Vector processes. Basics

Extension of basic definitions and results to vector processes. Fairly easy, though in some cases caution is required.

Vector stochastic processes. Given the probability space $\mathcal{S} = (\Omega, \mathcal{F}, P)$, an n-dimensional stochastic variable is a function

$$y: \; \Omega \longrightarrow \mathbb{R}^n$$

which is $\mathcal{F}$-measurable. An n-dimensional stochastic process on $\mathcal{S}$ is a function

$$t \longrightarrow x_t = (x_{1t} \; x_{2t} \; \cdots \; x_{nt})'$$

where x_t is an n-dimensional stochastic variable on $\mathcal{S}$.

Vector processes. Stationarity

The n-dimensional process x_t is weakly stationary if

1. $\mathsf{E}(x_t)$ does not depend on t. Thus the notation $\mu = \mathsf{E}(x_t)$.
2. The second moments

$$\mathsf{E}(x_t x'_{t-k}) = \mathsf{E}\left[\begin{pmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{nt} \end{pmatrix} \begin{pmatrix} x_{1t-k} & x_{2t-k} & \cdots & x_{nt-k} \end{pmatrix}\right]$$

do not depend on t. (1) and (2) imply that the covariance matrix between x_t and x_{t-k} is independent of t. Hence the notation $\Gamma_k = \mathsf{E}\left[(x_t - \mu)(x_{t-k} - \mu)'\right]$.

Stationarity implies

$$\mathsf{E}(x_t x'_{t-k}) = \mathsf{E}(x_{t+k} x'_t) = \mathsf{E}(x_t x'_{t+k})' = \mathsf{E}(x_t x_{t-(-k)})'$$

That is

$$\Gamma_k = \Gamma'_{-k}$$

Vector processes. Stationarity

Important: Stationarity of each of the components of x_t does not imply stationarity of the vector x_t.

Example. The process x_t is the first number drawn for Rome in the weekly Italian Lotto, while y_t is the corresponding number for Paris in the French Lotto. The vector $(x_t \; y_t)'$ seems stationary. However, suppose that by tradition in the first week of the year Italian and French Lottos draw their numbers together in a small village at the border between the two countries. The correlation between x_t and y_t would be zero for all weeks of the year, except for the first, when it would be one.

Stationarity in the vector case requires that the components of the vector are stationary and co-stationary.

Vector processes. White noise

White noise. The n-dimensional stationary process x_t is a white noise if $\mathsf{E}(x_t) = 0$ and

$$\Gamma_k^x = 0 \qquad \text{for all } k \neq 0$$

Important: A vector whose components are white noise is not necessarily a white noise.

Example. Let u_t be a scalar white noise. Consider $x_t = (u_t \; u_{t-1})'$. We have

$$\Gamma_0 = \begin{pmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_u^2 \end{pmatrix}, \qquad \Gamma_1 = \begin{pmatrix} 0 & 0 \\ \sigma_u^2 & 0 \end{pmatrix}$$

Vector processes. White noise

Consider, for $n = 2$, $x_t = (x_{1t}\ x_{2t})'$, the covariance matrix

$$\Gamma_0^x = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

The definition of white noise has no implication on Γ_0^x. We are not requiring that $\sigma_{12} = 0$.

The definition of white noise does not imply that Γ_0^x has maximum rank either. This however will be assumed, if x_t is a white noise then Γ_0 is non-singular.

Vector processes. Moving averages

Given the white noise $u_t = (u_{1t}\ u_{2t}\ \cdots\ u_{nt})'$, a moving average of u_t is defined as

$$x_t = \sum_{k=-m}^{m} A_k u_{t-k} = A_m u_{t-m} + \cdots + A_1 u_{t-1} + A_0 u_t + A_{-1} u_{t+1} + \cdots + A_{-m} u_{t+m}$$

where A_k is an $n \times n$ matrix.

Therefore, each component of x_t depends on all the components of u_t.

Example:

$$x_t = A_0 u_t + A_1 u_{t-1}, \quad A_0 = \begin{pmatrix} 1 & 0.5 \\ -1 & 0.7 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 1 \\ 0.8 & 0 \end{pmatrix}$$

that is

$$\begin{aligned} x_{1t} &= u_{1t} + 0.5 u_{2t} + u_{2,t-1} \\ x_{2t} &= -u_{1t} + 0.7 u_{2t} + 0.8 u_{1,t-1} \end{aligned}$$

Vector processes. Moving averages

We assume that A_0 is non singular. As a consequence the moving average can be rewritten as

$$
\begin{aligned}
x_t &= A_m u_{t-m} + \cdots + A_1 u_{t-1} + A_0 u_t + A_{-1} u_{t+1} + \cdots + A_{-m} u_{t+m} \\
&= B_m v_{t-m} + \cdots + B_1 v_{t-1} + v_t + B_{-1} v_{t+1} + \cdots + B_{-m} v_{t+m}
\end{aligned}
$$

where $v_t = A_0 u_t$, $B_k = A_k A_0^{-1}$.

Prove that $A_0 u_t$ is a white noise (assuming that u_t is a white noise).

Vector processes. Moving averages

Given the moving average

$$x_t = A_m u_{t-m} + \cdots + A_1 u_{t-1} + u_t + A_{-1} u_{t+1} + \cdots + A_{-m} u_{t+m}$$

the covariances can be computed just in the same way as in the scalar case. For example, for $k = 1$

$$x_{t-1} = A_m u_{t-m-1} + \cdots + A_1 u_{t-2} + u_{t-1} + A_{-1} u_t + \cdots + A_{-m} u_{t+m-1}$$

so that

$$\Gamma_1 = A_m \Sigma_u A'_{m-1} + \cdots + A_2 \Sigma_u A'_1 + A_1 \Sigma_u + \Sigma_u A'_{-1} + A_{-1} \Sigma_u A'_{-2} + A_{-m+1} \Sigma_u A'_{-m}$$

Vector processes. Moving averages

Infinite moving averages:

$$x_t = \sum_{k=-\infty}^{\infty} A_k u_{t-k}$$

that is

$$x_{jt} = \sum_{k=-\infty}^{\infty} \sum_{h=1}^{n} a_{jh,k} u_{h,t-k}$$

for $j = 1, \ldots, n$.

The condition for a finite variance is

$$\sum_{k=-\infty}^{\infty} \sum_{h=1}^{n} \sum_{j=1}^{n} a_{jh,k}^2 < \infty$$

Vector processes. Autoregressive equations

A motivation for infinite moving averages is hardly necessary:

$$\begin{aligned} x_{1t} &= a_{11}x_{1,t-1} + a_{12}x_{2,t-1} + u_{1t} \\ x_{2t} &= a_{21}x_{1,t-1} + a_{22}x_{2,t-1} + u_{2t} \end{aligned}$$

that is

$$x_t = Ax_{t-1} + u_t$$

We can reproduce the iteration procedure

$$\begin{aligned} x_t &= u_t + A[u_{t-1} + Ax_{t-2}] = u_t + Au_{t-1} + A^2x_{t-2} = \cdots \\ &= u_t + Au_{t-1} + A^2u_{t-2} + \cdots \end{aligned}$$

But what is the condition for convergence of this series?

Eigenvalues and eigenvectors

Given the square $n \times n$ matrix A, a vector $v \neq 0$ is called an eigenvector of A if

$$Av = \lambda v$$

where λ is a real number, i.e. if v and the transformed vector Av are parallel. The number λ, which is uniquely determined by v is the eigenvalue associated with v.

Alternatively, λ is an eigenvalue of A if

$$\det(A - \lambda I) = 0 \qquad (*)$$

Since the matrix $A - \lambda I$ is singular, there exists $v \neq 0$ such that $(A - \lambda I)v = 0$, i.e. such that $Av = \lambda v$.
Equation $(*)$ is an algebraic equation of degree n, hence the matrix A has n eigenvalues and eigenvectors (we do not want here to consider the possibility of multiple roots of $(*)$).

Eigenvalues and eigenvectors

Thus, for $j = 1, \ldots, n$, we have

$$Av_j = \lambda_j v_j$$

This can be rewritten as

$$AV = V\Lambda, \qquad V = (v_1 \; v_2 \; \cdots \; v_n), \qquad \Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

That is

$$A = V\Lambda V^{-1}$$

Eigenvalues and eigenvectors

Rewrite

$$A = V \Lambda V^{-1} \qquad (*)$$

This is known as diagonalization of a matrix. Every matrix A (we assume here that the eigenvalues are distinct) is equivalent to a diagonal matrix, with the eigenvalues of A on the diagonal, "equivalence" being defined in $(*)$.

Note that

$$A^s = \underbrace{V \Lambda V^{-1} V \Lambda V^{-1} \cdots V \Lambda V^{-1}}_{s \text{ times}} = V \Lambda^s V^{-1}$$

where

$$\Lambda^s = \begin{pmatrix} \lambda_1^s & 0 & \cdots & 0 \\ 0 & \lambda_2^s & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & \lambda_n^s \end{pmatrix}$$

Eigenvalues and eigenvectors

Example:

$$A = \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix}$$

The eigenvalue equation is

$$\det(A - \lambda I) = \det \begin{pmatrix} -\lambda & \alpha \\ \beta & -\lambda \end{pmatrix} = 0, \quad \text{that is} \quad \lambda^2 - \alpha\beta = 0$$

The roots are $\pm\sqrt{\alpha\beta}$. Moreover,

$$V = \begin{pmatrix} \sqrt{\alpha/\beta} & -\sqrt{\alpha/\beta} \\ 1 & 1 \end{pmatrix}$$

Check that

$$A^2 = \begin{pmatrix} \alpha\beta & 0 \\ 0 & \alpha\beta \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & \alpha^2\beta \\ \alpha\beta^2 & 0 \end{pmatrix} \quad \cdots$$

Vector processes. Autoregressive equations

Back to the series

$$x_t = u_t + Au_{t-1} + A^2 u_{t-2} + \cdots \qquad (*)$$

This can be written as

$$x_t = u_t + V\Lambda V^{-1} u_{t-1} + V\Lambda^2 V^{-1} u_{t-2} + \cdots$$

or, defining $y_t = V^{-1}x_t$ and $v_t = V^{-1}u_t$,

$$y_t = v_t + \Lambda v_{t-1} + \Lambda^2 v_{t-2} + \cdots \qquad (**)$$

Obviously x_t has finite variance if and only if y_t has finite variance.

Thus x_t has finite variance if and only if all the eigenvalues of A are smaller than unity in modulus.

It is important to fully appreciate how diagonalization transforms a vector problem into a collection of scalar problems, from $(*)$ to $(**)$.

Observation. Be sure that you understand that if x_t is stationary then $z_t = Bx_t$ is stationary and that if x_t is a white noise then z_t is a white noise.

Indeed

$$\Gamma_k^z = B\Gamma_k^x B'$$

Vector processes. Autoregressive equations

A different approach to the solution of the autoregressive equation:

$$x_t = A_1 x_{t-1} + A_2 x_{t-2} + \cdots + A_p x_{t-p} + u_t$$

that is

$$(I - A_1 L - A_2 L^2 - \cdots - A_p L^p) x_t = u_t$$

The entries of the matrix

$$A(L) = I - A_1 L - A_2 L^2 - \cdots - A_p L^p$$

are polynomials in L. For example

$$A_{11}(L) = 1 - a_{11,1} L - a_{11,2} L^2 - \cdots - a_{11,p} L^p, \quad A_{12}(L) = -a_{12,1} L - a_{12,2} L^2 - \cdots - a_{12,p} L^p$$

Vector processes. Autoregressive equations

For the matrix polynomial $A(L)$ we can define $\det A(L)$ and $A_{\mathrm{ad}}(L)$. You can easily check that

$$A(L)A_{\mathrm{ad}}(L) = [\det A(L)]I_n$$

Thus, given the autoregressive equation

$$A(L)x_t = u_t$$

we can transform it into

$$\det A(L)x_t = A_{\mathrm{ad}}(L)u_t$$

Example.

$$\begin{pmatrix} 1 - a_{11}L & -a_{12}L \\ -a_{21}L & 1 - a_{22}L \end{pmatrix} \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

This becomes

$$\det A(L) \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} 1 - a_{22}L & a_{12}L \\ a_{21}L & 1 - a_{11}L \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}, \quad \det A(L) = (1-a_{11}L)(1-a_{22}L)-a_{12}a_{21}L^2$$

that is

$$\begin{aligned} \det A(L)x_{1t} &= u_{1t} - a_{22}u_{1,t-1} + a_{12}u_{2,t-1} \\ \det A(L)x_{2t} &= u_{2t} - a_{11}u_{2,t-1} + a_{21}u_{1,t-1} \end{aligned}$$

Again, the autoregressive matrix $A(L)$ has been transformed into n (which is 2 here) autoregressive polynomials $\det A(L)$.

Vector processes. Autoregressive equations

Back to the general case

$$A(L)x_t = u_t \qquad (*)$$

which is transformed into

$$\det A(L)x_t = A_{\mathrm{ad}}(L)u_t$$

If the roots of $\det A(L) = 0$ lie outside the unit circle, then $\det A(L)$ can be inverted backwards and the solution to $(*)$ is

$$x_t = [\det A(L)]^{-1}A_{\mathrm{ad}}(L)]u_t, \quad \text{we also write} \quad x_t = A(L)^{-1}u_t$$

For $p = 1$, the roots of $\det(I - AL)$ are the reciprocals of the roots of $\det(A - \lambda I)$, that is the eigenvalues of A. The latter lie within the unit circle if and only if the former lie without (the results obtained with the different approaches are consistent).

Vector processes. Autoregressive equations

Third approach. Start with an example. Suppose that $n = p = 2$.

$$x_t = A_1 x_{t-1} + A_2 x_{t-2} + u_t \qquad (*)$$

Define

$$y_t = x_{t-1}$$

and rewrite $(*)$ as

$$\begin{aligned} x_t &= A_1 x_{t-1} + A_2 y_{t-1} + u_t \\ y_t &= x_{t-1} \end{aligned}$$

Defining

$$A = \begin{pmatrix} A_1 & A_2 \\ I_2 & 0_2 \end{pmatrix}, \quad z_t = (x_t' \; y_t')', \quad v_t = (u_t' \;\; 0_{[1,2]})'$$

equation $(*)$ becomes

$$z_t = A z_{t-1} + v_t$$

Note that now $p = 1$ but $n = 4$.

Vector processes. Autoregressive equations

In general, given

$$x_t = A_1 x_{t-1} + A_2 x_{t-2} + \cdots + A_p x_{t-p} + u_t \qquad (*)$$

defining $y_{1t} = x_{t-1},\ y_{2t} = y_{1,t-1},\ \cdots,\ y_{p-1,t} = y_{p-2,t-1}$, we obtain

$$z_t = A z_{t-1} + v_t$$

where

$$A = \begin{pmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I_n & 0_n & \cdots & 0_n & 0_n \\ 0_n & I_n & \cdots & 0_n & 0_n \\ & & \ddots & & \\ 0_n & 0_n & \cdots & I_n & 0_n \end{pmatrix}, \quad z_t = (x_t' \, y_{1t}' \, \cdots \, y_{p-1,t}')', \quad v_t = (u_t' \; 0_{[1,n(p-1)]})'$$

In conclusion, equation $(*)$, which has dimension n and order p, can be transformed into an equation of dimension np and order 1. The AR(1) equation in z_t is called the companion equation.

Vector processes. Autoregressive equations

From $x_t - (A_1 x_{t-1} + A_2 x_{t-2} + \cdots + A_p x_{t-p}) = u_t$ we have obtained $z_t - Az_{t-1} = v_t$, where

$$A = \begin{pmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I_n & 0_n & \cdots & 0_n & 0_n \\ 0_n & I_n & \cdots & 0_n & 0_n \\ & & \ddots & & \\ 0_n & 0_n & \cdots & I_n & 0_n \end{pmatrix}, \quad z_t = (x_t' \, y_{1t}' \, \cdots \, y_{p-1,t}')', \quad v_t = (u_t' \; 0_{[1,n(p-1)]})'$$

Note that (easy)

$$\det(I - AL) = \det(I - A_1 L - \cdots - A_p L^p)$$

Note also that v_t is singular (many of its components are zero).

The simple construction of $(I - AL)z_t = v_t$ from $(I - A_1 L - \cdots - A_p L^p)x_t = u_t$ is the basis for the typical presentation of dynamic system as vector equations of order 1.

Vector processes. VARMA

Vector ARMA processes, VARMA, are the stationary solutions of equations of the form

$$(I - A_1L - A_2L^2 - \cdots - A_pL^p)x_t = (I + B_1L + B_2L^2 + \cdots + B_qL^q)u_t \quad (*)$$

Under the assumptions that the roots of

$$\det(I - A_1L - A_2L^2 - \cdots - A_pL^p) = 0$$

lie outside of the unit circle, the stationary solution of $(*)$ is

$$x_t = C(L)u_t = A(L)^{-1}B(L)u_t = [\det A(L)]^{-1}B(L)u_t$$

The coefficients of $C(L)$ tend to zero with a speed determined by the worst root of $\det A(L)$, precisely like in the scalar case.

Vector processes. VARMA

If the roots of

$$\det(I + B_1 L + B_2 L^2 + \cdots + B_q L^q) = 0$$

we say that the VARMA is invertible.

In that case

$$u_t = B(L)^{-1} A(L) x_t$$

so that u_t is a linear combination of present and past values of x_t. This implies, as we are going to see, that u_t is the innovation of x_t.

Vector processes. Prediction

Now consider the vector $(x_{1t}\ x_{2t})'$, where the first variable is the inflation rate, the second is interest rate. We want to predict each of the variables by using the past of both variables. In particular, we want to predict inflation using inflation and interest rate. Given what we have discussed in the scalar case, the solution of this problem is obvious: We project x_{jt}, $j = 1, 2$, on 1 and $x_{j,t-k}$, $j = 1, 2$, $k > 0$,

$$\begin{aligned} x_{1t} &= [a_{10} + a_{11,1}x_{1,t-1} + a_{12,1}x_{2,t-1} + a_{11,2}x_{1,t-2} + a_{12,2}x_{2,t-2} + \cdots] + e_{1t} \\ x_{2t} &= [a_{20} + a_{21,1}x_{1,t-1} + a_{22,1}x_{2,t-1} + a_{21,2}x_{1,t-2} + a_{22,2}x_{2,t-2} + \cdots] + e_{2t} \end{aligned}$$

In vector notation

$$x_t = [A_0 + A_1 x_{t-1} + A_2 x_{t-2} + \cdots] + e_t$$

where A_0 is a 2×1 vector while A_k, for $k \geq 1$, is 2×2.

Vector processes. Prediction

Rewrite

$$
\begin{aligned}
x_{1t} &= [a_{10} + a_{11,1}x_{1,t-1} + a_{12,1}x_{2,t-1} + a_{11,2}x_{1,t-2} + a_{12,2}x_{2,t-2} + \cdots] + e_{1t} \\
x_{2t} &= [a_{20} + a_{21,1}x_{1,t-1} + a_{22,1}x_{2,t-1} + a_{21,2}x_{1,t-2} + a_{22,2}x_{2,t-2} + \cdots] + e_{2t}
\end{aligned}
$$

Using the same argument employed in the scalar case we obtain the result that e_t is a vector white noise, that is e_{1t} is orthogonal to past values of both e_{1t} and e_{2t}, and the same for e_{2t}.

For, e_{1t} is orthogonal to $1,\ x_{1,t-1},\ x_{2,t-1},\ \cdots$. But $e_{1,t-1}$ and $e_{2,t-1}$ are linear combinations of $1,\ x_{1,t-1},\ x_{2,t-1},\ \cdots$. Etc.

Vector processes. Prediction

In general, if x_t is n-dimensional the best linear prediction of the components of x_t is obtained by projecting each of them on 1 and past values of all the components of x_t:

$$x_t = A_0 + A_1 x_{t-1} + A_2 x_{t-2} + \cdots + e_t$$

where A_0 is $n \times 1$, A_k is $n \times n$ for $k \geq 1$.

The result that e_t is an n-dimensional white noise is obtained by an obvious generalization of the argument used for $n = 2$.

Vector processes. Wold representation

Now project each component of x_t on 1 and all components of e_{t-k}, $k \geq 0$,

$$x_t = [B + e_t + B_1 e_{t-1} + B_2 e_{t-2} + \cdots] + D_t$$

where B is $n \times 1$, B_k is $n \times n$ for $k > 0$. Moreover, the residual of the projection D_t is an n-dimensional vector that is predictable without error given its past.

Vector processes. Wold representation

As an exercise you can check that if

$$x_t = Ax_{t-1} + u_t$$

where x_t is n-dimensional and the roots of $\det(I - AL)$ lie outside the unit circle, then Ax_{t-1} is the projection etc.

Similar exercise for

$$x_t = u_t + Bu_{t-1}$$

under the assumption of invertibility.

And for

$$x_t = [A_1 x_{t-1} + \cdots + a_p x_{t-p}] + [B_1 u_{t-1} + \cdots + B_q u_{t-q}] + u_t$$

under stability and invertibility.

Vector processes. Wold representation

For reasons that I will discuss later on, vector ARMA processes had little success in macroeconomics. The prevailing model was a Vector AutoRegression, usually called VAR.