A General Theory of Rank Testing

Majid M. Al-Sadoon
Universitat Pompeu Fabra & BGSE

26/09/2015
NBER-NSF Time Series Conference

The Literature

- Rank tests are pervasive in economics: cointegration, factor models, identification, subspace Granger causality analysis, etc.

The Literature

- Rank tests are pervasive in economics: cointegration, factor models, identification, subspace Granger causality analysis, etc.
- There is a large literature in numerical analysts on the “effective” rank of a matrix (e.g. Golub & Van Loan (1996), Björck (1996), Hansen (1998)).

The Literature

- Rank tests are pervasive in economics: cointegration, factor models, identification, subspace Granger causality analysis, etc.
- There is a large literature in numerical analysts on the "effective" rank of a matrix (e.g. Golub & Van Loan (1996), Björck (1996), Hansen (1998)).
- Multivariate statistics mainly considers the case of i.i.d. normal data: Hotelling (1936), Bartlett (1947), Anderson (1951), Izenman (1975), Anderson (1999).

The Literature

- Rank tests are pervasive in economics: cointegration, factor models, identification, subspace Granger causality analysis, etc.
- There is a large literature in numerical analysts on the "effective" rank of a matrix (e.g. Golub & Van Loan (1996), Björck (1996), Hansen (1998)).
- Multivariate statistics mainly considers the case of i.i.d. normal data: Hotelling (1936), Bartlett (1947), Anderson (1951), Izenman (1975), Anderson (1999).
- The econometrics literature (surveyed in Camba-Mendez & Kapetanios (2009)) considers rank testing under general conditions. However...

The Literature

- Rank tests are pervasive in economics: cointegration, factor models, identification, subspace Granger causality analysis, etc.
- There is a large literature in numerical analysts on the “effective” rank of a matrix (e.g. Golub & Van Loan (1996), Björck (1996), Hansen (1998)).
- Multivariate statistics mainly considers the case of i.i.d. normal data: Hotelling (1936), Bartlett (1947), Anderson (1951), Izenman (1975), Anderson (1999).
- The econometrics literature (surveyed in Camba-Mendez & Kapetanios (2009)) considers rank testing under general conditions. However...
  1. The asymptotics are difficult (sometimes plainly wrong!).

The Literature

- Rank tests are pervasive in economics: cointegration, factor models, identification, subspace Granger causality analysis, etc.
- There is a large literature in numerical analysts on the "effective" rank of a matrix (e.g. Golub & Van Loan (1996), Björck (1996), Hansen (1998)).
- Multivariate statistics mainly considers the case of i.i.d. normal data: Hotelling (1936), Bartlett (1947), Anderson (1951), Izenman (1975), Anderson (1999).
- The econometrics literature (surveyed in Camba-Mendez & Kapetanios (2009)) considers rank testing under general conditions. However...
  1. The asymptotics are difficult (sometimes plainly wrong!).
  2. It is not clear what relationships exist between the various rank testing statistics.

The Literature

- Rank tests are pervasive in economics: cointegration, factor models, identification, subspace Granger causality analysis, etc.
- There is a large literature in numerical analysts on the "effective" rank of a matrix (e.g. Golub & Van Loan (1996), Björck (1996), Hansen (1998)).
- Multivariate statistics mainly considers the case of i.i.d. normal data: Hotelling (1936), Bartlett (1947), Anderson (1951), Izenman (1975), Anderson (1999).
- The econometrics literature (surveyed in Camba-Mendez & Kapetanios (2009)) considers rank testing under general conditions. However...
  1. The asymptotics are difficult (sometimes plainly wrong!).
  2. It is not clear what relationships exist between the various rank testing statistics.
  3. It does not take full advantage of the numerical analysis literature.

The Literature

- Rank tests are pervasive in economics: cointegration, factor models, identification, subspace Granger causality analysis, etc.
- There is a large literature in numerical analysts on the "effective" rank of a matrix (e.g. Golub & Van Loan (1996), Björck (1996), Hansen (1998)).
- Multivariate statistics mainly considers the case of i.i.d. normal data: Hotelling (1936), Bartlett (1947), Anderson (1951), Izenman (1975), Anderson (1999).
- The econometrics literature (surveyed in Camba-Mendez & Kapetanios (2009)) considers rank testing under general conditions. However...
  1. The asymptotics are difficult (sometimes plainly wrong!).
  2. It is not clear what relationships exist between the various rank testing statistics.
  3. It does not take full advantage of the numerical analysis literature.
  4. There is no fixed–b theory of rank testing (Kiefer et. al. (2000), Vogelsang (2001), Kiefer & Vogelsang (2002a,b, 2005)).

The Main Contribution

- The general structure of every rank testing statistic is:

$$T^{\theta}\tau\left(\underbrace{\{x_1,\ldots,x_T\}}_{\text{data}},\ \underbrace{P_{\widehat{N}_r},\quad P_{\widehat{M}_r}}_{\text{null space estimators}}\right).$$

The Main Contribution

- The general structure of every rank testing statistic is:

$$T^{\theta}\tau\left(\underbrace{\{x_1,\ldots,x_T\}}_{\text{data}},\ \underbrace{P_{\widehat{N}_r},\quad P_{\widehat{M}_r}}_{\text{null space estimators}}\right).$$

- In fact, the vast majority are of the form: $T^{\theta}\tau\left(\widehat{B},\widehat{\Omega},P_{\widehat{N}_r},P_{\widehat{M}_r}\right)$.

The Main Contribution

- The general structure of every rank testing statistic is:

$$T^\theta\tau\left(\underbrace{\{x_1,\ldots,x_T\}}_{\text{data}},\ \underbrace{P_{\widehat{N}_r},\quad P_{\widehat{M}_r}}_{\text{null space estimators}}\right).$$

- In fact, the vast majority are of the form: $T^\theta\tau\left(\widehat{B},\widehat{\Omega},P_{\widehat{N}_r},P_{\widehat{M}_r}\right)$.
- The asymptotic behaviour of all rank test statistics is identical to:

$$T^\theta\tau\left(\{x_1,\ldots,x_T\},\ \underbrace{P_{N_{rT}},\quad P_{M_{rT}}}_{\text{population analogues}}\right).$$

The Main Contribution

- The general structure of every rank testing statistic is:

$$T^{\theta}\tau\left(\underbrace{\{x_1,\ldots,x_T\}}_{\text{data}},\ \underbrace{P_{\widehat{N}_r},\quad P_{\widehat{M}_r}}_{\text{null space estimators}}\right).$$

- In fact, the vast majority are of the form: $T^{\theta}\tau\left(\widehat{B},\widehat{\Omega},P_{\widehat{N}_r},P_{\widehat{M}_r}\right)$.
- The asymptotic behaviour of all rank test statistics is identical to:

$$T^{\theta}\tau\left(\{x_1,\ldots,x_T\},\ \underbrace{P_{N_{rT}},\quad P_{M_{rT}}}_{\text{population analogues}}\right).$$

- This is termed the **plug–in principle** for rank testing statistics.

Corollaries to the Main Contribution

- The asymptotics of the infeasible statistics are much simpler.

Corollaries to the Main Contribution

- The asymptotics of the infeasible statistics are much simpler.
  1. We can easily obtain the behaviour of rank testing statistic under the null and the local and global alternatives.

Corollaries to the Main Contribution

- The asymptotics of the infeasible statistics are much simpler.
    1. We can easily obtain the behaviour of rank testing statistic under the null and the local and global alternatives.
    2. We can easily obtain the behaviour of rank test statistics under misspecification (White, 1994).

Corollaries to the Main Contribution

- The asymptotics of the infeasible statistics are much simpler.
  1. We can easily obtain the behaviour of rank testing statistic under the null and the local and global alternatives.
  2. We can easily obtain the behaviour of rank test statistics under misspecification (White, 1994).
- We obtain many new rank testing statistics based on different functional forms, different matrix decompositions (e.g. QR and Cholesky), and fixed–b asymptotics.

Corollaries to the Main Contribution

- The asymptotics of the infeasible statistics are much simpler.
  1. We can easily obtain the behaviour of rank testing statistic under the null and the local and global alternatives.
  2. We can easily obtain the behaviour of rank test statistics under misspecification (White, 1994).
- We obtain many new rank testing statistics based on different functional forms, different matrix decompositions (e.g. QR and Cholesky), and fixed–b asymptotics.
- We clarify the relationships between the various rank tests in the literature.
  1. Example: the Cragg & Donald (1996,1997) and Kleibergen & Paap (2006) statistics have the same local power.

Corollaries to the Main Contribution

- The asymptotics of the infeasible statistics are much simpler.
  1. We can easily obtain the behaviour of rank testing statistic under the null and the local and global alternatives.
  2. We can easily obtain the behaviour of rank test statistics under misspecification (White, 1994).
- We obtain many new rank testing statistics based on different functional forms, different matrix decompositions (e.g. QR and Cholesky), and fixed–b asymptotics.
- We clarify the relationships between the various rank tests in the literature.
  1. Example: the Cragg & Donald (1996,1997) and Kleibergen & Paap (2006) statistics have the same local power.
  2. Example: the Johansen (1988) and Kleibergen & van Dijk (1994) statistics have the same local power.

Examples

- Consider the regression model

$$y_t = \underset{n\times m}{B} x_t + \varepsilon_t, \qquad t = 1, \ldots, T,$$

under the usual assumptions.

Examples

- Consider the regression model

$$y_t = \underset{n \times m}{B} x_t + \varepsilon_t, \qquad t = 1, \ldots, T,$$

under the usual assumptions.

- We would like to test

$$H_0(r) : B = B^*, \qquad \mathrm{rank}(B^*) = r$$

Examples

- Consider the regression model

$$y_t = \underset{n\times m}{B} x_t + \varepsilon_t, \qquad t = 1, \ldots, T,$$

under the usual assumptions.

- We would like to test

$$H_0(r) : B = B^*, \qquad \mathrm{rank}(B^*) = r$$

against the local alternative

$$H_T(r) : B = B^* + D/\sqrt{T}, \qquad \mathrm{rank}(B^*) = r$$

Examples

- Consider the regression model

$$y_t = \underset{n \times m}{B} x_t + \varepsilon_t, \qquad t = 1, \dots, T,$$

under the usual assumptions.

- We would like to test

$$H_0(r) : B = B^*, \qquad \operatorname{rank}(B^*) = r$$

against the local alternative

$$H_T(r) : B = B^* + D/\sqrt{T}, \qquad \operatorname{rank}(B^*) = r$$

and the global alternative

$$H_1(r) : B = B^*, \qquad \operatorname{rank}(B^*) > r.$$

Examples

- Let $\widehat{B}$ and $\widehat{\Omega}$ be MLEs of B and the long run variance of $\widehat{B}$.

Examples

- Let $\widehat{B}$ and $\widehat{\Omega}$ be MLEs of B and the long run variance of $\widehat{B}$. The Anderson (1951) trace statistic for testing the hypothesis $H_0(r)$ can be expressed as

Examples

- Let $\widehat{B}$ and $\widehat{\Omega}$ be MLEs of B and the long run variance of $\widehat{B}$. The Anderson (1951) trace statistic for testing the hypothesis $H_0(r)$ can be expressed as

$$F = T\text{vec}'(P_{\widehat{N}_r}\widehat{B}P_{\widehat{M}_r})\{(P_{\widehat{M}_r} \otimes P_{\widehat{N}_r})\widehat{\Omega}(P_{\widehat{M}_r} \otimes P_{\widehat{N}_r})\}^{\dagger}\text{vec}(P_{\widehat{N}_r}\widehat{B}P_{\widehat{M}_r}).$$

where $\widehat{N}_r$ and $\widehat{M}_r$ are related to the canonical variates of y and x.

Examples

- Let $\widehat{B}$ and $\widehat{\Omega}$ be MLEs of B and the long run variance of $\widehat{B}$. The Anderson (1951) trace statistic for testing the hypothesis $H_0(r)$ can be expressed as

$$F = T\mathrm{vec}'(P_{\widehat{N}_r}\widehat{B}P_{\widehat{M}_r})\{(P_{\widehat{M}_r}\otimes P_{\widehat{N}_r})\widehat{\Omega}(P_{\widehat{M}_r}\otimes P_{\widehat{N}_r})\}^{\dagger}\mathrm{vec}(P_{\widehat{N}_r}\widehat{B}P_{\widehat{M}_r}).$$

  where $\widehat{N}_r$ and $\widehat{M}_r$ are related to the canonical variates of y and x.
- The Johansen (1988), Cragg & Donald (1996,1997), Robin & Smith (2000), Kleibergen and Paap (2006), and Donald, Fortuna, & Pipiras (2007) statistics are all of this form.

Examples

- Let $\widehat{B}$ and $\widehat{\Omega}$ be MLEs of B and the long run variance of $\widehat{B}$. The Anderson (1951) trace statistic for testing the hypothesis $H_0(r)$ can be expressed as

$$F = T\mathrm{vec}'(P_{\widehat{N}_r}\widehat{B}P_{\widehat{M}_r})\{(P_{\widehat{M}_r}\otimes P_{\widehat{N}_r})\widehat{\Omega}(P_{\widehat{M}_r}\otimes P_{\widehat{N}_r})\}^{\dagger}\mathrm{vec}(P_{\widehat{N}_r}\widehat{B}P_{\widehat{M}_r}).$$

  where $\widehat{N}_r$ and $\widehat{M}_r$ are related to the canonical variates of y and x.
- The Johansen (1988), Cragg & Donald (1996,1997), Robin & Smith (2000), Kleibergen and Paap (2006), and Donald, Fortuna, & Pipiras (2007) statistics are all of this form.
- They differ only in their constructions of $P_{\widehat{N}_r}$ and $P_{\widehat{M}_r}$.

Examples

- For symmetric positive definite B, Donald, Fortuna, & Pipiras (2007) have proposed:

$$t = \frac{\sqrt{T}\mathrm{tr}(P_{\widehat{M_r}}\widehat{B}P_{\widehat{M_r}})}{\sqrt{\mathrm{vec}'(I_m)(P_{\widehat{M_r}} \otimes P_{\widehat{M_r}})\widehat{\Omega}(P_{\widehat{M_r}} \otimes P_{\widehat{M_r}})\mathrm{vec}(I_m)}}.$$

Examples

- For symmetric positive definite B, Donald, Fortuna, & Pipiras (2007) have proposed:

$$t = \frac{\sqrt{T}\mathrm{tr}(P_{\widehat{M}_r}\widehat{B}P_{\widehat{M}_r})}{\sqrt{\mathrm{vec}'(I_m)(P_{\widehat{M}_r}\otimes P_{\widehat{M}_r})\widehat{\Omega}(P_{\widehat{M}_r}\otimes P_{\widehat{M}_r})\mathrm{vec}(I_m)}}.$$

- All of the above (and many more) have the form $T^{\theta}\tau(\widehat{B},\widehat{\Omega},P_{\widehat{N}_r},P_{\widehat{M}_r}) = T^{\theta}\kappa(P_{\widehat{N}_r}\widehat{B}P_{\widehat{M}_r},(P_{\widehat{M}_r}\otimes P_{\widehat{N}_r})\widehat{\Omega}(P_{\widehat{M}_r}\otimes P_{\widehat{N}_r}))$.

Null Spaces Estimation: the Basic Idea

- Numerical analysts have long known that there are problems with the concept of rank in practice.

Null Spaces Estimation: the Basic Idea

- Numerical analysts have long known that there are problems with the concept of rank in practice.
- For example

$$\widehat{B} = \begin{bmatrix} 1 & 1 \\ 0 & \varepsilon \end{bmatrix}$$

is of rank 2 for all $\varepsilon \neq 0$ but for small values of ε its rank is "effectively" 1.

Null Spaces Estimation: the Basic Idea

- Numerical analysts have long known that there are problems with the concept of rank in practice.
- For example

$$\widehat{B} = \begin{bmatrix} 1 & 1 \\ 0 & \varepsilon \end{bmatrix}$$

is of rank 2 for all $\varepsilon \neq 0$ but for small values of ε its rank is "effectively" 1.
- The reduced rank approximation (RRA)

$$\widehat{B} \approx \widehat{B}_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Null Spaces Estimation: the Basic Idea

- Numerical analysts have long known that there are problems with the concept of rank in practice.
- For example

$$\widehat{B} = \begin{bmatrix} 1 & 1 \\ 0 & \varepsilon \end{bmatrix}$$

is of rank 2 for all $\varepsilon \neq 0$ but for small values of ε its rank is "effectively" 1.
- The reduced rank approximation (RRA)

$$\widehat{B} \approx \widehat{B}_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

allows us to estimate the null spaces of $\widehat{B}$ as

$$\widehat{N}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \widehat{M}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Null Spaces Estimation: the Basic Idea

- Numerical analysts have long known that there are problems with the concept of rank in practice.
- For example

$$\widehat{B} = \begin{bmatrix} 1 & 1 \\ 0 & \varepsilon \end{bmatrix}$$

is of rank 2 for all $\varepsilon \neq 0$ but for small values of ε its rank is "effectively" 1.

- The reduced rank approximation (RRA)

$$\widehat{B} \approx \widehat{B}_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

allows us to estimate the null spaces of $\widehat{B}$ as

$$\widehat{N}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \widehat{M}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

- There are numerous algorithms in the literature that can identify the effective rank of a matrix (SVD, GSVD, WLRA, LU, QR, etc.).

Null Spaces Estimation: the Basic Idea

In summary, for a given $\widehat{B}$ that is converging to B^*. We may estimate the null spaces of B^* by the following steps:

Null Spaces Estimation: the Basic Idea

In summary, for a given $\widehat{B}$ that is converging to B^*. We may estimate the null spaces of B^* by the following steps:

- Approximate $\widehat{B}$ by $\widehat{B}_r^{RRA}$ of rank r.

Null Spaces Estimation: the Basic Idea

In summary, for a given $\widehat{B}$ that is converging to B^*. We may estimate the null spaces of B^* by the following steps:

- Approximate $\widehat{B}$ by $\widehat{B}_r^{RRA}$ of rank r.
- Take $\widehat{N}_r$ to span the left null space of $\widehat{B}_r^{RRA}$.

Null Spaces Estimation: the Basic Idea

In summary, for a given $\widehat{B}$ that is converging to B^*. We may estimate the null spaces of B^* by the following steps:

- Approximate $\widehat{B}$ by $\widehat{B}_r^{RRA}$ of rank r.
- Take $\widehat{N}_r$ to span the left null space of $\widehat{B}_r^{RRA}$.
- Take $\widehat{M}_r$ to span the right null space of $\widehat{B}_r^{RRA}$.

Null Spaces Estimation: the Basic Idea

In summary, for a given $\widehat{B}$ that is converging to B^*. We may estimate the null spaces of B^* by the following steps:

- Approximate $\widehat{B}$ by $\widehat{B}_r^{RRA}$ of rank r.
- Take $\widehat{N}_r$ to span the left null space of $\widehat{B}_r^{RRA}$.
- Take $\widehat{M}_r$ to span the right null space of $\widehat{B}_r^{RRA}$.

Lemma

Suppose $\sqrt{T}(\widehat{B} - B^*) = O_p(1)$, $r = \mathrm{rank}(B^*)$, *then*

Null Spaces Estimation: the Basic Idea

In summary, for a given $\widehat{B}$ that is converging to B^*. We may estimate the null spaces of B^* by the following steps:

- Approximate $\widehat{B}$ by $\widehat{B}_r^{RRA}$ of rank r.
- Take $\widehat{N}_r$ to span the left null space of $\widehat{B}_r^{RRA}$.
- Take $\widehat{M}_r$ to span the right null space of $\widehat{B}_r^{RRA}$.

Lemma

Suppose $\sqrt{T}(\widehat{B} - B^*) = O_p(1)$, $r = \mathrm{rank}(B^*)$, *then*

1. $\sqrt{T}(P_{\widehat{N}_r} - P_{N_r})$ *and* $\sqrt{T}(P_{\widehat{M}_r} - P_{M_r})$ *are* $O_p(1)$.

Null Spaces Estimation: the Basic Idea

In summary, for a given $\widehat{B}$ that is converging to B^*. We may estimate the null spaces of B^* by the following steps:

- Approximate $\widehat{B}$ by $\widehat{B}_r^{RRA}$ of rank r.
- Take $\widehat{N}_r$ to span the left null space of $\widehat{B}_r^{RRA}$.
- Take $\widehat{M}_r$ to span the right null space of $\widehat{B}_r^{RRA}$.

Lemma

Suppose $\sqrt{T}(\widehat{B} - B^*) = O_p(1)$, $r = \operatorname{rank}(B^*)$, *then*

1. $\sqrt{T}(P_{\widehat{N}_r} - P_{N_r})$ *and* $\sqrt{T}(P_{\widehat{M}_r} - P_{M_r})$ *are* $O_p(1)$.
2. *If* $0 \leq i < r$ *then* $P_{\widehat{N}_i} \widehat{B} P_{\widehat{M}_i}$ *is bounded away from zero in probability.*

Null Spaces Estimation: the Basic Idea

In summary, for a given $\widehat{B}$ that is converging to B^*. We may estimate the null spaces of B^* by the following steps:

- Approximate $\widehat{B}$ by $\widehat{B}_r^{RRA}$ of rank r.
- Take $\widehat{N}_r$ to span the left null space of $\widehat{B}_r^{RRA}$.
- Take $\widehat{M}_r$ to span the right null space of $\widehat{B}_r^{RRA}$.

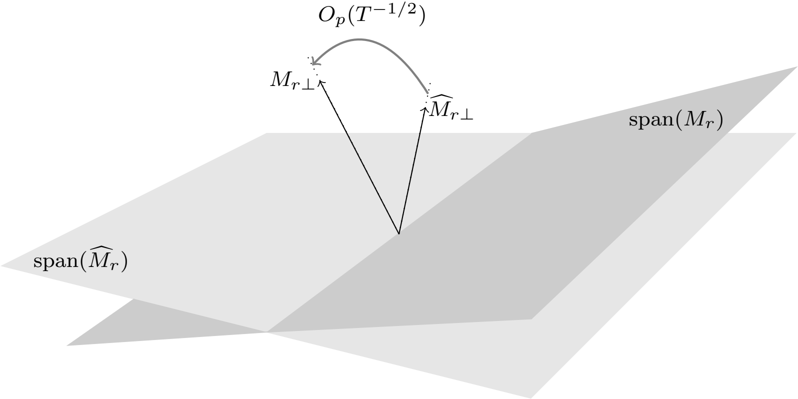
Lemma

Suppose $\sqrt{T}(\widehat{B} - B^*) = O_p(1)$, $r = \operatorname{rank}(B^*)$, *then*

1. $\sqrt{T}(P_{\widehat{N}_r} - P_{N_r})$ *and* $\sqrt{T}(P_{\widehat{M}_r} - P_{M_r})$ *are* $O_p(1)$.
2. *If* $0 \leq i < r$ *then* $P_{\widehat{N}_i} \widehat{B} P_{\widehat{M}_i}$ *is bounded away from zero in probability.*
3. *If* $0 \leq i < r$ *and the RRA is continuous at* B^*, *then* $P_{\widehat{N}_i} \widehat{B} P_{\widehat{M}_i}$ *converges in probability.*

Null Space Estimation in General

Figure: Convergence of a Two Dimensional Null Space Estimator in $\mathbb{R}^3$.



The Basic Set of Assumptions

The basic set of assumptions is:

1. $\widehat{B} \in \mathbb{R}^{n\times m}$ and $\widehat{\Omega} \in \mathbb{R}^{nm\times nm}$ are estimators indexed by T.

The Basic Set of Assumptions

The basic set of assumptions is:

1. $\widehat{B} \in \mathbb{R}^{n \times m}$ and $\widehat{\Omega} \in \mathbb{R}^{nm \times nm}$ are estimators indexed by T.
2. Each $\mathrm{vec}(\widehat{B}) \in \mathbb{R}^{nm}$ is a non–degenerate random vector.

The Basic Set of Assumptions

The basic set of assumptions is:

1. $\widehat{B} \in \mathbb{R}^{n\times m}$ and $\widehat{\Omega} \in \mathbb{R}^{nm\times nm}$ are estimators indexed by T.
2. Each $\mathrm{vec}(\widehat{B}) \in \mathbb{R}^{nm}$ is a non–degenerate random vector.
3. $\widehat{\Omega}$ is symmetric positive definite almost surely.

The Basic Set of Assumptions

The basic set of assumptions is:

1. $\widehat{B} \in \mathbb{R}^{n \times m}$ and $\widehat{\Omega} \in \mathbb{R}^{nm \times nm}$ are estimators indexed by T.
2. Each $\mathrm{vec}(\widehat{B}) \in \mathbb{R}^{nm}$ is a non–degenerate random vector.
3. $\widehat{\Omega}$ is symmetric positive definite almost surely.
4. $\sqrt{T}(\widehat{B} - B^*)$, $\widehat{\Omega}$, and $\widehat{\Omega}^{-1}$ are $O_p(1)$.

The Basic Set of Assumptions

The basic set of assumptions is:

1. $\widehat{B} \in \mathbb{R}^{n\times m}$ and $\widehat{\Omega} \in \mathbb{R}^{nm\times nm}$ are estimators indexed by T.
2. Each $\mathrm{vec}(\widehat{B}) \in \mathbb{R}^{nm}$ is a non–degenerate random vector.
3. $\widehat{\Omega}$ is symmetric positive definite almost surely.
4. $\sqrt{T}(\widehat{B} - B^*)$, $\widehat{\Omega}$, and $\widehat{\Omega}^{-1}$ are $O_p(1)$.

Similar assumptions are made for testing the rank of symmetric matrices.

The Feasible and Infeasible Statistics

Let $\widehat{N}_r$ and $\widehat{M_r}$ be null space estimators based on $\widehat{B}$. The **feasible** rank test statistic is:

$$T^{\theta}\tau\left(\widehat{B},\widehat{\Omega},P_{\widehat{N}_r},P_{\widehat{M_r}}\right)$$

The Feasible and Infeasible Statistics

Let $\widehat{N}_r$ and $\widehat{M}_r$ be null space estimators based on $\widehat{B}$. The **feasible** rank test statistic is:

$$T^{\theta}\tau\left(\widehat{B}, \widehat{\Omega}, P_{\widehat{N}_r}, P_{\widehat{M}_r}\right)$$

Let N_r and M_r span the null spaces of B^*. The **infeasible** rank test statistic is:

$$T^{\theta}\tau\left(\widehat{B}, \widehat{\Omega}, P_{N_r}, P_{M_r}\right)$$

The Plug–in Principle

The **weak plug–in principle** for rank test statistics is said to hold for a rank test statistic if:

The Plug–in Principle

The **weak plug–in principle** for rank test statistics is said to hold for a rank test statistic if:

1. Under either $H_0(r)$ or $H_T(r)$, the difference between the feasible and infeasible statistics is $O_p(T^{-1/2})$.

The Plug–in Principle

The **weak plug–in principle** for rank test statistics is said to hold for a rank test statistic if:

1. Under either $H_0(r)$ or $H_T(r)$, the difference between the feasible and infeasible statistics is $O_p(T^{-1/2})$.
2. Under $H_1(r)$, if the infeasible statistic has power, then so does the feasible statistic.

The Plug–in Principle

The **weak plug–in principle** for rank test statistics is said to hold for a rank test statistic if:

1. Under either $H_0(r)$ or $H_T(r)$, the difference between the feasible and infeasible statistics is $O_p(T^{-1/2})$.
2. Under $H_1(r)$, if the infeasible statistic has power, then so does the feasible statistic.

   It is said to satisfy the **strong plug–in principle** relative to the null spaces of B^* if additionally

3. Under $H_1(r)$, the feasible and infeasible statistics diverge at the same rate.

Testing Rank Hypotheses for General Matrices

Theorem

Under weak regularity conditions on τ*,* $T^{\theta}\tau(\widehat{B}, \widehat{\Omega}, P_{\widehat{N}_r}, P_{\widehat{M}_r})$ *satisfies the weak plug–in principle for rank test statistics. If, additionally, the RRA is continuous at* B^**, then the statistic satisfies the strong plug–in principle.*

Testing Rank Hypotheses for General Matrices

Theorem

Under weak regularity conditions on τ, $T^{\theta}\tau(\widehat{B}, \widehat{\Omega}, P_{\widehat{N}_r}, P_{\widehat{M}_r})$ satisfies the weak plug–in principle for rank test statistics. If, additionally, the RRA is continuous at B^*, then the statistic satisfies the strong plug–in principle.

- All of the rank testing statistics of the literature (and many more) satisfy the weak plug–in principle.

Testing Rank Hypotheses for General Matrices

Theorem

Under weak regularity conditions on τ, $T^{\theta}\tau(\widehat{B}, \widehat{\Omega}, P_{\widehat{N}_r}, P_{\widehat{M_r}})$ *satisfies the weak plug–in principle for rank test statistics. If, additionally, the RRA is continuous at* B^*, *then the statistic satisfies the strong plug–in principle.*

- All of the rank testing statistics of the literature (and many more) satisfy the weak plug–in principle. When the underlying RRA is continuous at the population matrix (a feature we have conjectured to be generic for all RRAs), also satisfy the strong plug–in principle.

Testing Rank Hypotheses for General Matrices

Theorem

Under weak regularity conditions on τ, $T^{\theta}\tau(\widehat{B}, \widehat{\Omega}, P_{\widehat{N}_r}, P_{\widehat{M}_r})$ *satisfies the weak plug–in principle for rank test statistics. If, additionally, the RRA is continuous at* B^*, *then the statistic satisfies the strong plug–in principle.*

- All of the rank testing statistics of the literature (and many more) satisfy the weak plug–in principle. When the underlying RRA is continuous at the population matrix (a feature we have conjectured to be generic for all RRAs), also satisfy the strong plug–in principle.
- The Cragg & Donald (1996,1997) and Kleibergen and Paap (2006) statistics are asymptotically equivalent.

Testing Rank Hypotheses for General Matrices

Theorem

Under weak regularity conditions on τ, $T^{\theta}\tau(\widehat{B}, \widehat{\Omega}, P_{\widehat{N}_r}, P_{\widehat{M}_r})$ *satisfies the weak plug–in principle for rank test statistics. If, additionally, the RRA is continuous at* B^*, *then the statistic satisfies the strong plug–in principle.*

- All of the rank testing statistics of the literature (and many more) satisfy the weak plug–in principle. When the underlying RRA is continuous at the population matrix (a feature we have conjectured to be generic for all RRAs), also satisfy the strong plug–in principle.
- The Cragg & Donald (1996,1997) and Kleibergen and Paap (2006) statistics are asymptotically equivalent. When $\widehat{\Omega}$ is a Kronecker product, we may add to the list the statistics of Anderson (1951) and Robin & Smith (2000).

Testing Rank Hypotheses for General Matrices

Theorem

Under weak regularity conditions on τ, $T^{\theta}\tau(\widehat{B}, \widehat{\Omega}, P_{\widehat{N}_r}, P_{\widehat{M}_r})$ satisfies the weak plug–in principle for rank test statistics. If, additionally, the RRA is continuous at B^*, then the statistic satisfies the strong plug–in principle.

- All of the rank testing statistics of the literature (and many more) satisfy the weak plug–in principle. When the underlying RRA is continuous at the population matrix (a feature we have conjectured to be generic for all RRAs), also satisfy the strong plug–in principle.
- The Cragg & Donald (1996,1997) and Kleibergen and Paap (2006) statistics are asymptotically equivalent. When $\widehat{\Omega}$ is a Kronecker product, we may add to the list the statistics of Anderson (1951) and Robin & Smith (2000).
- As there are no first–order differences between these statistics, we must look for either higher–order difference or Monte Carlo performance for guidance.

Corollary

Corollary

Under $H_0(r)$ or $H_T(r)$, let $N_r \in \mathbb{G}^{n\times(n-r)}$ and $M_r \in \mathbb{G}^{m\times(m-r)}$ span the left and right null spaces of B^*. If

$$\left(\sqrt{T}\mathrm{vec}(N_r'\widehat{B}M_r), (M_r \otimes N_r)'\widehat{\Omega}(M_r \otimes N_r)\right) \xrightarrow{d} (\xi_r, \Omega_r),$$

then we have

$$F \xrightarrow{d} \xi_r'\Omega_r^\dagger\xi_r \qquad\qquad t \xrightarrow{d} \frac{\mathrm{tr}(\mathrm{mat}(D_{m-r}\xi_r))}{(\mathrm{vec}'(I_{m-r})\Omega_r\mathrm{vec}(I_{m-r}))^{1/2}}.$$

Corollary

Corollary

Under $H_0(r)$ *or* $H_T(r)$, *let* $N_r \in \mathbb{G}^{n\times(n-r)}$ *and* $M_r \in \mathbb{G}^{m\times(m-r)}$ *span the left and right null spaces of* B^*. *If*

$$\left(\sqrt{T}\mathrm{vec}(N_r'\widehat{B}M_r), (M_r \otimes N_r)'\widehat{\Omega}(M_r \otimes N_r)\right) \xrightarrow{d} (\xi_r, \Omega_r),$$

then we have

$$F \xrightarrow{d} \xi_r'\Omega_r^\dagger\xi_r \qquad\qquad t \xrightarrow{d} \frac{\mathrm{tr}(\mathrm{mat}(D_{m-r}\xi_r))}{(\mathrm{vec}'(I_{m-r})\Omega_r\mathrm{vec}(I_{m-r}))^{1/2}}.$$

- Local power is simple to determine. Generalizes Cragg & Donald (1997).

Corollary

Corollary

Under $H_0(r)$ or $H_T(r)$, let $N_r \in \mathbb{G}^{n\times(n-r)}$ and $M_r \in \mathbb{G}^{m\times(m-r)}$ span the left and right null spaces of B^*. If

$$\left(\sqrt{T}\mathrm{vec}(N_r'\widehat{B}M_r), (M_r \otimes N_r)'\widehat{\Omega}(M_r \otimes N_r)\right) \xrightarrow{d} (\xi_r, \Omega_r),$$

then we have

$$F \xrightarrow{d} \xi_r'\Omega_r^{\dagger}\xi_r \qquad\qquad t \xrightarrow{d} \frac{\mathrm{tr}(\mathrm{mat}(D_{m-r}\xi_r))}{(\mathrm{vec}'(I_{m-r})\Omega_r\mathrm{vec}(I_{m-r}))^{1/2}}.$$

- Local power is simple to determine. Generalizes Cragg & Donald (1997).
- Misspecification asymptotics follow from White (1994). Generalizes Robin & Smith (2000).

Corollary

Corollary

Under $H_0(r)$ or $H_T(r)$, let $N_r \in \mathbb{G}^{n\times(n-r)}$ and $M_r \in \mathbb{G}^{m\times(m-r)}$ span the left and right null spaces of B^*. If

$$\left(\sqrt{T}\mathrm{vec}(N_r'\widehat{B}M_r), (M_r \otimes N_r)'\widehat{\Omega}(M_r \otimes N_r)\right) \xrightarrow{d} (\xi_r, \Omega_r),$$

then we have

$$F \xrightarrow{d} \xi_r'\Omega_r^{\dagger}\xi_r \qquad\qquad t \xrightarrow{d} \frac{\mathrm{tr}(\mathrm{mat}(D_{m-r}\xi_r))}{(\mathrm{vec}'(I_{m-r})\Omega_r\mathrm{vec}(I_{m-r}))^{1/2}}.$$

- Local power is simple to determine. Generalizes Cragg & Donald (1997).
- Misspecification asymptotics follow from White (1994). Generalizes Robin & Smith (2000).
- $\widehat{\Omega}$ can be a fixed–b estimator.

Cointegration Caveat I: Johansen (1988,1995)

Consider the following VAR(1)

$$\Delta y_t = B y_{t-1} + \varepsilon_t, \qquad t = 1, \ldots, T.$$

Suppose $\{y_t\}$ is at most $I(2)$ and $\widehat{B}$ is the OLS estimator of B.

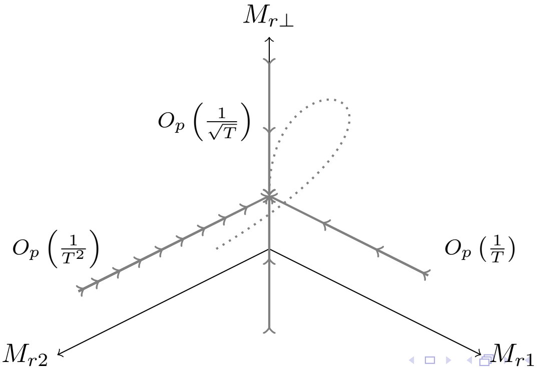
Cointegration Caveat I: Johansen (1988,1995)

Consider the following VAR(1)

$$\Delta y_t = B y_{t-1} + \varepsilon_t, \qquad t = 1, \ldots, T.$$

Suppose $\{y_t\}$ is at most $I(2)$ and $\widehat{B}$ is the OLS estimator of B.

Figure: Convergence of the Columns of $\widehat{B}$.



Cointegration Caveat II: Nyblom & Harvey (2000)

Consider the multivariate local level model

$$y_t = x_t + \varepsilon_t \qquad x_t = x_{t-1} + u_t, \qquad t = 1, \ldots, T.$$

Nyblom & Harvey test the rank of $\widehat{B} = \widehat{\Sigma}^{-1}\widehat{\Gamma}$ (analogue of the KPSS statistic).

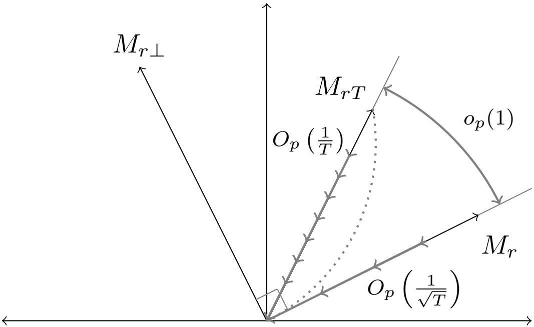
Cointegration Caveat II: Nyblom & Harvey (2000)

Consider the multivariate local level model

$$y_t = x_t + \varepsilon_t \qquad x_t = x_{t-1} + u_t, \qquad t = 1, \ldots, T.$$

Nyblom & Harvey test the rank of $\widehat{B} = \widehat{\Sigma}^{-1}\widehat{\Gamma}$ (analogue of the KPSS statistic).

Figure: Convergence of the Columns of $\widehat{B}$.



Cointegration Summary

- Null space estimation must take into account:
  1. Accelerated and heterogenous rates of convergence to zero.

Cointegration Summary

- Null space estimation must take into account:
  1. Accelerated and heterogenous rates of convergence to zero.
  2. The directions of accelerated convergence can be random and T–dependent.

Cointegration Summary

- Null space estimation must take into account:
  1. Accelerated and heterogenous rates of convergence to zero.
  2. The directions of accelerated convergence can be random and T–dependent.
- Null space estimators will exhibit accelerated and heterogenous rates of convergence under $H_0(r)$ and $H_T(r)$ and continue to detect the non–vanishing components of $\widehat{B}$ under $H_1(r)$.

Cointegration Summary

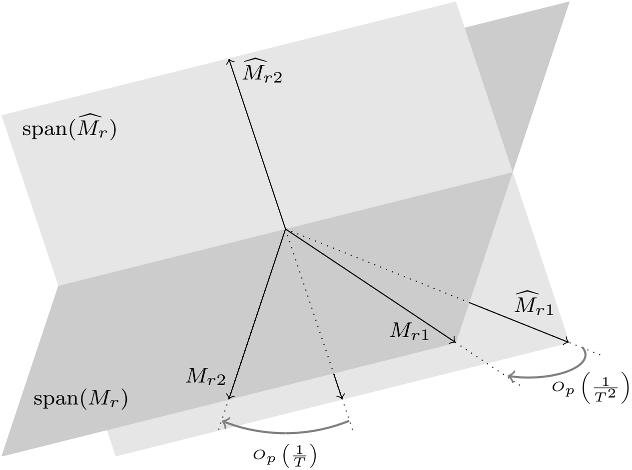
- Null space estimation must take into account:
    1. Accelerated and heterogenous rates of convergence to zero.
    2. The directions of accelerated convergence can be random and T–dependent.
- Null space estimators will exhibit accelerated and heterogenous rates of convergence under $H_0(r)$ and $H_T(r)$ and continue to detect the non–vanishing components of $\widehat{B}$ under $H_1(r)$.
- The general plug–in principle is relative to a random sequence of matrices.

Cointegration Summary

- Null space estimation must take into account:
  1. Accelerated and heterogenous rates of convergence to zero.
  2. The directions of accelerated convergence can be random and T–dependent.
- Null space estimators will exhibit accelerated and heterogenous rates of convergence under $H_0(r)$ and $H_T(r)$ and continue to detect the non–vanishing components of $\widehat{B}$ under $H_1(r)$.
- The general plug–in principle is relative to a random sequence of matrices.
- The general theory nests the standard asymptotics case, as well as the polynomial regression and the vast majority of cointegration settings.

Null Space Estimation in Cointegration

Figure: Accelerated & Heterogeneous Rates of Null Space Convergence in $\mathbb{R}^3$.



Corollaries of the Plug–in Principle for Cointegration

- (Correct Specification). The limiting behaviour of all of the statistics in:

  Johansen (1988), Johansen (1991), Kleibergen & van Dijk (1994), Yang & Bewley (1996), Quintos (1998), Gonzalo & Pitarakis (1999), Lutkepohl & Saikkonen (1999), Kleibergen & Paap (2006), Avarucci & Velasco (2009), Cavaliere et al. (2010a),...

  follow from Corollaries 3 and 4 of the paper.

Corollaries of the Plug–in Principle for Cointegration

- (Correct Specification). The limiting behaviour of all of the statistics in:

  Johansen (1988), Johansen (1991), Kleibergen & van Dijk (1994), Yang & Bewley (1996), Quintos (1998), Gonzalo & Pitarakis (1999), Lutkepohl & Saikkonen (1999), Kleibergen & Paap (2006), Avarucci & Velasco (2009), Cavaliere et al. (2010a),. . .

  follow from Corollaries 3 and 4 of the paper.
- (Misspecification). The F statistics proposed by Johansen (1988), Kleibergen & van Dijk (1994), and Kleibergen & Paap (2006) have the exact same behaviour under the misspecification conditions of Caner (1998) (infinite variance shocks), Cavaliere et al. (2010b) (heteroskedastic shocks), and Aznar & Salvador (2002) and Cavaliere et al. (2014) (misspecified lag length).

Example 1

- Let the data be generated as

$$
\begin{aligned}
y_t &= Bx_t + \varepsilon_t \\
\varepsilon_t &= 0.5\varepsilon_{t-1} + u_t \quad \text{(stationary)} \\
\{(x_t', u_t')'\} &\text{ i.i.d. } N(0, I_8) \\
B &= \begin{bmatrix} 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.75 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.
\end{aligned}
$$

Example 1

- Let the data be generated as

$$
\begin{aligned}
y_t &= Bx_t + \varepsilon_t \\
\varepsilon_t &= 0.5\varepsilon_{t-1} + u_t \quad \text{(stationary)} \\
\{(x_t', u_t')'\} &\text{ i.i.d. } N(0, I_8) \\
B &= \begin{bmatrix} 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.75 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.
\end{aligned}
$$

- The sample consists of $\{(y_t', x_t')' : t = 1, \ldots, 50\}$.

Example 1

- Let the data be generated as

$$
\begin{aligned}
y_t &= Bx_t + \varepsilon_t \\
\varepsilon_t &= 0.5\varepsilon_{t-1} + u_t \quad \text{(stationary)} \\
\{(x_t', u_t')'\} &\text{ i.i.d. } N(0, I_8) \\
B &= \begin{bmatrix} 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.75 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.
\end{aligned}
$$

- The sample consists of $\{(y_t', x_t')' : t = 1, \ldots, 50\}$.
- Estimate B by OLS and the variance of $\widehat{B}$ by the Newey–West estimator with bandwidth $\lfloor 4(50/100)^{1/4} \rfloor = 3$ for the small–b case and bandwidth 50 for the fixed–b case.

Example 1

- Let the data be generated as

$$\begin{aligned} y_t &= Bx_t + \varepsilon_t \\ \varepsilon_t &= 0.5\varepsilon_{t-1} + u_t \quad \text{(stationary)} \\ \{(x_t', u_t')'\} &\text{ i.i.d. } N(0, I_8) \\ B &= \begin{bmatrix} 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.75 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

- The sample consists of $\{(y_t', x_t')' : t = 1, \ldots, 50\}$.
- Estimate B by OLS and the variance of $\widehat{B}$ by the Newey–West estimator with bandwidth $\lfloor 4(50/100)^{1/4} \rfloor = 3$ for the small–b case and bandwidth 50 for the fixed–b case.
- The number of replications is set to 2000.

Example 1

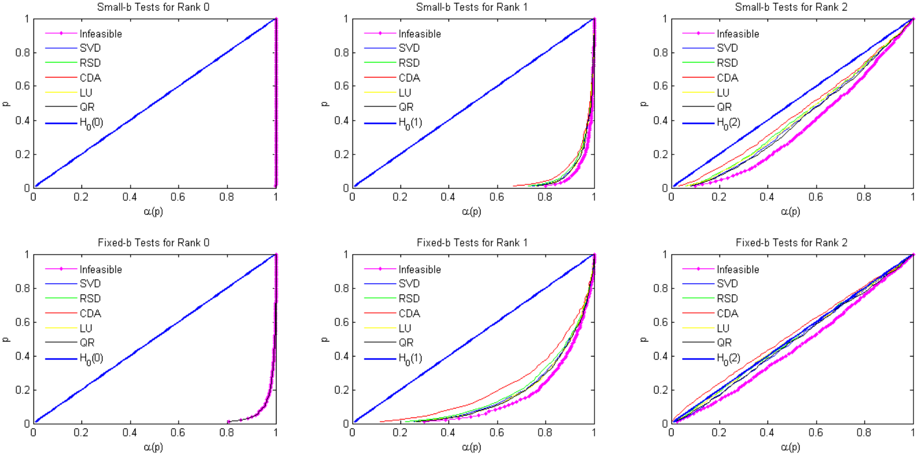
- Let the data be generated as

$$\begin{aligned} y_t &= Bx_t + \varepsilon_t \\ \varepsilon_t &= 0.5\varepsilon_{t-1} + u_t \quad \text{(stationary)} \\ \{(x_t', u_t')'\} &\text{ i.i.d. } N(0, I_8) \\ B &= \begin{bmatrix} 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.75 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

- The sample consists of $\{(y_t', x_t')' : t = 1, \ldots, 50\}$.
- Estimate B by OLS and the variance of $\widehat{B}$ by the Newey–West estimator with bandwidth $\lfloor 4(50/100)^{1/4} \rfloor = 3$ for the small–b case and bandwidth 50 for the fixed–b case.
- The number of replications is set to 2000.
- PP plots are reported for F.

Example 1

Figure: PP Plots for the F Statistic of Example 1.



Example 2

- Consider changing the matrix to

$$B = \begin{bmatrix} 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Example 2

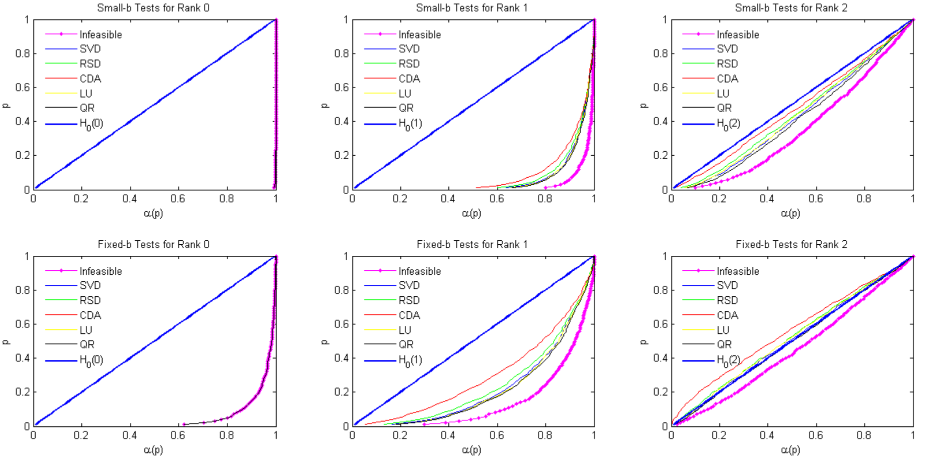
- Consider changing the matrix to

$$B = \begin{bmatrix} 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

- Rank–1 RRAs are discontinuous at B. Therefore, we expect the different F statistics for rank–1 to diverge at different rates.

Example 2

Figure: PP Plots for the F Statistic of Example 2.



Conclusion

- This paper has developed a general and simple theory of rank testing, which nests all existing rank tests as special cases and motivates many new ones.

Conclusion

- This paper has developed a general and simple theory of rank testing, which nests all existing rank tests as special cases and motivates many new ones.
- Practical take away:
  - Blue print for constructing custom rank test statistics.

Conclusion

- This paper has developed a general and simple theory of rank testing, which nests all existing rank tests as special cases and motivates many new ones.
- Practical take away:
  - Blue print for constructing custom rank test statistics.
  - Choice of statistic: whatever is convenient.

Conclusion

- This paper has developed a general and simple theory of rank testing, which nests all existing rank tests as special cases and motivates many new ones.
- Practical take away:
  - Blue print for constructing custom rank test statistics.
  - Choice of statistic: whatever is convenient.
  - QR and LU statistics are recommended when bootstrapping.

Conclusion

- This paper has developed a general and simple theory of rank testing, which nests all existing rank tests as special cases and motivates many new ones.
- Practical take away:
  - Blue print for constructing custom rank test statistics.
  - Choice of statistic: whatever is convenient.
  - QR and LU statistics are recommended when bootstrapping.
  - Fixed–b rank test statistics recommended when over–rejection is a problem.

Conclusion

- This paper has developed a general and simple theory of rank testing, which nests all existing rank tests as special cases and motivates many new ones.
- Practical take away:
  - Blue print for constructing custom rank test statistics.
  - Choice of statistic: whatever is convenient.
  - QR and LU statistics are recommended when bootstrapping.
  - Fixed–b rank test statistics recommended when over–rejection is a problem.
- Future research to focus on high-dimensional data.