Time Series lecture 9

Multivariate Time Series

Sofia Olhede

EPFL

April 16, 2025

Lecture outline

1. Multivariate Time Series
2. Multivariate spectra
3. Vector Autoregression: VAR

Multivariate Time Series

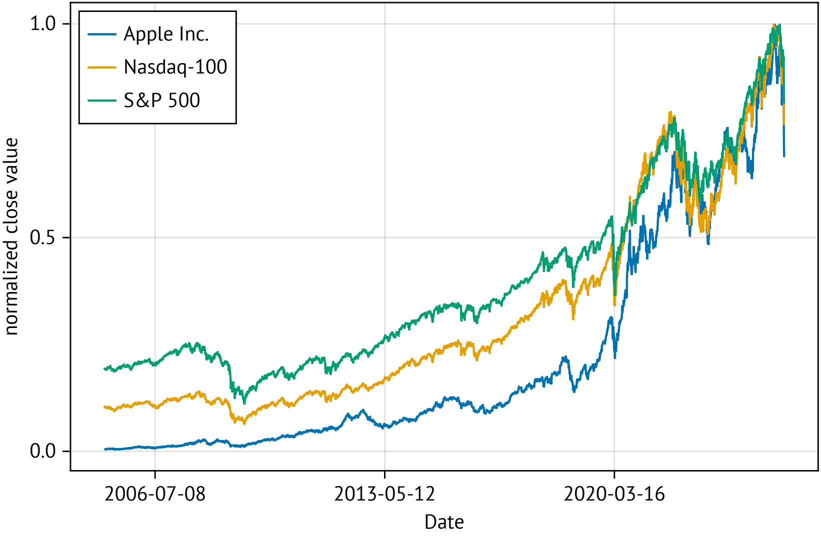
Multivariate time series

A multivariate time series is a time series which takes values in $\mathbb{R}^d$ instead of $\mathbb{R}$.
In other words, $\{\mathbf{X}_t\}$ denotes a real d-vector-valued discrete time stochastic process with

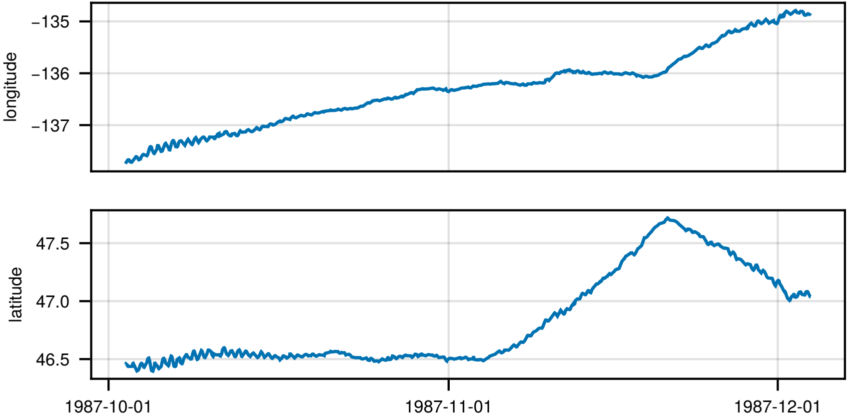
$$\mathbf{X}_t = \begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \\ \vdots \\ X_t^{(d)} \end{pmatrix}, \quad t \in \mathbb{Z},$$

where each of the marginal processes $\left\{X_t^{(1)}\right\}, \ldots, \left\{X_t^{(d)}\right\}$ are themselves univariate time series.

Stocks example



Drifter example



NOAA - hourly position, current, and sea surface temperature from drifters was accessed on 8/4/25 from `https://registry.opendata.aws/noaa-oar-hourly-gdp`.

Second order stationarity

Definition 9.1 (Multivariate second-order stationary)

For all $t, s, \tau \in \mathbb{Z}$

$$\begin{aligned} \mathbb{E}[\mathbf{X}_t] &= \mathbb{E}[\mathbf{X}_s], \\ \operatorname{Cov}(\mathbf{X}_{t+\tau}, \mathbf{X}_t) &= \operatorname{Cov}(\mathbf{X}_{s+\tau}, \mathbf{X}_s), \\ \operatorname{trace}(\operatorname{Var}(\mathbf{X}_t)) &< \infty. \end{aligned}$$

- Equivalently, we require that each of the univariate processes $\left\{X_t^{(1)}\right\}, \ldots, \left\{X_t^{(d)}\right\}$ are second-order stationary, and

$$\operatorname{Cov}\left(X_{t+\tau}^{(j)}, X_t^{(k)}\right) = \operatorname{Cov}\left(X_{s+\tau}^{(j)}, X_s^{(k)}\right),$$

for all $1 \le j, k \le d$ and for all $t, s, \tau \in \mathbb{Z}$.

Joint autocovariance sequence structure

Definition 9.2

The autocovariance sequence of a second-order stationary time series is the matrix valued sequence

$$\Gamma_\tau = \mathrm{Cov}\left(\mathbf{X}_{t+\tau}, \mathbf{X}_t\right), \quad \tau \in \mathbb{Z}.$$

- Notice that the lag is in the first argument. This does not matter in the univariate (real-valued) case, but does matter in the multivariate case!
- The jk^{th} element of Γ_τ is $\gamma_\tau^{(j,k)} = \mathrm{Cov}\left(X_{t+\tau}^{(j)}, X_t^{(k)}\right)$.
- When $j = k$, this is the usual autocovariance sequence of the univariate process.
- When $j \neq k$, this is called the cross-covariance sequence.

Properties of the autocovariance sequence (multivariate)

- From stationarity, we can see that

$$\begin{aligned}
\gamma_\tau^{(j,k)} = \mathrm{Cov}\left(X_{t+\tau}^{(j)}, X_t^{(k)}\right) &= \mathrm{Cov}\left(X_t^{(k)}, X_{t+\tau}^{(j)}\right) && \text{(symmetry)} \\
&= \mathrm{Cov}\left(X_{s-\tau}^{(k)}, X_s^{(j)}\right) && \text{(stationarity)} \\
&= \gamma_{-\tau}^{(k,j)}.
\end{aligned}$$

- Therefore we have $\Gamma_{-\tau} = \Gamma_\tau^T$.
- Notice that the cross-covariance sequences need not be symmetric in τ, unlike the univariate autocovariance sequences.
- $\{\Gamma_\tau\}$ is positive semi-definite, i.e. $\forall n \in \mathbb{N}$, $\forall \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^d$

$$\sum_{j=1}^n \sum_{k=1}^n \boldsymbol{a}_j^T \Gamma_{k-j} \boldsymbol{a}_k \geq 0.$$

Cross-correlation sequence

Definition 9.3 (Cross-correlation sequence)

We define the cross-correlation sequence (CCS) $\rho_\tau^{(j,k)}$ as

$$\rho_\tau^{(j,k)} = \frac{\gamma_\tau^{(j,k)}}{\sqrt{\gamma_0^{(j,j)} \cdot \gamma_0^{(k,k)}}}$$

- For $j \neq k$, we have that

$$\rho_\tau^{(j,k)} = \rho_{-\tau}^{(k,j)}.$$

Multivariate White Noise

Definition 9.4 (Multivariate White Noise)

The d-variate series $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is said to be white noise with zero-mean and covariance matrix Σ, written

$$\{\varepsilon_t\} \sim \mathrm{WN}(\mathbf{0}, \Sigma)$$

if and only if $\{\varepsilon_t\}$ is stationary with mean vector $\mathbf{0}$ and

$$\Gamma_\tau^{(\varepsilon)} = \begin{cases} \Sigma & \text{if } \tau = 0 \\ 0 & \text{otherwise.} \end{cases}$$

- Notice that there can be correlation between the different processes at lag 0.
- However, sometimes it is useful to assume that Σ is a diagonal matrix.

Example 9.5 (Uncorrelated processes)

Two jointly second-order processes $\left\{X_t^{(1)}\right\}$ and $\left\{X_t^{(2)}\right\}$ are said to be uncorrelated if

$$\gamma_\tau^{(1,2)} = 0$$

for all $\tau \in \mathbb{Z}$.

Example 9.6 (Contemporaneous correlation)

Say that $\mathbf{X}_t = (X_t^{(1)}, X_t^{(2)})^T$ is second-order stationary and $\{\epsilon_t\}$ is a univariate mean-zero white noise process which is independent of $\{\mathbf{X}_t\}$. Define

$$\boldsymbol{Y}_t = \mathbf{X}_t + \epsilon_t = \begin{pmatrix} X_t^{(1)} + \epsilon_t \\ X_t^{(2)} + \epsilon_t \end{pmatrix}$$

Then we have

$$\begin{aligned} \Gamma_\tau^{(\boldsymbol{Y})} &= \mathrm{Cov}\left(\boldsymbol{Y}_{t+\tau}, \boldsymbol{Y}_t\right) \\ &= \mathrm{Cov}\left(\boldsymbol{X}_{t+\tau}, \boldsymbol{X}_t\right) + \sigma_\epsilon^2 \delta_{\tau,0} \\ &= \Gamma_\tau^{(\mathbf{X})} + \sigma_\epsilon^2 \delta_{\tau,0} \end{aligned}$$

- Note that there is therefore always lag zero correlation between $Y_t^{(1)}$ and $Y_t^{(2)}$, irrespective of the correlation structure in $\mathbf{X}_t$.

Multivariate spectra

Spectral density matrix function

Definition 9.7 (Spectral density matrix function)

Assume that $\left\{\gamma_\tau^{(j,k)}\right\}_{\tau\in\mathbb{Z}} \in \ell^1$ for all $1 \le j,k \le d$. Then the spectral density matrix function is given by

$$\boldsymbol{S}(f) = \sum_{\tau\in\mathbb{Z}} \Gamma_\tau e^{-2\pi i\tau f}, \quad f \in [-1/2, 1/2].$$

- The jk^{th} element of the spectral matrix at frequency f is denoted $S_{j,k}(f)$.
- When $j = k$ it is the usual spectral density function.
- When $j \neq k$, it is called the cross-spectral density function.
- Note that $S_{j,k}(f)$ is in $\mathbb{C}$ in general, but if $j = k$ it is in $\mathbb{R}$.

Properties of the spectral density matrix function

- We have for any $f \in [-1/2, 1/2]$, for all $1 \le j, k \le d$,

$$S_{j,k}(f) = S_{k,j}(f)^*, \quad S_{j,k}(f) = S_{j,k}(-f)^*.$$

- As a consequence, we get the matrix results

$$\boldsymbol{S}(f) = \boldsymbol{S}(f)^H, \quad \boldsymbol{S}(f) = \boldsymbol{S}(-f)^T$$

  for any $f \in [-1/2, 1/2]$.
- Furthermore, we have that the matrix $\boldsymbol{S}(f)$ should be positive semi-definite.
- We also have the inverse relations

$$\Gamma_\tau = \int_{-1/2}^{1/2} \boldsymbol{S}(f)\, e^{2\pi i f \tau} \mathrm{d}f.$$

Theorem 9.8 (Multivariate spectral representation theorem)

Let $\{\mathbf{X}_t\}$ be a vector-valued discrete time stationary process with mean $\boldsymbol{\mu}$. Then there exists a vector-valued orthogonal increment process $\{\boldsymbol{Z}(f)\}$ on $\left[-\frac{1}{2}, \frac{1}{2}\right]$ such that

$$\mathbf{X}_t = \mu + \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i f t} \mathrm{d}\boldsymbol{Z}(f). \tag{9.1}$$

Assuming that $\left\{\gamma_\tau^{(j,k)}\right\}_{\tau \in \mathbb{Z}} \in \ell^1$, then for $f, f' \in \left[-\frac{1}{2}, \frac{1}{2}\right]$

1. $\mathbb{E}\left[\mathrm{d}\boldsymbol{Z}(f)\right] = 0$,
2. $\mathrm{Var}\left(\mathrm{d}\boldsymbol{Z}(f)\right) = \boldsymbol{S}(f)\,\mathrm{d}f$,
3. $\mathrm{Cov}\left(\mathrm{d}\boldsymbol{Z}(f), \mathrm{d}\boldsymbol{Z}(f')\right) = 0$ *if* $f \neq f'$.

- Note that $\boldsymbol{S}(f)$ here is a matrix,
- and that 0 in the third property refers to the zero matrix.

Coherence

Definition 9.9

The coherence between the j^{th} and k^{th} processes is given by

$$r_{j,k}(f) = \frac{S_{j,k}(f)}{\sqrt{S_{j,j}(f)\, S_{k,k}(f)}},$$

and is in essence the correlation of $\mathrm{d}Z_j$ and $\mathrm{d}Z_k$

- Coherence is a complex-valued quantity.
- Typically, we consider its magnitude or magnitude square as one statistic.
- We then look at its argument as an notion of the “sign” of the correlation (in the complex plane), which is called group delay.

Example 9.10 (Why is it called group delay?)

Consider a stationary process $\{Y_t\}$, and the bivariate process $\{\mathbf{X}_t\}$ s.t.

$$\mathbf{X}_t = \left(Y_t, \quad Y_{t-\nu}\right)^T, \quad t \in \mathbb{Z}$$

for some $\nu \in \mathbb{Z}$, then we have that

$$S_{1,1}(f) = S_{2,2}(f) = S_Y(f).$$

Furthermore, we have $\gamma_\tau^{(1,2)} = \mathrm{Cov}\left(Y_{t+\tau}, Y_{t-\nu}\right) = \gamma_{\tau+\nu}^{(Y)}$ and therefore

$$S_{1,2}(f) = S_Y(f)\, e^{2\pi i f \nu}.$$

So the coherence is

$$r_{1,2}(f) = e^{2\pi i f \nu}.$$

This has magnitude one and group delay $2\pi\nu f$.

Vector Autoregression: VAR

From univariate to multivariate models

- We treated vector-valued time series in terms of pair-wise relationships.
- But how do we propose models for them?
- The simplest framework is to start from AR processes

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t$$

- How would we generalize this to a vector-valued process?

Vector Autoregressive Process, VAR(p)

Definition 9.11 (Vector autoregression)

Let $\{\varepsilon_t\}$ be a d-dimensional multivariate white noise process with zero-mean. A process $\{\mathbf{X}_t\}$ is called a vector auto-regressive process of order p, VAR(p), if

$$\underset{d\times 1}{\mathbf{X}_t} = \underset{d\times d}{\Phi_1}\underset{d\times 1}{\mathbf{X}_{t-1}} + \underset{d\times d}{\Phi_2}\underset{d\times 1}{\mathbf{X}_{t-2}} + \cdots + \underset{d\times d}{\Phi_p}\underset{d\times 1}{\mathbf{X}_{t-p}} + \underset{d\times 1}{\varepsilon_t}$$

where the $\Phi_j \in \mathbb{R}^{d\times d}$ are matrices such that $\Phi_p \neq 0$.

- Note that we have the same regressors in the equation, namely past values of $\mathbf{X}_t$.
- Defining the polynomial $\Phi(z) = 1 - \sum_{j=1}^{p} \Phi_j z^j$, then we write

$$\Phi(\mathrm{B})\mathbf{X}_t = \varepsilon_t.$$

Everything is a VAR(1)

Consider a VAR(p), $\{\mathbf{X}_t\}$, as before. Now we want to find a way of writing this as

$$\underset{dp\times 1}{\mathbf{Y}_t} = \underset{dp\times dp}{F}\ \underset{dp\times 1}{\mathbf{Y}_{t-1}} + \underset{dp\times 1}{\mathbf{U}_t}$$

for some White noise process $\{\mathbf{U}_t\}$ and where

$$\mathbf{Y}_t = \begin{pmatrix} \mathbf{X}_t \\ \mathbf{X}_{t-1} \\ \vdots \\ \mathbf{X}_{t-p+1} \end{pmatrix}$$

- If we can do this, then we can study a VAR(1) process, and then recover the VAR(p) processes properties from it.

The companion form for a VAR(p)

Notice that

$$\mathbf{Y}_t = \begin{pmatrix} \mathbf{X}_t \\ \mathbf{X}_{t-1} \\ \vdots \\ \mathbf{X}_{t-p+1} \end{pmatrix} = \begin{pmatrix} \Phi_1 \mathbf{X}_{t-1} + \Phi_2 \mathbf{X}_{t-2} + \cdots + \Phi_p \mathbf{X}_{t-p} + \varepsilon_t \\ \mathbf{X}_{t-1} \\ \vdots \\ \mathbf{X}_{t-p+1} \end{pmatrix}$$

$$= \begin{pmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_{p-1} & \Phi_p \\ I_d & 0 & \dots & 0 & 0 \\ 0 & I_d & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & I_d & 0 \end{pmatrix} \begin{pmatrix} \mathbf{X}_{t-1} \\ \mathbf{X}_{t-2} \\ \vdots \\ \mathbf{X}_{t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Therefore, we have the companion form for a VAR(p) process as

$$\underset{dp\times 1}{\mathbf{Y}_t} = \underset{dp\times dp}{F}\,\underset{dp\times 1}{\mathbf{Y}_{t-1}} + \underset{dp\times 1}{\mathbf{U}_t}$$

where

$$\mathbf{Y}_t = \begin{pmatrix} \mathbf{X}_t \\ \mathbf{X}_{t-1} \\ \vdots \\ \mathbf{X}_{t-p+1} \end{pmatrix}, \quad F = \begin{pmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_{p-1} & \Phi_p \\ I_d & 0 & \dots & 0 & 0 \\ 0 & I_d & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & I_d & 0 \end{pmatrix}, \quad \mathbf{U}_t = \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Conditions for stationarity

The VAR(p) model is called stable if the roots of

$$\det\left\{I_d - \Phi_1 z - \cdots - \Phi_p z^p\right\} = 0$$

all lie outside the complex unit circle or equivalently we can require

$$F = \begin{pmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_{p-1} & \Phi_p \\ I_d & 0 & \dots & 0 & 0 \\ 0 & I_d & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & I_d & 0 \end{pmatrix}$$

has eigenvalues with modulus less than one.
Stability implies stationarity, but not the other way around in general.

Infinite moving average representation

Assume that the VAR(1) is stable. We can write the VAR(1) model as

$$\begin{aligned}
\mathbf{X}_t &= \Phi_1 \mathbf{X}_{t-1} + \varepsilon_t \\
&= \Phi_1 \left(\Phi_1 \mathbf{X}_{t-2} + \varepsilon_{t-1} \right) + \varepsilon_t \\
&= \Phi_1^2 \mathbf{X}_{t-2} + \Phi_1 \varepsilon_{t-1} + \varepsilon_t \\
&= \cdots \\
&= \sum_{j=0}^{\infty} \Phi_1^j \varepsilon_{t-j}
\end{aligned}$$

This manipulation can be shown to be formally correct, but is beyond the scope of the course.

Covariance

The covariance of a VAR(1) process is then given by

$$
\begin{aligned}
\Gamma_\tau &= \mathrm{Cov}\left(\mathbf{X}_{t+\tau}, \mathbf{X}_t\right) \\
&= \mathrm{Cov}\left(\sum_{j=0}^{\infty} \Phi_1^j \varepsilon_{t-j+\tau}, \sum_{k=0}^{\infty} \Phi_1^k \varepsilon_{t-k}\right) \\
&= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \Phi_1^j \,\mathrm{Cov}\left(\varepsilon_{t-j+\tau}, \varepsilon_{t-k}\right)\left(\Phi_1^k\right)^T \\
&= \sum_{k=0}^{\infty} \Phi_1^{k+\tau} \Sigma (\Phi_1^k)^T
\end{aligned}
$$

Recovering the VAR(p)

To recover the VAR(p) process from its companion form, we can note that

$$\underset{d\times 1}{\mathbf{X}_t} = \underset{d\times dp}{G}\ \underset{dp\times 1}{\mathbf{Y}_t}$$

where

$$G = \begin{pmatrix} I_d & 0 & \cdots & 0 \end{pmatrix}.$$

Therefore, we have that

$$\begin{aligned}
\Gamma_\tau^{(\mathbf{X})} &= \mathrm{Cov}\left(\mathbf{X}_{t+\tau}, \mathbf{X}_t\right) \\
&= \mathrm{Cov}\left(G\mathbf{Y}_{t+\tau}, G\mathbf{Y}_t\right) \\
&= G\Gamma_\tau^{(\mathbf{Y})} G^T \\
&= G\left(\sum_{k=0}^{\infty} F^{k+\tau}\Sigma(F^k)^T\right) G^T
\end{aligned}$$

Example 9.12 (bivariate-VAR(1))

We consider the bivariate or (two-dimensional) process of

$$\mathbf{X}_t = \boldsymbol{\nu} + \begin{pmatrix} 0.5 & 0.1 \\ 0.4 & 0.5 \end{pmatrix} \mathbf{X}_{t-1} + \begin{pmatrix} 0 & 0 \\ 0.25 & 0 \end{pmatrix} \mathbf{X}_{t-2} + \varepsilon_t.$$

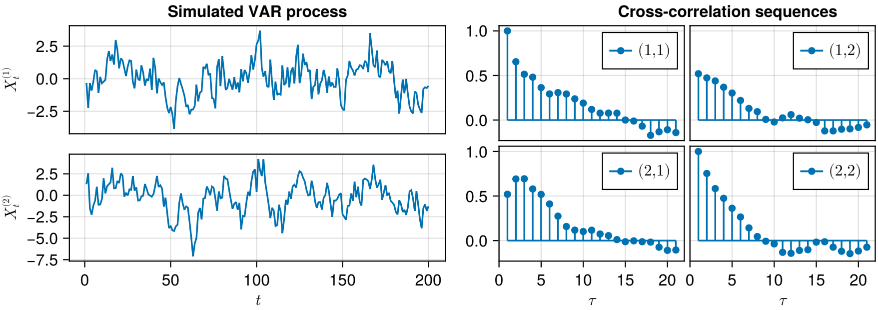
For this process the reverse characteristic polynomial is

$$\det \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.5 & 0.1 \\ 0.4 & 0.5 \end{pmatrix} z - \begin{pmatrix} 0 & 0 \\ 0.25 & 0 \end{pmatrix} z^2 \right\}$$
$$= 1 - z + 0.21z^2 - 0.025z^3.$$

The roots are now instead

$$z_1 = 1.3, \quad z_2 = 3.55 + 4.26i, \quad z_3 = 3.55 - 4.26i.$$

Clearly $|z_1| > 1$ and also $|z_2|^2 = |z_3|^2 = 3.55^2 + 4.26^2$. Taking the squareroot of the latter quantity we get $|z_2| = |z_3| = 5.545$



$$\mathbf{X}_t = \boldsymbol{\nu} + \begin{pmatrix} 0.5 & 0.1 \\ 0.4 & 0.5 \end{pmatrix} \mathbf{X}_{t-1} + \begin{pmatrix} 0 & 0 \\ 0.25 & 0 \end{pmatrix} \mathbf{X}_{t-2} + \varepsilon_t.$$

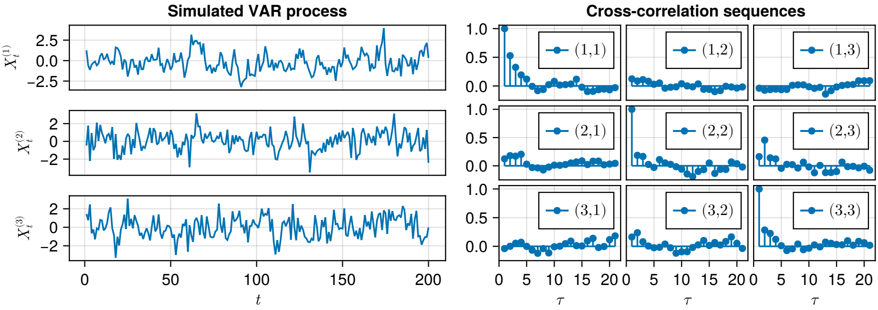
Example 9.13 (trivariate-VAR(1))

$$\mathbf{X}_t = \boldsymbol{\nu} + \begin{pmatrix} 0.5 & 0 & 0 \\ 0.1 & 0.1 & 0.3 \\ 0 & 0.2 & 0.3 \end{pmatrix} \mathbf{X}_{t-1} + \varepsilon_t$$

For this process the reverse characteristic polynomial is

$$\det\left\{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.5 & 0 & 0 \\ 0.1 & 0.1 & 0.3 \\ 0 & 0.2 & 0.3 \end{pmatrix} z\right\}$$
$$= (1 - 0.5z)\left(1 - 0.4z - 0.03z^2\right).$$

The roots are $z_1 = 2$, $z_2 \approx 2.1525$, $z_3 \approx -15.4858$. These are obviously greater than unity in magnitude, and so the VAR(1) is stable.



$$\mathbf{X}_t = \begin{pmatrix} 0.5 & 0 & 0 \\ 0.1 & 0.1 & 0.3 \\ 0 & 0.2 & 0.3 \end{pmatrix} \mathbf{X}_{t-1} + \varepsilon_t$$

Yule-Walker equations

We can derive the Yule-Walker equations also for the VAR process.

$$\mathbf{X}_t = \Phi_1 \mathbf{X}_{t-1} + \varepsilon_t.$$

It therefore follows that

$$\begin{aligned} \mathrm{Cov}\left(\mathbf{X}_{t+\tau}, \mathbf{X}_t\right) &= \mathrm{Cov}\left(\mathbf{X}_{t+\tau}, \Phi_1 \mathbf{X}_{t-1} + \varepsilon_t\right) \\ &= \Phi_1 \Gamma_{\tau-1} + \Sigma \delta_{\tau,0}. \end{aligned}$$

We therefore get

$$\Gamma_\tau = \Phi_1 \Gamma_{\tau-1} + \Sigma \delta_{\tau,0}.$$

If we know Σ and Γ_0 then we obtain the form of the other auto-covariance matrices $\{\Gamma_\tau\}$ iteratively.

Vector Autoregression

For a higher order VAR(p) we have

$$\mathbf{X}_t = \Phi_1 \mathbf{X}_{t-1} + \cdots + \Phi_p \mathbf{X}_{t-p} + \varepsilon_t.$$

Therefore for $\tau \geq 0$ we have

$$\begin{aligned}\Gamma_\tau &= \operatorname{Cov}\left(\mathbf{X}_{t+\tau}, \mathbf{X}_t\right)\\ &= \operatorname{Cov}\left(\mathbf{X}_{t+\tau}, \Phi_1 \mathbf{X}_{t-1} + \cdots + \Phi_p \mathbf{X}_{t-p} + \varepsilon_t\right)\\ &= \Phi_1 \Gamma_{\tau-1} + \cdots + \Phi_p \Gamma_{\tau-p} + \delta_{\tau,0}\Sigma.\end{aligned}$$

We can solve the first $p+1$ of these to obtain $\Gamma_0, \ldots, \Gamma_p$ in terms of Σ and the Φ_j matrices. The remaining Γ_τ for $\tau > p$ can be obtained recursively from the above equation.