VAR Models and Cointegration

The Granger representation theorem links cointegration to error correction models. In a series of important papers and in a marvelous textbook, Soren Johansen firmly roots cointegration and error correction models in a vector autoregression framework. This section outlines Johansen's approach to cointegration modeling.

The Cointegrated VAR

Consider the levels VAR(p) for the $(n \times 1)$ vector $\mathbf{Y}_t$

$$\begin{aligned}
\mathbf{Y}_t &= \mathbf{\Phi D}_t + \mathbf{\Pi}_1 \mathbf{Y}_{t-1} + \cdots + \mathbf{\Pi}_p \mathbf{Y}_{t-p} + \boldsymbol{\varepsilon}_t, \\
t &= 1, \ldots, T, \\
\mathbf{D}_t &= \text{ deterministic terms}
\end{aligned}$$

Remarks:

- The VAR(p) model is stable if

$$\det(\mathbf{I}_n - \mathbf{\Pi}_1 z - \cdots - \mathbf{\Pi}_p z^p) = 0$$

  has all roots outside the complex unit circle.

- If there are roots on the unit circle then some or all of the variables in $\mathbf{Y}_t$ are $I(1)$ and they may also be cointegrated.

- If Y_t is cointegrated then the VAR representation is not the most suitable representation for analysis because the cointegrating relations are not explicitly apparent.

The cointegrating relations become apparent if the levels VAR is transformed to the *vector error correction model* (VECM)

$$\begin{gathered}
\mathbf{\Delta Y}_t = \mathbf{\Phi D}_t + \mathbf{\Pi Y}_{t-1} + \mathbf{\Gamma}_1 \mathbf{\Delta Y}_{t-1} \\
+ \cdots + \mathbf{\Gamma}_{p-1} \mathbf{\Delta Y}_{t-p+1} + \varepsilon_t \\
\mathbf{\Pi} = \mathbf{\Pi}_1 + \cdots + \mathbf{\Pi}_p - \mathbf{I}_n \\
\mathbf{\Gamma}_k = - \sum_{j=k+1}^{p} \mathbf{\Pi}_j, \ k = 1, \ldots, p-1
\end{gathered}$$

- In the VECM, $\mathbf{\Delta Y}_t$ and its lags are $I(0)$.

- The term $\mathbf{\Pi Y}_{t-1}$ is the only one which includes potential $I(1)$ variables and for $\mathbf{\Delta Y}_t$ to be $I(0)$ it must be the case that $\mathbf{\Pi Y}_{t-1}$ is also $I(0)$. Therefore, $\mathbf{\Pi Y}_{t-1}$ must contain the cointegrating relations if they exist.

If the VAR(p) process has unit roots ($z = 1$) then

$$\begin{gathered}\det(\mathbf{I}_n - \mathbf{\Pi}_1 - \cdots - \mathbf{\Pi}_p) = 0 \\ \Rightarrow \det(\mathbf{\Pi}) = 0 \\ \Rightarrow \mathbf{\Pi} \text{ is singular}\end{gathered}$$

If $\mathbf{\Pi}$ is singular then it has *reduced rank*; that is $rank(\mathbf{\Pi}) = r < n$.

There are two cases to consider:

1. $rank(\mathbf{\Pi}) = 0$. This implies that

$$\begin{gathered}\mathbf{\Pi} = \mathbf{0} \\ \mathbf{Y}_t \sim I(1) \text{ and not cointegrated}\end{gathered}$$

The VECM reduces to a VAR($p-1$) in first differences

$$\Delta\mathbf{Y}_t = \mathbf{\Phi}\mathbf{D}_t + \mathbf{\Gamma}_1\Delta\mathbf{Y}_{t-1} + \cdots + \mathbf{\Gamma}_{p-1}\Delta\mathbf{Y}_{t-p+1} + \boldsymbol{\varepsilon}_t.$$

2. $0 < rank(\mathbf{\Pi}) = r < n$. This implies that $\mathbf{Y}_t$ is $I(1)$ with r linearly independent cointegrating vectors and $n-r$ common stochastic trends (unit roots). Since $\mathbf{\Pi}$ has rank r it can be written as the product

$$\underset{(n\times n)}{\mathbf{\Pi}} = \underset{(n\times r)}{\boldsymbol{\alpha}} \underset{(r\times n)}{\boldsymbol{\beta}}'$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $(n \times r)$ matrices with $rank(\boldsymbol{\alpha}) = rank(\boldsymbol{\beta}) = r$. The rows of $\boldsymbol{\beta}'$ form a basis for the r cointegrating vectors and the elements of $\boldsymbol{\alpha}$ distribute the impact of the cointegrating vectors to the evolution of $\Delta\mathbf{Y}_t$. The VECM becomes

$$\begin{aligned}\Delta\mathbf{Y}_t &= \mathbf{\Phi}\mathbf{D}_t + \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{Y}_{t-1} + \mathbf{\Gamma}_1\Delta\mathbf{Y}_{t-1} \\ &\quad + \cdots + \mathbf{\Gamma}_{p-1}\Delta\mathbf{Y}_{t-p+1} + \boldsymbol{\varepsilon}_t,\end{aligned}$$

where $\boldsymbol{\beta}'\mathbf{Y}_{t-1} \sim I(0)$ since $\boldsymbol{\beta}'$ is a matrix of cointegrating vectors.

Non-uniqueness

It is important to recognize that the factorization $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$ is not unique since for any $r \times r$ nonsingular matrix $\mathbf{H}$ we have

$$\begin{aligned} \boldsymbol{\alpha}\boldsymbol{\beta}' &= \boldsymbol{\alpha}\mathbf{H}\mathbf{H}^{-1}\boldsymbol{\beta}' = (\mathbf{a}\mathbf{H})(\boldsymbol{\beta}\mathbf{H}^{-1\prime})' = \mathbf{a}^{*}\boldsymbol{\beta}^{*\prime} \\ \mathbf{a}^{*} &= \mathbf{a}\mathbf{H},\ \boldsymbol{\beta}^{*} = \boldsymbol{\beta}\mathbf{H}^{-1\prime} \end{aligned}$$

Hence the factorization $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$ only identifies the space spanned by the cointegrating relations. To obtain unique values of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}'$ requires further restrictions on the model.

Example: Consider the bivariate VAR(1) model for $\mathbf{Y}_t = (y_{1t}, y_{2t})'$

$$\mathbf{Y}_t = \mathbf{\Pi}_1 \mathbf{Y}_{t-1} + \boldsymbol{\epsilon}_t.$$

The VECM is

$$\Delta \mathbf{Y}_t = \mathbf{\Pi} \mathbf{Y}_{t-1} + \varepsilon_t$$
$$\mathbf{\Pi} = \mathbf{\Pi}_1 - \mathbf{I}_2$$

Assuming $\mathbf{Y}_t$ is cointegrated there exists a 2×1 vector $\boldsymbol{\beta} = (\beta_1, \beta_2)'$ such that

$$\boldsymbol{\beta}' \mathbf{Y}_t = \beta_1 y_{1t} + \beta_2 y_{2t} \sim I(0)$$

Using the normalization $\beta_1 = 1$ and $\beta_2 = -\beta$ the cointegrating relation becomes

$$\boldsymbol{\beta}' \mathbf{Y}_t = y_{1t} - \beta y_{2t}$$

This normalization suggests the stochastic long-run equilibrium relation

$$y_{1t} = \beta y_{2t} + u_t$$

Since $\mathbf{Y}_t$ is cointegrated with one cointegrating vector, $rank(\mathbf{\Pi}) = 1$ so that

$$\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}' = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \begin{pmatrix} 1 & -\beta \end{pmatrix} = \begin{pmatrix} \alpha_1 & -\alpha_1\beta \\ \alpha_2 & -\alpha_2\beta \end{pmatrix}.$$

The elements in the vector $\boldsymbol{\alpha}$ are interpreted as *speed of adjustment* coefficients. The cointegrated VECM for $\Delta\mathbf{Y}_t$ may be rewritten as

$$\Delta\mathbf{Y}_t = \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{Y}_{t-1} + \boldsymbol{\varepsilon}_t.$$

Writing the VECM equation by equation gives

$$\begin{aligned} \Delta y_{1t} &= \alpha_1(y_{1t-1} - \beta y_{2t-1}) + \varepsilon_{1t}, \\ \Delta y_{2t} &= \alpha_2(y_{1t-1} - \beta y_{2t-1}) + \varepsilon_{2t}. \end{aligned}$$

The stability conditions for the bivariate VECM are related to the stability conditions for the disequilibrium error $\boldsymbol{\beta}'\mathbf{Y}_t$.

It is straightforward to show that $\boldsymbol{\beta}'\mathbf{Y}_t$ follows an AR(1) process

$$\boldsymbol{\beta}'\mathbf{Y}_t = (\mathbf{1} + \boldsymbol{\beta}'\boldsymbol{\alpha})\boldsymbol{\beta}'\mathbf{Y}_{t-1} + \boldsymbol{\beta}'\boldsymbol{\varepsilon}_t$$

or

$$\begin{aligned} u_t &= \phi u_{t-1} + v_t, \; u_t = \boldsymbol{\beta}'\mathbf{Y}_t \\ \phi &= 1 + \boldsymbol{\beta}'\boldsymbol{\alpha} = 1 + (\alpha_1 - \beta\alpha_2) \\ v_t &= \boldsymbol{\beta}'\boldsymbol{\varepsilon}_t = u_{1t} - \beta u_{2t} \end{aligned}$$

The AR(1) model for u_t is stable as long as

$$|\phi| = |1 + (\alpha_1 - \beta\alpha_2)| < 1$$

For example, suppose $\beta = 1$. Then the stability condition is

$$|\phi| = |1 + (\alpha_1 - \alpha_2)| < 1$$

which is satisfied if

$$\alpha_1 - \alpha_2 < 0 \text{ and } \alpha_1 - \alpha_2 > -2.$$

Johansen's Methodology for Modeling Cointegration

The basic steps in Johansen's methodology are:

1. Specify and estimate a VAR(p) model for $\mathbf{Y}_t$.

2. Construct likelihood ratio tests for the rank of $\mathbf{\Pi}$ to determine the number of cointegrating vectors.

3. If necessary, impose normalization and identifying restrictions on the cointegrating vectors.

4. Given the normalized cointegrating vectors estimate the resulting cointegrated VECM by maximum likelihood.

Likelihood Ratio Tests for the Number of Cointegrating Vectors

The unrestricted cointegrated VECM is denoted $H(r)$. The $I(1)$ model $H(r)$ can be formulated as the condition that the rank of $\mathbf{\Pi}$ is less than or equal to r. This creates a nested set of models

$$
\begin{aligned}
H(0) &\subset \cdots \subset H(r) \subset \cdots \subset H(n) \\
H(0) &= \text{non-cointegrated VAR} \\
H(n) &= \text{stationary VAR(p)}
\end{aligned}
$$

This nested formulation is convenient for developing a sequential procedure to test for the number r of cointegrating relationships.

Remarks:

- Since the rank of the long-run impact matrix $\mathbf{\Pi}$ gives the number of cointegrating relationships in $\mathbf{Y}_t$, Johansen formulates likelihood ratio (LR) statistics for the number of cointegrating relationships as LR statistics for determining the rank of $\mathbf{\Pi}$.

- These LR tests are based on the estimated eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > \cdots > \hat{\lambda}_n$ of the matrix $\mathbf{\Pi}$. These eigenvalues also happen to equal the squared *canonical correlations* between $\Delta\mathbf{Y}_t$ and $\mathbf{Y}_{t-1}$ corrected for lagged $\Delta\mathbf{Y}_t$ and $\mathbf{D}_t$ and so lie between 0 and 1.

- Recall, the rank of $\mathbf{\Pi}$ is equal to the number of non-zero eigenvalues of $\mathbf{\Pi}$.

Johansen's Trace Statistic

Johansen's LR statistic tests the nested hypotheses

$$H_0(r) : r = r_0 \text{ vs. } H_1(r_0) : r > r_0$$

The LR statistic, called the *trace statistic*, is given by

$$LR_{trace}(r_0) = -T \sum_{i=r_0+1}^{n} \ln(1 - \hat{\lambda}_i)$$

- If $rank(\mathbf{\Pi}) = r_0$ then $\hat{\lambda}_{r_0+1}, \ldots, \hat{\lambda}_n$ should all be close to zero and $LR_{trace}(r_0)$ should be small since $\ln(1 - \hat{\lambda}_i) \approx 0$ for $i > r_0$.

- In contrast, if $rank(\mathbf{\Pi}) > r_0$ then some of $\hat{\lambda}_{r_0+1}, \ldots, \hat{\lambda}_n$ will be nonzero (but less than 1) and $LR_{trace}(r_0)$ should be large since $\ln(1 - \hat{\lambda}_i) << 0$ for some $i > r_0$.

Result: The asymptotic null distribution of $LR_{trace}(r_0)$ is not chi-square but instead is a multivariate version of the Dickey-Fuller unit root distribution which depends on the dimension $n-r_0$ and the specification of the deterministic terms. Critical values for this distribution are tabulated in Osterwald-Lenum (1992) for $n-r_0=1,\ldots,10$.

Sequential Procedure for Determining the Number of Cointegrating Vectors

1. First test $H_0(r_0 = 0)$ against $H_1(r_0 > 0)$. If this null is not rejected then it is concluded that there are no cointegrating vectors among the n variables in $\mathbf{Y}_t$.

2. If $H_0(r_0 = 0)$ is rejected then it is concluded that there is at least one cointegrating vector and proceed to test $H_0(r_0 = 1)$ against $H_1(r_0 > 1)$. If this null is not rejected then it is concluded that there is only one cointegrating vector.

3. If the $H_0(r_0 = 1)$ is rejected then it is concluded that there is at least two cointegrating vectors.

4. The sequential procedure is continued until the null is not rejected.

Johansen's Maximum Eigenvalue Statistic

Johansen also derives a LR statistic for the hypotheses

$$H_0(r_0) : r = r_0 \text{ vs. } H_1(r_0) : r_0 = r_0 + 1$$

The LR statistic, called the maximum eigenvalue statistic, is given by

$$LR_{\max}(r_0) = -T \ln(1 - \hat{\lambda}_{r_0+1})$$

As with the trace statistic, the asymptotic null distribution of $LR_{\max}(r_0)$ is not chi-square but instead is a complicated function of Brownian motion, which depends on

- the dimension $n - r_0$
- the specification of the deterministic terms.

Critical values for this distribution are tabulated in Osterwald-Lenum (1992) for $n - r_0 = 1, \ldots, 10$.

Specification of Deterministic Terms

Following Johansen (1995), the deterministic terms in are restricted to the form

$$\mathbf{\Phi}\mathbf{D}_t = \boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t$$

If the deterministic terms are unrestricted then the time series in $\mathbf{Y}_t$ may exhibit quadratic trends and there may be a linear trend term in the cointegrating relationships. Restricted versions of the trend parameters $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ limit the trending nature of the series in $\mathbf{Y}_t$. The trend behavior of $\mathbf{Y}_t$ can be classified into five cases:

1. Model $H_2(r)$: $\boldsymbol{\mu}_t = 0$ (no constant):

$$\begin{aligned} \Delta \mathbf{Y}_t &= \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{Y}_{t-1} \\ &+ \boldsymbol{\Gamma}_1 \Delta \mathbf{Y}_{t-1} + \cdots + \boldsymbol{\Gamma}_{p-1} \Delta \mathbf{Y}_{t-p+1} + \boldsymbol{\varepsilon}_t, \end{aligned}$$

   and all the series in $\mathbf{Y}_t$ are $I(1)$ without drift and the cointegrating relations $\boldsymbol{\beta}'\mathbf{Y}_t$ have mean zero.

2. Model $H_1^*(r)$: $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 = \boldsymbol{\alpha}\boldsymbol{\rho}_0$ (restricted constant):

$$\begin{aligned} \Delta \mathbf{Y}_t &= \boldsymbol{\alpha}(\boldsymbol{\beta}'\mathbf{Y}_{t-1} + \boldsymbol{\rho}_0) \\ &+ \boldsymbol{\Gamma}_1 \Delta \mathbf{Y}_{t-1} + \cdots + \boldsymbol{\Gamma}_{p-1} \Delta \mathbf{Y}_{t-p+1} + \boldsymbol{\varepsilon}_t, \end{aligned}$$

   the series in $\mathbf{Y}_t$ are $I(1)$ without drift and the cointegrating relations $\boldsymbol{\beta}'\mathbf{Y}_t$ have non-zero means $\boldsymbol{\rho}_0$.

3. Model $H_1(r)$: $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0$ (unrestricted constant):

$$\Delta \mathbf{Y}_t = \boldsymbol{\mu}_0 + \boldsymbol{\alpha}\boldsymbol{\beta}' \mathbf{Y}_{t-1}$$
$$+ \boldsymbol{\Gamma}_1 \Delta \mathbf{Y}_{t-1} + \cdots + \boldsymbol{\Gamma}_{p-1} \Delta \mathbf{Y}_{t-p+1} + \boldsymbol{\varepsilon}_t$$

the series in $\mathbf{Y}_t$ are $I(1)$ with drift vector $\boldsymbol{\mu}_0$ and the cointegrating relations $\boldsymbol{\beta}'\mathbf{Y}_t$ may have a non-zero mean.

4. Model $H^*(r)$: $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\alpha}\boldsymbol{\rho}_1 t$ (restricted trend). The restricted VECM is

$$\Delta \mathbf{Y}_t = \boldsymbol{\mu}_0 + \boldsymbol{\alpha}(\boldsymbol{\beta}' \mathbf{Y}_{t-1} + \boldsymbol{\rho}_1 t)$$
$$+ \boldsymbol{\Gamma}_1 \Delta \mathbf{Y}_{t-1} + \cdots + \boldsymbol{\Gamma}_{p-1} \Delta \mathbf{Y}_{t-p+1} + \boldsymbol{\varepsilon}_t$$

the series in $\mathbf{Y}_t$ are $I(1)$ with drift vector $\boldsymbol{\mu}_0$ and the cointegrating relations $\boldsymbol{\beta}'\mathbf{Y}_t$ have a linear trend term $\boldsymbol{\rho}_1 t$.

5. Model $H(r)$: $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t$ (unrestricted constant and trend). The unrestricted VECM is

$$\begin{aligned} \Delta \mathbf{Y}_t &= \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t + \boldsymbol{\alpha}\boldsymbol{\beta}' \mathbf{Y}_{t-1} \\ &\quad + \boldsymbol{\Gamma}_1 \Delta \mathbf{Y}_{t-1} + \cdots + \boldsymbol{\Gamma}_{p-1} \Delta \mathbf{Y}_{t-p+1} + \boldsymbol{\varepsilon}_t, \end{aligned}$$

the series in $\mathbf{Y}_t$ are $I(1)$ with a linear trend (quadratic trend in levels) and the cointegrating relations $\boldsymbol{\beta}'\mathbf{Y}_t$ have a linear trend.

Maximum Likelihood Estimation of the Cointegrated VECM

If it is found that $rank(\mathbf{\Pi}) = r$, $0 < r < n$, then the cointegrated VECM

$$\begin{aligned}\Delta \mathbf{Y}_t &= \mathbf{\Phi}\mathbf{D}_t + \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{Y}_{t-1} + \mathbf{\Gamma}_1\Delta\mathbf{Y}_{t-1} \\ &\quad + \cdots + \mathbf{\Gamma}_{p-1}\Delta\mathbf{Y}_{t-p+1} + \boldsymbol{\varepsilon}_t,\end{aligned}$$

becomes a reduced rank multivariate regression. Johansen derived the maximum likelihood estimation of the parametes under the reduced rank restriction $rank(\mathbf{\Pi}) = r$ (see Hamilton for details). He shows that

- $\hat{\boldsymbol{\beta}}_{mle} = (\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_r)$,where $\hat{\mathbf{v}}_i$ are the eigenvectors associated with the eigenvalues $\hat{\lambda}_i$,

- The MLEs of the remaining parameters are obtained by least squares estimation of

$$\begin{aligned}\Delta \mathbf{Y}_t &= \mathbf{\Phi}\mathbf{D}_t + \boldsymbol{\alpha}\hat{\boldsymbol{\beta}}'_{mle}\mathbf{Y}_{t-1} + \mathbf{\Gamma}_1\Delta\mathbf{Y}_{t-1} \\ &\quad + \cdots + \mathbf{\Gamma}_{p-1}\Delta\mathbf{Y}_{t-p+1} + \boldsymbol{\varepsilon}_t,\end{aligned}$$

Normalized Estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$

- The factorization

$$\hat{\mathbf{\Pi}}_{mle} = \hat{\boldsymbol{\alpha}}_{mle}\hat{\boldsymbol{\beta}}'_{mle}$$

  is not unique

- The columns of $\hat{\boldsymbol{\beta}}_{mle}$ may be interpreted as linear combinations of the underlying cointegrating relations.

- For interpretations, it is often convenient to normalize or identify the cointegrating vectors by choosing a specific coordinate system in which to express the variables.

Johansen's normalized MLE

- An arbitrary normalization, suggested by Johansen, is to solve for the triangular representation of the cointegrated system (default method in Eviews). The resulting normalized cointegrating vector is denoted $\hat{\boldsymbol{\beta}}_{c,mle}$. The normalization of the MLE for $\boldsymbol{\beta}$ to $\hat{\boldsymbol{\beta}}_{c,mle}$ will affect the MLE of $\boldsymbol{\alpha}$ but not the MLEs of the other parameters in the VECM.

- Let $\hat{\boldsymbol{\beta}}_{c,mle}$ denote the MLE of the normalized cointegrating matrix $\boldsymbol{\beta}_c$. Johansen (1995) showed that

$$T(\text{vec}(\hat{\boldsymbol{\beta}}_{c,mle}) - \text{vec}(\boldsymbol{\beta}_c))$$

is asymptotically (mixed) normally distributed

- $\hat{\boldsymbol{\beta}}_{c,mle}$ is super consistent

Testing Linear Restrictions on $\boldsymbol{\beta}$

The Johansen MLE procedure only produces an estimate of the basis for the space of cointegrating vectors. It is often of interest to test if some hypothesized cointegrating vector lies in the space spanned by the estimated basis:

$$
\begin{aligned}
H_0 &: \underset{(r\times n)}{\boldsymbol{\beta}}' = \begin{pmatrix} \boldsymbol{\beta}_0' \\ \boldsymbol{\phi}' \end{pmatrix} \\
\boldsymbol{\beta}_0' &= s \times n \text{ matrix of hypothesized cv's} \\
\boldsymbol{\phi}' &= (r-s) \times n \text{ matrix of remaining unspecified cv's}
\end{aligned}
$$

Result: Johansen (1995) showed that a likelihood ratio statistic can be computed, which is asymptotically distributed as a χ^2 with $s(n-r)$ degrees of freedom.

Cointegration and the BN Decomposition

- The Granger Representation Theorem (GRT) provides an explicit link between the VECM form of a cointegrated VAR and the Wold or moving average representation.

- The GRT also provides insight into the Beveridge-Nelson decomposition of a cointegrated time series.

Let y_t be cointegrated with r cointegrating vectors captured in the $r \times n$ matrix $\boldsymbol{\beta}'$ so that $\boldsymbol{\beta}'\mathbf{y}_t$ is $I(0)$. Suppose $\Delta\mathbf{y}_t$ has the Wold representation

$$
\begin{aligned}
\Delta y_t &= \mu + \Psi(L)u_t \\
\Psi(L) &= \sum_{k=0}^{\infty} \Psi_k L^k \text{ and } \Psi_0 = I_n
\end{aligned}
$$

Using $\Psi(L) = \Psi(1) + (1 - L)\tilde{\Psi}(L)$, The BN decomposition of y_t is given by

$$
\begin{aligned}
y_t &= y_0 + \mu t + \Psi(1) \sum_{k=1}^{t} u_t + \tilde{u}_t - \tilde{u}_0 \\
\widetilde{u_t} &= \tilde{\Psi}(L) u_t
\end{aligned}
$$

Multiply both sides by β' to give

$$
\beta' y_t = \beta' \mu t + \beta' \Psi(1) \sum_{k=1}^{t} u_t + \beta'(y_0 + \tilde{u}_t - \tilde{u}_0)
$$

Since $\beta' y_t$ is $I(0)$ we must have that

$$
\beta' \Psi(1) = 0
$$

$\Psi(1)$ is singular and has rank $n - r$

The singularity of $\Psi(1)$ implies that the long-run covariance of Δy_t

$$
\Psi(1) \Sigma \Psi(1)'
$$

is singular and has rank $n - r$.

Now suppose that y_t has the VECM representation

$$\begin{aligned} \Gamma(L)\Delta y_t &= c + \Pi y_{t-1} + u_t \\ \Pi &= \alpha\beta' \end{aligned}$$

where the $n \times r$ matrices α and β both have rank r. The Granger Representation Theorem (GRT) gives an explicit mapping from the BN decomposition to the parameters of the VECM. Define the $n \times (n-r)$ full rank matrices $\alpha_\perp$ and $\beta_\perp$ such that

- $\alpha'\alpha_\perp = 0, \beta'\beta_\perp = 0$
- $rank(\alpha, \alpha_\perp) = n, rank(\beta, \beta_\perp) = n$
- $(\alpha'_\perp \Gamma(1)\beta_\perp)^{-1}$ exists where $\Gamma(1) = I_n - \sum_{i=1}^{p-1} \Gamma_i$
- $\beta_\perp(\alpha'_\perp \Gamma(1)\beta_\perp)^{-1}\alpha'_\perp + \alpha'(\alpha'_\perp \Gamma(1)\beta_\perp)^{-1}\beta' = I_n$

Theorem (GRT). If $\det(A(z)) = 0$ implies that $|z| > 1$ or $z = 1$ and $rank(\Pi) = r < n$, then there exist $n \times r$ matrices α and β of rank r such that

$$\Pi = \alpha\beta'.$$

A necessary and sufficient for $\beta' y_t$ to be $I(0)$ is that

$$\alpha'_{\perp}\Gamma(1)\beta_{\perp}$$

has full rank. Then the BN decomposition of y_t has the representation

$$y_t = \mu t + \Psi(1)\sum_{k=1}^{t} u_t + y_0 + \tilde{u}_t - \tilde{u}_0$$

where

$$\Psi(1) = \beta_{\perp}(\alpha'_{\perp}\Gamma\beta_{\perp})^{-1}\alpha'_{\perp}$$

y_t is a cointegrated process with cointegrating vectors given by the rows of β'.

The main part of the GRT is the explicit representation for $\Psi(1)$:

$$\Psi(1) = \beta_\perp(\alpha'_\perp \Gamma(1)\beta_\perp)^{-1}\alpha'_\perp.$$

Notice that

$$\begin{aligned} \beta'\Psi(1) &= \beta'\beta_\perp(\alpha'_\perp \Gamma(1)\beta_\perp)^{-1}\alpha'_\perp = 0 \\ \Psi(1)\alpha &= \beta_\perp(\alpha'_\perp \Gamma(1)\beta_\perp)^{-1}\alpha'_\perp\alpha = 0 \end{aligned}$$

The common trends in y_t are extracted using

$$\begin{aligned} TS_t &= \Psi(1)\sum_{k=1}^{t} u_t \\ &= \beta_\perp(\alpha'_\perp \Gamma(1)\beta_\perp)^{-1}\alpha'_\perp \sum_{k=1}^{t} u_t \\ &= \xi\alpha'_\perp \sum_{k=1}^{t} u_t \end{aligned}$$

where $\xi = \beta_\perp(\alpha'_\perp \Gamma(1)\beta_\perp)^{-1}$. Hence the common trends are the linear combinations

$$\alpha'_\perp \sum_{k=1}^{t} u_t$$

The GRT in a Cointegrated Bivariate VAR(1) Model

To illustrate the GRT, consider the simple cointegrated bivariate VECM

$$\Delta y_t = \alpha\beta' y_{t-1} + u_t.$$

where $\alpha = (-0.1, 0.1)'$ and $\beta = (1, -1)'$. Here there is one cointegrating vector and one common trend. It may be easily deduced that

$$\begin{aligned}
\Gamma(1) &= I_2, \\
\alpha_\perp &= (1, 1)', \\
\beta_\perp &= (1, 1)', \\
\alpha'_\perp \Gamma(1)\beta_\perp &= 2.
\end{aligned}$$

The common trend is then given by

$$TS_t = \alpha'_\perp \sum_{k=1}^{t} u_t = (1, 1)\begin{pmatrix} \sum_{k=1}^{t} u_{1t} \\ \sum_{k=1}^{t} u_{2t} \end{pmatrix} = \sum_{k=1}^{t} u_{1t} + \sum_{k=1}^{t} u_{2t}$$

and the loadings on the common trend are

$$\xi = \beta_\perp(\alpha'_\perp \Gamma(1)\beta_\perp)^{-1} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

Now suppose that $\alpha = (-0.1, 0)'$ so that y_{2t} is weakly and strongly exogenous. The VECM has the simplified form

$$\begin{aligned}
\Delta y_{1t} &= -0.1 \cdot \beta' y_{t-1} + u_{1t} \\
\Delta y_{2t} &= u_{2t}
\end{aligned}$$

Then

$$\begin{aligned}
\Gamma(1) &= I_2, \\
\alpha_\perp &= (0, 1)', \\
\beta_\perp &= (1, 1)', \\
\alpha'_\perp \Gamma(1) \beta_\perp &= 1.
\end{aligned}$$

Interestingly, the common trend is simply y_{2t} :

$$TS_t = \alpha'_\perp \sum_{k=1}^{t} u_t = (0, 1) \begin{pmatrix} \sum_{k=1}^{t} u_{1t} \\ \sum_{k=1}^{t} u_{2t} \end{pmatrix} = \sum_{k=1}^{t} u_{2t} = y_{2t}$$

The loadings on the common trend are

$$\xi = \beta_\perp (\alpha'_\perp \Gamma(1) \beta_\perp)^{-1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$