Basic Multivariate Time Series Concepts

September 3, 2024

The material in this set of notes is based on S&S 1.1-1.6.

Suppose we observe a multivariate time-series, i.e. an $n \times r$ matrix of r time series observed simultanously:

$$\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{y}_1 & \dots & \boldsymbol{y}_r \end{pmatrix} = \begin{pmatrix} y_{11} & \dots & y_{1r} \\ \vdots & \ddots & \vdots \\ y_{n1} & \dots & y_{nr} \end{pmatrix} = \boldsymbol{M}_y + \boldsymbol{W},$$

where $\boldsymbol{M}_x$ is a fixed but unknown mean, $\boldsymbol{W}$ are random errors, and elements of each column of $\boldsymbol{Y}$ denoted by $\boldsymbol{y}_i$ are ordered in time.

A multivariate time series is characterized by its mean function $m_{y,ij} = \mathbb{E}\left[y_{ij}\right]$ and covariance function $\gamma_{ij}\left(s, t\right) = \text{Cov}\left(y_{si}, y_{tj}\right)$.

- When $i = j$, this is the autocovariance function of time series $\boldsymbol{y}_i$.
- When $i \neq j$, we call this the **cross-covariance function** of the time series $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$.

The correlation function can be derived from the covariance function: $\rho_{ij}\left(s, t\right) = \frac{\gamma_{ij}(s,t)}{\sqrt{\gamma_{ii}(s,s)\gamma_{jj}(t,t)}}$, like its univariate counterpart the correlation function's values are between -1 and 1.

- When $i \neq j$, we call this the **cross-correlation function** of the time series $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$.

As in the univariate case, characterizing a time series in this way is too complicated and involves too many parameters, because the mean and covariance functions depend on the values s and t themselves. This leads us back to the idea of **stationarity**. A multivariate time series is **jointly stationary** if:

- The second moments of y_{ti} are finite for all of the time series, i.e. $\mathbb{E}\left[y_{ti}^2\right] < \infty$ for all t and $i = 1, \dots, r$.
- The mean function is constant for each time series and does not depend on time, $m_{y,ti} = m_{y,i}$.
- The autocovariance function $\gamma_{ii}(s,t)$ depends on s and t only through their absolute difference $h = |s - t|$ for all $i = 1, \dots, r$.
- The cross-covariance function $\gamma_{ij}(s,t)$ depends on s and t only through their difference $h = s - t$ for all $i = 1, \dots, r$.

As in the univariate case, when a time series is stationary, its autocovariance and autocorrelation functions can be written as functions of a single variable h. For this reason, we will drop the second arguments of the autocovariance and autocorrelation functions when a time series is stationary, writing $\gamma_{ij}\left(h\right) = \gamma_{ij}\left(h, 0\right)$ and $\rho_{ij}\left(h\right) = \rho_{ij}\left(h, 0\right)$.

When we observe a time series $\boldsymbol{Y}$, we do not know the mean, autocovariance, or autocorrelation functions a priori - we need to estimate them. When $\boldsymbol{Y}$ is stationary we can compute:

- The **sample mean** function:

$$\hat{m}_{y,i} = \bar{y}_i = \sum_{t=1}^{n} y_{ti}/n. \tag{1}$$

- The **sample auto-covariance function**:

$$\hat{\gamma}_{ii}\left(h\right) = \frac{1}{n}\sum_{t=1}^{n-h}\left(y_{t+h,i} - \hat{m}_{y,i}\right)\left(y_{ti} - \hat{m}_{y,i}\right), \tag{2}$$

  with $\hat{\gamma}_{ii}\left(h\right) = \hat{\gamma}_{ii}\left(-h\right)$ for $h = 0, 1, \dots, n-1$.

- The **sample cross-covariance function**:

$$\hat{\gamma}_{ij}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (y_{t+h,i} - \hat{m}_{y,i})(y_{tj} - \hat{m}_{y,j}), \tag{3}$$

  with $\hat{\gamma}_{ij}(h) = \hat{\gamma}_{ji}(-h)$ for $h = 0, 1, \ldots, n-1$.

- The **sample autocorrelation function**:

$$\hat{\rho}_{ii}(h) = \frac{\hat{\gamma}_{ii}(h)}{\hat{\gamma}_{ii}(0)}. \tag{4}$$

- The **sample cross-correlation function**:

$$\hat{\rho}_{ij}(h) = \frac{\hat{\gamma}_{ij}(h)}{\sqrt{\hat{\gamma}_{ii}(0)\,\hat{\gamma}_{jj}(0)}} \tag{5}$$

In practice, we might want to ask how different our estimates of the sample cross-correlation function $\hat{\rho}_{ij}(h)$ are from what we would expect if either $\boldsymbol{y}_i$ or $\boldsymbol{y}_j$ are **white noise** time series with no autocorrelation at all, i.e. if $\rho_{ij}(h) = 0$ for all $h \neq 0$. We can get a handle on this using the following result:

> If $y_{ti} = w_{ti}$ where $w_{ti} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_{w,ii})$ or $y_{tj} = w_{tj}$ where $w_{tj} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_{w,jj})$ for $t = 1, \ldots n$, then $\hat{\rho}_{ij}(h) \approx v/\sqrt{n}$, for $h = 1, \ldots H$, where $v \sim \mathcal{N}(0,1)$ and H is fixed but arbitrary.

This result allows us to perform an approximate test of the null hypothesis that $\rho_{ij}(h) = 0$ for any $h > 1$ and any pair of time series, $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$!