8.3: Estimating the Mean and Covariance Function

Taylor

University of Virginia

Motivation

In this section we introduce estimators, for a stationary m-variate time series $\{X_t\}$, of the components μ_j, $\gamma_{ij}(h)$, and $\rho_{ij}(h)$ of $\boldsymbol{\mu}$, $\Gamma(h)$, and $R(h)$, respectively. We also examine the large-sample properties of these estimators.

Estimating $\boldsymbol{\mu}$

One estimator for the mean is

$$\bar{\boldsymbol{X}}_n = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{X}_i$$

Proposition 8.3.1

If $\{\boldsymbol{X}_t\}$ is a stationary multivariate time series with mean $\boldsymbol{\mu}$ and covariance function $\Gamma(\cdot)$, then as $n \to \infty$

$$E(\boldsymbol{X}_n - \boldsymbol{\mu})'(\boldsymbol{X}_n - \boldsymbol{\mu}) \to 0 \qquad \text{if } \gamma_{ii}(n) \to 0 \qquad 1 \le i \le n$$

and

$$nE(\boldsymbol{X}_n - \boldsymbol{\mu})'(\boldsymbol{X}_n - \boldsymbol{\mu}) \to \sum_{i=1}^{m}\sum_{h=\infty}^{\infty} \gamma_{ii}(h) \quad \text{if } \sum_{h=\infty}^{\infty} |\gamma_{ii}(h)| < \infty \quad 1 \le i \le n$$

Estimating $\boldsymbol{\mu}$

The mean $\bar{\boldsymbol{X}}_n = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{X}_i$ is asymptotically Normal, but we are not going to look at the formula for it's asymptotic covariance matrix. If you want to make individual confidence intervals or tests for elements of $\boldsymbol{\mu}$, use Proposition 2.4.1.

Estimating $\Gamma(\cdot)$

A natural estimator for $\Gamma(\cdot)$ is

$$\hat{\Gamma}(h) = \begin{cases} n^{-1} \sum_{t=1}^{n-h} (\mathbf{X}_{t+h} - \bar{\mathbf{X}})(\mathbf{X}_t - \bar{\mathbf{X}})' & 0 \le h \le n-1 \\ -\hat{\Gamma}'(-h) & -(n-1) \le h < 0 \end{cases}$$

Note the typo in the book.

Cross-correlations can be estimated with $\hat{\rho}_{ij}(h) = \hat{\gamma}_{ij}(h)[\hat{\gamma}_{ii}(0)\hat{\gamma}_{jj}(0)]^{-1/2}$

Theorem 8.3.1

Let $\{\mathbf{X}_t\}$ be the bivariate time series whose components are defined by

$$X_{t1} = \sum_{k=-\infty}^{\infty} \alpha_k Z_{t-k,1}, \qquad \{Z_{t1}\} \sim \text{IID}(0, \sigma_1^2)$$

and

$$X_{t2} = \sum_{k=-\infty}^{\infty} \beta_k Z_{t-k,2}, \qquad \{Z_{t2}\} \sim \text{IID}(0, \sigma_2^2)$$

where the two sequences $\{Z_{t1}\}$ and $\{Z_{t2}\}$ are independent, $\sum_k |\alpha_k| < \infty$, and $\sum_k |\beta_k| < \infty$.
Then for all integers h, k, $h \neq k$, $\sqrt{n}(\hat{\rho}_{12}(h), \hat{\rho}_{12}(k))'$ is asymptotically Normal, with mean $\mathbf{0}$ and covariance matrix

$$\begin{bmatrix} \sum_{j=-\infty}^{\infty} \rho_{11}(j)\rho_{22}(j) & \sum_{j=-\infty}^{\infty} \rho_{11}(j)\rho_{22}(j+k-h) \\ \sum_{j=-\infty}^{\infty} \rho_{11}(j+k-h)\rho_{22}(j) & \sum_{j=-\infty}^{\infty} \rho_{11}(j)\rho_{22}(j) \end{bmatrix}$$

Theorem 8.3.1: Special case 1

Assume everything from the last slide. Then, in particular,

$$\sqrt{n}\hat{\rho}_{12}(h) \xrightarrow{D} \mathrm{N}\left(0, \sum_{j=-\infty}^{\infty} \rho_{11}(j)\rho_{22}(j)\right).$$

Practically speaking: you can still observe large cross-correlations for independent time series!

Theorem 8.3.1: Special case 2

Assume everything from the last slide. Assume further that atleast one of X_{t1} and $X_{t,2}$ is white noise. Then, in particular,

$$\sqrt{n}\hat{\rho}_{12}(h) \xrightarrow{D} \mathrm{N}\left(0, 1\right).$$

Option 1: prewhitening

If

$$\phi^{(1)}(B)X_{t1} = \theta^{(1)}(B)Z_{t1}$$
$$\phi^{(2)}(B)X_{t2} = \theta^{(2)}(B)Z_{t2}$$

then

$$\frac{\phi^{(1)}(B)}{\theta^{(2)}(B)}X_{t1} = \frac{\theta^{(1)}(B)}{\theta^{(2)}(B)}Z_{t1}$$
$$\frac{\phi^{(2)}(B)}{\theta^{(2)}(B)}X_{t2} = Z_{t2}.$$

This means, estimate a univariate model on one of the series. And then use it to filter both series.

Option 2

If

$$\phi^{(1)}(B)X_{t1} = \theta^{(1)}(B)Z_{t1}$$
$$\phi^{(2)}(B)X_{t2} = \theta^{(2)}(B)Z_{t2}$$

then estimate both models, and examine the CCF of the residuals $\{\hat{W}_{t1}\}$ and $\{\hat{W}_{t2}\}$ (section 5.3).

Bartlett's Formula

All of this so far assumes that the two series are independent! If they are not, then we can use Bartlett's formula.

Assume $\{\mathbf{X}_t\}$ is a bivariate *Gaussian* time series with covariances satisfying $\sum_{h=-\infty}^{\infty} |\gamma_{ij}(h)| < \infty$, $i, j = 1, 2$. Notice that these two series are not necessarily independent. Bartlet's formula is an approximation

$$n \operatorname{Cov}(\hat{\rho}_{12}(h), \hat{\rho}_{12}(k))$$

It is not reproduced here. But if you need it, it is on page 242.

Example: Sales with A Leading Indicator

We assume

$$X_{t1} - \mu_1 = (1 + \theta_{11} B) Z_{t1}$$

and

$$(1 - \phi_{11} B)(X_{t2} - \mu_2) = (1 + \theta_{21} B) Z_{t2}$$

We look at the residuals, which is kind of like looking at

$$\frac{1}{(1 + \theta_{11} B)} (X_{t1} - \mu_1) = Z_{t2}$$
$$\frac{(1 - \phi_{11} B)}{(1 + \theta_{21} B)} (X_{t2} - \mu_2) = Z_{t2}$$

Example: Sales with A Leading Indicator (continued)

$$\frac{1}{(1+\theta_{11}B)}(X_{t1}-\mu_1) = Z_{t1}$$
$$\frac{(1-\phi_{11}B)}{(1+\theta_{21}B)}(X_{t2}-\mu_2) = Z_{t2}$$

Looking at the CCF, we find $\text{Cov}(Z_{t-3,1}, Z_{t,2}) \neq 0$.

$$Z_{t,2} = \beta Z_{t-3,1} + N_t$$
$$\frac{(1-\phi_{11}B)}{(1+\theta_{21}B)}(X_{t2}-\mu_2) = \frac{1}{(1+\theta_{11}B)}(X_{t-3,1}-\mu_1) + N_t$$

...messy. See 8.3.r for more details.