econometrics

Article

Cointegration and Adjustment in the CVAR(∞) Representation of Some Partially Observed CVAR(1) Models

Søren Johansen

Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, DK-1353 Copenhagen K, Denmark; Soren.Johansen@econ.ku.dk; Tel.: +45-35323071

Received: 18 September 2018; Accepted: 8 January 2019; Published: 10 January 2019

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 for the cointegrating vectors in the approximating finite order CVAR are derived. The results are illustrated by two simple examples of relevance for modelling causal graphs.

Keywords: adjustment coefficients; cointegrating coefficients; CVAR; causal models

JEL Classification: C32

---

1. Introduction

In a conceptual exploration of long-run causal order, Hoover (2018) applies the CVAR(1) model for the processes $X_t = (x_{1t}, \ldots, x_{pt})'$ and $T_t = (T_{1t}, \ldots, T_{mt})'$, to model a causal graph. The process $(X_t'; T_t')'$ is a solution to the equations

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

where the error terms ε_t are independent identically distributed (i.i.d.) Gaussian variables with mean 0 and variance $\Omega_\varepsilon = \text{diag}(\omega_{11}, \ldots, \omega_{pp}) > 0$, and are independent of the errors η_t, which are (i.i.d.) Gaussian with mean 0 and variance Ω_η.

Thus, the stochastic trends, T_t are nonstationary random walks and conditions will be given below for X_t to be $I(1)$, that is, nonstationary, but ΔX_t stationary. This will imply that $MX_t + CT_t$ is stationary, so that X_t and T_t cointegrate.

The entry $M_{ij} \neq 0$ means that x_j causes x_i, which is written $x_j \to x_i$, and $C_{ij} \neq 0$ means that $T_j \to x_i$, and it is further assumed that $M_{ii} \neq 0$. Note that the model assumes that there are no causal links from X_t to T_t, so that T_t is strongly exogenous.

A simple example for three variables, x_1, x_2, x_3, and a trend T, is the graph

$$T \to x_1 \to x_2 \to x_3,$$

where the matrices are given by

$$M = \begin{pmatrix} * & 0 & 0 \\ * & * & 0 \\ 0 & * & * \end{pmatrix}, C = \begin{pmatrix} * \\ 0 \\ 0 \end{pmatrix}$$

where $*$ indicates a nonzero coefficient.

Provided that $I_p + M$ has all eigenvalues in the open unit disk, it is seen that

$$MX_{t+1} + CT_{t+1} = (I_p + M)(MX_t + CT_t) + M\varepsilon_{t+1} + C\eta_{t+1},$$

determines a stationary process defined for all t. We define a nonstationary solution to (1) for $t = 0, 1, \ldots$ by

$$X_t = -M^{-1}C\sum_{i=1}^{t}\eta_i + M^{-1}\sum_{i=0}^{\infty}(I_p + M)^i(M\varepsilon_{t-i} + C\eta_{t-i}) \text{ and } T_t = \sum_{i=1}^{t}\eta_i. \tag{2}$$

Note that the starting values are

$$X_0 = M^{-1}\sum_{i=0}^{\infty}(I_p + M)^i(M\varepsilon_{-i} + C\eta_{-i}) \text{ and } T_0 = 0.$$

It is seen that ΔX_{t+1}, ΔT_{t+1} and $MX_t + CT_t$ are stationary processes for all t, and that $(X_t'; T_t{}')'$ is a solution to Equation (1). In the following, we assume that $(X_t'; T_t')'$ is defined by (2) for $t = 0, 1, \ldots$

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 unobserved, that is, $X_t = (X_{1t}'; X_{2t}')'$ 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, $p = p_1 + p_2$. It is assumed that there are at least as many observations as trends, that is $p_1 \geq m$.

Model (1) is therefore rewritten as

$$\begin{aligned} \Delta X_{1,t+1} &= M_{11}X_{1t} + M_{12}X_{2t} + C_1T_t + \varepsilon_{1,t+1}, \\ \Delta X_{2,t+1} &= M_{21}X_{1t} + M_{22}X_{2t} + C_2T_t + \varepsilon_{2,t+1}, \\ \Delta T_{t+1} &= \eta_{t+1}. \end{aligned} \tag{3}$$

Note that there is now a causal link from the observed process X_{1t} to the unobserved process X_{2t} if $M_{21} \neq 0$.

It follows from (3) that X_{1t} is $I(1)$ and cointegrated with $p_1 - m$ cointegrating vectors β, see Theorem 1. Therefore, ΔX_{1t} has an infinite order autoregressive representation, see (Johansen and Juselius 2014, Lemma 2), which is written as

$$\Delta X_{1,t+1} = \alpha\beta' X_{1t} + \sum_{i=1}^{\infty}\Gamma_i\Delta X_{1,t+1-i} + \nu_{t+1}^{\beta}, \tag{4}$$

where the operator norm $||\Gamma_i|| = \lambda_{\max}^{1/2}(\Gamma_i'\Gamma_i)$ is $O(\rho^i)$ for some $0 < \rho < 1$. The matrices α and β are $p_1 \times m$ of rank m, and $\nu_{t+1}^{\beta} = \Delta X_{1,t+1} - E(\Delta X_{1,t+1}|\mathcal{F}_t^{\beta})$, where $\mathcal{F}_t^{\beta} = \sigma(\Delta X_{1s}, s \leq t, \beta' X_{1t})$. Thus, X_{1t} is not measurable with respect to $\mathcal{F}_t^{\beta}$, but $\beta' X_{1t}$ is measurable with respect to $\mathcal{F}_t^{\beta}$. Here, the prediction errors ν_{t+1}^{β} are i.i.d. $N_{p_1}(0, \Sigma)$, where Σ is calculated below. The representation of X_{1t}, similar to (2), is

$$X_{1t} = \beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp'\sum_{i=1}^{t}\nu_i^{\beta} + \sum_{i=0}^{\infty}C_i\nu_{t-i}^{\beta}, t = 0, 1, \ldots \tag{5}$$

where $\Gamma = I_{p_1} - \sum_{i=1}^{\infty}\Gamma_i$ and $||C_i|| = O(\rho^i)$. Here, $\beta_\perp$ is a $p_1 \times (p_1 - m)$ matrix of full rank for which $\beta'\beta_\perp = 0$, and similarly for $\alpha_\perp$. This shows that X_{1t} is a cointegrated $I(1)$ process, that is, X_{1t} is nonstationary, while $\beta' X_{1t}$ and ΔX_{1t} are stationary.

A statistical analysis, including estimation of α, β, and Γ, can be conducted for the observations X_{1t}, $t = 1, \ldots T$, using an approximating finite order CVAR, see Saikkonen (1992) and Saikkonen and Lütkepohl (1996).

Hoover (2018) investigates, in particular, whether weak exogeneity for β in the approximating finite order CVAR, that is, a zero row in α, is a useful tool for 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 (4) for the observation X_{1t}, as functions of the parameters in model (3), 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 Assumptions and Main Results

First, some definitions and assumptions are given, then the main results on α and β are presented and proved in Theorems 1 and 2. These results rely on Theorem A1 on the solution of an algebraic Riccati equation, which is given and proved in the Appendix A.

In the following, a $k \times k$ matrix is called stable, if all eigenvalues are contained in the open unit disk. If A is a $k_1 \times k_2$ matrix of rank $k \leq \min(k_1, k_2)$, an orthogonal complement, $A_\perp$, is defined as a $k_1 \times (k_1 - k)$ matrix of rank $k_1 - k$ for which $A'_\perp A = 0$. If $k_1 = k$, $A_\perp = 0$. Note that $A_\perp$ is only defined up to multiplication from the right by a $(k_1 - k) \times (k_1 - k)$ matrix of full rank. Throughout, $E_t(.)$ and $Var_t(.)$ denote conditional expectation and variance given the sigma-field $\mathcal{F}_{0,t} = \sigma\{X_{1,s}, 0 \leq s \leq t\}$, generated by the observations.

Assumption 1. *In Equation (3), it is assumed that*

(i) ε_{1t}, ε_{2t}, and η_t are mutually independent and 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) $I_{p_1} + M_{11}$, $I_{p_2} + M_{22}$ and $I_p + M$ are stable,

(iii) $C_{1.2} = C_1 - M_{12}M_{22}^{-1}C_2$ has full rank m.

Let $(X'_{1t}; X'_{2t}; T'_t)'$, $0 = 1, \ldots, n$, be the solution to (3) given in (2), such that ΔX_t and $MX_t + CT_t$ are stationary.

Assumption 1(ii) on M_{11}, M_{22} and M is 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 (3) 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 M is nonsingular, because $I_p + M$ is stable, and similarly for M_{11} and M_{22}. Moreover $M_{11.2} = M_{11} - M_{12}M_{22}^{-1}M_{21}$ is nonsingular because

$$\det M = \det M_{22} \det M_{11.2} \neq 0.$$

The Main Results

The first result on β is a simple consequence of model (3).

Theorem 1. *Assumption 1 implies that the cointegrating rank is $r = p_1 - m$, and that the coefficients β and $\beta_\perp$ in the CVAR(∞) representation for X_{1t}, see (4), are given for $p_1 > m$ as*

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

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

Proof of Theorem of 1. From the model Equation (3), it follows, by eliminating X_{2t} from the first two equations, that

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

Solving for the nonstationary terms gives

$$M_{11.2}X_{1t} + C_{1.2}T_t = \Delta X_{1,t+1} - M_{12}M_{22}^{-1}\Delta X_{2,t+1} - \varepsilon_{1t+1}, + M_{12}M_{22}^{-1}\varepsilon_{2,t+1}. \tag{7}$$

Multiplying by $\beta' M_{11.2}^{-1}$, it is seen that $\beta' X_{1t}$ is stationary, if $\beta' M_{11.2}^{-1} C_{1.2} = 0$. By Assumption 1(i), $C_{1.2}$ has rank m, so that β has rank $p_1 - m$, which proves (6). □

The result for α is more involved and is given in Theorem 2. The proof is a further analysis of (7) and involves first, the representation X_{1t} in terms of a sum of prediction errors $\nu_t^{\beta} = \Delta X_{1t} - E(\Delta X_{1t}|\mathcal{F}_{t-1}^{\beta})$, see (5), and second, a representation of $E(T_t|\mathcal{F}_{0,t}) = E(T_t|X_{10}, \ldots, X_{1t})$ as the (weighted) sum of the prediction errors $\nu_{0t} = \Delta X_{1t} - E(\Delta X_{1t}|\mathcal{F}_{0,t-1})$. The second representation requires a result from control theory on the solution of an algebraic Riccati equation, together with some results based on the Kalman filter for the calculation of the conditional mean and variance of the unobserved processes X_{2t}, T_t given the observations X_{0s}, $0 \le s \le t$. These are collected as Theorem A1 in the Appendix A.

For the discussion of these results, it is useful to reformulate (3) by defining the unobserved variables and errors

$$T_t^* = \begin{pmatrix} X_{2t} \\ T_t \end{pmatrix}, \eta_t^* = \begin{pmatrix} \varepsilon_{2t} \\ \eta_t \end{pmatrix}, \Omega^* = Var(\eta_t^*) = \begin{pmatrix} \Omega_2 & 0 \\ 0 & \Omega_\eta \end{pmatrix} \tag{8}$$

and the matrices

$$Q^* = \begin{pmatrix} I_{p_2} + M_{22} & C_2 \\ 0 & I_m \end{pmatrix}, M_{21}^* = \begin{pmatrix} M_{21} \\ 0 \end{pmatrix}, C^* = (M_{12}; C_1). \tag{9}$$

Then, (3) becomes

$$\begin{aligned} X_{1,t+1} &= (I_{p_1} + M_{11})X_{1t} + C^* T_t^* + \varepsilon_{1,t+1}, \\ T_{t+1}^* &= M_{21}^* X_{1t} + Q^* T_t^* + \eta_{t+1}^*. \end{aligned} \tag{10}$$

One can then show, see Theorem A1, that based on properties of the Gaussian distribution, a recursion can be found for the calculation of $V_t = Var_t(T_t^*)$ and $E_t = E_t(T_t^*) = E_t(T_t^*|\mathcal{F}_{0t})$ and $V_t = Var_t(T_t^*) = Var_t(T_t^*|\mathcal{F}_{0t})$, using the matrices in (8) and (9), by the equations Some

$$V_{t+1} = Q^* V_t Q^{*\prime} + \Omega^* - Q^* V_t C^{*\prime}(C^* V_t C^{*\prime} + \Omega_1)^{-1} C^* V_t Q^{*\prime}, \tag{11}$$

$$E_{t+1} = M_{21}^* X_{1t} + Q^* E_t + Q^* V_t C^{*\prime}(C^* V_t C^{*\prime} + \Omega_1)^{-1} \nu_{0t+1}. \tag{12}$$

It then follows from results from control theory, that $V = \lim_{t\to\infty} Var_t(T_t^*)$ exists and satisfies the algebraic Riccati equation

$$V = Q^* V Q^{*\prime} + \Omega^* - Q^* V C^{*\prime}(C^* V C^{*\prime} + \Omega_1)^{-1} C^* V Q^{*\prime}. \tag{13}$$

Moreover, the prediction errors $\nu_{0t} = \Delta X_{1t} - E(\Delta X_{1t}|\mathcal{F}_{0,t-1})$ are independent $N_{p_1}(0, \Sigma_t)$ for $\Sigma_t = C^* V_t C^{*\prime} + \Omega_1$, and the prediction errors $\nu_t^{\beta} = \Delta X_{1t} - E(\Delta X_{1t}|\mathcal{F}_{t-1}^{\beta})$ are independent identically distributed $N_{p_1}(0, \Sigma)$ for $\Sigma = C^* V C^{*\prime} + \Omega_1$. Finally, $E_t(T_t)$ has the representation in the prediction errors, ν_{0i},

$$E_t(T_t) = E_0(T_0) + (0; I_m) \sum_{i=1}^{t} V_i C^{*\prime} \Sigma_i^{-1} \nu_{0i}, \tag{14}$$

where $E_0(T_0) = E(T_0|X_{10}) = 0$.

Comparing the representation (5) for X_{1t} and (14) for $E_t(T_t)$ gives a more precise relation between the coefficients of the nonstationary terms in (7). The main result of the paper is to show how this leads to expressions for the coefficients α and $\alpha_\perp$ as functions of the parameters in model (3).

Theorem 2. *Assumption 1 implies, that the coefficients α and $\alpha_\perp$ in the CVAR(∞) representation of X_{1t} are given for $p_1 > m$ as*

$$\alpha_\perp = \Sigma^{-1}(M_{12}V_{2T} + C_1V_{TT}),\ \alpha = \Sigma(M_{12}V_{2T} + C_1V_{TT})_\perp, \tag{15}$$

where

$$\Sigma = Var(\nu_t^\beta) = C^*VC^{*\prime} + \Omega_1 = (M_{12}; C_1)\begin{pmatrix} V_{22} & V_{2T} \\ V_{T2} & V_{TT} \end{pmatrix}(M_{12}; C_1)' + \Omega_1. \tag{16}$$

Proof of Theorem 2. The left hand side of (7) has two nonstationary terms. The observation X_{1t} is represented in (5) in terms of a random walk in the prediction errors ν_i^β, plus a stationary term, and T_t is a random walk in η_i. Calculating the conditional expectation given the sigma-field $\mathcal{F}_{0,t}$, T_t is replaced by $E_t(T_t)$, which in (14) is represented as a weighted sum of ν_{0i}. Thus, the conditional expectation of (7) gives

$$M_{11.2}X_{1t} + C_{1.2}E_t(T_t) = E_t(\Delta X_{1t+1} - M_{12}M_{22}^{-1}\Delta X_{2,t+1}), \tag{17}$$

where the right hand side is bounded in mean:

$$E|E_t(\Delta X_{1,t+1} - M_{12}M_{22}^{-1}\Delta X_{2,t+1})| \le c\{E|\Delta X_{1,t+1}| + |\Delta X_{2,t+1}|\} \le c.$$

Setting $t = [nu]$ and dividing by $n^{1/2}$, it follows from (5) that

$$n^{-1/2}X_{1[nu]} \xrightarrow{\mathrm{D}} \beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp' W_\nu(u), \tag{18}$$

where $W_\nu(u)$ is the Brownian motion generated by the i.i.d. prediction errors ν_t^β.

From (14), it can be proved that

$$n^{-1/2}E_{[nu]}(T_{[nu]}) = (0; I_m)n^{-1/2}\sum_{t=1}^{[nu]} V_tC^{*\prime}\Sigma_t^{-1}\nu_{0t} \xrightarrow{\mathrm{D}} (0; I_m)VC^{*\prime}\Sigma^{-1}W_\nu(u). \tag{19}$$

This follows by replacing V_t, Σ_t by V, Σ, because for $\delta_t' = V_tC^{*\prime}\Sigma_t^{-1} - VC^{*\prime}\Sigma^{-1} \to 0$, it holds that

$$Var(n^{-1/2}\sum_{t=1}^{[nu]}\delta_t'\nu_{0t}) = n^{-1}\sum_{t=1}^{[nu]}\delta_t'\Sigma_t\delta_t \to 0,\ n \to \infty.$$

Next we can replace ν_{0t} by ν_t^β as follows: For $t = 0, 1, \ldots$ the sum

$$\alpha\beta'X_{1t} + \sum_{i=1}^{t}\Gamma_i\Delta X_{1,t+1-i} = \alpha\beta'X_{1t} + \Gamma_1\Delta X_{1t} + \cdots + \Gamma_t\Delta X_{11},$$

is measurable with respect to both $\mathcal{F}_t^\beta$ and $\mathcal{F}_{0t}$, such that

$$\nu_{0,t+1} - \nu_{t+1}^\beta = -E(\sum_{i=t+1}^{\infty}\Gamma_i\Delta X_{1,t+1-i}|\mathcal{F}_{0,t}) + \sum_{i=t+1}^{\infty}\Gamma_i\Delta X_{1,t+1-i}.$$

Then

$$E|\nu_{0,t+1} - \nu_{t+1}^\beta| \le c\sum_{i=t+1}^{\infty}\rho^iE|\Delta X_{1,t+1-i}| = O(\rho^t),$$

and therefore

$$E|n^{-1/2}\sum_{i=1}^{[nu]}(v^{\beta}_{t+1}-v_{0,t+1})| \leq n^{-1/2}\sum_{i=1}^{[nu]}E|v^{\beta}_{t+1}-v_{0,t+1}| \leq cn^{-1/2}\sum_{i=1}^{[nu]}\rho^i \to 0, n\to\infty,$$

which proves (19).

Finally, setting $t=[nu]$ and normalizing (17) by $n^{-1/2}$, it follows that in the limit

$$M_{11.2}\beta_\perp(\alpha'_\perp\Gamma\beta_\perp)^{-1}\alpha'_\perp W_v(u)+C_{1.2}(0;I_m)VC^{*\prime}\Sigma^{-1}W_v(u)=0 \text{ for } u\in[0,1].$$

This relation shows that the coefficient to $W_v(u)$ is zero, so that $\alpha_\perp$ can be chosen as

$$\alpha_\perp=\Sigma^{-1}C^*V(0;I_m)'=\Sigma^{-1}(M_{12}V_{2T}+C_1V_{TT})$$

and therefore $\alpha=\Sigma(M_{12}V_{2T}+C_1V_{TT})_\perp$ which proves (15). □

3. Two Examples of Simplifying Assumptions

It follows from Theorem 2 that in order to investigate a zero row in α, the matrix V is needed. This is easy to calculate from the recursion (11), for a given value of the parameters, but the properties of V are more difficult to evaluate. In general, α does not contain a zero row, but if $M_{12}V_{2T}=0$, the expressions for α and $\alpha_\perp$ simplify, so that simple conditions on M_{12} and C_1 imply a zero row in α and hence give weak exogeneity in the statistical analysis of the approximating finite order CVAR. This extra condition, $M_{12}V_{2T}=0$, implies that

$$\Sigma=(M_{12};C_1)V(M_{12};C_1)'+\Omega_1=M_{12}V_{22}M'_{12}+C_1V_{TT}C'_1+\Omega_1,$$

and

$$(M_{12}V_{2T}+C_1V_{TT})_\perp=(C_1V_{TT})_\perp=C_{1\perp},$$

such that α simplifies to

$$\alpha=(M_{12}V_{22}M'_{12}+C_1V_{TT}C'_1+\Omega_1)C_{1\perp}=(M_{12}V_{22}M'_{12}+\Omega_1)C_{1\perp}.$$

Thus, a condition for a zero row in α is

$$e'_i\alpha=e'_iM_{12}V_{22}M'_{12}C_{1\perp}+\omega_ie'_iC_{1\perp}=0 \quad (20)$$

because $\Omega_1=\text{diag}(\omega_1,\dots,\omega_{p_1})$. This is simple to check by inspecting the matrices M_{12} and $C_{1\perp}$ in model (3). In the next section, two cases are given, where such a simple solution is available.

Case 1 ($M_{12}=0$). *If the unobserved process X_{2t} does not cause the observation X_{1t}, then $M_{12}=0$. Therefore, $M_{12}V_{2T}=0$ and from (20) it follows that*

$$e'_i\alpha=\omega_ie'_iC_{1\perp}=0.$$

Thus, α has a zero row if $C_{1\perp}$ has a zero row.

An example of $M_{12}=0$ is the chain $T\to x_1\to x_2\to x_3$, where $X_1=\{x_1,x_2,x_3\}$ is observed and $X_2=0$, and hence $M_{12}=0$ and $C_2=0$. Then, because $T\to x_1$

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

Thus, the first row of $C_{1\perp}$ is a zero row, such that x_1 is weakly exogenous.

To formulate the next case, a definition of strong orthogonality of two matrices is introduced.

Definition 1. *Let A be a $k \times k_1$ matrix and B a $k \times k_2$ matrix. Then, A and B are called strongly orthogonal if $A'DB = 0$ for all diagonal matrices D, or equivalently if $A_{ji}B_{j\ell} = 0$ for all i, j, ℓ.*

Thus, if $A_{ji} \neq 0$, we assume that row j of B is zero, and if $B_{j\ell} \neq 0$, row j of A is zero. A simple example is

$$A = \begin{pmatrix} * & * \\ 0 & * \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 \\ 0 \\ * \end{pmatrix}.$$

Thus, the definition means that if two matrices are strongly orthogonal, it is due to the positions of the zeros and not to linear combination of nonzero numbers being zero.

Thus, in particular if M_{12} and C_1 are strongly orthogonal, and if T causes a variable in X_1, then X_2 does not cause that variable. The expression for V simplifies in the following case.

Lemma 1. *If $C_2 = 0$, and $M_{12}'\Omega_1^{-1}C_1 = 0$, then $Q^* = \text{blockdiag}(I_{p_2} + M_{22}; I_m)$, and $V_{2T} = 0$ such that $V = \text{blockdiag}(V_{22}; V_{TT})$.*

Proof of Lemma 1. We first prove that V_t is blockdiagonal for $t = 0$. From (2), it follows that

$$\begin{pmatrix} X_{10} \\ X_{20} \end{pmatrix} = M^{-1} \sum_{i=0}^{\infty} (I_p + M)^i (M\varepsilon_{-i} + C\eta_{-i}) \text{ and } T_0 = 0.$$

Thus, if Φ denotes the variance of $(X_{10}'; X_{20}')'$, then

$$V_0 = Var\left(\begin{pmatrix} X_{20} \\ T_0 \end{pmatrix} |X_{10}\right) = \begin{pmatrix} \Phi_{22.1} & 0 \\ 0 & 0 \end{pmatrix},$$

and hence blockdiagonal. Assume, therefore, that V_t =blockdiag$(V_{t22}; V_{tTT})$ and consider the expression for V_{t+1}, see (11). In this expression, Q^* is block diagonal (because $C_2 = 0$) and $Q^* V_t Q^{*\prime}$ and Ω^* are block diagonal, and the same holds for $Q^* V_t^{1/2}$. Thus, it is enough to show that

$$V_t^{1/2} C^{*\prime} \{C^* V_t C^{*\prime} + \Omega_1\}^{-1} C^* V_t^{1/2},$$

is block diagonal. To simplify the notation, define the normalized matrices

$$\check{M} = \Omega_1^{-1/2} M_{12} V_{t22}^{1/2} \text{ and } \check{C} = \Omega_1^{-1/2} C_1 V_{tTT}^{1/2}.$$

Then, by assumption,

$$\check{M}'\check{C} = V_{t22}^{1/2} M_{12}' \Omega_1^{-1} C_1 V_{tTT}^{1/2} = 0,$$

so that, using $V_{t2T} = 0$,

$$V_t^{1/2} C^{*\prime} (C^* V_t C^{*\prime} + \Omega_1)^{-1} C^* V_t^{1/2} = (\check{M}, \check{C})' (\check{M}\check{M}' + \check{C}\check{C}' + I_{p_1})^{-1} (\check{M}, \check{C}).$$

A direct calculation shows that

$$(\check{M}\check{M}' + \check{C}\check{C}' + I_{p_1})^{-1} = I_{p_1} - \check{M}(I_{p_2} + \check{M}'\check{M})^{-1}\check{M}' - \check{C}(I_{p_2} + \check{C}'\check{C})^{-1}\check{C}',$$

and that

$$\check{M}'\{I_{p_1} - \check{M}(I_{p_2} + \check{M}'\check{M})^{-1}\check{M}' - \check{C}(I_{p_2} + \check{C}'\check{C})^{-1}\check{C}'\}\check{C} = 0$$

such that $(\check{M}, \check{C})'(\check{M}\check{M}' + \check{C}\check{C}' + I_{p_1})^{-1}(\check{M}, \check{C})$ is block diagonal.

Then, $V_t^{1/2}C^{*\prime}\{C^* V_t C^{*\prime} + \Omega_1\}^{-1}C^* V_t^{1/2}$ and hence V_{t+1} are block diagonal. Taking the limit for $t \to \infty$, it is seen that also V is block diagonal. □

Case 2 ($C_2 = 0$, and M_{12} and C_1 are strongly orthogonal)**.** *Because $C_2 = 0$ and $M_{21}'\Omega_1^{-1}C_1 = 0$, Lemma 1 shows that $V_{2T} = 0$, so that the condition $M_{12}V_{2T} = 0$ and (20) hold. Moreover, strong orthogonality also implies that $M_{12}'C_1 = 0$ such that $M_{12} = C_{1\perp}\xi$ for some ξ. Hence*

$$e_i'\alpha = e_i' M_{12} V_{22} M_{12}' C_{1\perp} + \omega_i e_i' C_{1\perp} = e_i' C_{1\perp}(\xi V_{22} M_{12}' C_{1\perp} + \omega_i I_{p_1 - m}), \tag{21}$$

and therefore, a zero row in $C_{1\perp}$ gives a zero row in α.

Consider again the chain $T \to x_1 \to x_2 \to x_3$, but assume now that x_2 is not observed. Thus, $X_1 = \{x_1, x_3\}$ and $X_2 = \{x_2\}$. Here, T causes x_1, and x_2 causes x_3, so that

$$M_{12} = \begin{pmatrix} 0 \\ * \end{pmatrix}, C_1 = \begin{pmatrix} * \\ 0 \end{pmatrix}, C_2 = 0.$$

Note that $M_{12}'DC_1 = 0$ for all diagonal D because T and X_2 cause disjoint subsets of X_1. This, together with $C_2 = 0$, implies that V is block diagonal and that (21) holds. Thus, x_i is weakly exogenous, $e_i'\alpha = 0$, if

$$e_i' C_{1\perp} = e_i' \begin{pmatrix} 0 \\ * \end{pmatrix} = 0.$$

4. Conclusions

This paper investigates the problem of finding adjustment and cointegrating coefficients for the infinite order CVAR representation of a partially observed simple CVAR(1) model. The main tools are some classical results for the solution of the algebraic Riccati equation, and the results are exemplified by an analysis of CVAR(1) models for causal graphs in two cases where simple conditions for weak exogeneity are derived in terms of the parameters of the CVAR(1) model.

Funding: This research received no external funding

Acknowledgments: The author would like to thank Kevin Hoover for long discussions on the problem and its solution, and Massimo Franchi for reading a first version of the paper and for pointing out the excellent book by Lancaster and Rodman, and two anonymous referees who helped clarify some of the proofs.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A.

The next Theorem shows how the Kalman filter can be used to calculate $Var_t(T_t^*)$ and $E_t(T_t^*)$ using the same technique as for the common trends model and proves the existence of the limit of V_t. The last result follows from the theory of the algebraic Riccati equation, see Lancaster and Rodman (1995), in the following LR(1995).

Theorem A1. *Let X_{1t} and T_t^* be given by model (10) and let Assumption 1 be satisfied. Then, $V_t = Var_t(T_t^*)$ and $E_t = E_t(T_t^*)$ are given recursively, using the starting values E_0 and V_0 by*

$$V_{t+1} = Q^* V_t Q^{*\prime} + \Omega^* - Q^* V_t C^{*\prime} \Sigma_t^{-1} C^* V_t Q^{*\prime}, \tag{A1}$$

$$E_{t+1} = M_{21}^* X_{1t} + Q^* E_t + Q^* V_t C^{*\prime} \Sigma_t^{-1} v_{0,t+1}, \tag{A2}$$

where

$$\Sigma_t = C^* V_t C^{*\prime} + \Omega_1, \tag{A3}$$

and the prediction errors

$$\nu_{0,t+1} = X_{1,t+1} - E_t(X_{1,t+1}) \tag{A4}$$

are independent $N_{p_1}(0, \Sigma_t)$.

The sequence V_t *starting with* V_0, *converges to a finite positive limit* V, *which satisfies the algebraic Riccati equation,*

$$V = Q^* V Q^{*\prime} + \Omega^* - Q^* V C^{*\prime} \Sigma^{-1} C^* V Q^{*\prime}, \quad \Sigma = C^* V C^{*\prime} + \Omega_1. \tag{A5}$$

Furthermore,

$$Q^* - Q^* V C^{*\prime} \Sigma^{-1} C^* \tag{A6}$$

is stable, and $E_t(T_t)$ *satisfies the equation*

$$E_{t+1}(T_{t+1}) = E_t(T_t) + (0; I_m) V_t C^{*\prime} \Sigma_t^{-1} \nu_{0,t+1}. \tag{A7}$$

Proof of Theorem A1. The variance $V_t = Var_t(T_t^*)$ can be calculated recursively, using the properties of the Gaussian distribution, as

$$\begin{aligned} Var_{t+1}(T_{t+1}^*) &= Var_t(T_{t+1}^* | X_{1,t+1}) \\ &= Var_t(T_{t+1}^*) - Cov_t(T_{t+1}^*; X_{1,t+1}) Var_t(X_{1,t+1})^{-1} Cov_t(X_{1,t+1}; T_{t+1}^*). \end{aligned} \tag{A8}$$

From the model Equation (10), it follows that

$$Var_t(T_{t+1}^*) = Var_t\{M_{21}^* X_{1t} + Q^* T_t^* + \eta_{t+1}^*\} = Q^* Var_t(T_t^*) Q^{*\prime} + \Omega^*, \tag{A9}$$

$$Cov_t(T_{t+1}^*; X_{1,t+1}) = Cov_t\{T_{t+1}^*; (I_{p_1} + M_{11}) X_{1t} + C^* T_t^* + \varepsilon_{1t+1}\} = Q^* Var_t(T_t^*) C^{*\prime}, \tag{A10}$$

$$Var_t(X_{1,t+1}) = Var_t\{(I_{p_1} + M_{11}) X_{1t} + C^* T_t^* + \varepsilon_{1t+1}\} = C^* Var_t(T_t^*) C^{*\prime} + \Omega_1. \tag{A11}$$

Then, (A8)–(A11) give the recursion for $V_t = Var_t(T_t^*)$ in (A1). Similarly, for the conditional mean, it is seen that

$$\begin{aligned} E_{t+1}(T_{t+1}^*) &= E_t(T_{t+1}^* | X_{1,t+1}) = E_t(T_{t+1}^*) + Cov_t(T_{t+1}^*; X_{1,t+1}) Var_t(X_{1,t+1})^{-1} \nu_{0,t+1}, \\ E_t(T_{t+1}^*) &= M_{21}^* X_{1t} + Q^* E_t(T_t^*), \end{aligned}$$

which implies (A2) with prediction error $\nu_{0,t+1} = \Delta X_{1,t+1} - E_t(\Delta X_{1,t+1})$.

Note that (A1) is the usual recursion from the Kalman filter equations for the state space model obtained from (10) for $M_{21}^* = 0$, see Durbin and Koopman (2012). Note also, however, that (A2) is not the usual recursion from the common trends model, because of the first term containing M_{21}^*. It is seen from (A1) that if V_t converges to V, then V has to satisfy the algebraic Riccati equation (A5) and Σ is given as indicated.

The result that V_t converges to a finite positive limit follows from LR (1995, Theorem 17.5.3), where the assumptions, in the present notation, are

a.1 $(Q^*; I_{p_2+m})$ is controllable,
a.2 $(Q^*; I_{p_2+m})$ is stabilizable,
a.3 $(C^*; Q^*)$ is detectable.

Before giving the proof, some definitions from control theory are given, which are needed for checking the conditions of the results in LR(1995).

Let A be a $k \times k$ matrix and B be a $k \times k_1$ matrix.

d.1 The pair $\{A, B\}$ is called *controllable* if

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

LR(1995, (4.1.3)).

d.2 The pair $\{A; B\}$ is *stabilizable* if there is a $k_1 \times k$ matrix K, such that $A + BK$ is stable LR(1995, page 90, line 5-).

d.3 Finally $\{B; A\}$ is *detectable* means that $\{A'; B'\}$ is stabilizable, LR(1995, page 91 line 6-).

The first assumption, *a*.1, is easy to check: The pair $(Q^*; I_{p_2+m})$ is controllable, see *d*.1, means that

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

The second assumption, *a*.2, follows because controllability implies stabilizability, see LR (1995, Theorem 4.4.2).

Finally, *d*.3 shows that $(C^*; Q^*)$ detectable means $(Q^{*\prime}; C^{*\prime})$ stabilizable, and LR(1995, Theorem 4.5.6 (b)), see also Hautus (1969), shows that $(Q^{*\prime}; C^{*\prime})$ is stabilizable, if and only if

$$\text{rank}(Q^{*\prime} - \lambda I_{p_2+m}; C^{*\prime}) = \text{rank} \begin{pmatrix} M_{12} & C_1 \\ I_{p_2} + M_{22} - \lambda I_{p_2} & C_2 \\ 0 & I_m - \lambda I_m \end{pmatrix} = p_2 + m \text{ for all } |\lambda| \geq 1.$$

For $\lambda = 1$, using $C_{1.2} = C_1 - M_{12} M_{22}^{-1} C_2$ and Assumption 1, it follows that

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

For $|\lambda| > 1$, using Assumption 1(ii), it is seen that

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

because λ is not an eigenvalue of the stable matrix $I_{p_2} + M_{22}$, when $|\lambda| > 1$.

Thus, $(Q^{*\prime}; C^{*\prime})$ is stabilizable, and assumptions *a*.1, *a*.2, *a*.3 hold, such that and LR (1995, Theorem 17.5.3) applies. This proves that limit $V = \lim_{t\to\infty} V_t$ exists and (A6) holds.

Multiplying (A2) by $(0; I_m)$, it is seen, using $(0; I_m)Q^* = (0; I_m)$, and $(0; I_m)M_{21}^* = 0$, that a recursion for $E_t(T_t)$ is given by (A7). □

References

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

Hautus, Malo L. J. 1969. Controllability and observability conditions of linear autonomous systems. *Koninklijke Nederlandse Akademie van Wetenschappen. Indagationes Mathematicae* 12: 443–48.

Hoover, Kevin D. 2018. *Long-Run Causal Order: A Preliminary Investigation*. Durham: Department of Economics and Department of Philosophy, Duke University.

Johansen, Søren, and Katarina Juselius. 2014. An asymptotic invariance property of the common trends under linear transformations of the data. *Journal of Econometrics* 17: 310–15. [CrossRef]

Lancaster, Peter, and Leiba Rodman. 1995. *Algebraic Riccati Equations*. Oxford: Clarendon Press.

Saikkonen, Pentti. 1992. Estimation and testing of cointegrated systems by an autoregressive approximation. *Econometric Theory* 8: 1–27. [CrossRef]

Saikkonen, Pentti, and Helmut Lütkepohl. 1996. Infinite order cointegrated vector autoregressive processes. Estimation and Inference. *Econometric [CrossRef] Theory* 12: 814–44.