Estimation of matrix rank: historical overview and more recent developments

Vladas Pipiras
CEMAT, IST and University of North Carolina, Chapel Hill

UTL Probability and Statistics Seminar in Lisbon
November 7, 2008
Outline:


2. Several examples.

3. Available rank tests.

4. The case of symmetric matrices (joint work with S. Donald and N. Fortuna).
1. General problem
**General framework:**

$M: p \times q$ unknown matrix with unknown rank $\text{rk}\{M\}$.

$\widehat{M} = \widehat{M}(N)$: matrix estimator such that

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

There is $\widehat{W}$ such that $\widehat{W} \xrightarrow{p} W$.

**Interested in testing for matrix rank:**

$H_0 : \text{rk}\{M\} = r,$

$H_1 : \text{rk}\{M\} > r,$

where $r = 0, \ldots, \min\{p, q\} - 1$. The resulting tests are referred to as (matrix) rank tests.
Remarks:

- Matrix rank is estimated, for example, through sequential testing. This will NOT be discussed further. But here is a typical output for a $4 \times 5$ matrix:

<table>
<thead>
<tr>
<th>rank $r$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>test statistic</td>
<td>24677</td>
<td>2944</td>
<td>97.928</td>
<td>3.415</td>
</tr>
<tr>
<td>critical value</td>
<td>31.41</td>
<td>21.026</td>
<td>12.592</td>
<td>5.9915</td>
</tr>
<tr>
<td>$P$-value</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.18132</td>
</tr>
</tbody>
</table>

- Focus here is NOT on estimation of a reduced rank matrix, which can be done in a number of ways.
2. Several examples
**Example 1:** (Reduced rank regression) Consider a multivariate linear regression model

\[
\begin{align*}
\begin{pmatrix} y_i \end{pmatrix}_{p \times 1} &= \begin{pmatrix} A \end{pmatrix}_{p \times q} \begin{pmatrix} x_i \end{pmatrix}_{q \times 1} + \begin{pmatrix} \epsilon_i \end{pmatrix}_{p \times 1}, \quad i = 1, \ldots, n,
\end{align*}
\]

with response variables \( y_i \in \mathbb{R}^p \), regressors (predictors) \( x_i \in \mathbb{R}^q \), and error terms \( \epsilon_i \in \mathbb{R}^p \). Interested in testing for \( \text{rk}\{A\} \). For a standard, least squares estimator \( \hat{A} = \sum_{i=1}^{n} y_i x_i' (\sum_{i=1}^{n} x_i x_i')^{-1} \), under mild assumptions,

\[
\sqrt{n} \text{vec}(\hat{A} - A) \overset{d}{\rightarrow} \mathcal{N}(0, W),
\]

where

\[
W = E(x_i x_i') \otimes E(\epsilon_i \epsilon_i')
\]

has the so-called Kronecker product form, and can be estimated consistently through \( \hat{W} = \frac{1}{n} \sum_{i=1}^{n} x_i x_i' \otimes \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{A} x_i)(y_i - \hat{A} x_i)' \). Note: non-Kronecker product would arise with heteroscedastic error terms.
Example 2: (Number of factors in nonparametric relationship) Consider a multivariate nonparametric model

\[ y_i = F(x_i) + \epsilon_i, \quad i = 1, \ldots, n, \]

where \( F \) is an unknown but otherwise smooth function. Interested in estimating the maximum number of linearly independent component functions of \( F \), denoted as \( \text{rk}\{F\} \).

One can show that

\[ \text{rk}\{F\} = \text{rk}\{M\}, \quad M = Ef(x_i)F(x_i)F(x_i)'. \]

Note that \( M \) is symmetric (and positive semi-definite). This will be discussed further and asymptotically normal estimators will be given below.
Some other examples:


One general view:

Model dimension $\Leftrightarrow$ Matrix rank
3. Available rank tests and other work
Econometrics literature:

• LDU rank test: Gill and Lewbel (1992), Cragg and Donald (1996).

• MINCHI2 rank test: Cragg and Donald (1997).

• CHART rank test: Robin and Smith (2000).


• Camba-Méndez, Kapetanios and others.

Statistics literature:

• Reduced rank regression: goes back to the Anderson (1951) and has been studied extensively since by many others. Natural connections to canonical correlation analysis and extensions to VAR series (cointegration).

• Asymptotics of eigenvalues under quite general assumptions: series of papers by Tyler and coauthors in the 1980’s.
**Example: MINCHI2 rank test:** based on the test statistic

\[ \hat{\xi}_{\text{minchi2}}(r) = N \min_{\text{rk}\{M\} \leq r} \text{vec}(\hat{M} - M)' \hat{W}^{-1} \text{vec}(\hat{M} - M). \]

One can show that under \( H_0: \text{rk}\{M\} = r \): \( \hat{\xi}_{\text{minchi2}}(r) \xrightarrow{d} \chi^2((p - r)(q - r)) \), under \( H_1: \text{rk}\{M\} > r \): \( \hat{\xi}_{\text{minchi2}}(r) \xrightarrow{P} +\infty \).

**Remark:** Here and in most available rank tests, one assumes that the covariance matrix \( W \) is nonsingular.

**Remark:** If \( W = \Psi^{-1} \otimes \Sigma^{-1} \) has a Kronecker product structure, then

\[ \hat{\xi}_{\text{minchi2}}(r) = N \sum_{i=1}^{p-r} \hat{\lambda}_i^2, \]

is a classical test statistic, where \( 0 \leq \hat{\lambda}_1^2 \leq \ldots \leq \hat{\lambda}_p^2 \) are the ordered eigenvalues of \( \hat{\Sigma} \hat{M} \hat{\Psi} \hat{M}' \). The same happens for the test statistics corresponding to CHART and SVD tests.
4. The case of symmetric matrices
General framework:

\[ M, \hat{M} = \hat{M}(N) : \text{symmetric, } p \times p \text{ matrices such that} \]

\[ \sqrt{N} \text{vech}(\hat{M} - M) \xrightarrow{d} \mathcal{N}(0, V) \]

or, in a weaker form, \( \hat{M} = \hat{M}_1 + \hat{M}_2 \) where \( Mu = 0 \) if and only if \( \hat{M}_1 u = 0 \), \( M - \hat{M}_1 = O_p(1/\sqrt{N}) \), and

\[ \sqrt{N} \text{vech}(\hat{M}_2) \xrightarrow{d} \mathcal{N}(0, V). \]

There is \( \hat{V} \) such that \( \hat{V} \xrightarrow{p} V \).

2 different cases:

**ID case:** Indefinite \( \hat{M} \): indefinite or semidefinite \( M \): \( V \) is nonsingular,

**SD case:** Semidefinite \( \hat{M} \): semidefinite \( M \): \( V \) is singular under reduced rank of \( M \).
1-dimensional case ($1 \times 1$ matrices):

Rank deficiency $\iff \text{rk}\{M\} = 0 \iff M = 0$.

Under rank deficiency (or $H_0$): $\sqrt{N} \text{vech}(\widehat{M} - M) = \sqrt{N} \widehat{M} \xrightarrow{d} \mathcal{N}(0, V)$.

SD case: $\widehat{M}$ positive semidefinite $\iff \widehat{M} \succeq 0 : V = 0$.

ID case: $\widehat{M}$ indefinite $\iff \widehat{M} \in \mathbb{R} : V > 0$.

Related work:


Earlier example 2: (Number of factors in nonparametric relationship) In the model
\[ y_i = F(x_i) + \epsilon_i, \quad i = 1, \ldots, n, \]
interested in estimating \( \text{rk}\{F\} \), the number of linearly independent component functions of \( F \). One can show that \( \text{rk}\{F\} = \text{rk}\{M\} \), where \( M = Ef(x_i)F(x_i)F(x_i)' \).

**Semidefinite estimator (DFP (2008)):**
\[ \hat{M}_S = \frac{1}{n} \sum_{i=1}^{n} \hat{f}(x_i)\hat{F}(x_i)\hat{F}(x_i)'. \]
Then,
\[ \sqrt{n} \text{vech}(\hat{M}_S - M) \xrightarrow{d} \mathcal{N}(0, V_S), \] \( V_S \) singular.

**Indefinite estimator (Donald (1997)):**
\[ \hat{M}_I = \frac{1}{n(n-1)} \sum_{i \neq j} y_i y_j'K_h(x_i - x_j). \]
Then, \( \hat{M}_I = \hat{M}_{I,1} + \hat{M}_{I,2} \) with \( M_I u = 0 \iff \hat{M}_{I,1} u = 0 \), \( \hat{M}_{I,1} - M = O_p((nh^{q/2})^{-1}) \), and \( nh^{q/2} \text{vech}(\hat{M}_{I,2}) \xrightarrow{d} \mathcal{N}(0, V_I), \) \( V_I \) nonsingular.

Some other examples: ID case: reduced rank regression with unknown symmetric matrix (Robin and Smith (2000)). SD, ID cases: multiple index mean regression model (Donkers and Schafgans (2005)). ID estimator due to DFP (2008). SD case: spectral density matrices in multivariate time series, cointegration (Camba-Méndez and Kapetanios (2005)).
**ID case (DFP(2007))**: adaptations of available rank tests (LDU, MINCHI2, SVD) to symmetric matrices, and a new rank test (EIG) based on eigenvalues. **Notes on results in ID case**: EIG test: for symmetric (square) matrices can talk about their eigenvalues directly. LDU test: using symmetric pivoting. MINCHI2 and other tests: adjustment for degrees of freedom in limiting $\chi^2$ distribution: $(p - r)(p - r + 1)/2$ instead of $(p - r)^2$. Simulation study: SVD appears to be superior over other rank tests.

**Main point for ID case**: many rank tests are available with well developed theory.

**SD case (DFP(2008))**: Available rank tests for singular $W$: for example, conditions on $W$ in Robin and Smith (2000) are not satisfied; conditions for using generalized inverses of $W$ as in Camba-Méndez and Kapetanios (2008) are also not satisfied.

**Main point for SD case**: no satisfactory rank tests are available. **But**: understand where difficulties lie and what possible alternatives are.