Macroeconometrics

Lecture 7 Vector Autoregressions

Tomasz Woźniak

Department of Economics
University of Melbourne

Vector autoregressions

Properties

Granger causality

Maximum likelihood estimation

Useful reading:

Kilian & Lütkepohl (2017) Chapter 2: Vector Autoregressive Models, Structural Vector Autoregressive Analysis

Objectives.

- To introduce vector autoregressions and their properties
- To derive unconditional moments of the process
- To derive maximum likelihood estimator of the parameters and to understand its properties

Learning outcomes.

- Presenting the dynamic properties of the process
- Applying algebraic transformations to derive interpretable matrix-valued results
- To apply derivatives wrt matrices in optimisation problems

Macroeconomic Forecasting with Fat Data

Vector autoregressions

Vector autoregressions

VAR(p) model.

$$
\begin{aligned}
y_t &= \mu_0 + A_1 y_{t-1} + \cdots + A_p y_{t-p} + \epsilon_t \\
\epsilon_t | Y_{t-1} &\sim iid\left(\mathbf{0}_N, \Sigma\right)
\end{aligned}
$$

for $t = 1, \ldots, T$, where :

y_t – $N \times 1$ vector of observations at time t

μ_0 – $N \times 1$ vector of constant terms

A_i – $N \times N$ matrix of autoregressive slope parameters

ϵ_t – $N \times 1$ vector of error terms – a multivariate white noise process

Y_{t-1} – information set collecting observations on y up to time $t-1$

Σ –$N \times N$ covariance matrix of the error term

Vector autoregressions

A bivariate VAR(p) model.

$$\begin{bmatrix} y_{1.t} \\ y_{2.t} \end{bmatrix} = \begin{bmatrix} \mu_{0.1} \\ \mu_{0.2} \end{bmatrix} + \begin{bmatrix} A_{1.11} & A_{1.12} \\ A_{1.21} & A_{1.22} \end{bmatrix} \begin{bmatrix} y_{1.t-1} \\ y_{2.t-1} \end{bmatrix} + \cdots + \begin{bmatrix} A_{p.11} & A_{p.12} \\ A_{p.21} & A_{p.22} \end{bmatrix} \begin{bmatrix} y_{1.t-p} \\ y_{2.t-p} \end{bmatrix} + \begin{bmatrix} \epsilon_{1.t} \\ \epsilon_{2.t} \end{bmatrix}$$

$$= \begin{bmatrix} \mu_{0.1} + A_{1.11} y_{1.t-1} + A_{1.12} y_{2.t-1} + \cdots + A_{p.11} y_{1.t-p} + A_{p.12} y_{2.t-p} + \epsilon_{1.t} \\ \mu_{0.2} + A_{1.21} y_{1.t-1} + A_{1.22} y_{2.t-1} + \cdots + A_{p.21} y_{1.t-p} + A_{p.22} y_{2.t-p} + \epsilon_{2.t} \end{bmatrix}$$

Vector autoregressions properties

Vector autoregressions: notation using lag polynomial

$$
\begin{aligned}
y_t &= \mu_0 + A_1 y_{t-1} + \cdots + A_p y_{t-p} + \epsilon_t \\
y_t - A_1 y_{t-1} - \cdots - A_p y_{t-p} &= \mu_0 + \epsilon_t \\
\left(I_N - A_1 L - \cdots - A_p L^p\right) y_t &= \mu_0 + \epsilon_t \\
A\left(L\right) y_t &= \mu_0 + \epsilon_t
\end{aligned}
$$

Stationarity condition for the VAR(p) process.

$$\det\left(A\left(z\right)\right) = 0 \qquad |z| > 1 \quad \forall z \in \mathbb{C}$$

The VAR(p) process is stationary if the roots of the characteristic polynomial are outside of the complex unit circle.

$\det(X)$ is a determinant of matrix X.

Vector autoregressions: VAR(1) representation

$$Y_t = \mathbf{m} + \mathbf{A} Y_{t-1} + E_t$$

$$\underset{(Np\times 1)}{Y_t} = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix} \qquad \underset{(Np\times 1)}{\mathbf{m}} = \begin{bmatrix} \mu_0 \\ \mathbf{0}_N \\ \vdots \\ \mathbf{0}_N \end{bmatrix} \qquad \underset{(Np\times 1)}{E_t} = \begin{bmatrix} \epsilon_t \\ \mathbf{0}_N \\ \vdots \\ \mathbf{0}_N \end{bmatrix}$$

$$\underset{(Np\times Np)}{\mathbf{A}} = \begin{bmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I_N & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_N & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & I_N & \mathbf{0} \end{bmatrix}$$

Vector autoregressions: VAR(1) representation

$$Y_t = \mathbf{m} + \mathbf{A} Y_{t-1} + E_t$$

Transform the model back to VAR(p) representation using matrix

$$\underset{(N \times Np)}{J} = \begin{bmatrix} I_n & \mathbf{0}_{N \times N(p-1)} \end{bmatrix}$$

$$JY_t = y_t \qquad J\mathbf{m} = \mu_0 \qquad JE_t = \epsilon_t$$

$$J\mathbf{A} Y_{t-1} = A_1 y_{t-1} + \cdots + A_p y_{t-p}$$

Vector autoregressions: unconditional moments

Stationarity condition.

$$\det(I_{Np} - \mathbf{A}z) = 0 \qquad \text{for } |z| > 1 \quad \forall z \in \mathbb{C}$$

Assume stationarity

Unconditional mean.

$$\begin{aligned}
\mathbb{E}[Y_t] &= \mathbf{m} + \mathbf{A}\mathbb{E}[Y_{t-1}] + \mathbb{E}[E_t] \\
&\overset{LIE}{=} \mathbf{m} + \mathbf{A}\mathbb{E}[Y_{t-1}] \\
(I_{Np} - \mathbf{A})\mathbb{E}[Y_t] &= \mathbf{m} \qquad \text{– by stationarity } \mathbb{E}[Y_t] = \mathbb{E}[Y_{t-1}]
\end{aligned}$$

$$\mathbb{E}[Y_t] = (I_{Np} - \mathbf{A})^{-1}\mathbf{m} = \mathbf{M}$$

$$\mu = \mathbb{E}[y_t] = \mathbb{E}[JY_t] = J\mathbb{E}[Y_t] = J\mathbf{M} = J(I_{np} - \mathbf{A})^{-1}\mathbf{m}$$

Vector autoregressions: unconditional moments

Assume stationarity

Autocovariances.

$$\underset{N\times N}{\gamma_s} = \mathbb{E}\left[(y_t - \mu)(y_{t-s} - \mu)'\right]$$

$$\gamma_s \neq \gamma_s' \qquad \text{– an asymmetric matrix}$$
$$\gamma_s = \gamma_{-s}'$$

$$\underset{Np\times Np}{\Gamma_s} = \mathbb{E}\left[\left(Y_t - \mathbf{M}\right)\left(Y_{t-s} - \mathbf{M}\right)'\right]$$

$$\Gamma_s \neq \Gamma_s' \qquad \text{– an asymmetric matrix}$$
$$\Gamma_s = \Gamma_{-s}'$$

$$\gamma_s = J\Gamma_s J'$$

Vector autoregressions: unconditional moments

Autocovariances.

Step 1: Write the intercept as $\mathbf{m} = (I_{np} - \mathbf{A})\,\mathbf{M}$ to rewrite

$$(Y_t - \mathbf{M}) = \mathbf{A}\,(Y_{t-1} - \mathbf{M}) + E_t$$

Step 2: Multiply by $(Y_{t-s} - \mathbf{M})'$

$$(Y_t - \mathbf{M})\,(Y_{t-s} - \mathbf{M})' = \mathbf{A}\,(Y_{t-1} - \mathbf{M})\,(Y_{t-s} - \mathbf{M})' + E_t\,(Y_{t-s} - \mathbf{M})'$$

Step 3: Take the expectations

$$\boldsymbol{\Gamma}_s = \mathbf{A}\boldsymbol{\Gamma}_{s-1} + \mathbb{E}\left[E_t\,(Y_{t-s} - \mathbf{M})'\right]$$

Step 4: Solve the system of equations given $\boldsymbol{\Sigma}_E = \mathbb{E}[E_t E_t']$

$$\begin{aligned} \boldsymbol{\Gamma}_0 &= \mathbf{A}\boldsymbol{\Gamma}_1' + \boldsymbol{\Sigma}_E \\ \boldsymbol{\Gamma}_1 &= \mathbf{A}\boldsymbol{\Gamma}_0 \end{aligned}$$

$$\boldsymbol{\Gamma}_s = \mathbf{A}\boldsymbol{\Gamma}_{s-1} \qquad \text{for } s \geq 1$$

Vector autoregressions: unconditional moments

Step 4: Solve the system of equations:

$$\Gamma_0 = \mathbf{A}\Gamma_1' + \Sigma_E$$
$$\Gamma_1 = \mathbf{A}\Gamma_0$$

$$\begin{aligned}
\Gamma_0 &= \mathbf{A}\Gamma_0\mathbf{A}' + \Sigma_E \\
\text{vec}(\Gamma_0) &= \text{vec}(\mathbf{A}\Gamma_0\mathbf{A}') + \text{vec}(\Sigma_E) \\
\text{vec}(\Gamma_0) &= (\mathbf{A} \otimes \mathbf{A})\text{vec}(\Gamma_0) + \text{vec}(\Sigma_E) \\
(I_{N^2p^2} - \mathbf{A} \otimes \mathbf{A})\text{vec}(\Gamma_0) &= \text{vec}(\Sigma_E) \\
\text{vec}(\Gamma_0) &= (I_{N^2p^2} - \mathbf{A} \otimes \mathbf{A})^{-1}\text{vec}(\Sigma_E)
\end{aligned}$$

$$\Gamma_s = \mathbf{A}\Gamma_{s-1} \qquad \text{for } s \geq 1.$$

Useful matrix transformations.

$(AB)' = B'A'$

$\text{vec}(ABC) = (C' \otimes A)\text{vec}(B)$

$\text{vec}\underset{(m\times n)}{(X)}$ – is an $mn \times 1$ vector stacking columns of X one by one

$\otimes$ – is a Kronecker product

Vector autoregressions: Vector Moving Average form

VMA(∞) representation of the VAR(p) process in VAR(1) form.

$$
\begin{aligned}
Y_t &= \mathbf{m} + \mathbf{A} Y_{t-1} + E_t \\
(I_{Np} - \mathbf{A}L) Y_t &= \mathbf{m} + E_t \\
Y_t &= (I_{Np} - \mathbf{A}L)^{-1}\mathbf{m} + (I_{Np} - \mathbf{A}L)^{-1} E_t \\
&= \mathbf{M} + E_t + \mathbf{A}E_{t-1} + \mathbf{A}^2 E_{t-2} + \dots
\end{aligned}
$$

VMA(∞) representation of the VAR(p) process.

A stationary VAR(p) process has a VMA(∞) representation:

$$
\begin{aligned}
y_t = JY_t &= J\mathbf{M} + \sum_{i=0}^{\infty} J\mathbf{A}^i E_{t-i} \\
&= J\mathbf{M} + \sum_{i=0}^{\infty} J\mathbf{A}^i J' J E_{t-i} \\
&= \mu + \epsilon_t + \Phi_1 \epsilon_{t-1} + \Phi_2 \epsilon_{t-2} + \dots
\end{aligned}
$$

where $\mathbf{A}^2 = \mathbf{A}\mathbf{A}$, and $\Phi_i = J\mathbf{A}^i J'$, since $E_t = J'JE_t$ and $JE_t = \epsilon_t$

Granger causality

Granger causality

Partition an $N \times 1$ vector y_t into sub-vectors $y_{1.t}$, $y_{2.t}$ and $y_{3.t}$, of dimensions N_1, N_2, N_3 respectively, and $N_1 + N_2 + N_3 = N$

$$y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix}$$

Notation.

Y_t – information set with observations on y_t up to period t

$Y_{-1.t}$ – constrained information set with observations on $y_{2.t}$ and $y_{3.t}$ up to period t

$\mathbb{E}[y_{2.t+1}|\mathcal{I}_t]$ – a 1-step ahead predictor of $y_{2.t}$ based on set $\mathcal{I}_t$

$\mathbf{e}_{2.t+1}$ – the corresponding forecast error

$\mathbb{V}\text{ar}[\mathbf{e}_{2.t+1}|\mathcal{I}_t]$ – the corresponding forecast error variance

Granger causality

Definition.

Process $y_{1.t}$ is said not to Granger cause $y_{2.t}$ if:

$$\mathbb{V}\text{ar}[\mathbf{e}_{2.t+1}|Y_t] = \mathbb{V}\text{ar}[\mathbf{e}_{2.t+1}|Y_{-1.t}]$$

Equivalent definition for linear models.

$$\mathbb{E}\left[y_{2.t+1}|Y_t\right] = \mathbb{E}\left[y_{2.t+1}|Y_{-1.t}\right]$$

References.

Granger (1969, ECTA), Sims (1972, AER)
Lütkepohl (2005, Chapter 2) New Introduction to Multiple Time Series Analysis

Granger causality and VARs

Consider a VAR(p) process:

$$\begin{aligned}(I_n - A_1 L - \cdots - A_p L^p) y_t &= \mu_0 + u_t \\ A(L) y_t &= \mu_0 + u_t\end{aligned}$$

Rewrite the VAR(p) model respecting the partitioning of y_t:

$$\begin{bmatrix} A_{11}(L) & A_{12}(L) & A_{13}(L) \\ A_{21}(L) & A_{22}(L) & A_{23}(L) \\ A_{31}(L) & A_{32}(L) & A_{33}(L) \end{bmatrix} \begin{bmatrix} y_{1.t} \\ y_{2.t} \\ y_{3.t} \end{bmatrix} = \begin{bmatrix} \mu_{0.1} \\ \mu_{0.2} \\ \mu_{0.3} \end{bmatrix} + \begin{bmatrix} \epsilon_{1.t} \\ \epsilon_{2.t} \\ \epsilon_{3.t} \end{bmatrix},$$

where:

$$\underset{(N_i \times N_i)}{A_{ii}(L)} = I_{N_i} - A_{ii.1} L - \cdots - A_{ii.p} L^p \qquad \text{for } i = 1, 2, 3$$

$$\underset{(N_i \times N_j)}{A_{ij}(L)} = -A_{ij.1} L - \cdots - A_{ij.p} L^p \qquad \text{for } i \neq j$$

Granger causality in VAR models

Granger causality.
The process y_{1t} does not Granger cause y_{2t} iff:

$$A_{21}(z) = 0 \qquad \forall z \in \mathbb{C}$$

which results in the restrictions on the parameters:

$$A_{21.1} = \cdots = A_{21.p} = 0$$

Useful distribution

Matrix-variate normal distribution

A $K \times N$ matrix A is said to follow a matrix-variate normal distribution:

$$A \sim \mathcal{MN}_{K \times N}(M, Q, P),$$

where M is a $K \times N$ matrix and

Q $N \times N$ row-specific covariance matrix

P $K \times K$ column-specific covariance matrix

if $\text{vec}(A)$ is multivariate normal:

$$\text{vec}(A) \sim \mathcal{N}_{KN}(\text{vec}(M), Q \otimes P)$$

Density function.

$$\mathcal{MN}_{K \times N}(M, Q, P) = c_{mn}^{-1} \exp\left\{-\frac{1}{2}\text{tr}\left[Q^{-1}(A - M)' P^{-1}(A - M)\right]\right\}$$

$$c_{mn} = (2\pi)^{\frac{KN}{2}} \det(Q)^{\frac{K}{2}} \det(P)^{\frac{N}{2}}$$

Maximum likelihood estimation

Estimation: matrix notation

VAR(p) model.

$$
\begin{aligned}
y_t &= \mu_0 + A_1 y_{t-1} + \cdots + A_p y_{t-p} + \epsilon_t \\
\epsilon_t | Y_{t-1} &\sim iid\mathcal{N}_N\left(\mathbf{0}_N, \Sigma\right)
\end{aligned}
$$

Matrix notation (multivariate linear regression form).

$$
\begin{aligned}
Y &= XA + E \\
E|X &\sim \mathcal{MN}_{T\times N}\left(\mathbf{0}, \Sigma, I_T\right)
\end{aligned}
$$

$$
\underset{(K\times N)}{A} = \begin{bmatrix} \mu_0' \\ A_1' \\ \vdots \\ A_p' \end{bmatrix} \quad \underset{(T\times N)}{Y} = \begin{bmatrix} y_1' \\ y_2' \\ \vdots \\ y_T' \end{bmatrix} \quad \underset{(K\times 1)}{x_t} = \begin{bmatrix} 1 \\ y_{t-1} \\ \vdots \\ y_{t-p} \end{bmatrix} \quad \underset{(T\times K)}{X} = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_T' \end{bmatrix} \quad \underset{(T\times N)}{E} = \begin{bmatrix} \epsilon_1' \\ \epsilon_2' \\ \vdots \\ \epsilon_T' \end{bmatrix}
$$

where $K = 1 + pN$

Likelihood Function

$$
\begin{aligned}
L\left(A, \boldsymbol{\Sigma} \mid Y, X\right) &= (2\pi)^{-\frac{NT}{2}} \det(\boldsymbol{\Sigma})^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T} \epsilon_t' \boldsymbol{\Sigma}^{-1} \epsilon_t\right\} \\
&= (2\pi)^{-\frac{NT}{2}} \det(\boldsymbol{\Sigma})^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T} \left(y_t - A' x_t\right)' \boldsymbol{\Sigma}^{-1} \left(y_t - A' x_t\right)\right\} \\
&= (2\pi)^{-\frac{NT}{2}} \det(\boldsymbol{\Sigma})^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\text{vec}\left((Y - XA)'\right)' \left(I_T \otimes \boldsymbol{\Sigma}^{-1}\right) \text{vec}\left((Y - XA)'\right)\right\} \\
&= (2\pi)^{-\frac{NT}{2}} \det(\boldsymbol{\Sigma})^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\text{tr}\left[\boldsymbol{\Sigma}^{-1}(Y - XA)' I_T (Y - XA)\right]\right\} \\
&= (2\pi)^{-\frac{NT}{2}} \det(\boldsymbol{\Sigma})^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\text{tr}\left[\boldsymbol{\Sigma}^{-1}(Y - XA)'(Y - XA)\right]\right\}
\end{aligned}
$$

The trace.

$\text{tr}(X) = \sum_{n=1}^{N} X_{nn}$ for a square $N \times N$ matrix X

$\text{tr}\left[ABCD\right] = \text{vec}\left(D'\right)' \left(C' \otimes A\right) \text{vec}\left(B\right)$

Maximum likelihood estimation

$$\hat{A}, \hat{\Sigma} = \operatorname{argmax}\, l\left(A, \Sigma | Y, X\right)$$

$$\begin{aligned} l\left(A, \Sigma | Y, X\right) \propto \frac{T}{2} \log\det\left(\Sigma^{-1}\right) - \frac{1}{2}\mathrm{tr}\left[\Sigma^{-1}(Y - XA)'(Y - XA)\right] \\ = \frac{T}{2} \log\det\left(\Sigma^{-1}\right) - \frac{1}{2}\mathrm{tr}\left[\Sigma^{-1}\left(Y'Y - 2A'X'Y + A'X'XA\right)\right] \end{aligned}$$

To derive the MLE use calculus

Useful formulae.

$$\det\left(X^{-1}\right) = \det\left(X\right)^{-1} \qquad \frac{\partial \det(X)}{\partial X} = \det(X)(X')^{-1}$$

$$\frac{\partial \mathrm{tr}(X)}{\partial X} = I_N \qquad \frac{\partial \mathrm{tr}(AXB)}{\partial X} = A'B'$$

$$\frac{\partial \mathrm{tr}(AX'B)}{\partial X} = BA \qquad \frac{\partial \mathrm{tr}(AX'BX)}{\partial X} = (B + B')XA$$

Maximum likelihood estimation

$$\frac{\partial l\left(A, \Sigma | Y, X\right)}{\partial A} = -\frac{1}{2}\left[-2X'Y\Sigma^{-1} + 2X'XA\Sigma^{-1}\right]$$
$$= X'Y\Sigma^{-1} - X'XA\Sigma^{-1}$$

$$\frac{\partial l\left(A, \Sigma | Y, X\right)}{\partial \hat{A}} = \mathbf{0}_{K\times N}$$
$$X'Y\hat{\Sigma}^{-1} - X'X\hat{A}\hat{\Sigma}^{-1} = \mathbf{0}_{K\times N}$$
$$X'X\hat{A}\hat{\Sigma}^{-1} = X'Y\hat{\Sigma}^{-1} \quad \Big/ \times \hat{\Sigma}$$
$$(X'X)^{-1} \times \Big/ \quad X'X\hat{A} = X'Y$$

$$\hat{A} = (X'X)^{-1}X'Y$$

Maximum likelihood estimation

$$
\begin{aligned}
\frac{\partial l\left(A, \Sigma | Y, X\right)}{\partial \Sigma^{-1}} &= \frac{T}{2} \frac{1}{\det\left(\Sigma^{-1}\right)} \det\left(\Sigma^{-1}\right)\left(\Sigma^{-1}\right)^{-1} - \frac{1}{2}(Y - XA)'(Y - XA) \\
&= \frac{T}{2}\Sigma - \frac{1}{2}(Y - XA)'(Y - XA)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial l\left(A, \Sigma | Y, X\right)}{\partial \hat{\Sigma}^{-1}} &= \mathbf{0}_{N\times N} \\
\mathbf{0}_{N\times N} &= \frac{T}{2}\hat{\Sigma} - \frac{1}{2}(Y - X\hat{A})'(Y - X\hat{A}) \\
\frac{T}{2}\hat{\Sigma} &= \frac{1}{2}(Y - X\hat{A})'(Y - X\hat{A}) \\
\hat{\Sigma} &= \frac{1}{T}(Y - X\hat{A})'(Y - X\hat{A})
\end{aligned}
$$

Maximum likelihood estimator: asymptotic properties

Assumption.
A white noise process $\epsilon_t = (\epsilon_{1.t}, \ldots, \epsilon_{N.t})$ is called a standard white noise if ϵ_t are continuous random vectors satisfying $\mathbb{E}[\epsilon_t] = 0$, $\Sigma = \mathbb{E}[\epsilon_t \epsilon_t']$ is nonsingular, ϵ_t and ϵ_{t-s} are independent for $s \neq 0$, and, for some positive finite constant c:

$$\mathbb{E}[|\epsilon_{i.t}\epsilon_{j.t}\epsilon_{k.t}\epsilon_{m.t}|] \leq c \qquad \text{for } i, j, k, m = 1, \ldots, N, \text{ and all } t$$

The assumption states that all fourth moments exist and are bounded.

Maximum likelihood estimator: asymptotic properties

Asymptotic properties of the MLE.
Based on Lütkepohl (2005, p.74)

Let y_t be a stationary N-dimensional VAR(p) process with standard white noise residuals. Then the MLE is

consistent:

$$\text{plim } \hat{A} = A$$

asymptotically normally distributed:

$$\sqrt{T}\left(\hat{A} - A\right) \xrightarrow{d} \mathcal{MN}_{K\times N}\left(\mathbf{0}_{K\times N}, \boldsymbol{\Sigma}, T(X'X)^{-1}\right)$$

Maximum likelihood estimator: asymptotic properties

Asymptotic properties of the MLE.

Based on Harris, Hurn, Martin (2012, Chapter 16)

Let y_t be a unit-root nonstationary N-dimensional VAR(p) process with standard white noise residuals $y_t = y_{t-1} + \epsilon_t$
Then the MLE is

consistent:

$$\text{plim } \hat{A} = A$$

asymptotically non-normally distributed:

$$T\left(\hat{A} - A\right) \xrightarrow{d} \left[\int_0^1 B_K(s)B_K(s)'ds\right]^{-1} \int_0^1 B(s)_K B_N(s)'ds$$

$B_N(s)$ – denotes an N-dimensional Brownian motion

Four representations.

lag polynomial – stationarity conditions

VAR(1) – unconditional moments

VMA(∞) – effects of the shocks

multivariate regression – estimation