CREATES

Center for Research in Econometric Analysis of Time Series

CREATES Research Paper 2008-2

Reduced-Rank Regression: A Useful Determinant Identity

Peter Reinhard Hansen

SOLIDUM PETIT IN PROFUNDIS · UNIVERSITAS ARHUSIENSIS

School of Economics and Management
University of Aarhus
Building 1322, DK-8000 Aarhus C
Denmark

Aarhus School of Business
University of Aarhus
Handelshøjskolen
Aarhus Universitet

UNIVERSITY OF COPENHAGEN

Reduced-Rank Regression:

A Useful Determinant Identity

Peter Reinhard Hansen†

Stanford University

Department of Economics

579 Serra Mall

Stanford CA 94305

Abstract

We derive an identity for the determinant of a product involving non-squared matrices. The identity can be used to derive the maximum likelihood estimator in reduced-rank regressions with Gaussian innovations. Furthermore, the identity sheds light on the structure of the estimation problem that arises when the reduced-rank parameters are subject to additional constraints.

JEL Classification: C3, C32

Keywords: Determinant Identity; Reduced Rank Regression; Least Squares

†The author is also affiliated with CREATES at the University of Aarhus, a research center funded by Danish National Research Foundation.

1 Introduction

Several seminal papers came out of Ted's Ph.D. dissertation, Anderson (1945). One of these paper is Anderson (1951), in which he established the theory of reduced-rank regressions. This paper was, in part, based on Anderson (1945, pp. 137-141), where Ted related the reduced-rank regression problem to earlier work by Fisher (1938) before giving the first rigorous treatment of the problem.

A cornerstone of the econometrics literature is Anderson and Rubin (1949), which introduced the *limited information maximum likelihood* estimator. A special case of the reduced-rank regression result was given in this paper, while the general form of reduced-rank regression was published in Anderson (1951). I was first exposed to Ted's research as an undergraduate student at the University of Copenhagen. The name *T.W. Anderson* frequently appeared in my lecture notes, which were authored by faculty members at the Institute of Mathematical Statistics. Ted's research had left a big footprint on the statistics and econometrics curriculum in Copenhagen, which continues to be true today. In recent years, I have been fortunate to interact with Ted at Stanford and I am grateful for the opportunity to contribution to this issue.

While the reduced-rank regression (RRR) can be attributed to Ted, he did not coin the catchy phrase. The reduced-rank terminology was first used in Burket (1964) who compared a number of reduced-rank methods for the purpose of prediction, while Izenman (1975) bonded the maximum likelihood estimator of Anderson (1951) with reduced-rank regression. See also Tso (1981), Davies and Tso (1982), and Anderson (1984).

Let $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$ be a $p \times p$ diagonal matrix and let $y \in \mathbb{R}^{p\times r}$ be a $p \times r$ matrix that is such that $y'y = I_r$, where I_r denotes the $r \times r$ identity matrix. The contribution of this paper is an identity that ties the determinant, $|y'\Lambda y|$, to an inner product with a convenient structure. Specifically

$$|y'\Lambda y| = \langle \xi, \theta \rangle,$$

where $\xi = \xi(\Lambda) \in \mathbb{R}^n$ is a vector that depends only on Λ, and $\theta = \theta(y) \in \mathbb{R}^n$ is a vector that depends only on y. In fact, θ is such that $\sum_{i=1}^n \theta_i = 1$, and $\theta_i \geq 0$, so that $|y'\Lambda y|$ is simply a convex combination of $(\xi_1, \ldots, \xi_n)$. An implication is that the following problem,

$$\max_{y \in \mathbb{R}^{p \times r}} |y'\Lambda y| / |y'y|,$$

which appears in the RRR estimation problem, is rather simple to solve.

A reduced-rank regression model take the form,

$$Y_t = \Pi X_t + \Psi Z_t + \varepsilon_t, \qquad t = 1, \ldots, T, \tag{1}$$

where Y_t, X_t and Z_t are of dimension p, q, and s, respectively. What make a RRR distinct from a standard regression model is a requirement that Π has reduced rank,

$$\text{rank}(\Pi) \leq r, \qquad \text{where } 0 \leq r < \min(p, q).$$

Naturally, a standard least squares regression model emerges when $r = \min(p, q)$.

The RRR estimation problem is given by,

$$\min_{\Pi, \Psi} \left| \sum_{t=1}^{T} \varepsilon_t \varepsilon_t' \right|, \qquad \text{subject to} \quad \text{rank}(\Pi) \leq r.$$

The reduced-rank condition makes the estimation problem more complex than the standard least squares regression problem. However, the RRR estimation problem can be simplified to $\max_{x \in \mathbb{R}^{p \times r}} |x'Mx| / |x'Nx|$, where M and N are data-dependent matrices. The difficult step is to show that $\hat{x} = (\hat{v}_1, \ldots, \hat{v}_r)$ is the solution to this problem, where $\hat{v}_1, \ldots, \hat{v}_r$ are the eigenvectors of $|\lambda N - M| = 0$ that corresponds to the r largest eigenvalues. This result can be obtained by a second order Taylor expansion of $\log |x'Mx| / |x'Nx|$, as in Johansen (1996); by reference to Poincare's theorem, see Magnus and Neudecker (1988); or by making use of the determinant identity presented in this paper.

When the reduced-rank parameters are subject to additional restrictions, it will typically result in more complex estimation problems that do not have closed-form solutions. Problems of this kind have been considered by Ahn and Reinsel (1990), Johansen and Juselius (1992), Boswijk (1995), Hansen and Johansen (1998), and Hansen (2003). Several dexterous algorithms to solve various estimation problems of this kind are proposed in these papers. The determinant identity of this paper could potentially be useful for estimation problems of this sort, because the identity uncovers more of the underlying structure.

The RRR model is interesting because it appears in several econometric models, including the analysis of multivariate time-series, see Velu, Reinsel, and Wichern (1986) and Velu and Reinsel (1987), and the analysis of cointegrated variables in the vector autoregressive framework, see Johansen (1988, 1991, 1996). The book by Reinsel and Velu (1998) gives an excellent exposition of reduced-rank regression analysis and its relations to many econometric models. It is remarkable that Ted, more that half a century later, continues to contribute to this line of research. For instance, Anderson (1999) generalizes the results in Anderson (1951) and Anderson (2002) discuss the properties of RRR and ordinary least squares estimators in both a stationary and non-stationary setting.

The rest of this paper is organized as follows. We derive the determinant identity in Section 2 and use the identity to prove a result that is useful for the RRR estimation problem. Reduced-rank regressions are discussed in Section 3, and Section 4 concludes.

2 A Determinant Identity

We introduce the following notation. Let $\mathbb{D}_p^r$ denote the set of all possible subsets of $J \subset \{1, \ldots, p\}$ with r distinct integers ($r \leq p$). For a given subset $J \in \mathbb{D}_p^r$, a $p \times r$ matrix y, and a $p \times p$ matrix Λ, we define the $r \times r$ sub-matrix, $y_J = \{y_{ij}\}_{i \in J, j=1,\ldots,r}$ and $\Lambda_J = \{\Lambda_{ij}\}_{i,j \in J}$. We use $\mathrm{diag}(a_1, \ldots, a_p)$ to denote the $p \times p$ diagonal matrix with diagonal elements: $a_1, \ldots, a_p$.

Example 1 *Consider the case where $p = 3$ and $r = 2$, so that*

$$y = \begin{pmatrix} y_{11} & y_{21} & y_{31} \\ y_{12} & y_{22} & y_{32} \end{pmatrix}' \qquad \textit{and} \qquad \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3).$$

For this case we have $\mathbb{D}_3^2 = \{\{1,2\},\{1,3\},\{2,3\}\}$. If we take $J = \{1,2\}$ then

$$y_J = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} \qquad \textit{and} \qquad \Lambda_J = \mathrm{diag}(\lambda_1, \lambda_2).$$

The main contribution of this paper is the following Theorem that formulates a useful expression for a determinant of non-square matrices.

Theorem 1 *Let* $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_p)$ *and let* y *be a* $p \times r$ *matrix, where* $p \geq r$*. Then*

$$|y'\Lambda y| = \sum_{J \in \mathbb{D}_p^r} |y_J' \Lambda_J y_J| = \sum_{J \in \mathbb{D}_p^r} |y_J' y_J| \, \Pi_{i \in J} \lambda_i = \sum_{J \in \mathbb{D}_p^r} |y_J|^2 \, \Pi_{i \in J} \lambda_i. \quad (2)$$

An immediate consequence of Theorem 1 is that $|y'\Lambda y|$ can be written as an inner product of the two vectors

$$\tilde{\theta} = (|y_{J_1}|^2, \ldots, |y_{Jn}|^2)' \quad \text{and} \quad \xi = (\Pi_{i \in J_1} \lambda_i, \ldots, \Pi_{i \in J_n} \lambda_i)',$$

where $n = \binom{p}{r}$. If we define $\theta = \tilde{\theta}/|y'y|$, then $\theta_i \geq 0$ for all i, and by applying Theorem 1 with $\Lambda = I$, it follows that $\sum_{i=1}^n \theta_i = 1$. So $\theta \in \Delta^{n-1}$ where Δ^{n-1} is the $(n-1)$-simplex in $\mathbb{R}^n$. This shows that

$$|y'\Lambda y|/|y'y| = \sum_{i=1}^n \theta_i \xi_i, \quad (3)$$

is simply a convex linear combination of the elements in ξ.

Proof of Theorem 1. The second and third equality follow from the fact that $|AB| = |A|\,|B|$ for matrices of proper dimensions. The first identity of (2) holds trivially when $r = 1$ and $p = r$. These cases are indicated by checkmarks in the follows scheme.

$p \backslash r$	1	2	3	4	$\cdots$
1	✓	–	–	–	
2	✓	✓	–	–	
3	✓	?	✓	–	
4	✓	?	?	✓	
$\vdots$	$\vdots$				$\ddots$

We now complete the proof by induction, as we show that (2) holds for (p, r) when the identity is assumed to hold for the two cases: $(p-1, r-1)$ and $(p-1, r)$.

Let $\tilde{\Lambda} \equiv \text{diag}(\lambda_1, \ldots, \lambda_{p-1})$. First we consider the special case where the pth row of y is a row of zeroes, i.e. $(y_{p1}, \ldots, y_{pr}) = (0, \ldots, 0)$. We let $\tilde{y}$ denote the $p-1 \times r$ sub-matrix of y that results from deleting the pth row of y. Using the assumption that (2) holds for the case, $(p-1,r)$, we have

$$\begin{aligned} |y'\Lambda y| &= |\tilde{y}'\tilde{\Lambda}\tilde{y}| = \sum_{J\in\mathbb{D}^r_{p-1}} |y'_J y_J| \cdot \Pi_{i\in J}\lambda_i \\ &= \sum_{J\in\mathbb{D}^r_p, p\notin J} |y'_J y_J| \cdot \Pi_{i\in J}\lambda_i + \sum_{J\in\mathbb{D}^r_p, p\in J} |y'_J y_J| \cdot \Pi_{i\in J}\lambda_i, \end{aligned}$$

where the last term is zero, because the pth row of y only consists of zeroes. So we have established the result for this special case.

Consider next, the situation where $(y_{p1},\ldots,y_{pr}) \neq (0,\ldots,0)$. Here we choose a full rank $r \times r$-matrix, Q, that satisfies $(y_{p1},\ldots,y_{pr})Q = (0,\ldots,0,1)$. Given our choice for Q, we define $\tilde{z}$ to be the $p-1 \times r-1$ matrix given by the first $r-1$ columns of $\tilde{y}Q$. We have

$$|Q|^2 \, |y'\Lambda y| = \left| Q'\tilde{y}'\tilde{\Lambda}\tilde{y}Q + \begin{pmatrix} 0_{r-1\times r-1} & 0 \\ 0 & \lambda_p \end{pmatrix} \right| = |Q'\tilde{y}'\tilde{\Lambda}\tilde{y}Q| + |\tilde{z}'\tilde{\Lambda}\tilde{z}|\lambda_p. \tag{4}$$

Since we have assumed that the identity holds for the case $(p-1,r)$, the first term can be expressed as

$$|Q'\tilde{y}'\tilde{\Lambda}\tilde{y}Q| = |Q|^2 \sum_{J\in\mathbb{D}^r_{p-1}} |\tilde{y}'_J\tilde{\Lambda}_J\tilde{y}_J| = |Q|^2 \sum_{J\in\mathbb{D}^r_p, p\notin J} |y'_J\Lambda_J y_J|. \tag{5}$$

Finally, we turn to the last term of (4), $|\tilde{z}'\tilde{\Lambda}\tilde{z}|\lambda_p$. For $J \in \mathbb{D}^{r-1}_{p-1}$ we have

$$|\tilde{z}_J| = \left| \begin{pmatrix} \tilde{z}_J & 0 \\ 0 & 1 \end{pmatrix} \right| = |y_{\tilde{J}}Q|,$$

and $\lambda_p|\tilde{\Lambda}_J| = |\Lambda_{\tilde{J}}|$, where $\tilde{J} = \{J \cup \{p\}\} \in \mathbb{D}^r_p$. So the second term of (4) can be expressed as

$$|\tilde{z}'\tilde{\Lambda}\tilde{z}|\lambda_p = |Q|^2 \sum_{\substack{J\in\mathbb{D}^{r-1}_p \\ p\in J}} |y'_J\Lambda_J y_J|, \tag{6}$$

where we made use of the assumption that the identity holds when $(p-1,r-1)$. Combining the identities (4)–(6), we have shown

$$|Q|^2|y'\Lambda y| = |Q|^2 \sum_{J\in\mathbb{D}^r_p, p\notin J} |y'_J\Lambda_J y_J| + |Q|^2 \sum_{J\in\mathbb{D}^r_p, p\in J} |y'_J\Lambda_J y_J| = |Q|^2 \sum_{J\in\mathbb{D}^r_p} |y'_J\Lambda_J y_J|,$$

which completes the proof. ■

Corollary 2 *Let* $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_p)$ *where* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ *and consider the function*

$$g(y) = |y'\Lambda y| \, / \, |y'y| \, .$$

Then

$$\max_{y \in \mathbb{R}^{p \times r}} g(y) = g(\hat{y}) = \prod_{i=1}^{r} \lambda_i \quad \text{and} \quad \min_{y \in \mathbb{R}^{p \times r}} g(y) = g(\check{y}) = \prod_{i=p-r+1}^{p} \lambda_i,$$

where $\hat{y} = (I_r, 0_{r \times p-r})'$ *and* $\check{y} = (0_{r \times p-r}, I_r)'$.

Proof. From our discussion that followed Theorem 1, we have that $g(y) = |y'\Lambda y| \, / \, |y'y|$ is a convex combination over the elements in ξ, $\prod_{i \in J} \lambda_i$, $J \in \mathbb{D}_p^r$. So we have

$$\prod_{i=p-r+1}^{p} \lambda_i \leq g(y) \leq \prod_{i=1}^{r} \lambda_i.$$

Since these two bounds are attained for $y = \check{y}$ and $y = \hat{y}$, respectively, the proof is complete. ■

The following Corollary is a key to the estimation problem in reduced-rank regressions that we discuss in the next Section.

Corollary 3 *Let* M *and* N *be symmetric* $p \times p$ *matrices, where* M *is positive semi-definite and* N *is positive definite. Consider the matrix function*

$$f(x) = |x'Mx| \, / \, |x'Nx| \, ,$$

where $x \in \mathbb{R}^{p \times r}$. *Then*

$$\sup_x f(x) = f(\hat{x}) = \prod_{i=1}^{r} \lambda_i \qquad \text{and} \qquad \inf_x f(x) = f(\check{x}) = \prod_{i=p-r+1}^{p} \lambda_i,$$

where $\hat{x} = (v_1, \ldots, v_r)$ *and* $\check{x} = (v_{p-r+1}, \ldots, v_p)$ *and* $\lambda_1, \ldots, \lambda_p$ *are the eigenvalues of* $|\lambda N - M| = 0$, *ordered in descending order, and* $v_1, \ldots, v_p$ *the corresponding eigenvectors.*

A partial proof of the result in Corollary 3 was given in Bellman (1960, theorem 10, p. 129), and the result can also be found in Rao (1973) along with many related results. Johansen (1988) establishes the result using a second order Taylor expansion of $\log f(x)$, see also Johansen (1996). Here we present a proof that is based on the determinant identity of Theorem 1.

Proof of Corollary 3. Since N is symmetric and positive definite it can be diagonalized as $N = V'DV$, where D is a diagonal matrix with positive entries and $V'V = I$. So $N^{-\frac{1}{2}} = V'D^{-\frac{1}{2}}V$ is well-defined. It follows that the matrix $N^{-\frac{1}{2}}MN^{-\frac{1}{2}}$ is symmetric and positive semi-definite and can be diagonalized, $N^{-\frac{1}{2}}MN^{-\frac{1}{2}} = Q'\Lambda Q$, where $Q'Q = I$ and $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_p)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$. With $y = QN^{\frac{1}{2}}x$ we have $|x'Mx| \,/\, |x'Nx| = |y'\Lambda y| \,/\, |y'y|$. By Corollary 2, this term is maximized (minimized) by $\hat{y} = (I_r, 0)'$ ($\check{y} = (0, I_r)'$), so $f(x)$ is maximized (minimized) by $\hat{x} = N^{-\frac{1}{2}}Q'\hat{y}$ ($\check{x} = N^{-\frac{1}{2}}Q'\tilde{y}$). ■

3 Reduced Rank Regression

In this Section, we discuss the RRR estimation problem. A convenient way to impose the rank condition, $\text{rank}(\Pi) \leq r$, is to rewrite Π as the product, $\Pi = \alpha\beta'$, where α and β have dimensions $p \times r$ and $q \times r$, respectively. Then the reduced-rank regression model, (1), becomes

$$Y_t = \alpha\beta' X_t + \Psi Z_t + \varepsilon_t, \qquad t = 1, \ldots, T,$$

and the RRR estimators of α, β, and Ψ are defined from the solution to

$$\min_{\alpha,\beta,\Psi} \left| \sum_{t=1}^{T} \varepsilon_t \varepsilon_t' \right|. \tag{7}$$

In matrix notation a RRR take the form $\mathbf{Y} = \mathbf{X}\beta\alpha' + \mathbf{Z}\Psi' + \boldsymbol{\varepsilon}$, where the tth row of $\mathbf{Y}$, $\mathbf{X}$, $\mathbf{Z}$, and $\boldsymbol{\varepsilon}$ are given by Y_t', X_t', Z_t', and ε_t' respectively. Thus $\text{var}(\boldsymbol{\varepsilon}') = I_T \otimes \Omega$.

We define the moment matrix, $M_{yx} = \mathbf{Y}'\mathbf{X}/T$, and define the matrices, M_{yy}, M_{yz}, M_{xx}, etc. similarly. Finally, define

$$S_{yy} = M_{yy} - M_{yz}M_{zz}^{-1}M_{zy}, \qquad S_{yx} = M_{yx} - M_{yz}M_{zz}^{-1}M_{zx},$$

and define S_{xx} and $S_{xy} = S_{yx}'$ in a similar manner.

Theorem 4 (Reduced Rank Regression) *The estimators that solve (7) are given by*

$$\begin{aligned}\hat{\beta} &= (\hat{v}_1, \ldots, \hat{v}_r)\phi, \\ \hat{\alpha} &= S_{yx}\hat{\beta}(\hat{\beta}' S_{xx}\hat{\beta})^{-1}, \\ \hat{\Psi} &= M_{yz}M_{zz}^{-1} - \hat{\alpha}\hat{\beta}' M_{xz}M_{zz}^{-1},\end{aligned}$$

where $(\hat{v}_1, \ldots, \hat{v}_r)$ are the eigenvectors corresponding to the r largest eigenvalues of

$$|\lambda S_{xx} - S_{xy}S_{yy}^{-1}S_{yx}| = 0,$$

and where ϕ is an arbitrary $r \times r$ matrix with full rank.

Remark 1 *The parameters α and β are not identified. However, the $r \times r$ matrix, ϕ, can be used as a normalization device. E.g. if the normalization $\beta = (I_r, \beta_2')'$ is desired, one can choose ϕ to be the inverse of the matrix that consists of the first r rows of $(\hat{v}_1, \ldots, \hat{v}_r)$.*

Remark 2 *The eigenvectors in Theorem 4 are easy to obtain with standard software, such as Ox, Gauss, or Matlab, because $(\hat{v}_1, \ldots, \hat{v}_p) = S_{xx}^{-1/2}(x_1, \ldots, x_p)$, where $(x_1, \ldots, x_p)$ are the eigenvectors of the matrix $S_{xx}^{-1/2}S_{xy}S_{yy}^{-1}S_{yx}S_{xx}^{-1/2}$. The eigenvalues of the two problems are identical and the eigenvectors in Theorem 4 satisfy $S_{xy}S_{yy}^{-1}S_{yx}\hat{v}_i = \lambda_i S_{xx}\hat{v}_i$, $\hat{v}_i' S_{xx}\hat{v}_j = \delta_{ij}$, where δ_{ij} is the Kronecker delta.*

Proof of Theorem 4. The objective is to minimize $m_0(\alpha, \beta, \Psi)$, where

$$m_0(\alpha, \beta, \Psi) = \left| T^{-1}\sum_{t=1}^{T} \varepsilon_t \varepsilon_t' \right|, \qquad \varepsilon_t = Y_t - \alpha\beta' X_t - \Psi Z_t.$$

It is simple to verify that $\arg\min_\Psi m(\alpha, \beta, \Psi) = \hat{\Psi}(\alpha, \beta) \equiv M_{yz}M_{zz}^{-1} - \alpha\beta' M_{xz}M_{zz}^{-1}$, as it follows from the standard least squares result. By defining the auxiliary variables, $\tilde{Y}_t = Y_t - M_{yz}M_{zz}^{-1}Z_t$ and $\tilde{X}_t = X_t - M_{xz}M_{zz}^{-1}Z_t$, the estimation problem is simplified to minimizing

$$m_1(\alpha, \beta) = m_0(\alpha, \beta, \Psi(\alpha, \beta)) = |T^{-1}\sum_{t=1}^{T} \tilde{\varepsilon}_t \tilde{\varepsilon}_t'|,$$

where $\tilde{\varepsilon}_t = \tilde{Y}_t - \alpha\beta' \tilde{X}_t$.

For a fixed value of β, the estimation problem (of α) is a least squares problem, so that $\arg\min_\alpha m_1(\alpha, \beta) = \hat{\alpha}(\beta) \equiv S_{yx}\beta(\beta' S_{xx}\beta)^{-1}$. The residual problem is now to minimize

$m_2(\beta) = m_0\left(\alpha(\beta), \beta, \Psi(\alpha(\beta), \beta)\right)$, where

$$\begin{aligned} m_2(\beta) &= |T^{-1}\sum_{t=1}^{T}(\tilde{Y}_t - \hat{\alpha}(\beta)\beta'\tilde{X}_t)(\tilde{Y}_t - \hat{\alpha}(\beta)\beta'\tilde{X}_t)'| \\ &= |S_{yy} - S_{yx}\beta(\beta' S_{xx}\beta)^{-1}\beta' S_{xy}| = |S_{yy}|\frac{|\beta'(S_{xx} - S_{xy}S_{yy}^{-1}S_{yx})\beta|}{|\beta' S_{xx}\beta|}. \end{aligned}$$

Let $0 \leq \hat{\rho}_1 \leq \cdots \leq \hat{\rho}_p$ be the eigenvalues of $|\rho S_{xx} - (S_{xx} - S_{xy}S_{yy}^{-1}S_{yx})| = 0$ and $\hat{v}_1, \ldots, \hat{v}_p$ the corresponding eigenvectors. By Corollary 3 it follows that $\hat{\beta} = (v_1, \ldots, v_r)$ minimizes $m_2(\beta)$. The eigenvectors satisfy $(S_{xx} - S_{xy}S_{yy}^{-1}S_{yx})v_i = \rho_i S_{xx} v_i$, such that $S_{xy}S_{yy}^{-1}S_{yx}v_i = (1-\hat{\rho}_i)S_{xx}v_i$ $i = 1, \ldots, p$. So v_i is also an eigenvector of the eigenvalue problem, $|\lambda S_{xx} - S_{xy}S_{yy}^{-1}S_{yx}| = 0$, with eigenvalue $\hat{\lambda}_i = 1 - \hat{\rho}_i$. It now follows that the solution to $\min_\beta m_2(\beta)$ is given from the r eigenvectors that are associated with the r largest eigenvalues of $|\lambda S_{xx} - S_{xy}S_{yy}^{-1}S_{yx}| = 0$. The results for $\hat{\alpha}$ and $\hat{\Psi}$ now follows by substituting $\hat{\beta}$ into $\hat{\alpha}(\beta)$ and $(\hat{\alpha}, \hat{\beta})$ into $\hat{\Psi}(\alpha, \beta)$. ■

3.1 Maximum Likelihood Estimators by RRR

The estimators derived in Theorem 4 were simply defined to be the solution to least squares problem, (7). However, the RRR estimators, $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\Psi}$, are the maximum likelihood estimators under the assumptions that: X_t and Z_t are deterministic or predetermined; and ε_t is independent and identically Gaussian distributed, $\varepsilon_t \sim N(0, \Omega)$.

In this situation the maximum likelihood estimator of Ω is given by

$$\hat{\Omega} = S_{yy} - S_{yx}\hat{\beta}(\hat{\beta}' S_{xx}\hat{\beta})^{-1}\hat{\beta}' S_{xy},$$

and the maximum value of the likelihood function, $L(\hat{\alpha}, \hat{\beta}, \hat{\Psi}, \hat{\Omega})$, is

$$L_{\max}^{-2/T} = (2\pi e)^p \, |S_{yy}| \prod_{i=1}^{r}(1 - \hat{\lambda}_i).$$

So we have

$$2\log L_{\max} \propto T\sum_{i=1}^{r}\log(1 - \hat{\lambda}_i),$$

and this leads to the the likelihood ratio test of the rank of Π, by Anderson (1951), as the likelihood ratio statistic for the hypothesis: H_0 : rank$(\Pi) \leq r$ against rank$(\Pi) = p \wedge q$, is simply $T\sum_{i=r+1}^{q}\log(1-\hat{\lambda}_i)$. The asymptotic distribution of this likelihood ratio statistic is χ^2 in a setting with stationary variables, whereas non-standard distributions arise when the underlying variables are non-stationary, see Johansen (1991, 1996) and Anderson (2002).

Empirical studies using the cointegrated vector autoregressive model by Johansen (1988), is probably the area where reduced-rank regression analysis is used most extensively. In the special case with a vector autoregressive model of order two, this model can be written as the, so-called, error correction model:

$$\Delta V_t = \alpha\beta' V_{t-1} + \Gamma_1 \Delta V_{t-1} + \mu + \varepsilon_t,$$

where $\Delta V_t = V_t - V_{t-1}$, μ is a vector of constants, and $\{\varepsilon_t\}$ is a sequence of iid Gaussian random variables, $\varepsilon_t \sim N(0,\Omega)$. The RRR structure is evident in this model. Simply set $Y_t = \Delta V_t$, $X_t = V_{t-1}$ and $Z_t = (\Delta V_{t-1}', 1)'$. See Johansen (1996) for a comprehensive treatment of this model.

3.2 RRR with Additional Restrictions

Many econometric problems lead to a reduced-rank regression where the parameters are subject to additional restrictions, see Ahn and Reinsel (1990), Johansen and Juselius (1992), Boswijk (1995), Hansen and Johansen (1998), and Hansen (2003). Estimation problems of this kind will typically have a complex structure, that cannot be solved with the methods discusses above.

In the proof of Theorem 4 we saw that the estimation problem for β could be expresses as the problem $\min_y |y'\Lambda y|/|y'y|$, and the estimator of β is tied to that of y by the relation $y = QS_{xx}^{1/2}\beta$, where Q is given from the orthogonal decomposition: $Q'\Lambda Q = S_{xx}^{-1/2} S_{xy} S_{yy}^{-1} S_{yx} S_{xx}^{-1/2}$. This shows that restrictions on β translate into restrictions on y through $y(\beta) = QS_{xx}^{1/2}\beta$. From (3) we have that

$$|y'\Lambda y|/|y'y| = \langle\theta, \xi\rangle,$$

where $\theta(y)$ is a vector that depends only on y, and $\xi = \xi(\Lambda)$ is a vector that depends only on Λ. So restrictions on β, $\beta \in B_0$ say, can be translated into restrictions on the vector θ, $\theta \in \Theta_0$ say, in which case the estimation problem can be expressed as $\max_{\theta\in\Theta}\theta'\xi$. This suggests a general framework for analyzing RRR subject to additional restrictions on β. We shall not pursue this issue further in this paper.

4 Conclusion

This paper derived a determinant identity, that can be used to solve the estimation problem in the reduced-rank regression model by Anderson (1951). The expression for the determinant provides additional insight about the estimation problem in reduced-rank regressions where the reduced-rank parameters are subject to additional restrictions.

References

Ahn, S. K., and G. C. Reinsel (1990): "Estimation for Partially Non-Stationary Multivariate Autoregressive Models," *Journal of the American Statistical Association*, 85, 813–823.

Anderson, T., and H. Rubin (1949): "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, 20, 46–63.

Anderson, T. W. (1945): "The Non-Central Wishart Distribution and its Application to Problems in Multivariate Statistics," Ph.D. thesis, Princeton University.

Anderson, T. W. (1951): "Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions," *Annals of Mathematical Statistics*, 22, 327–351.

Anderson, T. W. (1984): *An Introduction to Multivariate Statistical Analysis*. John Wiley and Sons, New York, 2nd edn.

Anderson, T. W. (1999): "Asymptotic Distribution of the Reduced Rank Regression Estimator under General Conditions," *Annals of Statistics*, 27, 1141–1154.

——— (2002): "Reduced Rank Regression in Cointegrated Models," *Journal of Econometrics*, 106, 203–216.

Bellman, R. (1960): *Introduction to Matrix Analysis*. McGraw-Hill, New York.

Boswijk, P. H. (1995): "Identifiability of Cointegrated Systems," *Working paper, Tinbergen Institute*.

Burket, G. R. (1964): "A Study of Reduced Rank Models for Multiple Prediction," *Psychometric Monographs*, No. 12.

Davies, P. T., and M. K.-S. Tso (1982): "Procedures for Reduced-Rank Regression," *Applied Statistics*, 31, 244–255.

Fisher, R. A. (1938): "The Statistical Utilization of Multiple Measurements," *Annals of Eugenics*, 8, 376–386.

Hansen, P. R. (2003): "Structural Changes in the Cointegrated Vector Autoregressive Model," *Journal of Econometrics*, 114, 261–295.

Hansen, P. R., and S. Johansen (1998): *Workbook on Cointegration.* Oxford University Press, Oxford.

Izenman, A. J. (1975): "Reduced-Rank Regression for the Multivariate Linear Model," *Journal of Multivariate Analysis*, 5, 248–264.

Johansen, S. (1988): "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control*, 12, 231–254.

——— (1991): "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59, 1551–1580.

——— (1996): *Likelihood Based Inference in Cointegrated Vector Autoregressive Models.* Oxford University Press, Oxford, 2nd edn.

Johansen, S., and K. Juselius (1992): "Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and the UIP for UK," *Journal of Econometrics*, 53, 211–244.

Magnus, J. R., and H. Neudecker (1988): *Matrix Differential Calculus with Applications in Statistics and Economics.* Wiley, New York.

Rao, C. R. (1973): *Linear Statistical Inference and Its Applications.* John Wiley, New York, 2nd edn.

Reinsel, G. C., and R. P. Velu (1998): *Multivariate Reduced-Rank Regression.* Springer, New York.

Tso, M. K.-S. (1981): "Reduced-Rank Regression and Canonical Analysis," *Journal of Royal Statistical Society, series B*, 43, 183–189.

Velu, R. P., and G. C. Reinsel (1987): "Reduced Rank Regression with Autoregressive Errors," *Journal of Econometrics*, 35, 317–375.

Velu, R. P., G. C. Reinsel, and D. W. Wichern (1986): "Reduced Rank Models for Multiple Time Series," *Biometrika*, 73, 105–118.

Research Papers 2008

CREATES
Center for Research in Econometric Analysis of Time Series

2007-33 Søren Johansen and Morten Ørregaard Nielsen: Likelihood inference for a nonstationary fractional autoregressive model

2007-34 Charlotte Christiansen and Angelo Ranaldo: Extreme Coexceedances in New EU Member States' Stock Markets

2007-35 Søren Johansen: Correlation, regression, and cointegration of nonstationary economic time

2007-35 David F. Hendry, Søren Johansen and Carlos Santos: Selecting a Regression Saturated by Indicators

2007-37 Peter Christoffersen, Kris Jacobs and Karim Mimouni: Models for S&P500 Dynamics: Evidence from Realized Volatility, Daily Returns, and Option Prices

2007-38 Dennis Kristensen and Anders Rahbek: Likelihood-Based Inference in Nonlinear Error-Correction Models

2007-39 Peter Christoffersen, Kris Jacobs and Gregory Vainberg: Forward-Looking Betas

2007-40 Olav Bjerkholt: Trygve Haavelmo's visit in Aarhus 1938-39

2007-41 Søren Johansen and Anders Rygh Swensen: Exact rational expectations, cointegration, and reduced rank regression

2007-42 Ole E. Barndorff-Nielsen, José Manuel Corcuera and Mark Podolskij: Power variation for Gaussian processes with stationary increments

2007-43 Jean Jacod, Yingying Li, Per A. Mykland, Mark Podolskij and Mathias Vetter: Microstructure Noise in the Continuous Case: The Pre-Averaging Approach - JLMPV-9

2007-44 Charles S. Bos, Siem Jan Koopman and Marius Ooms: Long memory modelling of inflation with stochastic variance and structural breaks

2007-45 James Davidson and Nigar Hashimzade: Representation and Weak Convergence of Stochastic Integrals with Fractional Integrator Processes

2008-01 John A. Carlson, Christian M. Dahl, and Carol L. Osler: Short-run Exchange-Rate Dynamics: Theory and Evidence

2008-02 Peter Reinhard Hansen: Reduced-Rank Regression: A Useful Determinant Identity