)^{-1}f(x_j)^{-1}\epsilon_i\epsilon_j\frac{1}{h^{l+2}}K^{(2)}\left(\frac{x_i-x_j}{h}\right). \tag{8.4}$$

A simple analysis of second moments suggests that (with $i \neq j$)

$$\begin{aligned}
&E\mathrm{vech}(\widehat{M}_2)\mathrm{vech}(\widehat{M}_2)' \\
&= \frac{2\sigma^4}{n(n-1)}Ef(x_i)^{-2}f(x_j)^{-2}\frac{1}{h^{2l+4}}\mathrm{vech}\left(K^{(2)}\left(\frac{x_i-x_j}{h}\right)\right)\mathrm{vech}\left(K^{(2)}\left(\frac{x_i-x_j}{h}\right)\right)' \\
&\sim \frac{2\sigma^4}{n^2h^{l+4}}Ef(x_i)^{-2}f(x_j)^{-2}\frac{1}{h^l}\mathrm{diag}\left(\mathrm{vech}\left(K^{(2)}\left(\frac{x_i-x_j}{h}\right)\right)^2\right) \\
&\sim \frac{2\sigma^4}{n^2h^{l+4}}Ef(x_i)^{-3}\mathrm{diag}\left(\mathrm{vech}\left(\|K^{(2)}(u)\|_2^2\right)\right) =: \frac{1}{n^2h^{l+4}}V,
\end{aligned} \tag{8.5}$$

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

$$nh^{\frac{l}{2}+2}\mathrm{vech}(\widehat{M}_2) \xrightarrow{d} \mathcal{N}(0, V), \tag{8.6}$$

where V is nonsingular, and hence that the framework of Donald et al. (2007) can be used for $\widehat{M}_I$.

Yet another possibility is to postulate a related model

$$y_i = f(x_i)^{-1}g(x_i) + \epsilon_i, \quad i = 1, \ldots, n, \tag{8.7}$$

where $g(x)$ is given by (4.16). An indefinite matrix estimator of M in this case is

$$\widehat{M}_I = \frac{1}{n(n-1)}\sum_{i\neq j} y_i y_j \frac{1}{h^{l+2}}K^{(2)}\left(\frac{x_i-x_j}{h}\right), \tag{8.8}$$

which is much simpler to deal with.

Remark 8.1 Whether indefinite matrix estimators can be introduced naturally in Examples 4.1 and 4.2 is still an open question.

A Technical proofs

Proof of Proposition 4.1: 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 assumptions of the proposition, note that $\widehat{\Sigma}_3(x) = O_p(1/n)$, $\widehat{\Sigma}_4(x) = O_p(1/n)$. Using the properties of kernel function and $\Sigma(x)$, we also have $\widehat{\Sigma}_2(x) = O(h^r)$. It is then enough to show that, with $W(x)$ in (4.5),

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

or, equivalently, that $\sqrt{Nh^q}a'\mathrm{vec}(\widehat{\Sigma}_1(x)) \to_d \mathcal{N}(0, a'W(x)a)$ 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 Feller-Lindeberg 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 above condition is satisfied by the assumptions of the proposition. □

The following are basic assumptions for Proposition 4.5. Let $G(x) = F(x)f(x)$, where $f(x)$ is the density of x_i. Suppose $f(x)$ 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 f and G up to the order $\nu+1$ are bounded and $f, G \in H_{\nu+2}(U)$.

(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 $\{\|x\| \le 1\}$ 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 $l_1 + \dots + l_q = l$ with nonnegative integers l_i.

(C4) W in (4.32) is well-defined.

Proof of Proposition 4.5: We briefly outline the proof of Theorem 1 in Samarov (1993), p. 845. Note that

$$\widehat{T}_n := \mathrm{vec}(\widehat{M}_a) = \frac{1}{n}\sum_{i=1}^n \widehat{f}(x_i)^{-1}(\widehat{G}(x_i) \otimes \widehat{G}(x_i)).$$

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

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

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

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

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

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

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

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

Gathering all these observations, one concludes that

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

The asymptotic normality result now follows. □

Proof of Proposition 5.2: Let Q be an orthogonal matrix as in (2.7) and $D = \mathrm{diag}\{\upsilon_1,\dots,\upsilon_p\}$. Then,

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

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 $\sqrt{N}\mathrm{vec}(Q_2'\widehat{M}Q_2)$ is also asymptotically normal. By Lemma A.1 below, the only way this can happen is when $\sqrt{N}\mathrm{vec}(Q_2'\widehat{M}Q_2) \to_d 0$. Since $Q_2'\widehat{M}Q_2$ is of dimension $(p-r)^2$, this means that $(p-r)^2$ elements of $\sqrt{N}\mathrm{vec}(Q'\widehat{M}Q)$ are asymptotically 0. Hence, $\mathrm{rk}\{\widetilde{W}\} \le p^2 - (p-r)^2 = r(2p-r)$. □

The following elementary lemma was used in the proof above.

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

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

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 nonnegative definite. In particular,

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

for all a_i, where $\mathcal{X} = (x_{ij})_{i,j=1,\dots,m}$. 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 $\mathcal{X} = 0_{m\times m}$. □

Proof of Theorem 6.1: 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 5.2 above and $\mathcal{Y}$ appears in (2.3). Since $\sqrt{N}\widehat{\upsilon}_k \ge 0$, by using Lemma A.1, this can happen only when $Q_2'\mathcal{Y}Q_2 = 0$. This leads to (6.1). The relation (6.2) follows from $\widehat{\upsilon}_k \to_p \upsilon_k > 0$, $k = p-r+1,\dots,p$. □

Proof of Proposition 7.1: We only consider the case $p = 1$. Then, we may suppose that $Q_2 = 1$ and $\Sigma(x) = 0$. As in the proof of Proposition 4.1, $h^{-2}\widehat{\Sigma}(x) = h^{-2}\widehat{\Sigma}_1(x) + h^{-2}\widehat{\Sigma}_2(x) + O_p(n^{-1}h^{-2})$. Note that

$$E(h^{-2}\widehat{\Sigma}_1(x)) \leq \frac{1}{nh^4} E\epsilon_k^4 K_h(x - x_k)^2 = O_p\left(\frac{1}{nh^5}\right)$$

and

$$h^{-2}\widehat{\Sigma}_2(x) = h^{-2} E\epsilon_k^2 K_h(x - x_k) \to \frac{\Sigma(x)''}{6}. \ \square$$

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*a*), 'Estimating the rank of the spectral density matrix', *Journal of Time Series Analysis* **26**(1), 37–48.

Camba-Méndez, G. & Kapetanios, G. (2005*b*), Statistical tests of the rank of a matrix and their applications in econometric modelling, Working Paper 541, Department of Economics, Queen Mary, University of London, London, United Kingdom.

Camba-Méndez, G. & Kapetanios, G. (2008), Statistical tests and estimators of the rank of a matrix and their applications in econometric modelling, Working Paper 850, European Central Bank, Frankfurt, Germany.

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.

Donkers, B. & Schafgans, M. (2003), A derivative based estimator for semiparametric index models, Econometric Institute Report, 2003-08, Erasmus University Rotterdam, Rotterdam, The Netherlands.

Donkers, B. & Schafgans, M. (2005), A method of moments estimator for semiparametric index models, Discussion Paper EM/05/493, STICERD, London School of Economics, London, United Kingdom.

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.

Ratsimalahelo, Z. (2003), Strongly consistent determination of the rank of matrix, 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.