Discussion Papers
Department of Economics
University of Copenhagen

No. 18-05

Cointegration and adjustment in the infinite order CVAR representation of some partially observed CVAR(1) models

Søren Johansen

Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K., Denmark
Tel.: +45 35 32 30 01 – Fax: +45 35 32 30 00
http://www.econ.ku.dk

ISSN: 1601-2461 (E)

Cointegration and adjustment in the infinite order CVAR representation of some partially observed CVAR(1) models

Søren Johansen[*†]
University of Copenhagen and CREATES

May 15, 2018

Abstract

A multivariate CVAR(1) model for some observed variables and some unobserved variables is analysed using its infinite order CVAR representation of the observations. Cointegration and adjustment coefficients in the infinite order CVAR are found as functions of the parameters in the CVAR(1) model. Conditions for weak exogeneity of the cointegrating vectors in the approximating finite order CVAR are derived. The results are illustrated by a few simple examples of relevance for modelling causal graphs.

Keywords: Adjustment coefficients, cointegrating coefficients, CVAR, causal models.

JEL Classification: C32.

1 Introduction

Hoover (2018) applies the CVAR(1) model for the processes X_t and T_t of dimension p and m respectively, given by the equations

$$\begin{aligned} \Delta X_{t+1} &= MX_t + CT_t + \varepsilon_{t+1}, \\ \Delta T_{t+1} &= \eta_{t+1}, \end{aligned} \tag{1}$$

to model a causal graph for the p variables $X = \{X_1, \ldots, X_p\}$ and m trends $T = \{T_1, \ldots, T_m\}$. Here the entry $M_{ij} \neq 0$ means that X_j causes X_i, which we write $X_j \to X_i$, and $C_{ij} \neq 0$ means that $T_j \to X_i$, see Examples 1–2. Note that the model assumes that there are no causal links from X to T, so that T_t is strongly exogenous.

The paper by Hoover gives a detailed and general discussion of the problems of recovering causal structures from nonstationary observations X_t, or subsets of X_t, when T_t is

[*]Søren Johansen, Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, DK-1353, Copenhagen K, Denmark. Telephone: +45 35323071. Email: Soren.Johansen@econ.ku.dk

[†]The author would like to thank Kevin Hoover for long discussions of the problem and its solution, and Massimi Franchi for reading a first version of the paper and for references to the book by Lancaster and Rodman.

unobserved, that is, if $X_t = (X'_{1t}, X'_{2t})'$, where the observations X_{1t} are p_1-dimensional and the unobserved processes X_{2t} and T_t are p_2 and m-dimensional respectively, $p = p_1 + p_2$.

Model (1) is therefore rewritten as

$$\begin{aligned} \Delta X_{1t+1} &= M_{11}X_{1t} + M_{12}X_{2t} + C_1 T_t + \varepsilon_{1t+1}, \\ \Delta X_{2t+1} &= M_{21}X_{1t} + M_{22}X_{2t} + C_2 T_t + \varepsilon_{2t+1}, \\ \Delta T_{t+1} &= \eta_{t+1}. \end{aligned} \tag{2}$$

Note that there is now a causal link from the observed process X_{1t} to the unobserved processes, if the matrix $M_{21} \neq 0$.

The process X_{1t} is a linear transformation of $\{X_t, T_t\}$ and therefore allows, in steady state, a CVAR(∞) representation, see Johansen and Juselius (2014),

$$\Delta X_{1t+1} = \alpha\beta' X_{1t} + \sum_{i=1}^{\infty} \Gamma_i \Delta X_{1,t+1-i} + \nu_{t+1}, \tag{3}$$

where $\nu_{t+1} = X_{1t+1} - E_t X_{1t+1}$ is the prediction error for the observation X_{1t+1} given $X_{10}, \ldots, X_{1t}$.

Thus, a statistical analysis, including estimation of α and β, can be conducted for the observations X_{1t} using an approximating finite order CVAR, see Saikkonen (1992) and Saikkonen and Lütkepohl (1996).

Hoover (2018) investigates in particular if weak exogeneity for β in the approximating finite order CVAR, that is, a zero row in α, can help finding the causal structure in the graph. The present note solves the problem of finding expressions for the parameters α and β in the CVAR(∞) model (3) for the observation X_{1t}, as functions of the parameters in model (2), and finds conditions on these for the presence of a zero row in α, and hence weak exogeneity for β in the approximating finite order CVAR.

2 The model and a preliminary analysis

We first give some notation and then formulate the assumptions of the model.

If A is a $k_1 \times k_2$ matrix of rank $m \leq \min(k_1, k_2)$, we define $A_\perp$ as a $k_1 \times (k_1 - m)$ matrix of rank $k_1 - m$ for which $A'_\perp A = 0$. If $k_1 = m$, we define $A_\perp = 0$. The $k_1 \times k_1$ matrix A is stable if the eigenvalues are in the open unit circle.

Assumption 1 *We make the following assumptions*

(i) The processes $\{X_{1t}, X_{2t}, T_t\}$, $t = 1, \ldots, n$, are given by the equations (2) with starting values zero: $T_0 = 0$, $X_{10} = 0$, $X_{20} = 0$, and ε_{1t}, ε_{2t}, and η_t are i.i.d. Gaussian with mean zero and variances Ω_1, Ω_2, and Ω_η, where Ω_1 and Ω_2 are diagonal matrices.

(ii) The matrices $I_{p_1} + M_{11}$, $I_{p_2} + M_{22}$ and $I_p + M$ are stable.

(iii) The matrix $C_{1.2} = C_1 - M_{12}M_{22}^{-1}C_2$ satisfies

$$rank(C_{1.2}) = m. \tag{4}$$

Assumption 1 (ii) on the matrices M_{11}, M_{22} and M are taken from Hoover (2018) to ensure that for instance the process X_t given by the equations $X_t = (I_p + M)X_{t-1} + input$, is stationary if the input is stationary, such that the nonstationarity of X_t in model (2) is created by the trends T_t, and not by the own dynamics of X_t, as given by M. It follows from this assumption that the matrices M_{11}, M_{22}, M and therefore $M_{11.2} = M_{11} - M_{12}M_{22}^{-1}M_{21}$ are nonsingular.

Assumption 1 (iii) on the rank of $C_{1.2} = C_1 - M_{12}M_{22}^{-1}C_2$, ensures that the m trends T_t cause the observed variables X_{1t}, and that $E(T_t|X_{10},\ldots,X_{1t})$ can be represented as a random walk in the prediction errors of X_{1t}.

From the model equations (2) we find, by eliminating X_{2t} from the first two equations, that

$$\Delta X_{1t+1} - M_{12}M_{22}^{-1}\Delta X_{2t+1} = M_{11.2}X_{1t} + C_{1.2}T_t + \varepsilon_{1t+1} - M_{12}M_{22}^{-1}\varepsilon_{2t+1},$$

where $M_{11.2} = M_{11} - M_{12}M_{22}^{-1}M_{21}$ has full rank p_1 and $C_{1.2} = C_1 - M_{12}M_{22}^{-1}C_2$ has rank m by Assumption 1 (ii).

The terms $\Delta X_{1t+1} - M_{12}M_{22}^{-1}\Delta X_{2t+1}$ is asymptotically stationary and $\varepsilon_{1t+1} - M_{12}M_{22}^{-1}\varepsilon_{2t+1}$ is stationary, and therefore

$$M_{11.2}X_{1t} + C_{1.2}T_t \tag{5}$$

is asymptotically stationary. It follows that for $\beta' M_{11.2}X_{1t}$ to be stationary, it must hold that $\beta' M_{11.2}^{-1}C_{1.2} = 0$, or equivalently, up to multiplication from the right by a square matrix of full rank,

$$\beta_\perp = M_{11.2}^{-1}C_{1.2} \text{ and } \beta = M_{11.2}'(C_{1.2})_\perp. \tag{6}$$

Thus it is easy to find $\beta_\perp$ and β directly from the model formulation. The derivation of α is more complicated and will be dealt with below, using the theory for the solution of algebraic Riccati equations, based on Assumption 1 (iii).

It is convenient for the analysis to rewrite the equations (2) as follows. We define the unobserved processes

$$\tilde{T}_t = \begin{pmatrix} X_{2t} \\ T_t \end{pmatrix}, \; \tilde{\eta}_t = \begin{pmatrix} \varepsilon_{2t} \\ \eta_t \end{pmatrix},$$

where $\tilde{\eta}_t$ are i.i.d. $N_{p_1+m}(0, \tilde{\Omega})$, say, and the corresponding matrices

$$\tilde{Q} = \begin{pmatrix} I_{p_2} + M_{22} & C_2 \\ 0 & I_m \end{pmatrix}, \; \tilde{C} = (M_{12}, C_1).$$

Then (2) becomes

$$\begin{aligned} X_{1t+1} &= (I_{p_1} + M_{11})X_{1t} + \tilde{C}\tilde{T}_t + \varepsilon_{1t+1}, \\ \tilde{T}_{t+1} &= \begin{pmatrix} M_{21} \\ 0 \end{pmatrix} X_{1t} + \tilde{Q}\tilde{T}_t + \tilde{\eta}_{t+1}. \end{aligned} \tag{7}$$

Note that the observations X_{1t} cause the unobserved process $\tilde{T}_{t+1}$, with coefficients M_{21}, but for $M_{21} = 0$, we get the common trend state space model. The analysis of model (7) requires a representation of $E_t\tilde{T}_{t+1}$ in terms of the prediction errors of the observations,

$$\nu_{t+1} = X_{1t+1} - E_t(X_{1t+1}) = \Delta X_{1t+1} - M_{11}X_{1t} - \tilde{C}\tilde{T}_t,$$

and for that we need to calculate the variance $V_t = Var_t\tilde{T}_t$. Here E_t and Var_t indicate conditional mean and variance given $\{X_{10},\ldots,X_{1t}\}$. The variance can be calculated recursively as follows

$$\begin{aligned} Var_{t+1}(\tilde{T}_{t+1}) &= Var_t(\tilde{T}_{t+1}|X_{1t+1}) \\ &= Var_t(\tilde{T}_{t+1}) - Cov_t(\tilde{T}_{t+1}; X_{1t+1})Var_t(X_{1t+1})^{-1}Cov_t(X_{1t+1}; \tilde{T}_{t+1}). \end{aligned}$$

Here X_{1t} does not contribute to the conditional variance $Var_t(\tilde{T}_{t+1}|X_{1t+1})$, and

$$\begin{aligned} Var_t(\tilde{T}_{t+1}) &= \tilde{Q}V_t(\tilde{T}_t)\tilde{Q}' + \tilde{\Omega}, \\ Cov_t(\tilde{T}_{t+1}; X_{1t+1}) &= Cov_t(\tilde{T}_{t+1}; (I_{p_1} + M_{11})X_{1t} + \tilde{C}\tilde{T}_t + \varepsilon_{1t+1}) = \tilde{Q}Var_t(\tilde{T}_t)\tilde{C}', \\ Var_t(X_{1t+1}) &= Var_t((I_{p_1} + M_{11})X_{1t} + \tilde{C}\tilde{T}_t + \varepsilon_{1t+1}) = \tilde{C}Var_t(\tilde{T}_t)\tilde{C}' + \Omega_1. \end{aligned}$$

Thus, we find the recursion for $V_t = Var_t(\tilde{T}_t)$,

$$V_{t+1} = \tilde{Q}V_t\tilde{Q}' + \tilde{\Omega} - \tilde{Q}V_t\tilde{C}'\{\tilde{C}V_t\tilde{C}' + \Omega_1\}^{-1}\tilde{C}V_t\tilde{Q}'. \tag{8}$$

This is the usual recursion from the Kalman filter equations for state space models. If $V_0 = 0$ and $V_t \to V$, $t \to \infty$, the limit must satisfy the algebraic Riccati matrix equation,

$$V = \tilde{Q}V\tilde{Q}' + \tilde{\Omega} - \tilde{Q}V\tilde{C}'\{\tilde{C}V\tilde{C}' + \Omega_1\}^{-1}\tilde{C}V\tilde{Q}', \tag{9}$$

see Lancaster and Rodman (1995).

Note that the prediction errors ν_{t+1} are independent and distributed as $N_{p_1}(0, \Sigma_t)$, and for $t \to \infty$ we find

$$\Sigma_t = \tilde{C}V_t\tilde{C}' + \Omega_1 \to \tilde{C}V\tilde{C}' + \Omega_1 = \Sigma. \tag{10}$$

In the next section we give some results on the algebraic Riccati matrix equation, and show that under Assumption 1 we have $V_t \to V$, so that in steady state E_tT_t is a random walk in the prediction errors. We then use (5) to find the parameters α and β. The proofs are given in the Appendix.

3 Main results

An important result on the algebraic Riccati equation, see Hautus (1969) and Lancaster and Rodman (1995, Theorems 4.5.6 and 17.5.3), gives a simple condition for the existence of the limit of the recursively defined V_t, as the solution of (9). We have chosen a formulation that is directly applicable using simple matrix theory.

Theorem 1 *Let $V_0 = 0$, and let V_t be defined by (8). If*

$$rank\begin{pmatrix} \tilde{C} \\ \tilde{Q} - \lambda I_{p_2+m} \end{pmatrix} = p_2 + m \text{ for all } |\lambda| \geq 1, \tag{11}$$

then there exists a unique V, which is the largest solution to (9), and $V_t \to V$ as $t \to \infty$. Furthermore $0 < V < \infty$ and

$$\tilde{Q} - \tilde{Q}V\tilde{C}'(\tilde{C}V\tilde{C}' + \Omega_1)^{-1}\tilde{C}' \tag{12}$$

is stable.

For model (2) we can check the assumption (11) of Theorem 1 and find the following result.

Corollary 2 *Let X_{1t}, X_{2t}, T_t be given by model (2) and let $V_t = Var_t(X_{2t}, T_t)$. Then Assumption 1 implies that V_t, given by (8) and starting with $V_t = 0$, converges to a finite positive definite limit V, which solves the algebraic Riccati equation (9).*

For large t the process X_{1t} therefore approaches steady state, see Durbin and Koopman (2012), defined by $V_t = V$, and in steady state the prediction errors are i.i.d. $N_{p_1}(0, \Sigma)$, see (10).

We can now formulate our main result. We use the notation

$$V = Var_t \begin{pmatrix} X_{2t} \\ T_t \end{pmatrix} = \begin{pmatrix} V_{X_2,X_2} & V_{X_2,T} \\ V_{T,X_2} & V_{T,T} \end{pmatrix}.$$

Theorem 3 *Under Assumption 1, and for the process X_{1t} in steady state, the coefficients α and β in the CVAR(∞) representation of X_{1t} are given for $m < p_1$ as*

$$\beta_\perp = M_{11.2}^{-1} C_{1.2}, \qquad \text{and } \alpha_\perp = \Sigma^{-1}(M_{12} V_{X_2,T} + C_1 V_{T,T}), \tag{13}$$

$$\beta = M_{11.2}' (C_{1.2})_\perp \text{ and } \alpha = \Sigma (M_{12} V_{X_2,T} + C_1 V_{T,T})_\perp. \tag{14}$$

For $m = p_1$, $\beta_\perp$ has rank p_1, and there is no cointegration: $\alpha = \beta = 0$.

The result for β is simple to analyse in terms of the parameters of the model, see (6), but the expression for α is more complicated, because it involves the matrix V.

Let $W = M_{12} V_{X_2,T} + C_1 V_{T,T}$ so that $\alpha = \Sigma W_\perp$. Then,

$$\Sigma = \tilde{C} V \tilde{C}' + \Omega_1 = M_{12}(V_{X_2 X_2} M_{12}' + V_{X_2 T} C_1') + C_1 W' + \Omega_1,$$

such that

$$\alpha = \{M_{12}(V_{X_2 X_2} M_{12}' + V_{X_2 T} C_1') + \Omega_1\} W_\perp.$$

Thus in order to investigate a zero row in α we need to know the elements of V. The matrix V is easy to calculate from the recursion (8) for given value of the parameters, because the convergence is exponentially fast, but the properties of V are more difficult to evaluate. We therefore consider some simple examples.

Example 1. If $M_{12} = 0$, so that the unobserved process X_{2t} does not cause the observation X_{1t}, we have $C_{1.2} = C_1$, $\alpha_\perp = \Sigma^{-1} C_1 V_{T,T}$, and

$$\alpha = \Sigma C_{1\perp} = ((0, C_1) V (0, C_1)' + \Omega_1) C_{1\perp} = \Omega_1 C_{1\perp}.$$

Let e_i be a p_1-dimensional unit vector, then

$$e_i' \alpha = \omega_i e_i' C_{1\perp},$$

because $\Omega_1 = diag(\omega_1, \dots, \omega_{p_1})$. Thus α has a zero row if $C_{1\perp}$ has a zero row. Thus, the observations are not caused by X_{2t}, the observations that are directly caused by the trends T_t, are weakly exogenous.

An example of $M_{12} = 0$ is the chain $T \to A \to B \to C \to D$, where we observe $X_1 = \{A, B, C, D\}$ and $X_2 = 0$ such that $M_{12} = 0$, $C_2 = 0$. Then

$$C_1 = \begin{pmatrix} * \\ 0 \\ 0 \\ 0 \end{pmatrix}, C_{1\perp} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus, the first row is a zero row, such that A is weakly exogenous.

The matrices C_1,C_2, and M_{12} indicate the causation from T to X_1 and X_2, and from X_2 to X_1. We define a property which we call strong orthogonality between these matrices. We say that two matrices A and B are strongly orthogonal if $A'DB = 0$ for all diagonal matrices or equivalently if $A_{ji}B_{jk} = 0$ for all i, j, k. Thus, if M_{12} and C_1 are strongly orthogonal, and if the j'th row of M_{12} has a nonzero element, then the j'th row of C_1 is zero and vice versa. Thus under strong orthogonality of M_{12} and C_1, if T causes a variable in X_1, then X_2 does not cause that variable and vice versa.

Example 2. If M_{12} and C_1 are strongly orthogonal and $C_2 = 0$, another simplification occurs in the result for α, namely that $V_{X_2T} = 0$, such that V is block diagonal. We prove this by induction. First for $V_0 = 0$, we find $V_1 = \tilde{\Omega}$, which is block diagonal. Next assume V_t is block diagonal and consider the expression for V_{t+1}, see (8). In this expression we note that $C_2 = 0$ implies that Q is block diagonal, and therefore the same holds for $QV_t^{1/2}$. Thus, we only have to show block diagonality of $V_t^{1/2}\tilde{C}'\{\tilde{C}V_t\tilde{C}' + \Omega_1\}^{-1}\tilde{C}V_t^{1/2}$, see (8). To simplify the notation define the normalized parameters

$$\check{M} = \Omega_1^{-1/2}M_{12}(V_t^{1/2})_{X_2,X_2} \text{ and } \check{C} = \Omega_1^{-1/2}C_1(V_t^{1/2})_{T,T}.$$

Then $\check{M}'\check{C} = 0$ and

$$\begin{aligned} & V_t^{1/2}\tilde{C}'\{\tilde{C}V_t\tilde{C}' + \Omega_1\}^{-1}\tilde{C}V_t^{1/2} \\ & = (\check{M}, \check{C})'\{\check{M}\check{M}' + \check{C}\check{C}' + I_{p_1}\}^{-1}(\check{M}, \check{C}) \\ & = (\check{M}, \check{C})'\{I_{p_1} - \check{M}(I_{p_2} + \check{M}'\check{M})^{-1}\check{M}' - \check{C}(I_m + \check{C}'\check{C})^{-1}\check{C}'\}(\check{M}, \check{C}) \\ & = \begin{pmatrix} (I_{p_2} + \check{M}'\check{M})^{-1} & 0 \\ 0 & (I_m + \check{C}'\check{C})^{-1} \end{pmatrix}. \end{aligned}$$

Thus, V_t and therefore its limit, V, are block diagonal, such that $V_{X_2,T} = 0$.

It follows that $\alpha_\perp = \Sigma^{-1}(M_{12}V_{X_2,T} + C_1V_{T,T}) = \Sigma^{-1}C_1V_{T,T}$ so that, using $M_{12}'C_1 = 0$, so that $M_{12} = C_{1\perp}\xi$, we find

$$\begin{aligned} \alpha &= \Sigma C_{1\perp} = (\tilde{C}V\tilde{C}' + \Omega_1)C_{1\perp} = (M_{12}V_{X_2,X_2}M_{12}' + \Omega_1)C_{1\perp} \\ &= C_{1\perp}\xi V_{X_2,X_2}\xi' C_{1\perp}' C_{1\perp} + \Omega_1 C_{1\perp}, \end{aligned}$$

Thus

$$e_i'\alpha = e_i'C_{1\perp}\xi V_{X_2,X_2}\xi'C_{1\perp}'C_{1\perp} + \omega_i e_i'C_{1\perp} = e_i'C_{1\perp}(\xi V_{X_2,X_2}\xi'C_{1\perp}'C_{1\perp} + \omega_i I_{p_1-m}) = e_i'C_{1\perp}\phi,$$

say. Thus again a zero row in $C_{1\perp}$ gives a zero row in α.

An example of $M_{12} \neq 0$, is given by the chain $T \to A \to B \to C \to D$, where we only observe $X_1 = \{A, C, D\}$ such that $X_2 = \{B\}$. Here B causes C and T causes A so that, for some coefficients a and b

$$M_{12} = \begin{pmatrix} 0 \\ a \\ 0 \end{pmatrix}, C_1 = \begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix}.$$

Furthermore, $C_2 = 0$ because T does not cause B directly. Note that $M'_{12}DC_1 = 0$ for all D because T and X_2 cause disjoint subsets of X_1.This implies that V is block diagonal and

$$\alpha = \Sigma C_{1\perp} = e'_i \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \phi,$$

and it is seen that A is weakly exogenous.

4 Conclusion

This paper investigates the problem of finding adjustment and cointegrating coefficients for the infinite order CVAR representation of a partially observed CVAR(1) model. The main tools are some results for the solution of the algebraic Riccati equation, and the results are exemplified by an analyse of CVAR(1) models for causal graphs.

5 Appendix

Before giving the proof of Theorem 1, we give some definitions from control theory, which are useful for working with the results in Lancaster and Rodman (1995), subsequently LR(1995).

Definition 1

(i) Let A be $n \times n$ and B be $n \times m$. The pair $\{A, B\}$ is called controllable if

$$rank(B; AB; \ldots; A^{n-1}B) = n,$$

LR(1995, (4.1.3)).

(ii) The pair $\{A, B\}$ is stabilizable if there is an $m \times n$ matrix K, such that $A + BK$ is stable LR(1995, p. 90).

(iii) Finally $\{B, A\}$ is detectable means that $\{A', B'\}$ is stabilizable, LR(1995, page 91 line 5-).

Proof of Theorem 1. The result follows from LR (1995, Theorem 17.5.3), where the assumptions, in the present notation, are

1. $(\tilde{C}, \tilde{Q})$ is detectable,
2. $(\tilde{Q}, I_{p_2+m})$ is stabilizable,
3. $(\tilde{Q}, I_{p_2+m})$ controllable.

Now $(\tilde{Q}, I_{p_2+m})$ controllable implies $(\tilde{Q}, I_{p_2+m})$ stabilizable by LR (1995,Theorem 4.4.2), and is easily established, see Definition 1 (i), because

$$rank(I_{p_2+m}; \tilde{Q}I_{p_2+m}; \ldots; \tilde{Q}^{p_2+m-1}I_{p_2+m}) = p_2 + m.$$

Thus the assumptions 2 and 3. hold.

Definition 1 (iii) shows that $(\tilde{C}, \tilde{Q})$ detectable means $(\tilde{Q}', \tilde{C}')$ stabilizable, and LR(1995, Theorem 4.5.6 (b)), see also Hautus (1969), shows that $(\tilde{Q}', \tilde{C}')$ is stabilizable, if and only if

$$rank(\tilde{Q}' - \lambda I_{p_2+m}; \tilde{C}') = p_2 + m \text{ for all } |\lambda| \geq 1, \qquad (15)$$

which is the condition (11) in Theorem 1. Thus also condition 1. is satisfied, which proves Theorem 1. ■

Proof of Corollary 2. We verify that condition (11) in Theorem 1 follows from Assumption 1 (ii) and (iii). We define

$$M(\lambda) = \begin{pmatrix} \tilde{C} \\ \tilde{Q} - \lambda I_{p_2+m} \end{pmatrix} = \begin{pmatrix} M_{12} & C_1 \\ I_{p_2} + M_{22} - \lambda I_{p_2} & C_2 \\ 0 & I_m - \lambda I_m \end{pmatrix}.$$

For $\lambda = 1$ we get, using $C_{1.2} = C_1 - M_{12}M_{22}^{-1}C_2$ and Assumption 1 (ii) and (iii), that

$$\begin{aligned} rank(M(1)) &= rank \begin{pmatrix} M_{12} & C_1 \\ M_{22} & C_2 \end{pmatrix} = rank \begin{pmatrix} 0 & C_{1.2} \\ M_{22} & C_2 \end{pmatrix} \\ &= rank(C_{1.2}) + rank(M_{22}) = m + p_2. \end{aligned}$$

For $|\lambda| > 1$ we find, using Assumption 1 (ii),

$$rank(M(\lambda)) = rank(I_{p_2} + M_{22} - \lambda I_{p_2}) + rank(I_m - \lambda I_m) = p_2 + m,$$

because λ is not an eigenvalue of the stable matrix $I_{p_2} + M_{22}$, when $|\lambda| > 1$. Thus we can apply Theorem 1 which proves Corollary 2. ■

In the following we assume that the process X_{1t} is in steady state, where the prediction errors ν_{t+1} are i.i.d. $N_{p_1}(0, \Sigma)$ for $\Sigma = \tilde{C}V\tilde{C}' + \Omega_1$.

Proof of Theorem 3. We first prove that E_tT_t is a random walk in the prediction errors $\nu_{t+1} = X_{1t+1} - E_tX_{1t+1}$. We find a recursion for the calculation of $E_t(\tilde{T}_t)$ in terms of prediction errors,

$$\begin{aligned} E_{t+1}(\tilde{T}_{t+1}) &= E_t(\tilde{T}_{t+1}|X_{1t+1}) = E_t(\tilde{T}_{t+1}) + Cov(\tilde{T}_{t+1}; X_{1t+1})Var_t(X_{1t+1})^{-1}\nu_{t+1} \\ &= \begin{pmatrix} M_{21} \\ 0 \end{pmatrix} X_{1t} + \tilde{Q}E_t(\tilde{T}_t) + Cov(\tilde{T}_{t+1}; X_{1t+1})Var_t(X_{1t+1})^{-1}\nu_{t+1}. \end{aligned}$$

Note that this is not the usual recursion from the common trends model, because of the term with M_{21}. Still, multiplying by $(0, I_m)$ we find, using $(0, I_m)\tilde{Q} = (0, I_m)$, that M_{21} cancels and

$$\begin{aligned} E_{t+1}T_{t+1} &= (0, I_m)\tilde{Q}E_t(\tilde{T}_t) + (0, I_m)\tilde{Q}V\tilde{C}\Sigma^{-1}\nu_{t+1} \\ &= E_t(T_t) + (0, I_m)V\tilde{C}\Sigma^{-1}\nu_{t+1}, \end{aligned}$$

so that E_tT_t is a random walk in the prediction errors

$$E_tT_t = E_0T_0 + (0, I_m)V\tilde{C}\Sigma^{-1}\sum_{i=1}^{t}\nu_i. \tag{16}$$

From (5) we find, taking conditional expectations, that

$$M_{11.2}X_{1t} + C_{1.2}E_tT_t = \{M_{11.2}\beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp' + C_{1.2}(0, I_m)V\tilde{C}'\Sigma^{-1}\}\sum_{i=1}^{t}\nu_i$$

is stationary, such that

$$M_{11.2}\beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp' + C_{1.2}(0, I_m)V\tilde{C}'\Sigma^{-1} = 0.$$

We use this relation to find $\alpha_\perp$, up to multiplication from the right by a full rank $m \times m$ matrix,

$$\alpha_\perp = \Sigma^{-1}\tilde{C}V(0, I_m)' = \Sigma^{-1}(M_{12}V_{X_2,T} + C_1V_{T,T})$$

and therefore

$$\alpha = \Sigma(M_{12}V_{X_2,T} + C_1V_{T,T})_\perp.$$

■

6 References

Durbin, J. and Koopman, S.J. 2012. *Time Series Analysis by State Space Methods*, 2nd ed. Oxford: Oxford University Press.

Hautus, M.L.J. 1969. Controllability and observability conditions of linear autonomous systems. Koninklijke Nederlandse Akademie van Wetenschappen. Indagationes Mathematicae, 12, 443-448.

Hoover, K.D. 2018. Long-run Causal Order: A Preliminary Investigation. Department of Economics and Department of Philosophy, Duke University.

Johansen, S. and Juselius, K. 2014. An asymptotic invariance property of the common trends under linear transformations of the data. *Journal of Econometrics* 178, 310–315.

Lancaster, P. and Rodman, L. 1995. *Algebraic Riccati equations,* Oxford, Clarendon Press.

Saikkonen, P. 1992. Estimation and testing of cointegrated systems by an autoregressive approximation. *Econometric Theory* 8, 1–27.

Saikkonen, P. and Lütkepohl, H. 1996. Infinite order cointegrated vector autoregressive processes. Estimation and Inference. *Econometric Theory* 12, 814–844.