Vector autoregressions

Based on the book 'New Introduction to Multiple Time Series Analysis' by Helmut Lütkepohl'

Robert M. Kunst
robert.kunst@univie.ac.at

University of Vienna
and
Institute for Advanced Studies Vienna

November 23, 2011

Outline

Introduction

Stable VAR Processes
- Basic assumptions and properties
- Forecasting
- Structural VAR analysis

Vector error correction models
- Univariate integrated processes
- Integrated VAR processes
- Cointegrated VAR processes
- Deterministics in the VECM
- Causality and impulse response analysis

Review: univariate integrated AR processes

If a univariate autoregressive process

$$y_t = \nu + a_1 y_{t-1} + \ldots + a_p y_{t-p} + u_t$$

has inverted roots of its characteristic polynomial $\lambda_1, \ldots, \lambda_p$ such that d roots are exactly one and the remaining roots fulfil $|\lambda_j| < 1$, then it is unstable, whereas its d–th difference

$$\Delta^d y_t = (1 - L)^d y_t$$

is stable. It is then called *integrated of order d* or *I(d)*.

$d = 1$ is the most important case. The random walk is I(1).

Univariate integrated processes

More general definitions of integrated processes do not assume the process to be autoregressive. Then, a process is called $I(d)$ iff its d–th difference has a Wold-type MA representation

$$\Delta^d y_t = \sum_{j=1}^{\infty} \theta_j u_{t-j}$$

with $\theta(1) = \sum_{j=0}^{\infty} \theta_j \neq 0$ and

$$\sum_{j=0}^{\infty} j|\theta_j| < \infty.$$

The flanking conditions guarantee uniqueness of d and convergence of the so-called Beveridge-Nelson decomposition. Other authors may require different flanking conditions.

The Beveridge-Nelson decomposition

One can show that every I(1) process has a unique representation of the form

$$y_t = y_0 + \theta(1) \sum_{s=1}^{t} u_s + \sum_{j=0}^{\infty} \theta_j^* u_{t-j} - w_0^*,$$

with $\theta_j^* = -\sum_{i=j+1}^{\infty} \theta_i$ and $w_0^* = \sum_{j=0}^{\infty} \theta_j^* u_{-j}$ reflecting starting conditions.

Essentially, any I(1) process can be decomposed into a random walk and a stable remainder. A comparable decomposition for multivariate processes was established by the *Granger representation theorem*.

Integrated VAR processes: the concept

Consider a VAR without intercept

$$A(L)y_t = u_t$$

with the characteristic polynomial $A(z) = I_K - A_1 z - \ldots - A_p z^p$. Formally, any matrix A can be represented as a product of its determinant and an invertible matrix, whose inverse is called its adjoint A^{adj}. Thus,

$$|A(L)|y_t = A(L)^{adj} u_t$$

always works. If m roots of $|A(z)|$ are one and the remaining ones are larger than one, i.e.

$$|A(L)| = (1 - L)^m \alpha(L),$$

with $\alpha(.)$ invertible, the process y_t is called *integrated*. However, m need not be the desired d.

Two role-model examples for integrated VAR processes

The VAR(1)

$$y_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} y_{t-1} + u_t$$

has determinant $(1-z)^2$. Here, y_t should be I(1).

The VAR(2)

$$y_t = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} y_{t-1} + \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} y_{t-2} + u_t$$

has determinant $(1-z)^2$. Here, y_t should be I(2).

Integration order for vector autoregressions

Definition

A vector autoregressive process (y_t) (whose determinant has only roots at one or outside the unit circle) is called *integrated of order d* or *I(d)* iff $\Delta^d y_t$ is stable but $\Delta^{d-1} y_t$ is not stable.

Remark: Usually, the differenced processes do not have finite-order VAR representations. Some components often have individual lower integration order than d and may also be stationary.

Cointegrated VAR processes: the concept

Definition

Let (y_t) be a vector autoregressive process of order p that is integrated of order d. It is called *cointegrated of order* (d, b) or *CI(d,b)* iff there exists a linear combination $z_t = \beta' y_t$ with $\beta = (\beta_1, \ldots, \beta_K)' \neq 0$ such that z_t is $I(d - b)$.

Remarks:

- If a component of y_t, say the first, is integrated of lower order than d, the conditions are fulfilled for $\beta = (1, 0, \ldots, 0)'$: there is *self-cointegration*, although it may not correspond to the original idea;
- The case $CI(1, 1)$ is the most important one;
- The linear combination β is called a *cointegrating vector*.

Cointegrated VAR: role-model case

Consider the VAR(1)

$$y_t = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} y_{t-1} + u_t,$$

which has determinant $(1 - z)$ and is I(1). The combination $(1, -1)y_t = y_{1,t} - y_{2,t}$ is white noise, and $\beta = (1, -1)'$ is a cointegrating vector. Note the equivalent representation

$$\Delta y_t = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} y_{t-1} + u_t,$$

which shows the vector β in the coefficient matrix.

Properties of cointegrating vectors

- Cointegrating vectors are *not unique*. Even when there is only one cointegrating vector, all multiples will also cointegrate;
- If there are several linear independent cointegrating vectors, all their linear combinations also cointegrate. The cointegrating vectors can be seen as a basis of a linear *cointegrating space*;
- The dimension of the cointegrating space, the *cointegrating rank*, is unique.

The vector error-correction model

Any VAR(p) model $y_t = A_1 y_{t-1} + \ldots + A_p y_{t-p} + u_t$ can be re-written in the error-correction form (*vector error-correction model*, VECM)

$$\Delta y_t = \Pi y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + u_t,$$

with a one-to-one functional relation between the parameters $(A_1, \ldots, A_p)$ and $(\Pi, \Gamma_1, \ldots, \Gamma_{p-1})$. In particular,

$$\Pi = -(I - A_1 - \ldots - A_p), \quad \Gamma_j = -(A_{j+1} + \ldots + A_p).$$

Cointegration in the VECM

If (y_t) is integrated, Π will have reduced rank $r < K$. Any matrix of rank $r < K$ can be decomposed as

$$\Pi = \alpha\beta',$$

with α and β of rank r and dimension $K \times r$. Thus,

$$\Delta y_t = \alpha\beta' y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + u_t,$$

and β contains the cointegrating vectors if (y_t) is CI(1,1). Note that all terms $\Delta y_{t-j} \quad (j = 0, \ldots, p-1)$, u_t, and $\beta' y_{t-1}$ are stationary. α is called the *loading matrix*.

A suggested normalization

While Π is unique, the decomposition and hence α and β are not. A potential normalization is to impose unit length on all columns of β. Alternatively, normalize β to the form

$$\beta = \begin{pmatrix} I_r \\ \beta_{K-r} \end{pmatrix}.$$

With self-cointegration, the corresponding column in β is a unit vector. Stationary component variables of y_t should be among the first r positions.

The economic interpretation

- Economists see the vectors in β as dynamic equilibrium conditions: the stationary islands in a wild integrated sea;
- Loading matrices describe how component variables react to deviations from equilibrium: α_{ij} shows how component i reacts to deviations from equilibrium condition j;
- Γ_j represent short-run dynamics in the system (inertia).

Granger's Representation Theorem

Proposition

Suppose

$$\Delta y_t = \alpha\beta' y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + u_t, \quad t = 1, 2, \ldots,$$

with $y_t = u_t = 0, t \leq 0$ *and* u_t *is white noise for* $t > 0$*. The corresponding polynomial* $C(z)$ *is defined by*

$$C(z) = (1-z)I_K - \alpha\beta' z - \sum_{j=1}^{p-1} \Gamma_j (1-z) z^j.$$

Assume (a) $|C(z)| = 0 \therefore |z| > 1 \quad \text{or} \quad z = 1$*; (b) The number of unit roots is* $K - r$*; (c)* α *and* β *have dimension* $K \times r$ *and rank* r*. Then* y_t *can be represented as*

$$y_t = \Xi \sum_{j=1}^{t} u_j + \Xi^*(L) u_t + y_0^*.$$

Some details for the representation theorem

$$\Xi = \beta_\perp \left\{ \alpha'_\perp \left(I_K - \sum_{j=1}^{p-1} \Gamma_j \right) \beta_\perp \right\}^{-1} \alpha'_\perp$$

describes the 'weight' of the pure $(K-r)$–dimensional random walk in the multivariate Beveridge-Nelson decomposition. y_0^* contains initial values. $\Xi^*(L)u_t$ is a convergent stable process, with Ξ_j^* some (complex) function of the original parameters.

Deterministic terms in the VECM

Suppose

$$y_t = \mu_t + x_t,$$

with x_t non-deterministic and (1) $\mu_t = \mu_0$ or (2) $\mu_t = \mu_0 + \mu_1 t$. For (1), substitute into the VECM

$$\Delta y_t = \alpha\beta'(y_{t-1} - \mu_0) + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + u_t,$$

as $\Delta\mu_0 = 0$. The term $\beta'(y_{t-1} - \mu_0)$ can be re-written as

$$(\beta' : -\beta'\mu_0) \begin{pmatrix} y_{t-1} \\ 1 \end{pmatrix}.$$

Formally, this looks as if y were cointegrated with the constant 1. Note that y is not trending here.

The trending case of the VECM

Suppose case (2), i.e. $\mu_t = \mu_0 + \mu_1 t$. Substitute into the VECM

$$\Delta y_t - \mu_1 = \alpha\beta'\{y_{t-1} - \mu_0 - \mu_1(t-1)\} + \sum_{j=1}^{p-1} \Gamma_j(\Delta y_{t-j} - \mu_1) + u_t,$$

as $\Delta(\mu_1 t) = \mu_1$. The term $\beta'\{y_{t-1} - \mu_0 - \mu_1(t-1)\}$ can be re-written as

$$(\beta' : -\beta'\mu_1)\begin{pmatrix} y_{t-1} \\ t-1 \end{pmatrix} - \beta'\mu_0.$$

Formally, this looks as if y were cointegrated with the linear time trend. Note that y is I(1) plus linear trend here, while $\beta' y$ is trend-stationary but not stationary.

The trending case of the VECM without explicit t

If $\beta'\mu_1 = 0$, then the corresponding term disappears and the system can be written as

$$\Delta y_t = \nu + \alpha\beta' y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + u_t,$$

with $\nu = (I_K - \Gamma_1 - \ldots - \Gamma_{p-1})\mu_1 - \alpha\beta'\mu_0$. This is the often estimated VECM for trending variables. y is I(1) plus linear trend, while $\beta' y$ is stationary.

It is debatable whether the restriction relative to the previous model is natural in applications. A cointegrating vector that yields a trend-stationary variable instead of a stationary one may be of little interest.

Dynamic analysis in the VECM

Forecasting, causality, impulse response analysis: what changes in an integrated VAR relative to the stable VAR?

- Forecasting: the same calculus applies, the prediction variance increases without finite bounds as the horizon h increases, while the prediction variance in the cointegrating directions remains bounded;
- Causality: there are two sources of causality in a cointegrated VAR, short- and long-run causality;
- IRF: response functions approach a non-zero constant in all integrated directions, while they approach zero for cointegrated directions, including potential stationary components (self-cointegration).

Causality analysis in the VECM

For two sub-vectors $y = (z', x')'$, consider the partitioned representation of the VECM

$$\begin{bmatrix} \Delta z_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{bmatrix} \begin{bmatrix} z_{t-1} \\ x_{t-1} \end{bmatrix} + \sum_{j=1}^{p-1} \begin{bmatrix} \Gamma_{11,j} & \Gamma_{12,j} \\ \Gamma_{21,j} & \Gamma_{22,j} \end{bmatrix} \begin{bmatrix} \Delta z_{t-j} \\ \Delta x_{t-j} \end{bmatrix} + u_t.$$

Absence of causality from x to z requires $\Pi_{12} = 0$ *and* $\Gamma_{12,j} = 0, j = 1, \ldots, p-1$. The former event means that z does not adjust to the disequilibrium $\beta' y$ that may include x. The latter one means that Δz is unaffected by lagged 'short-run' effects in Δx.