Journal of Econometrics 199 (2017) 49–62

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

A unifying theory of tests of rank[☆]

Majid M. Al-Sadoon

Universitat Pompeu Fabra & Barcelona GSE, Spain

ARTICLE INFO

Article history:
Received 23 July 2015
Received in revised form 11 January 2017
Accepted 31 March 2017
Available online 17 April 2017

JEL classification:
C12
C13
C30

Keywords:
Tests of rank
Cointegration
Plug-in principle
Null space estimation
Reduced-rank approximation
Local power
Misspecification

ABSTRACT

The general principles underlying tests of matrix rank are investigated. It is demonstrated that statistics for such tests can be seen as implicit functions of null space estimators. In turn, the asymptotic behaviour of the null space estimators is shown to determine the asymptotic behaviour of the statistics through a plug-in principle. The theory simplifies the asymptotics under a variety of alternatives of empirical relevance as well as misspecification, clarifies the relationships between the various existing tests, makes use of important results in the numerical analysis literature, and motivates numerous new tests. A brief Monte Carlo study illustrates the results.

1. Introduction

The literature on tests of matrix rank has grown into a large and eminently applicable branch of econometrics since the seminal contribution by Anderson (1951) (see Camba-Mendez and Kapetanios (2009) for a survey). Much of this progress has taken place in spite of the difficulty of the asymptotics of these tests; indeed, statistics for testing the rank of a matrix often involve eigenvectors, inverses, and other discontinuous functions of matrices. Consequently, significant gaps have persisted in our knowledge of the relationships between the various tests in the literature. Relationships between the Anderson (1951), Johansen (1991), Robin and Smith (2000), and Kleibergen and Paap (2006) statistics are known, as are relationships between the Cragg and Donald (1996, 1997) statistics. However, there is as of yet no characterization of the general structure of statistics for testing the rank of a matrix. Another consequence, is that little is known about the behaviour of tests of rank under local alternatives or under misspecification. Local power is considered in Cragg and Donald (1997) and a handful of papers surveyed by Hubrich et al. (2001), while misspecification is considered in Robin and Smith (2000), Caner (1998), Cavaliere et al. (2010b), Aznar and Salvador (2002), and Cavaliere et al. (2014). All of these results relate to specific tests and there is as of yet no known general principle that unifies all of these results. The statistical and econometric literature has also made little use of the numerical analysis literature, which has made great strides in understanding and discovering effective matrix rank (Hansen, 1998).

Thus, the objective of this paper is to investigate the general principles underlying tests of matrix rank by: (i) characterizing the general structure of statistics for testing matrix rank, (ii) describing the behaviour of these statistics under a variety of alternatives of empirical relevance and misspecification, and (iii) making use of important insights from the numerical analysis literature. These intermediate objectives are achieved along the following steps.

[☆] Earlier versions of this paper were titled, ''A General Theory of Rank Testing'' and ''The Stochastic Wald Test''. I am grateful to the editor, the associated editor, and three anonymous referees for the many comments, suggestions, and corrections that have helped improved the manuscript. Thanks are also due to Sean Holly, M. Hashem Pesaran, Richard J. Smith, George Kapetanios, Rod McCrorie, Frank Kleibergen, Jesús Gonzalo, Vasco Carvalho, Kristoffer Nimark, Gábor Lugosi, Sergi Jiménez, Christian Brownlees, Geert Mesters, Barbara Rossi, Peter Boswijk, Herman van Dijk, Andrew Harvey, Luis Martins, Søren Johansen, Natercia Fortuna, Vladas Pipiras, Alan Izenman, Emre Aylar, Hanno Reuvers, Jean-Marie Dufour, Sara Nyman, and to seminar participants at Universitat Pompeu Fabra, Universidad Carlos III de Madrid, Universidade do Minho, Universidade do Porto, and Lund University for helpful comments and suggestions. The author takes sole responsibility for any remaining errors. Some of the results of this paper were part of the author's Ph.D. thesis. This paper is dedicated to T.W. Anderson (1918–2016), who made the seminal contribution to this field. Research for this paper was supported by Spanish Ministry of Economy and Competitiveness projects ECO2012-33247 and ECO2015-68136-P (MINECO/FEDER, UE) and Fundación BBVA Scientific Research Grant PR16-DAT-0043.

E-mail address: majid.alsadoon@gmail.com.

http://dx.doi.org/10.1016/j.jeconom.2017.03.002
0304-4076/

First, the paper shows that the general structure of statistics for tests of rank is of the form of an implicit function of estimators of the null spaces of the matrix in question (see Sections 3.1 and 4.3). This achieves intermediate objective (i) as it is demonstrated that most statistics in the literature have common functional forms although they may differ in the implicit null space estimators (see Table 1).

Next, the paper develops the theory of null space estimation based on reduced-rank approximations, which have been widely studied in the numerical analysis literature. This achieves intermediate objective (iii). Lemmas 1 and 2 provide a full characterization of the asymptotic properties of null space estimators under the various alternatives under study. These results generalize Dufour and Valery (2011) in that they apply to general matrices rather than just the positive semi-definite ones and are not restricted to eigenprojections. They also allow us to use any reduced-rank approximation to construct a statistical test of matrix rank; this is demonstrated by a number of new tests based on the QR and Cholesky decompositions (see Section 5).

Finally, it is demonstrated that the behaviour of statistics for tests of rank is completely governed by the implicit null space estimators. A plug-in principle is shown to hold, whereby every statistic mimics the asymptotic behaviour of an infeasible statistic that plugs in null spaces related to the population value of the matrix under study. This greatly simplifies the asymptotics of tests of rank under the various alternatives as well as misspecification. Under the null hypothesis or the local alternative, one can simply ignore the fact that the null spaces are estimated and derive the asymptotics as if the appropriate null spaces were known. Under the global alternative, the statistic diverges whenever the associated infeasible statistic diverges and under certain conditions (conjectured to be generic) both statistics are proven to diverge at the same rate. Thus, the plug-in principle allows us to achieve intermediate objective (ii). It also follows that statistics that have a common functional form but differ in their null space estimators are asymptotically equivalent, therefore establishing the asymptotic equivalence of a number of tests in the literature. Theorem 2 and Corollaries 3 and 4 are shown to imply the asymptotics of almost all tests of rank, with the handful of exceptions demonstrably satisfying a weaker form of the plug-in principle (see the discussion in Section 4.3).

It is important to emphasize several distinctive features of the approach of this paper. First, the approach is Waldian in that the primitives are taken to be a matrix estimator and a normalizing matrix; this allows it to encompass a much wider variety of tests than Reinsel and Velu (1998) and Massmann (2007), which nest some of the likelihood-based tests but miss a host of other tests. Second, it is based on orthogonal projection matrices, so that no identifying restrictions are imposed on the null space estimators; this allows for an elegant and compact description of their rates of convergence. Third, it encompasses both standard asymptotics and cointegration in a way that illuminates the continuity between the two strands of the literature. In this regard, the paper is developed gradually from the special case of standard asymptotics to the general case that allows for cointegration.

It is also important to note two aspects of the plug-in principle that have been well known in the literature. First, as far back as Stock and Watson (1988) and as recently as Boswijk et al. (2015), researchers have relied on the idea that the population cointegration relationship could be substituting in for a super-consistent estimator in working out the asymptotics of cointegration statistics. This paper demonstrates that this idea does not hold in general (see Example 4) and proposes the necessary modifications. Second, the proofs of the asymptotics of some tests sometimes involved an implicit use of the plug-in principle (e.g. Cragg and Donald (1996) and Robin and Smith (2000)). However, these instances concerned specific rather than generic tests and did not recognize the plug-in principle as an overarching framework that elucidates the asymptotics of tests of rank in general.

In terms of practical recommendations for practitioners, the following results emerge: (i) both theoretical and Monte Carlo results fail to point to an optimal test of rank, thus researchers can base their choice of test on other considerations, (ii) test statistics based on the QR and LU decompositions (e.g. the Cragg and Donald (1996) statistic) are recommended for high intensity computing such as the bootstrap as they are numerically less expensive than the alternatives (see Al-Sadoon (2016) for an illustration), and (iii) the paper proposes a number of new tests, which include robust extensions of the likelihood ratio test of Anderson (1951) and the maximum eigenvalue test of Johansen (1991) as well as tests based on the QR and Cholesky decompositions.[1]

The paper is organized as follows. Section 2 develops the notation of the paper. Section 3 develops the theory under standard asymptotics. Section 4 develops the theory under non-standard asymptotics. Section 5 provides Monte Carlo evidence. Section 6 concludes. Further Monte Carlo results and technical material as well as the proofs of the results can be found in the on-line appendix to the paper at http://dx.doi.org/10.1016/j.jeconom.2017.03.002.

2. Notation

$\mathbb{R}^{n\times m}$ denotes the set of $n \times m$ real valued matrices and $\mathbb{G}^{n\times m}$ is the subset of matrices of full rank. $\mathbb{P}^m_+ \subset \mathbb{P}^m \subset \mathbb{S}^m$ denote the set of positive definite, positive semi-definite, and symmetric matrices in $\mathbb{R}^{m\times m}$ respectively. $\mathrm{vec}(B)$ is the vector formed by vertically stacking the columns of B and $\mathrm{vech}(B)$ is the one formed by vertically stacking the elements below and including the diagonal elements of B. The mat operator is defined as the inverse to the vec operator (its range will be evident from the context). The Euclidean norm of $B \in \mathbb{R}^{n\times m}$ is defined as $\|B\| = (\mathrm{vec}'(B)\mathrm{vec}(B))^{1/2}$. The Mahalanobis norm is defined as $\|B\|_\Theta = (\mathrm{vec}'(B)\Theta^{-1}\mathrm{vec}(B))^{1/2}$ for $\Theta \in \mathbb{P}^{nm}_+$. The 2-norm is defined as $\|B\|_2 = \max_{x\in\mathbb{R}^m,\|x\|=1}\|Bx\|$. If $\mathcal{P} \subset \mathbb{R}^{n\times m}$, define $d(B,\mathcal{P}) = \inf_{X\in\mathcal{P}}\|B - X\|$. The singular values of B are denoted by $\sigma_1(B) \geq \sigma_2(B) \geq \cdots \geq \sigma_{\min\{n,m\}}(B)$. The condition number of B is defined as $\mathrm{cond}(B) = \sigma_1(B)/\sigma_r(B)$, where $r = \mathrm{rank}(B)$. The Moore–Penrose inverse of B is denoted by $B^\dagger$. For any $B \in \mathbb{G}^{n\times m}$ with $n > m$, an orthogonal complement $B_\perp$ is any matrix in $\mathbb{G}^{n\times(n-m)}$ satisfying $B'_\perp B = 0$. The column space of B is denoted by $\mathrm{span}(B)$. The orthogonal projection onto $\mathrm{span}(B)$ is denoted by P_B. The duplication matrix D_m is the mapping $\mathrm{vech}(B) \mapsto \mathrm{vec}(B)$ over $B \in \mathbb{S}^m$. For $B \in \mathbb{P}^m$, $B^{1/2}$ is the positive semi-definite square root matrix and $B^{\dagger/2} = (B^{1/2})^\dagger = (B^\dagger)^{1/2}$.

Finally, we say that a sequence of random matrices $X_T \in \mathbb{R}^{n\times m}$ indexed by T is bounded away from zero in probability and denote this by $X = O_p^{-1}(1)$ if for all $\varepsilon > 0$, there exist a $\delta_\varepsilon > 0$ and a $T_\varepsilon \geq 0$ such that the probability that $\|X_T\| > \delta_\varepsilon$ is at least $1-\varepsilon$ for all $T \geq T_\varepsilon$. It is easy to show that $\|X_T\|^{-1} = O_p^{-1}(1)$ if and only if $X_T = O_p(1)$ and $X_T = O_p^{-1}(1)$ if and only if $\|X_T\|^{-1} = O_p(1)$. Hence the notation, $O_p^{-1}(1)$. The product of two $O_p^{-1}(1)$ sequences is again $O_p^{-1}(1)$ and $a_T\|X_T\| \xrightarrow{p} \infty$ for any non-random sequence $a_T \to \infty$. The deterministic version, $O^{-1}(1)$, is defined similarly.

3. Tests of rank under standard asymptotics

This section lays the foundations of our study. First, the general structure of statistics for tests of rank is considered. It is shown

[1] Practitioners may also wish to consult the Matlab tutorial accompanying this paper, `tutorial.m`, which is included in the compressed file, `rank.rar`, available on the author's website.