A version of this paper was presented at the International Conference on Modelling & Simulation - ICMS'04 Valladolid (Spain), 22, 23 and 24 September 2004

State Space Modelling of Cointegrated Systems using Subspace Algorithms.

Segismundo Izquierdo Millán[1], Cesáreo Hernández Iglesias[2], Javier Pajares Gutiérrez[3]
Dpto. de Organización de Empresas y Comercialización e Investigación de Mercados, Universidad de Valladolid, E.T.S. Ingenieros Industriales, Pº cauce s/n. 47011 Valladolid
[1] segis@eis.uva.es
[2] cesareo@eis.uva.es
[3] pajares@eis.uva.es

SUMMARY

The use of subspace algorithms for the identification of non-stationary cointegrated stochastic systems is a promising technique that is currently under discussion. A revision of the literature provides two distinct algorithms: State Space Aoki Time Series (SSATS) identification algorithm (Aoki and Havenner 1991) and the Adapted Canonical Correlations Analysis (ACCA) of Bauer and Wagner (2002). Aoki's method is intuitively appealing, but lacks statistical foundation. In contrast, ACCA has a sound statistical basis, though intuition is somewhat lost. Both algorithms are revisited and commented. The study of the underlying ideas and properties of both previous algorithms leads us to propose a new method for subspace identification of non-stationary cointegrated stochastic systems, trying to combine the best features of each one. This new method provides a state space trend-cycle representation of a cointegrated system. Some preliminary simulation results are summarised, comparing these subspace methods with Johansen's maximum likelihood approach.

***Key words**: system identification, state space, subspace, cointegration, CCA.*

1. Introduction.

This article deals with stochastic system identification based on the "state space" formulation. System identification builds mathematical models of dynamical systems based on observed data from the system (Ljung 1999). A state space formulation constitutes one of the possible general mathematical ways to get a parametric representation of a linear, stochastic, multivariate system.

It is the object of this article to study the capabilities of the state-space formulation for the analysis of cointegrated systems, trying to throw some light as to whether it can be an alternative or a complement to the (most widely used) Johansen procedure. We will consider only discrete time and time invariant (constant parameters) systems. The two main previously proposed subspace algorithms for cointegrated systems will be revisited, and a new original algorithm will be proposed in this article. Some results from a separate comparative simulation study will be summarised.

It must be noted that there are different (though equivalent) ways to formulate a state space model, and throughout this article we will use the "innovations" state space form (Hannan and Deistler 1988), which is :

$$\mathbf{z}_{t+1} = \mathbf{A}\,\mathbf{z}_t + \mathbf{K}\,\mathbf{e}_t \qquad \text{State transition equation} \qquad [1]$$

$$\mathbf{y}_t = \mathbf{C}\,\mathbf{z}_t + \mathbf{e}_t \qquad \text{Observations equation} \qquad [2]$$

where $\mathbf{z}_t$ is a (n × 1) state vector, $\mathbf{y}_t$ is a (m × 1) observations vector, $\mathbf{e}_t$ is a (m × 1) white noise vector with $E(\mathbf{e}_t) = \mathbf{0}$ and $E(\mathbf{e}_t\, \mathbf{e}'_t) = \mathbf{R}$ (where $\mathbf{e}'_t$ stands for the transpose of $\mathbf{e}_t$) and **A**, **K**, **C** are constant matrices of coherent dimensions. The matrices {**A**, **K**, **C**, **R**} are known as system matrices.

Other popular options for the formulation and analysis of linear stochastic systems are transfer functions and VARMA formulations. If the model for a system is known, these different general formulations are equivalent, in the sense that we can interchangeably represent a same model (same object) using any of them. However, if the model is unknown, different identification methods and algorithms are associated with each different mathematical formulation, leading to (possibly) different models for the same data.

For the analysis of cointegrated systems, the most widely used method is probably Johansen's maximum likelihood method, based on a VAR formulation (Johansen 1988). We consider alternative approaches based on the state space formulation, given that this formulation offers some very interesting features for model interpretation (trend-cycle), estimation and specification.

There are several possibilities to carry out an analysis of a system based on a state space formulation. To start with, the best known and most extended econometric application of the state space formulation is quasi maximum likelihood estimation of parameters via Kalman filtering. This approach is based upon the following facts:

- Using the (unknown, parametric) state space system matrices, Kalman filtering recursively provides the (parametric) conditional mean and conditional variance of observations in a sample.

- Assuming gaussian innovations, conditional mean and variance are enough to calculate the (parametric) likelihood function of a sample.

- If we know the (parametric) likelihood function, we can search for the parameter values that maximize this function, obtaining maximum likelihood estimations of the system parameters.

This is a very general approach that can be considered a "universal method" to obtain quasi maximum likelihood estimations in linear systems. The main practical drawback of this method is that the likelihood function is usually highly non linear, and its maximum must be searched by numerical methods, which involves difficulties due to local optima and convergence issues.

Focusing on cointegrated systems, we can see that this Kalman filtering approach is a possible way to search for maximum likelihood parameter estimations. However, Johansen's method provides maximum likelihood estimations in a single step (from the residuals of several auxiliary autoregressions), and it does not require numerical optimization of a non linear function. This is the reason why, a priori, we consider that this Kalman filter - maximum likelihood approach could hardly compete with Johansen method (the only possible drawback of Johansen method in this comparison is that it requires a VAR formulation –approximation- of the generating process). After these considerations, we turn our view to other state space

based analysis tools, and, in particular, to subspace methods.

The article is structured as follows: in section 2 we describe the main underlying ideas and characteristics of subspace methods. Section 3 and section 4 provide a revision of the SSATS and ACCA algorithms, which, up to now, constitute the two most relevant subspace algorithms that have been proposed for the identification of cointegrated systems. In section 5 we present our original CCA2 algorithm. Section 6 summarises some results from a separate comparative simulation study. Finally, section 7 presents conclusions and future research.

2. Subspace algorithms.

"Subspace algorithms" make up one of the options to carry out a state-space-based system identification. The use of subspace algorithms for system identification in a stationary framework has been widely explored since the 80´s. In the non stationary framework, Aoki's SSATS (Aoki 1997) is considered, up to our knowledge, the first case of subspace algorithm specially designed to cope with non stationarities, and the only one to get some practical relevance until the recently proposed Adapted Canonical Correlation Analysis (ACCA) algorithm of Bauer and Wagner (2002). Both algorithms will be revisited in this article.

It must be said that the family of subspace methods present a number of algorithm options and variants that give rise to several different specific algorithms. Bauer and Ljung (2002) distinguish between Larimore type and MOESP type subspace algorithms. We shall focus on Larimore type algorithms, and, specially, on the Canonical Correlations Analysis (CCA) algorithm, which presents some optimal properties for stochastic identification (Larimore 1996, Bauer and Ljung 2002).

There are two main ideas exploited by Larimore type subspace algorithms:

1.- A sequence of (Kalman filter like) states can be estimated directly from observations.

2.- All system matrices can be estimated via least squares (provided that observations and a sequence of states are known).

We shall dwell briefly upon both ideas.

2.1 Estimation of a sequence of states

Estimation of a sequence of states from observations is possible by combining two facts:

a) Consider the set $\mathbf{y}_t^f$, made up by observation $\mathbf{y}_t$ plus next f-1 future observations

$$\mathbf{y}_t^f = [\mathbf{y}_t', \mathbf{y}_{t+1}', ..., \mathbf{y}_{t+f-1}']'$$

and the set $\mathbf{y}_{t-1}^p$ of p past observations

$$\mathbf{y}_{t-1}^p = [\mathbf{y}_{t-1}', \mathbf{y}_{t-2}', ..., \mathbf{y}_{t-p}']'$$

An expected value of $\mathbf{y}_t^f$ based on $\mathbf{y}_{t-1}^p$ can be calculated by the linear projection theorem: the orthogonal projection ($\mathbf{y}_t^f / \mathbf{y}_{t-1}^p$) of f future observations ($\mathbf{y}_t^f$) into p past

observations ($\mathbf{y}_{t-1}^{p}$) is

$$\mathbf{y}_t^f / \mathbf{y}_{t-1}^p = E(\mathbf{y}_t^f \ \mathbf{y}_{t-1}^{p\prime}) \, [E(\mathbf{y}_{t-1}^p \, \mathbf{y}_{t-1}^{p\prime})]^{-1} \, \mathbf{y}_{t-1}^p$$

where the projection matrix $E(\mathbf{y}_t^f \ \mathbf{y}_{t-1}^{p\prime}) \, [E(\mathbf{y}_{t-1}^p \ \mathbf{y}_{t-1}^{p\prime})]^{-1}$ can be, for stationary processes, consistently estimated directly from data.

b) Let $\hat{\mathbf{y}}_{t|t-1}$ be the expected value of $\mathbf{y}_t$ based on past values (up to time t-1). Consider a system that evolves according to state space equations [1] and [2]. At time t, the expected value ($\hat{\mathbf{y}}_{t|t-1}^{\ f}$) of $\mathbf{y}_t^f$ conditional on past observations is the product of an (extended) observability matrix by the expected value ($\check{\mathbf{z}}_{t|t-1}$) of the state vector at time t conditional on past observations. This fact can be seen by recursive substitution in the state space system equations:

$$\begin{aligned}
\hat{\mathbf{y}}_{t|t-1} &= \mathbf{C} \ \check{\mathbf{z}}_{t|t-1} \\
\hat{\mathbf{y}}_{t+1|t-1} &= \mathbf{C}\,\mathbf{A} \ \check{\mathbf{z}}_{t|t-1} \\
&\cdots \\
\hat{\mathbf{y}}_{t+f-1|t-1} &= \mathbf{C}\,\mathbf{A}^{f-1} \ \check{\mathbf{z}}_{t|t-1}
\end{aligned}$$

which can be expressed as

$$\hat{\mathbf{y}}_{t|t-1}^{\ f} = \mathbf{O}_f \, \check{\mathbf{z}}_{t|t-1}$$

where $\mathbf{O}_f$ is the (extended) observability matrix $\mathbf{O}_f = [\mathbf{C}' \ (\mathbf{C}\,\mathbf{A})' \ \ldots \ (\mathbf{C}\,\mathbf{A}^{f-1})' \ ]'$.

Combining the two previous facts, subspace algorithms decompose a matrix of estimated orthogonal projections (predictions) into the product of an (estimated) observability matrix plus an (estimated) "sequence of states" matrix. This decomposition is flexible in a certain way (it is not unique, and it fixes the coordinate basis for the states) and it can be made by means of a singular value decomposition (SVD) of the matrix of orthogonal projections.

Basically, subspace methods find (one of the possible combinations of) an observability matrix plus a sequence of states which provide almost the same predictions as those based on a linear data projection.

Importantly, the order of the matrix of projections is the order of the system (the state vector dimension in a minimal state space representation), and it can be estimated by a SVD of the projection matrix, thus allowing for a data-based system specification.

2.2 Estimation of the system matrices

The second main idea of subspace algorithms is that, if a sequence of states is known, all the system matrices {**A**, **K**, **C**, **R**} can be estimated by least squares. Note that, if $\mathbf{y}_t$ and $\mathbf{z}_t$ are known, the formula

$$\mathbf{y}_t = \mathbf{C}\,\mathbf{z}_t + \mathbf{e}_t$$

allows for a least squares estimation of **C**, $\mathbf{e}_t$ and **R**, and then, the formula

$$\mathbf{z}_{t+1} = \mathbf{A}\,\mathbf{z}_t + \mathbf{K}\,\mathbf{e}_t$$

allows for an estimation of **A** and **K**.

2.3 Subspace methods versus "classical" identification methods

One of the main advantages of subspace methods when compared with other "classical" identification methods, like Prediction Error Methods (Ljung, 1999), is that subspace algorithms are not iterative, so they are usually faster and do not present convergence problems (Van Oberschee and De Moor 1996). Final estimators can be compactly written in an elegant way as a function of a matrix decomposition of orthogonal projections obtained from original data (De Cock and De Moor, 2003).

The differences between Larimore type subspace algorithms and the Kalman filter – maximum likelihood approach to system identification can be seen in figure 1.

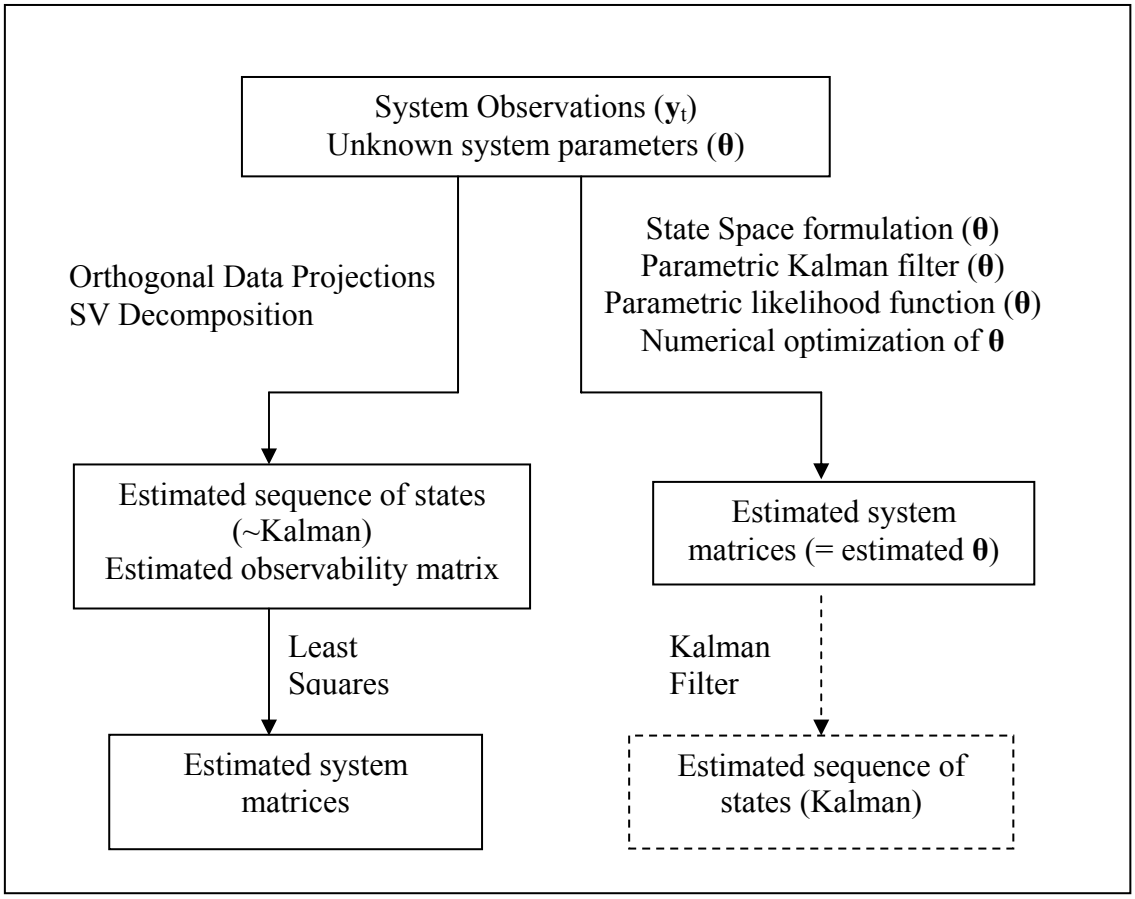

Figure 1: Two different approaches to system identification. Subspace methods (left) estimate a sequence of states first, directly from observed data, and then obtain estimates of the system matrices. Other classical identification procedures, like the one based on Kalman filtering plus maximum likelihood, estimate system parameters first, and estimates of the states may then be obtained, from system matrices plus observations. The figure is based on Van Oberschee and De Moor (1996).

3. State Space Aoki Time Series (SSATS) algorithm.

Aoki's SSATS method (Aoki 1997) was developed in 1983 as an instrumental variables

method, not as a subspace algorithm. However, several estimation alternatives have been developed since the initial algorithm (Vargas 1999), and, in fact, Aoki's method can be seen as a particular case of the subspace methods family, the one known as Principal Components (PC) (Van Oberschee and De Moor, 1996).

A very interesting idea proposed by Aoki is a two-stage application of his method to cope with non-stationary systems. The observed vector of m temporal series $\mathbf{y}_t$ is decomposed into a combination of long term (low frequency) components $\boldsymbol{\tau}_t$ (trends) plus m stationary cycles $\mathbf{y}^*_t$. Aoki proposes a first stage to estimate the trend model:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{C}_\tau \boldsymbol{\tau}_t + \mathbf{y}^*_t \\ \boldsymbol{\tau}_{t+1} &= \mathbf{A}_\tau \boldsymbol{\tau}_t + \mathbf{K}_\tau \mathbf{y}^*_t \end{aligned}$$

The residuals $\mathbf{y}^*_t$ from this first stage would not be white noise, but stationary cycles. A state space model for the cycles can be estimated in a second stage:

$$\begin{aligned} \mathbf{y}^*_t &= \mathbf{C}_\eta \boldsymbol{\eta}_t + \mathbf{e}_t \\ \boldsymbol{\eta}_{t+1} &= \mathbf{A}_\eta \boldsymbol{\eta}_t + \mathbf{K}_\eta \mathbf{e}_t \end{aligned}$$

Both models can then be stacked into a single one :

$$\begin{pmatrix} \boldsymbol{\tau}_{t+1} \\ \boldsymbol{\eta}_{t+1} \end{pmatrix} = \begin{bmatrix} \mathbf{A}_\tau & \mathbf{B}_\tau \mathbf{C}_\eta \\ \mathbf{0} & \mathbf{A}_\eta \end{bmatrix} \begin{pmatrix} \boldsymbol{\tau}_t \\ \boldsymbol{\eta}_t \end{pmatrix} + \begin{pmatrix} \mathbf{B}_\tau \\ \mathbf{B}_\eta \end{pmatrix} \mathbf{e}_t \qquad [3]$$

$$\mathbf{y}_t = \begin{bmatrix} \mathbf{C}_\tau & \mathbf{C}_\eta \end{bmatrix} \begin{pmatrix} \boldsymbol{\tau}_t \\ \boldsymbol{\eta}_t \end{pmatrix} + \mathbf{e}_t \qquad [4]$$

Note that, in this representation, trends are affected by the state of cycles, and cycles evolve independently of trends.

The expected advantage of this two-stage procedure is that short-term dynamics can be estimated without the effects caused by long term movements, which are removed in the first stage. If there are non stationarities, they should be captured in the trend model, and cointegration appears if the number of non stationary (common) trends is less than the number of non stationary observed series. Cointegration relations can be estimated because they are determined by the null space of $\mathbf{C}_\tau$.

In Aoki and Havenner (1991), a number of criteria are developed to estimate the orders of the trend (number of common trends) and cycle models. However, the method relies on heuristic arguments and it has been criticized for lack of statistical foundation. Vargas (2000) provides statistical tests for the number of common trends based on this procedure.

4. Adapted Canonical Correlations Analysis (ACCA) algorithm.

Bauer and Wagner (2002) prove that, for cointegrated processes, the standard CCA algorithm allows for consistent estimation of the cointegrating subspace, but consistent estimation of all other system parameters is not guaranteed. They propose a modification of the standard CCA algorithm in order to achieve consistent estimates of all system parameters of (co)integrated

processes of order 1. This modification is called adapted CCA or ACCA, and it chooses a canonical form in which the state transition matrix is block-diagonal, separating the dynamics into independent stationary (cycle) and non-stationary (trend) components. If desired, unit roots can be imposed for the trends, obtaining a representation of the form:

$$\begin{pmatrix} \boldsymbol{\tau}_{t+1} \\ \boldsymbol{\eta}_{t+1} \end{pmatrix} = \begin{bmatrix} \mathbf{I}_c & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_\eta \end{bmatrix} \begin{pmatrix} \boldsymbol{\tau}_t \\ \boldsymbol{\eta}_t \end{pmatrix} + \begin{pmatrix} \mathbf{B}_\tau \\ \mathbf{B}_\eta \end{pmatrix} \mathbf{e}_t$$

$$\mathbf{y}_t = \begin{bmatrix} \mathbf{C}_\tau & \mathbf{C}_\eta \end{bmatrix} \begin{pmatrix} \boldsymbol{\tau}_t \\ \boldsymbol{\eta}_t \end{pmatrix} + \mathbf{e}_t$$

Note that the dynamic matrix for the states is block diagonal, so, in this representation, trends are random walks and evolve independently of the cycles. Again, cointegration involves a number of common trends inferior to the number of observed I(1) series.

Consistent subspace estimation of all parameters of a cointegrated system is an important achievement, because it is a theoretical advantage over Johansen's method: for underlying VARMA cointegrated processes, Johansen's method applied with a fixed autoregressive lag length provides consistent estimates of the cointegrating subspace, but not consistent estimates of all system parameters (Wagner 1999). Apart from consistent estimation of all system parameters, Bauer and Wagner (2002) propose also several tests for the cointegrating rank, and a consistent estimation criterion for the system order. First results indicate that the proposed tests perform at least comparable to the Johansen method.

5. Two-stage Canonical Correlation Analysis (CCA2)

Following Aoki's ideas, we suggest a two-stage method for non stationary systems, where each stage will use the standard CCA algorithm. As stated previously, two different algorithms within the subspace family are the principal components (PC) algorithm and the Canonical Correlations Analysis (CCA) algorithm. While Aoki's SSATS is equivalent to a principal components subspace algorithm (Aoki 1997, Van Oberschee and De Moor 1996), the CCA algorithm presents optimal properties for stochastic identification, being better suited than the other subspace methods (Bauer and Ljung 2002).

So, for non stationary cointegrated systems, we take from Aoki the intuitive idea of removing low frequency trends in a first stage before modelling high frequency cycles, but we propose using the CCA subspace algorithm instead of Aoki's method for each of the modelling stages. We call this method "two-stage CCA" or CCA2.

Stacking the cycle and trend models would provide a representation like [3] and [4], in which trends are affected by the states of cycles. However, a representation with a different interpretation and a different relation between trends and cycles can be obtained if the state basis is rotated (similarity transformation).

The dimension of the state vector in the first stage (trend model) must equal the number of estimated common trends. Aoki suggested limiting the dimension of $\mathbf{y}_{t-1}^{p}$ (taking p =1) for this first stage, but that would not provide good results in our simulations. Instead, in order to

estimate the number of common trends we propose using a subspace based test (see Bauer and Wagner 2002 for a discussion of several possibilities) or any other classical test (like the Johansen test). To estimate the order of the second stage (cycle model) we selected the BA(n) criterion of Bauer and Wagner (2003).

6. Preliminary results.

We present some preliminary results from a simulation study that compares the performance of different identification methods for cointegrated systems: Johansen, CCA, ACCA and CCA2. The complete settings and results of these experiments make up a different study which is now being completed, and we just present here some general preliminary conclusions.

We considered VARMA (or state space) data generating processes (DGPs) for series of dimensions two and three, with one or two cointegrating relations. The DGPs included no deterministic trends, but constant terms in the cointegrating relations were considered (see Hamilton 1994). For Johansen method, the order of a VAR representation of the model was calculated by the Schwarz criterion (BIC). For subspace methods, the order of the system was calculated by the BA(n) criterion of Bauer and Wagner (2003).

For the models obtained by each different identification method, we compared the quality of estimation of the cointegrating subspace (see Gonzalo 1994 for related work) and the predictive performance. The cointegrating subspace was estimated assuming the right number of common trends, and quality of estimation was measured by the angle between real and estimated subspaces[*]. The predictive performance was evaluated for different prediction horizons by the mean square prediction error (in prediction samples different from the identification samples). Main results are:

- No identification method (Johansen, subspaces) is uniformly better than the other. The relative performance depends greatly on the specification and the parameter values of the DGP. For instance, consider the following DGP (DGP1):

$$\begin{bmatrix} \tau \\ \eta_1 \\ \eta_2 \end{bmatrix}_{t+1} = \begin{bmatrix} 1 & a & 0 \\ 0 & \varphi_1 & 0 \\ 0 & 0 & \varphi_2 \end{bmatrix} \begin{bmatrix} \tau \\ \eta_1 \\ \eta_2 \end{bmatrix}_t + \begin{bmatrix} 1 & 1 \\ \varphi_1 + \theta & 0 \\ 0 & \varphi_2 + \theta \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}_t$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_t = \left[\begin{array}{cc:c} \beta & 1 & 0 \\ 1 & 0 & 1 \end{array}\right] \begin{bmatrix} \frac{\sigma}{1+a}\tau \\ \eta_1 \\ \eta_2 \end{bmatrix}_t + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}_t + \begin{bmatrix} 0 \\ \delta \end{bmatrix}$$

[*] The angle between subspaces does not depend on any arbitrary normalization. For 2 and 3 dimensional series it has a clear interpretation, being an angle between two vectors in the Euclidean space, or an angle between the director vectors of two planes. The concept can be generalized to several variables (see Van Oberschee and De Moor 1996, and the function "subspace" of matlab).

where the observed series ($\mathbf{y}_t$) are composed by a common trend (τ_t) plus two stationary ARMA (1,1) processes.

Note that the parameter σ amplifies the trend component in relation to the cycle component, and the cointegration relation is determined by the first column vector of the matrix **C**, which determines the way in which the trend enters each observed series. The value δ includes a constant term in the cointegration relation. We studied the values $\beta = (1,-3)$, $\sigma = (.2, .5)$, $a = (0, 1)$, $\varphi_1 = (.5, .9)$, $\varphi_2 = (.5, .9)$, $\theta = (.5, .9)$, $\delta = 50$, plus independent noise (e_{1t}, e_{2t}) distributions (N(0, 1), t_5 , U(-2,2)), and sample sizes T = (50, 100, 300).

For most of the considered combinations of values of the parameters of DGP1, the predictive performance of models identified by subspace methods was superior to models identified by the Johansen procedure. However, for most 3 dimensional DGPs explored, the Johansen procedure outperformed subspace methods (but not uniformly, i.e., not for all combinations of system parameters).

- Although we have found some DGPs for which subspace methods perform better than Johansen's, according to our practical results we cannot defend subspace methods as being the substitute for the Johansen method (in the case of underlying VARMA generating processes), but they stand as a good complement or as an alternative.

- In general, CCA2 produced models with worse predictive performance than those provided by the standard CCA algorithm, or by ACCA. Though the 2-stage idea of CCA2 was intuitively appealing, it seems that estimating non stationary states first and stationary states later, and stacking both models, produces worse results than jointly estimating stationary and non stationary states.

- ACCA offers theoretical consistency estimation advantages over the standard CCA. However, in our experiments with different sample sizes (50, 100, 300), there were differences between ACCA and CCA models (both in prediction performance and in quality of estimation of the cointegrating subspace), but no clear advantages of any of the methods over the other. For instance, for DGP1, CCA models performed better (in predictive power and quality of estimation of the cointegrating subspace) than ACCA models, but it was the other way round in many other cases.

7. Conclusions and future research.

The two different subspace algorithms found in literature for non-stationary system identification (SSATS and ACCA) have been revisited and commented. Combining ideas and properties from both, a two-stage CCA (CCA2) algorithm for non-stationary systems has been proposed.

Preliminary simulation studies comparing CCA, ACCA, CCA2 and Johansen's ML method suggest that:

- The use of CCA2 does not seem justified when compared with the standard CCA.

- The characteristics of the generating process affect the relative performance of the

different algorithms (ML, CCA, ACCA) for model estimation.

Future research will focus on the exploration of the data features that can help us to make a decision for one or another identification procedure.

Acknowledgements

The authors want to thank Dr. J. del Hoyo from the U.A. Madrid for helpful comments, and Dr. D. Bauer for kindly providing matlab code for the ACCA algorithm. This work has been supported by the Spanish MICyT project number 2001-2108.

References

Aoki, M. (1997). "The SSATS Algorithm and Subspace Methods". *Applications of Computer Aided Time Series Modeling*, pp. 3-13, Springer-Verlag New York.

Aoki, M. and Havenner, A. (1991). "State Space Modeling of Multiple Time Series". *Econometric Reviews*, vol. 10 (1), 1-59.

Bauer, D. and Ljung, L. (2002). "Some Facts about the Choice of the Weighting Matrices in Larimore Type of Subspace Algorithms". *Automatica* 38, Issue 5, 763-773

Bauer, D. and Wagner, M. (2002). "Estimating Cointegrated Systems using Subspace Algorithms". *Journal of Econometrics* 111, 47-84.

Bauer, D. and Wagner, M. (2003). "The Performance of Subspace Algorithm Cointegration Analysis: A Simulation Study". Working paper, University of Bern.

De Cock, K. and De Moor, B. ``Subspace Identification Methods". In contribution to section 5.5, "Control systems robotics and automation" of EOLSS, UNESCO Encyclopedia of Life Support Systems, (Unbehauen H.D.), vol. 1 of 3, Eolss Publishers Co., Ltd. (Oxford, UK), 2003, pp. 933-979.

Gonzalo, J. (1994). "Five Alternative Methods of Estimating Long-run Equilibrium Relationships". *Journal of Econometrics*, 60, pp. 203-233.

Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press.

Hannan, E. and Deistler, M. (1988). *The Statistical Theory of Linear Systems*. Wiley, New York.

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

Larimore, W.E. (1996). "Statistical Optimality and Canonical Variate Analysis System Identification". *Signal Processing* 52, Issue 2, 131-144

Ljung, L. (1999). *System Identification. Theory for the User*. 2nd edition. Prentice Hall.

Van Oberschee, P. and De Moor, B. (1996). *Subspace Identification for Linear Systems: Theory – Implementation - Applications*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Vargas, M. (1999). "Modelización de Series Temporales Estacionarias en el Espacio de Estados". Working paper 2/1999/4. Fac. CC.EE. y EE. Albacete, U.C.L.M.

Vargas, M. (2000). "Análisis de Tendencias Comunes y Cointegración en el Espacio de Estados". Working paper 2/2000/3. Fac. CC.EE. y EE. Albacete, U.C.L.M.

Wagner, M. (1999). "VAR Cointegration in VARMA Models". *Economics Series* 65. Institute for Advanced Studies, Vienna.