SVERIGES RIKSBANK
WORKING PAPER SERIES

150

Bayes Estimators of the Cointegration Space

Mattias Villani

SEPTEMBER 2003

WORKING PAPERS ARE OBTAINABLE FROM

Sveriges Riksbank • Information Riksbank • SE-103 37 Stockholm
Fax international: +46 8 787 05 26
Telephone international: +46 8 787 01 00
E-mail: info@riksbank.se

Bayes Estimators of the Cointegration Space

Mattias Villani*

Sveriges Riksbank Working Paper Series
No. 150
September 2003

Abstract

A neglected aspect of the otherwise fairly well developed Bayesian analysis of cointegration is the point estimation of the cointegration space. It is pointed out here that, due to the well known non-identification of the cointegration vectors, the parameter space is not an inner product space and conventional Bayes estimators therefore stand without their usual decision theoretic foundation. We present a Bayes estimator of the cointegration space which takes the curved geometry of the parameter space into account. Contrary to many of the Bayes estimators used in the literature, this estimator is invariant to the ordering of the time series. A dimension invariant overall measure of cointegration space uncertainty is also proposed. A small simulation study shows that the Bayes estimator compares favorably to the maximum likelihood estimator.

Keywords: Bayesian inference, Cointegration analysis, Estimation, Grassman manifold, Subspaces.

JEL classification: C11, C13, C32.

*Sveriges Riksbank, Research Department, SE-103 37 Stockholm, Sweden, and Department of Statistics, Stockholm University. E-mail: mattias.villani@riksbank.se. This work has benefitted from valuable discussions with Daniel Thorburn. The author gratefully acknowledges financial support from the Swedish Council of Research in Humanities and Social Sciences (HSFR), grant no. F0582/1999, and the Swedish Research Council (Vetenskapsrådet), grant no. 412-2002-1007. The views expressed in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Executive Board of Sveriges Riksbank.

1. Introduction

Building on the work of Kleibergen and van Dijk (1994), Bauwens and Lubrano (1996) and Geweke (1996), the Bayesian analysis of cointegration has recently been developed by, for example, Kleibergen and Paap (2002), Strachan (2003) and Villani (2003) into a fairly complete alternative to the more established classical approaches in e.g. Phillips (1991) and Johansen (1995).

The focus in the above mentioned works has been on developing suitable prior distributions and deriving the corresponding posterior distributions. The efforts put into this activity have diverted attention from other important aspects of the analysis. One such aspect is the topic here: how to best summarize the posterior distribution of the cointegration vectors by a few well chosen quantities, such as measures of location and spread. Since the cointegration vectors are only identified up to arbitrary linear combinations, the parameter space is not an inner product space and the usual measures, such as the posterior mode, mean and median, cannot be given their usual decision theoretic motivation.

The paper is organized as follows. The next section discusses the geometry of the parameter space of the cointegration vectors and the problem with the currently available Bayes estimators. Section 3 proposes a new Bayes estimator of the cointegration space and the following section gives corresponding measures of cointegration space variation. Section 5 contains an empirical illustration and Section 6 a small simulation study where the new Bayes estimator is compared with the maximum likelihood estimator in Johansen (1995). The last section gives some concluding remarks.

2. The geometry of the parameter space in cointegration models

Following Johansen (1995), let β denote the $p \times r$ matrix of cointegration vectors and α the $p \times r$ matrix of adjustment coefficients. It is well known that the likelihood function is invariant under the class of transformations $T_Q : (\alpha, \beta) \to (\alpha Q'^{-1}, \beta Q)$, where Q is a non-singular $r \times r$ matrix; this is usually phrased as: only the space spanned by the columns of β, the *cointegration space*, is identified. In order to obtain a unique estimate of β some of its elements have to be restricted to known values, usually zero or one, in such a way that the mapping from the remaining unrestricted elements of β to the cointegration spaces is one-to-one. A particularly simple set of restrictions is $\beta' = (I_r, B')$, where B is the $(p-r) \times r$ matrix of unrestricted elements.

It should be clear, however, that the use of identifying restrictions does not change the fact that the real parameters of the model are *not* the unrestricted coefficients under a particular identifying scheme, but rather the cointegration space as a whole. Thus, the parameter space is not Euclidean, but the abstract space of all r-dimensional subspaces of $\mathbb{R}^p$, the *Grassman manifold* $\mathbb{G}_{r,p-r}$ (James, 1954). $\mathbb{G}_{r,p-r}$ is an (analytic) manifold of dimension $r(p-r)$. The current state of maximum likelihood analysis in cointegration models relies almost exclusively on asymptotic analysis. The fact that the Grassman manifold is locally Euclidean justifies the use of statistical theory for Euclidean spaces in the asymptotic analysis of β. Bayesian analysis, on the other hand, claims to be applicable for all sample sizes and as such is concerned with global properties for which we are no longer able to rely on theory developed for Euclidean spaces.

Despite the non-Euclidean geometry of the parameter space, all Bayesian applications have estimated the cointegration space by the posterior mean, mode or median of the unrestricted elements of β inserted into β. The most common approach is to insert the posterior mode or median of B into $\beta' = (I_r, B')$, the posterior mean of B does not exist (see Bauwens and Lubrano, 1996 or Kleibergen and van Dijk, 1994, for a proof). These estimators are not invariant to the way the variables are ordered, even if the posterior distribution itself exhibits this invariance property, which is illustrated empirically in Section 5. The estimator in Strachan (2003) is based on the identifying scheme in Anderson (1951) and Johansen (1995), which he terms the *non-ordinal normalization.* In the case of a single cointegrating relation, the non-ordinal normalization restricts β to lie on a hemisphere with a fixed radius. Strachan (2003) estimates the cointegration vector with the posterior mean of β, which exists in the non-ordinal normalization. Strachan's estimator is clearly invariant to the variable order, but, as we illustrate in Section 3, may generate counterintuitive results.

3. The posterior mean cointegration space estimator

We shall assume that a posterior distribution $p(\beta|\mathcal{D})$ for the matrix of cointegration vectors is available, perhaps obtained from one of the algorithms proposed in the references in the introduction. We may further assume, without loss of generality, that $\beta'\beta = I_r$. If the posterior of β has been obtained using a different normalization, one may simply make the transformation $\beta \to \beta(\beta'\beta)^{-1/2}$. The posterior distributions in all Bayesian analyses of cointegration to date have been evaluated numerically by sampling from the joint posterior distribution, so the transformation to orthonormality is conveniently performed in each draw.

We shall now consider the question: given a (posterior) distribution of β, what single point $\hat{\beta} \in \mathbb{G}_{r,p-r}$ is the, in some sense, best summary of this distribution? The solution to this problem is given by

$$\hat{\beta} \overset{def}{=} \underset{\tilde{\beta}\in\mathbb{G}_{r,p-r}}{\operatorname{argmin}} E[l(\beta,\tilde{\beta})],$$

where $E(\cdot)$ denotes the posterior expectation with respect to $p(\beta|\mathcal{D})$ and $l(\beta,\tilde{\beta})$ is a loss function on $\mathbb{G}_{r,p-r} \times \mathbb{G}_{r,p-r}$. Although there are many distances on the Grassman manifold (Edelman et al., 1998) which may be used to construct a loss function, we shall in this note restrict attention to the *projective Frobenius distance*

$$l(\beta,\tilde{\beta}) = ||\beta\beta' - \tilde{\beta}\tilde{\beta}'||,$$

where $||A|| = \operatorname{tr}(A'A)^{1/2}$ is the usual Frobenius norm for matrices. Note that $\beta\beta'$ is the projection matrix for the subspace $\operatorname{sp}\beta$ and recall that a subspace is uniquely determined by the orthogonal projection on it. $l(\beta,\tilde{\beta})$ is therefore obtained by embedding the Grassman manifold in the set of $p \times p$ projection matrices of rank r and then using the Frobenius norm. The projective Frobenius distance is one of the most widely used distances between subspaces and has the additional advantage of leading to a simple analytical expression for $\hat{\beta}$, as the next result shows. We will refer to $\hat{\beta}$ under the projective Frobenius distance as the posterior mean cointegration space (PMCS) estimator. The following theorem was proved independently by Srivastava (2000) and Villani (2000).

Theorem 3.1. *The posterior mean cointegration space estimator is*

$$\hat{\beta} = (v_1, ..., v_r),$$

where v_i is the eigenvector of $E(\beta\beta')$ corresponding to the ith largest eigenvalue.

A closed form expression for $E(\beta\beta')$ may not be available, but a numerical approximation may be used in its place. For example, importance sampling (Kloek and van Dijk, 1978) or the Gibbs sampler (Tierney, 1994) can be used to generate N draws from the distribution of β. These generated matrices can subsequently be made orthonormal and the following

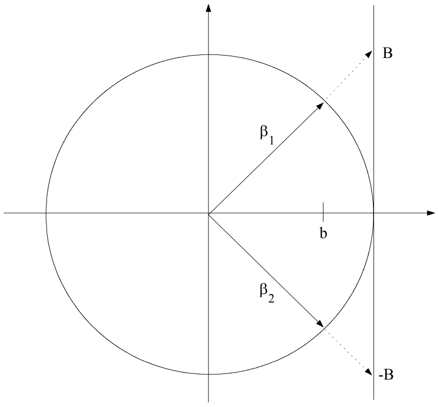

FIGURE 1. Illustration of the example in the text.

well-known result (Tierney, 1994) can be used to estimate $E(\beta\beta')$

$$\frac{1}{N}\sum_{i=1}^{N}\beta^{(i)}\beta^{(i)\prime} \overset{a.s.}{\to} E(\beta\beta'),$$

where $\beta^{(i)}$ denotes the ith sampled matrix after the transformation to orthonormality and $\overset{a.s.}{\to}$ denotes almost sure convergence.

We shall use a simple example for illustration where all posterior mass is distributed equally on the two vectors $\beta_1 = (b, \sqrt{1-b^2})'$ and $\beta_2 = (b, -\sqrt{1-b^2})'$, where $0 \leq b \leq 1$. Such a posterior distribution would clearly not be encountered in practice, but it caricatures less extreme situations that do occur in applications, at least when b is small. The situation is depicted in Figure 1, where also the normalization $\beta = (1, B)'$ is illustrated. For small values of b, we clearly have $\beta_1 \approx (0,1)'$ and $\beta_2 \approx (0,-1)'$, which both say that the second variable is stationary, since the sign of the second coefficient in β does not matter when the first coefficient is zero. On the other hand, as $b \to 1$, we have $\beta_1 \to (1,0)'$ and $\beta_2 \to (1,0)'$, both implying that the first variable is stationary. It is easy to see that the mode, median and mean plug-in estimators of β are all equal to $(1,0)'$, using the usual convention to handle ties. Strachan's (2003) posterior mean estimate of β is $(\beta_1 + \beta_1)/2 = (b, 0)'$, which after normalization becomes $(1,0)'$. Thus, all previously suggested estimators estimates β with $(1,0)'$, regardless of b, which goes against intuition. The PMCS estimator tells a completely different story. The two eigenvalues of $E(\beta\beta')$ are $2(1-b^2)$ and $2b^2$, with corresponding eigenvectors $(0,1)'$ and $(1,0)'$. Thus, $\hat{\beta} = (0,1)'$ if $b < 1/2$ and $\hat{\beta} = (1,0)'$ if $b > 1/2$, exactly as suggested by intuition. The crux of the matter is of course that the previously proposed estimators are based on distances for Euclidean spaces which fail to acknowledge that β_1 and β_2 become arbitrarily close as $b \to 0$.

4. MEASURES OF COINTEGRATION SPACE VARIATION

Although the usual measures of spread of the free coefficients in β are easily computed numerically by sampling from the distribution of β, their motivation comes from Euclidean space theory and may therefore be of limited value in assessing the variation of $\operatorname{sp}\beta$, at least for moderately informative distributions.

A quite different measure of variation suggests itself from Theorem 3.1. Let $\lambda_1 \geq \cdots \geq \lambda_p$ denote the eigenvalues of $E(\beta\beta')$. Since λ_i measures the variation of $\operatorname{sp}\beta$ in the direction determined by v_i, $\lambda_1, ..., \lambda_r$ can be used to assess the uncertainty regarding $\mathrm{sp}\,\beta$. A natural suggestion for an *overall* measure of variation of $\mathrm{sp}\,\beta$ to accompany the PMCS estimate is given in the following definition.

Definition 4.1. *The projective Frobenius span variation is defined as*

$$\tau^2_{\mathrm{sp}\,\beta} \overset{def}{=} \frac{E[l^2(\beta,\hat{\beta})]}{r(p-r)/p},$$

where $\hat{\beta}$ is the PMCS estimate of β and $l(\cdot,\cdot)$ is the projective Frobenius distance.

Lemma 4.1.

$$\tau^2_{\mathrm{sp}\,\beta} = \frac{r - \sum_{i=1}^{r} \lambda_i}{r(p-r)/p},$$

where λ_i is the ith largest eigenvalue of $E(\beta\beta')$.

Proof. Follows from Proposition A.4 in Lütkepohl (1991, Section A.14). □

The following theorem shows that the lower and upper bound of $\tau_{\mathrm{sp}\,\beta}$ do not depend on either the number of time series in the system or the cointegration rank, which facilitates comparisons across studies. In addition, the maximal value of $\tau_{\mathrm{sp}\,\beta}$ is obtained under what is usually referred to as the uniform distribution on the Grassman manifold (Mardia and Khatri, 1977).

Theorem 4.1. *$\tau_{\mathrm{sp}\,\beta}$ satisfies*

$$0 \leq \tau_{\mathrm{sp}\,\beta} \leq 1$$

and the upper bound is obtained when $\mathrm{sp}(\beta)$ *is distributed according to the (unique) invariant Haar distribution on $\mathbb{G}_{r,p-r}$.*

Proof. The non-negativity of $\tau_{\mathrm{sp}\,\beta}$ follows directly from Definition 4.1 and the non-negativity of the projective Frobenius distance. From Lemma 4.1, $\tau_{\mathrm{sp}\,\beta}$ is maximal when $\sum_{i=1}^{r} \lambda_i$ is minimal. Note that $\lambda_1, ..., \lambda_p$ are constrained to satisfy the equation

$$\sum_{i=1}^{p} \lambda_i = \mathrm{tr}[E(\beta\beta')] = E[\mathrm{tr}(\beta'\beta)] = E[\mathrm{tr}(I_r)] = r,$$

in addition to the order restriction. One immediate candidate as a minimizer of $\sum_{i=1}^{r} \lambda_i$ is

$$\lambda_i = \frac{r}{p}, \text{ for } i = 1, ..., r,$$

resulting in $\sum_{i=1}^{r} \lambda_i = r^2/p$; we shall now prove that this is indeed the minimum by first proving that $\sum_{i=1}^{r} \lambda_i > r^2/p$ for all sequences $\lambda_1 \geq ... \geq \lambda_p$ of eigenvalues, where $\lambda_{r+1} \neq r/p$. First, if $\lambda_{r+1} < r/p$, then $\lambda_i < r/p$ for $i = r+2, ..., p$, by the order restriction, with the result that $\sum_{i=1}^{r} \lambda_i = r - \sum_{i=r+1}^{p} \lambda_i > r - r(p-r)/p = r^2/p$. On the other hand, if $\lambda_{r+1} > r/p$, then $\lambda_i > r/p$ for $i = 1, ..., r$, again using the order restriction, with the result that $\sum_{i=1}^{r} \lambda_i > r^2/p$. Having established that λ_{r+1} must equal r/p in the minimizing sequence, it follows that $\lambda_i = r/p$, for $i = 1, ..., r$, minimizes $\sum_{i=1}^{r} \lambda_i$. The upper bound now follows from Lemma 4.1

$$\max \tau^2_{\mathrm{sp}\,\beta} = \frac{r - \sum_{i=1}^{r} r/p}{r(p-r)/p} = 1.$$

In remains to be shown that the upper bound of $\tau_{\mathrm{sp}\,\beta}$ is obtained when $\mathrm{sp}(\beta)$ follows the Haar invariant distribution on $\mathbb{G}_{r,p-r}$. Mardia and Khatri (1977) shows that $E(\beta\beta') = (r/p)I_p$ if $\mathrm{sp}(\beta)$ follows the Haar invariant distribution on $\mathbb{G}_{r,p-r}$ and thus $\lambda_i = r/p$, for $i = 1, ..., r$, which in turn implies that $\tau_{\mathrm{sp}\,\beta} = 1$. □

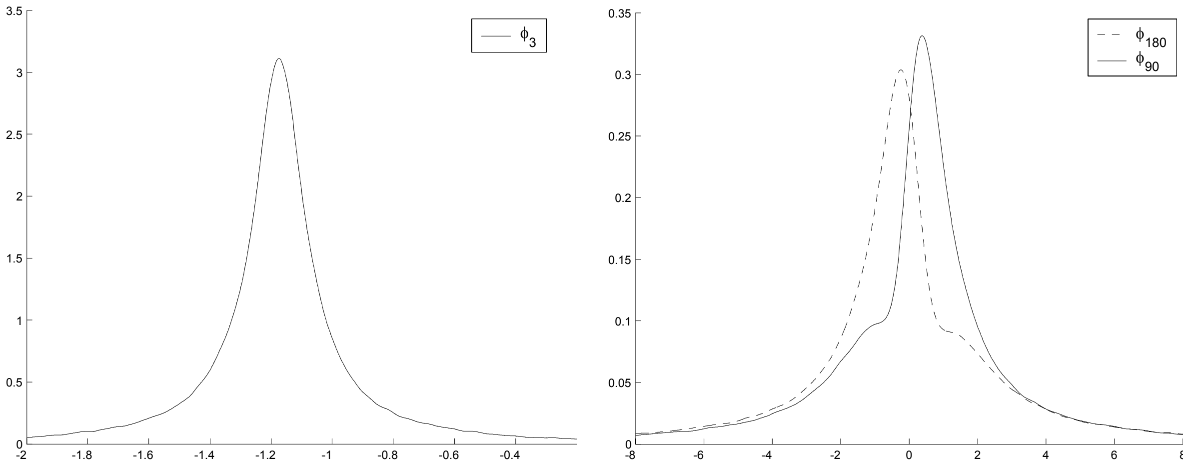

FIGURE 2. Posterior distribution of the cointegration vector $\beta = (1, \phi_3, \phi_{90}, \phi_{180})$.

5. EMPIRICAL ILLUSTRATION

The Australian interest rates data in Strachan (2003) will be used for illustration. The data consist of 94 monthly observations on four Australian interest rates of different maturity during the time period 1994:1-2001:10. Two of the rates are taken from the longer part of the yield curve with 5 and 3 years to maturity (i_5 and i_3) and the other two are shorter rates with 180 and 90 days to maturity (i_{180} and i_{90}), respectively. Following Strachan (2003), we condition our analysis on a single cointegration vector and two lagged differences in the error correction model.

The posterior distribution of the unrestricted elements of $\beta = (1, \phi_3, \phi_{180}, \phi_{90})'$ was computed using the Gibbs sampler in Villani (2003) under a flat prior. The results based on 200,000 draws (excluding 10,000 burn-in iterations) are displayed in Figure 2. Note how the marginal posterior of ϕ_{180} and ϕ_{90} are nearly mirror images of each other, reflecting that the two short rates most probably enter the cointegrating relation as a difference, but the coefficient on this spread relative the longer rates is much less clearly determined in the data. Table 1 gives the maximum likelihood and the PMCS estimates along with the median and mode plug-in estimates (see Section 2); Strachan's (2003) estimates are also presented. To show the effects of a different order of the time series, the estimates for a reversed order of the time series are also displayed. Although the PMCS estimator is invariant to the way variables have been ordered, it has been computed under both orders to show the magnitude of numerical error in the Gibbs sampling algorithm. The PMCS and maximum likelihood estimates are very close to each other and far off the other three estimates. The plug-in estimates are heavily dependent on the order of the variables, to the extent of completely over-turning the inferences.

The projective Frobenius span variation is $\tau_{\mathrm{sp}\,\beta} = 0.4223$, roughly half-way between the degenerate and the uniform distribution on the Grassman manifold, implying a rather uninformative posterior distribution of the cointegration vector.

			Order $i_5, i_3, i_{180}, i_{90}$			Order $i_{90}, i_{180}, i_3, i_5$		
Coefficient	ML	Strachan	PMCS	Median	Mode	PMCS	Median	Mode
ϕ_3	-1.026	-1.100	-1.014	-1.174	-1.175	-1.011	-0.900	0.800
ϕ_{180}	-2.022	-12.376	-2.174	-0.266	-0.240	-2.217	-3.712	-21.400
ϕ_{90}	2.040	12.263	2.176	0.420	0.380	2.216	3.591	20.000

Table 1. Estimates of the cointegration vector normalized on i_5.

6. Generalized likelihood estimators - a small simulation study

This section investigates the performance of the PMCS estimator in repeated sampling by simulation methods. In order to attract the attention of practitioners with preferences toward likelihood procedures, we will compute the expectation of $E(\beta\beta')$ with respect to the normalized likelihood function. The resulting PMCS estimator may be called the generalized likelihood estimator or the mean likelihood estimator.

The data generating process is chosen to be the bivariate VAR(1) with a single cointegration vector

$$\begin{aligned}\Delta x_{1t} &= \alpha_1(\beta_1 x_{1,t-1} + \beta_2 x_{2,t-1}) + \varepsilon_{1t}\\ \Delta x_{2t} &= \alpha_2(\beta_1 x_{1,t-1} + \beta_2 x_{2,t-1}) + \varepsilon_{2t},\end{aligned}$$

where Δ is the difference operator and $(\varepsilon_{1t}, \varepsilon_{2t})'$ independent bivariate normal vectors with zero mean and covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix}.$$

This simple process has two related advantages. First, the number of parameters is small enough to cover a relatively large part of the parameter space. Second, the marginal normalized likelihood function of the unrestricted elements of the cointegration vector is one-dimensional and may therefore be evaluated numerically over a grid without having to recourse to more advanced numerical procedures which would be prohibitively time-consuming in a simulation study.

We consider three different orthonormal cointegration vectors in the simulations

$$\beta = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \quad -\pi/2 < \theta \leq \pi/2,$$

where $\theta = 0, \pi/4$ and $\pi/2$, respectively. α_1 takes values in the set $\{-.25, -.15, -.10, -.05\}$, $\alpha_2 \in \{-.25, -.15, -.10, -.05, .00\}$, $\sigma \in \{.25, 1, 3\}$ and $\rho \in \{-.7, 0, .7\}$, all in all 180 combinations of parameters values. 50,000 data sets are simulated for each parameter setting and the ML and PMCS estimates computed for each generated data set. Two different sample sizes, $n = 25$ and $n = 50$ are used. These sample sizes are probably a fair representation of the information typically available in empirical studies where the sample sizes are usually larger, but the data much less 'tidy' than those resulting from our generating models.

The efficiency of the PMCS estimator relative the maximum likelihood estimator is measured by

$$RE = \frac{\text{Mean distance between PMCS estimate and true cointegration vector}}{\text{Mean distance between ML estimate and true cointegration vector}},$$

where the distance between an estimate of the cointegration vector and the true value is measured by the arc length distance (Edelman et al., 1998). Other distances, including the projective Frobenius distance, led to essentially the same results. $RE < 1$ indicates that the

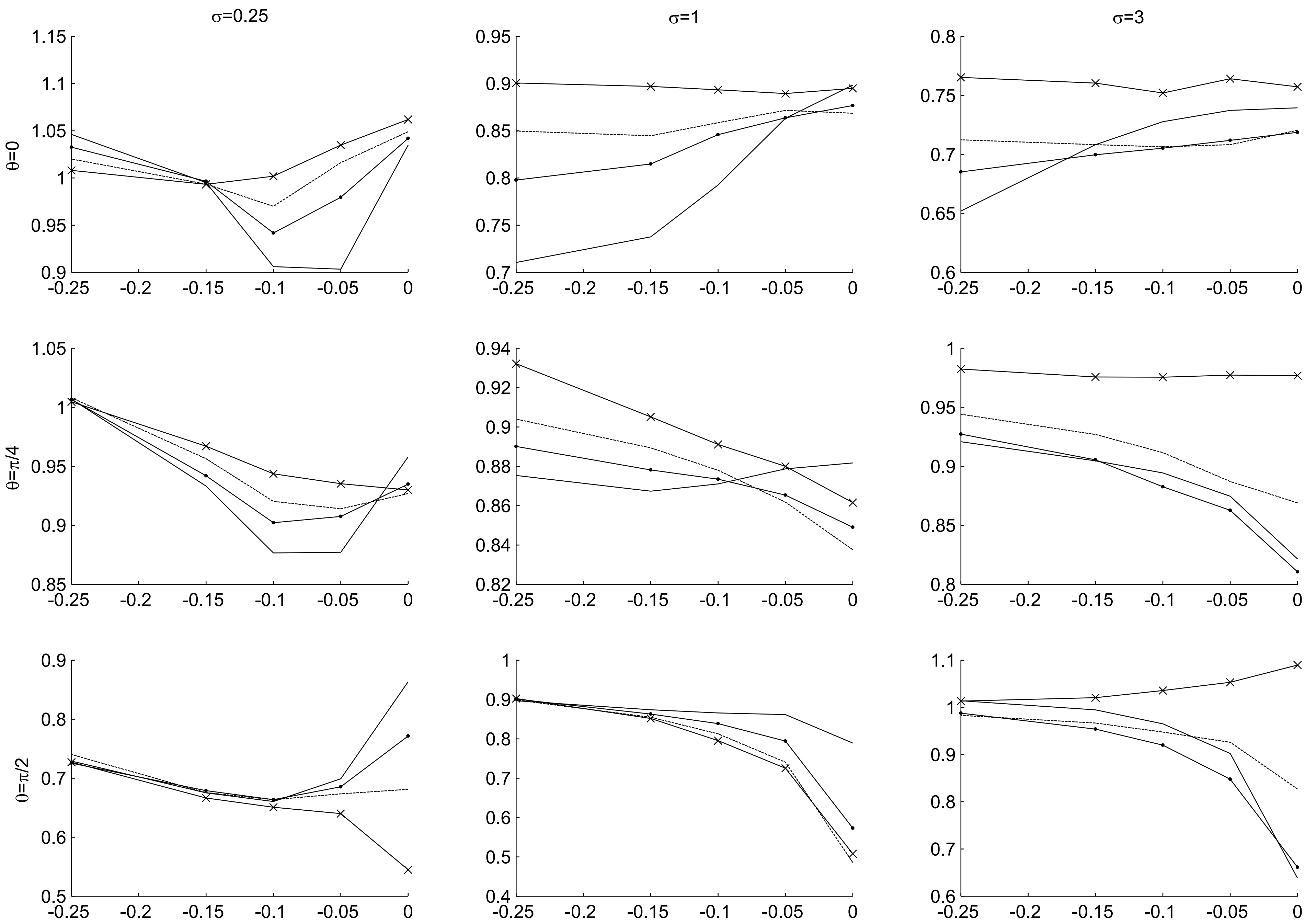

FIGURE 3. Relative efficiency (RE) of the PMCS estimator compared to the ML estimator for ρ=0 and sample size n=25. The horizontal axis displays the values of α_2. Each line graph corresponds to one of the four values of α_1: -.25 (—×—), -.15 (-----), -.1 (—●—) and -.05 (—).

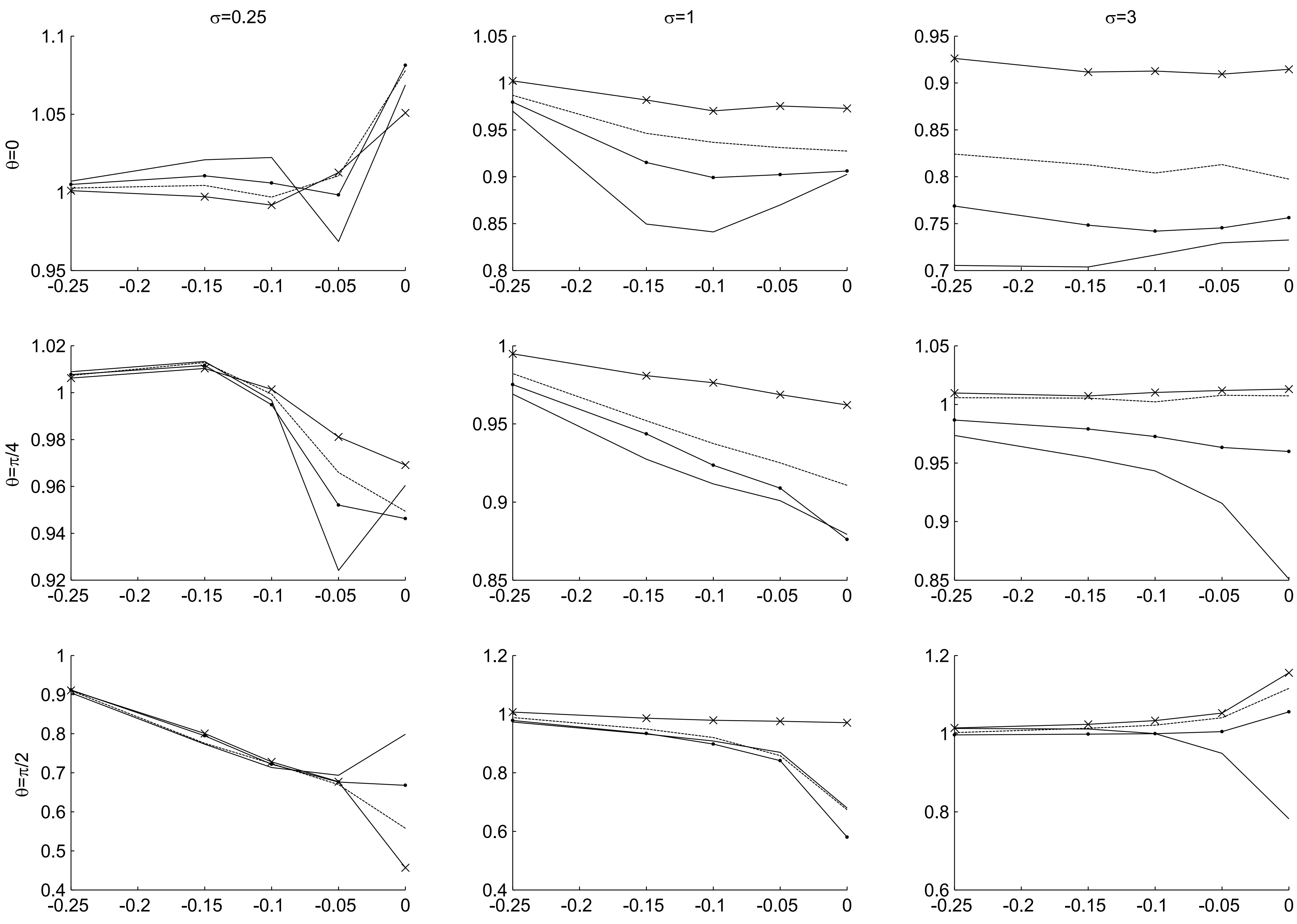

FIGURE 4. Relative efficiency (RE) of the PMCS estimator compared to the ML estimator for ρ=0 and sample size n=50. The horizontal axis displays the values of α_2. Each line graph corresponds to one of the four values of α_1: -.25 (—×—), -.15 (-----), -.1 (—●—) and -.05 (—).

PMCS estimator outperforms the ML estimator and for $RE > 1$ the opposite holds. We will only present the simulation results for the case $\rho = 0$; the results for $\rho = -0.7$ and $\rho = 0.7$ are qualitatively similar and may be obtained from the author by request. We shall let Figure 3 and 4 speak more or less for themselves and merely make a few comments on the results. A general observation is that the PMCS estimator outperforms the ML estimator in a large majority of the parameter settings, sometimes to the extent of a 50% improvement in RE. Furthermore, in those cases where the ML estimator performs better than the PMCS estimator the improvement is always modest. The differences between the two estimators diminishes as the sample increases from 25 to 50 observations, but the PMCS estimator is substantially better in some parameter settings even for the larger sample size.

7. Concluding remarks

It should be clear that the results here apply to any situation where a part of the parameter space is the Grassman manifold, e.g. subspace estimation problems. This includes of course the reduced rank regression model in Anderson (1951), but also the common factor model (Anderson, 1984) extensively used in psychometrics, the simultaneous equations model (Kleibergen and van Dijk, 1998), and many other widely used models in multivariate analysis.

We have focused here on the just-identified case, which is the starting point of most analyses. When the same over-identifying restrictions are imposed on all r cointegration vectors, the parameter space of the remaining unrestricted elements is a Grassman manifold of smaller dimension and the results here apply directly. This is a situation of substantial practical interest; it covers all linear restrictions when $r = 1$, the frequently occurring case where one or several variables are assumed not to enter any of the r cointegrating relations and many other situations. It would of course be nice to extend the results to general linear over-identifying restrictions, but this leads to complicated optimization problems which do not seem to have a closed form solution. It should be kept in mind, however, that it is less important to take the correct geometry into account when the cointegration vectors are heavily restricted.

It would be interesting to conduct the type of analysis presented here for other distance measures on the Grassman manifold and compare the resulting estimators and variation measures; Edelman et al. (1998) lists six common distances, including the projective Frobenius, but, at least in some cases, approximate or numerical solutions may be needed to solve the optimization problems.

References

[1] Anderson, T. W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions, *Annals of Mathematical Statistics*, **22**, 327-351.

[2] Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis* (2nd ed.), Wiley, New York.

[3] Bauwens, L. and Lubrano, M. (1996). Identification restrictions and posterior densities in cointegrated Gaussian VAR systems. In *Advances in Econometrics*, **11**, Part B, JAI Press, 3-28.

[4] Edelman, A., Arias, T. A. and Smith, S. T. (1998). The geometry of algorithms with orthogonality constraints, *SIAM Journal of Matrix Analysis and Applications*, **20** (2), 303-353.

[5] Geweke, J. (1996). Bayesian reduced rank regression in econometrics, *Journal of Econometrics*, **75**, 121-46.

[6] James, A. T. (1954). Normal multivariate analysis and the orthogonal group, *Annals of Mathematical Statistics*, **25**, 40-74.

[7] Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

[8] Kleibergen, F. and Paap, R. (2002). Priors, posteriors and bayes factors for a Bayesian analysis of cointegration. *Journal of Econometrics*, **111**, 223-249.

[9] Kleibergen, F. and van Dijk, H. K. (1994). On the shape of the likelihood/posterior in cointegration models, *Econometric Theory*, **10**, 514-51.

[10] Kleibergen, F. and van Dijk, H. K. (1998). Bayesian simultaneous equations analysis using reduced rank structures, *Econometric Theory*, **14**, 701-743.
[11] Kloek, T. and van Dijk, H. K. (1978). Bayesian estimates of equation system parameters: an application of integration by Monte Carlo, *Econometrica*, **46**, 1-19.
[12] Lütkepohl, H. (1993). *Introduction to Multiple Time Series Analysis*. Berlin: Springer-Verlag.
[13] Mardia, K. V. and Khatri, C. G. (1977). Uniform distribution on a Steifel manifold, *Journal of Multivariate Analysis*, **7**, 468-473.
[14] Phillips, P. C. B. (1991). Optimal inference in cointegrated systems, *Econometrica*, **59**, 283-306.
[15] Srivastava, A. (2000). A Bayesian approach to geometric subspace estimation, *IIIE Transactions on Signal Processing*, **48** (5), 1390-1400.
[16] Strachan, R. W. (2003). Valid Bayesian estimation of the cointegrating error correction model, *Journal of Business and Economic Statistics*, **21**, 185-195.
[17] Tierney, L. (1994). Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics*, **22**, 1701-1762.
[18] Villani, M. (2000). *Aspects of Bayesian Cointegration*. Doctoral thesis, Stockholm University.
[19] Villani, M. (2003). Bayesian reference analysis of cointegration. Revised version resubmitted to *Econometric Theory*. Available at http://www.statistics.su.se.

Earlier Working Papers:

Sveriges Riksbank
Visiting address: Brunkebergs torg 11
Mail address: se-103 37 Stockholm

Website: www.riksbank.se
Telephone: +46 8 787 00 00, Fax: +46 8 21 05 31
E-mail: registratorn@riksbank.se

ISSN 1402-9103 Riksbank Printing Office Stockholm 2003