11 Estimation of Mean and Covariance Function

11.1 Estimators and Their Asymptotic Distributions

We characterize the stationary process $\{X_t\}$ by its mean and its (matrix) covariance function. In the Gaussian case, this already characterizes the whole distribution. The estimation of these entities becomes crucial in the empirical analysis. As it turns out, the results from the univariate process carry over analogously to the multivariate case. If the process is observed over the periods $t = 1, 2, \ldots, T$, then a natural estimator for the mean μ is the arithmetic mean or sample average:

$$\hat{\mu} = \overline{X}_T = \frac{1}{T}(X_1 + \ldots + X_T) = \begin{pmatrix} \overline{X}_1 \\ \vdots \\ \overline{X}_n \end{pmatrix}.$$

We get a theorem analogously to Theorem 4.1 in Sect. 4.1.

Theorem 11.1. *Let $\{X_t\}$ be stationary process with mean μ and covariance function $\Gamma(h)$ then asymptotically, for $T \to \infty$, we get*

$$\mathbb{E}\left(\overline{X}_T - \mu\right)'\left(\overline{X}_T - \mu\right) \to 0, \ \textit{if}\, \gamma_{ii}(T) \to 0\, \textit{for all}\ 1 \le i \le n;$$

$$T\mathbb{E}\left(\overline{X}_T - \mu\right)'\left(\overline{X}_T - \mu\right) \to \sum_{i=1}^{n}\sum_{h=-\infty}^{\infty} \gamma_{ii}(h), \quad \textit{if} \sum_{h=-\infty}^{\infty} |\gamma_{ii}(h)| < \infty\, \textit{for all}\ 1 \le i \le n.$$

Proof. The Theorem can be established by applying Theorem 4.1 individually to each time series $\{X_{it}\}$, $i = 1, 2, \ldots, n$. □

K. Neusser, *Time Series Econometrics*, Springer Texts in Business and Economics,
DOI 10.1007/978-3-319-32862-1_11

Thus, the sample average converges in mean square and therefore also in probability to the true mean. Thereby the second condition is more restrictive than the first one. They are, in particular, fulfilled for all VARMA processes (see Chap. 12). As in the univariate case analyzed in Sect. 4.1, it can be shown with some mild additional assumptions that $\overline{X}_T$ is also asymptotically normally distributed.

Theorem 11.2. *For any stationary process* $\{X_t\}$

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \Psi_j Z_{t-j}$$

with $Z_t \sim \mathrm{IID}(0, \Sigma)$ *and* $\sum_{j=-\infty}^{\infty} \|\Psi_j\| < \infty$, *the arithmetic average* $\overline{X}_T$ *is asymptotically normal:*

$$\begin{aligned}\sqrt{T}\left(\overline{X}_T - \mu\right) \xrightarrow{d} \quad & \mathrm{N}\left(0, \sum_{h=-\infty}^{\infty} \Gamma(h)\right) \\ &= \mathrm{N}\left(0, \left(\sum_{j=-\infty}^{\infty} \Psi_j\right) \Sigma \left(\sum_{j=-\infty}^{\infty} \Psi_j'\right)\right) \\ &= \mathrm{N}\left(0, \Psi(1)\Sigma\Psi(1)'\right).\end{aligned}$$

Proof. The proof is a straightforward extension to the multivariate case of the one given for Theorem 4.2 of Sect. 4.1. □

The summability condition is quite general. It is, in particular, fulfilled by causal VARMA processes (see Chap. 12) as their coefficients matrices Ψ_j go exponentially fast to zero. Remarks similar to those following Theorem 4.2 apply also in the multivariate case.

The above formula can be used to construct confidence regions for μ. This turns out, however, to be relatively complicated in practice so that often univariate approximations are used instead (Brockwell and Davis 1996, 228–229).

As in the univariate case, a natural estimator for the covariance matrix function $\Gamma(h)$ is given by the corresponding empirical moments $\widehat{\Gamma}(h)$:

$$\widehat{\Gamma}(h) = \begin{cases} \frac{1}{T}\sum_{t=1}^{T-h}\left(X_{t+h} - \overline{X}_T\right)\left(X_t - \overline{X}_T\right)', & 0 \le h \le T-1; \\ \widehat{\Gamma}'(-h), & -T+1 \le h < 0. \end{cases}$$

The estimator of the covariance function can then be applied to derive an estimator for the correlation function:

$$\widehat{R}(h) = \hat{V}^{-1/2}\widehat{\Gamma}(h)\hat{V}^{-1/2}$$

where $\hat{V}^{1/2} = \text{diag}\left(\sqrt{\hat{\gamma}_{11}(0)}, \ldots, \sqrt{\hat{\gamma}_{nn}(0)}\right)$. Under the conditions given in Theorem 11.2 the estimator of the covariance matrix of order h, $\widehat{\Gamma}(h)$, converges to the true covariance matrix $\Gamma(h)$. Moreover, $\sqrt{T}\left(\widehat{\Gamma}(h) - \Gamma(h)\right)$ is asymptotically normally distributed. In particular, we can state the following Theorem:

Theorem 11.3. *Let $\{X_t\}$ be a stationary process with*

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \Psi_j Z_{t-j}$$

where $Z_t \sim \text{IID}(0, \Sigma)$, $\sum_{j=-\infty}^{\infty} \|\Psi_j\| < \infty$, and $\sum_{j=-\infty}^{\infty} \Psi_j \neq 0$. Then, for each fixed h, $\widehat{\Gamma}(h)$ converges in probability as $T \to \infty$ to $\Gamma(h)$:

$$\widehat{\Gamma}(h) \xrightarrow{p} \Gamma(h)$$

Proof. A proof can be given along the lines given in Proposition 13.1. □

As for the univariate case, we can define the long-run covariance matrix J as

$$J = \sum_{h=-\infty}^{\infty} \Gamma(h). \tag{11.1}$$

As a non-parametric estimator we can again consider the following class of estimators:

$$\hat{J}_T = \sum_{h=-T+1}^{T-1} k\left(\frac{h}{\ell_T}\right) \widehat{\Gamma}(h)$$

where $k(x)$ is a kernel function and where $\widehat{\Gamma}(h)$ is the corresponding estimate of the covariance matrix at lag h. For the choice of the kernel function and the lag truncation parameter the same principles apply as in the univariate case (see Sect. 4.4 and Haan and Levin (1997)).

11.2 Testing Cross-Correlations of Bivariate Time Series

The determination of the asymptotic distribution of $\widehat{\Gamma}(h)$ is complicated. We therefore restrict ourselves to the case of two time series.

Theorem 11.4. *Let $\{X_t\}$ be a bivariate stochastic process whose components can be described by*

$$X_{1t} = \sum_{j=-\infty}^{\infty} \alpha_j Z_{1,t-j} \quad \textit{with } Z_{1t} \sim \mathrm{IID}(0, \sigma_1^2)$$

$$X_{2t} = \sum_{j=-\infty}^{\infty} \beta_j Z_{2,t-j} \quad \textit{with } Z_{2t} \sim \mathrm{IID}(0, \sigma_2^2)$$

where $\{Z_{1t}\}$ and $\{Z_{2t}\}$ are independent from each other at all leads and lags and where $\sum_j |\alpha_j| < \infty$ and $\sum_j |\beta_j| < \infty$. Under these conditions the asymptotic distribution of the estimator of the cross-correlation function $\rho_{12}(h)$ between $\{X_{1t}\}$ and $\{X_{2t}\}$ is

$$\sqrt{T}\hat{\rho}_{12}(h) \xrightarrow{d} \mathrm{N}\left(0, \sum_{j=-\infty}^{\infty} \rho_{11}(j)\rho_{22}(j)\right), \qquad h \geq 0. \tag{11.2}$$

For all h and k with $h \neq k$, $(\sqrt{T}\hat{\rho}_{12}(h), \sqrt{T}\hat{\rho}_{12}(k))'$ converges in distribution to a bivariate normal distribution with mean zero and variances and covariances given by $\sum_{j=-\infty}^{\infty} \rho_{11}(j)\rho_{22}(j)$ and $\sum_{j=-\infty}^{\infty} \rho_{11}(j)\rho_{22}(j+k-h)$, respectively.

This result can be used to construct a test of independence, respectively uncorrelatedness, between two time series. The above theorem, however, shows that the asymptotic distribution of $\sqrt{T}\hat{\rho}_{12}(h)$ depends on $\rho_{11}(h)$ and $\rho_{22}(h)$ and is therefore unknown. Thus, the test cannot be based on the cross-correlation alone.[1]

This problem can, however, be overcome by implementing the following two-step procedure suggested by Haugh (1976).

First step: Estimate for each times series separately a univariate invertible ARMA model and compute the resulting residuals $\hat{Z}_{it}$ as $\hat{Z}_{it} = \sum_{j=0}^{\infty} \hat{\pi}_j^{(i)} X_{i,t-j}$, $i = 1, 2$. If the ARMA models correspond to the true ones, these residuals should approximately be white noise. This first step is called pre-whitening.

Second step: Under the null hypothesis the two time series $\{X_{1t}\}$ and $\{X_{2t}\}$ are uncorrelated with each other. This implies that the residuals $\{Z_{1t}\}$ and $\{Z_{2t}\}$ should also be uncorrelated with each other. The variance of the cross-correlations between $\{Z_{1t}\}$ and $\{Z_{2t}\}$ are therefore asymptotically equal to $1/T$ under the null hypothesis. Thus, one can apply the result of Theorem 11.4 to construct confidence intervals based on formula (11.2). A 95-% confidence interval is therefore given by $\pm 1.96 T^{-1/2}$. The Theorem may also be used to construct a test whether the two series are uncorrelated.

If one is not interested in modeling the two time series explicitly, the simplest way is to estimate a high order AR model in the first step. Thereby, the order should be chosen high enough to obtain white noise residuals in the first step. Instead

[1]The theorem may also be used to conduct a causality test between two times series (see Sect. 15.1).

of looking at each cross-correlation separately, one may also test the joint null hypothesis that all cross-correlations are simultaneously equal to zero. Such a test can be based on T times the sum of the squared cross-correlation coefficients. This statistic is distributed as a χ^2 with L degrees of freedom where L is the number of summands (see the Haugh-Pierce statistic (15.1) in Sect. 15.1).

11.3 Some Examples for Independence Tests

Two Independent AR Processes

Consider two AR(1) process $\{X_{1t}\}$ and $\{X_{2t}\}$ governed by the following stochastic difference equation $X_{it} = 0.8X_{i,t-1} + Z_{it}$, $i = 1, 2$. The two white noise processes $\{Z_{1t}\}$ and $\{Z_{2t}\}$ are such that they are independent from each other. $\{X_{1t}\}$ and $\{X_{2t}\}$ are therefore independent from each other too. We simulate realizations of these two processes over 400 periods. The estimated cross-correlation function of these so-generated processes are plotted in the upper panel of Fig. 11.1. From there one can see that many values are outside the 95 % confidence interval given by $\pm 1.96T^{-1/2} = 0.098$, despite the fact that by construction both series are independent of each other. The reason is that the so computed confidence interval is not correct because it does not take the autocorrelation of each series into account. The application of Theorem 11.4 leads to the much larger 95-% confidence interval of

$$\frac{\pm 1.96}{\sqrt{T}}\sqrt{\sum_{j=-\infty}^{\infty}\rho_{11}(j)\rho_{22}(j)} = \frac{\pm 1.96}{20}\sqrt{\sum_{j=-\infty}^{\infty}0.8^j 0.8^j}$$

$$= \frac{\pm 1.96}{20}\sqrt{1 + \frac{2 \times 0.64}{1 - 0.64}} = \pm 0.209$$

which is more than twice as large. This confidence interval then encompasses most the cross-correlations computed with respect to the original series.

If one follows the testing procedure outline above instead and fits an AR(10) model for each process and then estimates the cross-correlation function of the corresponding residual series (filtered or pre-whitened time series), the plot in the lower panel of Fig. 11.1 is obtained.[2] This figure shows no significant cross-correlation anymore so that one cannot reject the null hypothesis that both time series are independent from each other.

[2]The order of the AR processes are set arbitrarily equal to 10 which is more than enough to obtain white noise residuals.

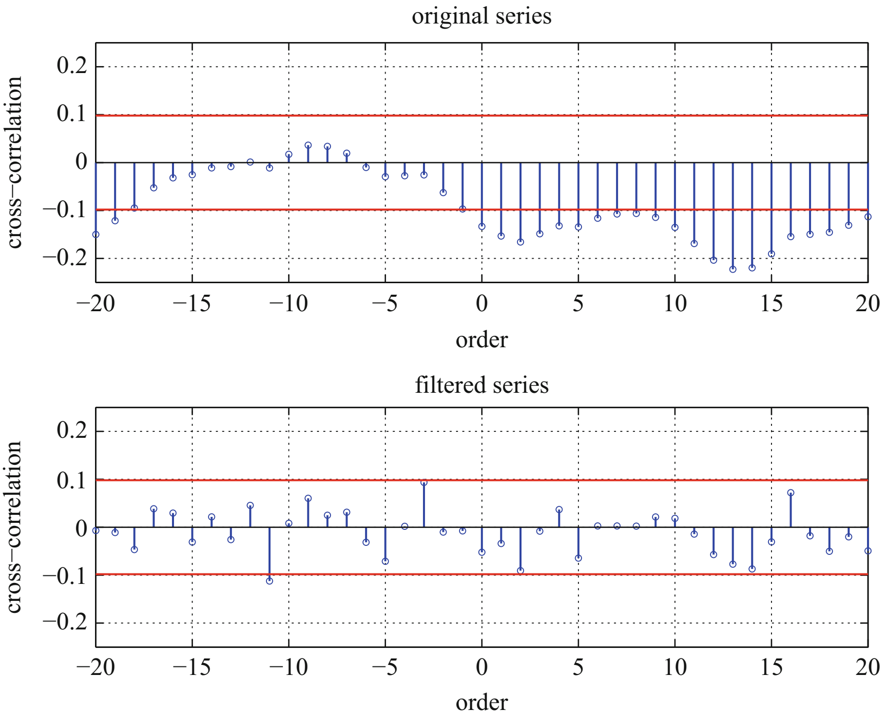

Fig. 11.1 Cross-correlations between two independent AR(1) processes with $\phi = 0.8$

Consumption Expenditure and Advertisement Expenses

This application focuses on the interaction between nominal aggregate private consumption expenditure and nominal aggregate advertisement expenditures. Such an investigation was first conducted by Ashley et al. (1980) for the United States.[3] The upper panel of Fig. 11.2 shows the raw cross-correlations between the two time series where the order h runs from -20 to $+20$. Although almost all cross-correlations are positive and outside the conventional confidence interval, it would be misleading to infer a statistically significant positive cross-correlation. In order to test for independence, we filter both time series by an AR(10) model and estimate the cross-correlations for the residuals.[4] These are displayed in the lower panel of Fig. 11.2. In this figure, only the correlations of order 0 and 16 fall outside the confidence interval and can thus be considered as statistically significant. Thus, we

[3]The quarterly data are taken from Berndt (1991). They cover the period from the first quarter 1956 to the fourth quarter 1975. In order to achieve stationarity, we work with first differences.

[4]The order of the AR processes are set arbitrarily equal to 10 which is more than enough to obtain white noise residuals.

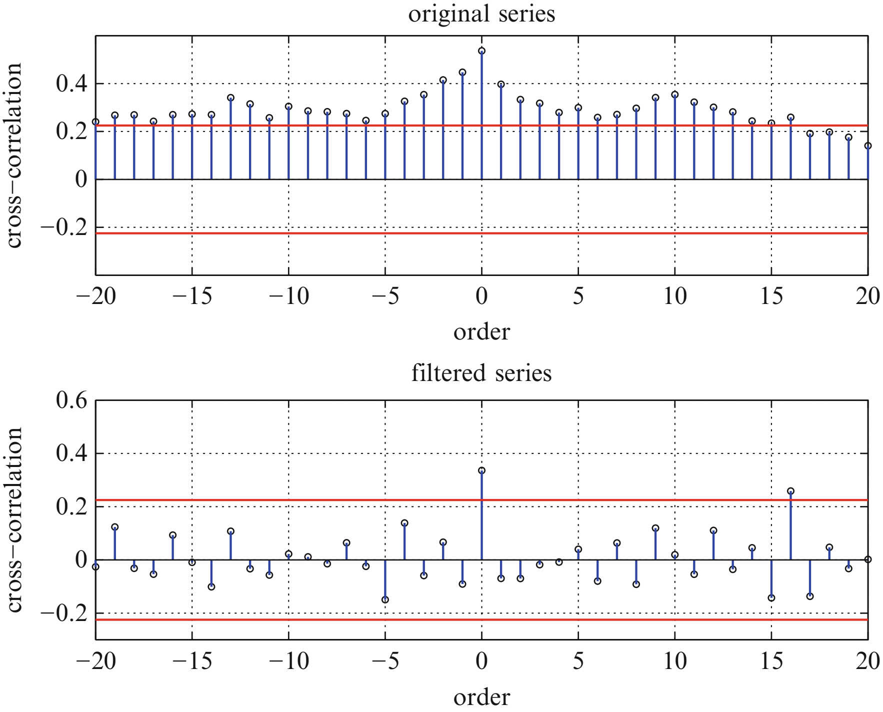

Fig. 11.2 Cross-correlations between aggregate nominal private consumption expenditures and aggregate nominal advertisement expenditures

can reject the null hypothesis of independence between the two series. However, most of the interdependence seems to come from the correlation within the same quarter. This is confirmed by a more detailed investigation in Berndt (1991) where no significant lead and/or lag relations are found.

Real Gross Domestic Product and Consumer Sentiment

The procedure outlined above can be used to examine whether one of the two time series is systematically leading the other one. This is, for example, important in the judgment of the current state of the economy because first provisional national accounting data are usually published with a lag of at least one quarter. However, in the conduct of monetary policy more up-to-date knowledge is necessary. Such a knowledge can be retrieved from leading indicator variables. These variables should be available more quickly and should be highly correlated with the variable of interest at a lead.

We investigate whether the Consumer Sentiment Index is a leading indicator for the percentage changes in real Gross Domestic Product (GDP).[5] The raw

[5]We use data for Switzerland as published by the State Secretariat for Economic Affairs SECO.

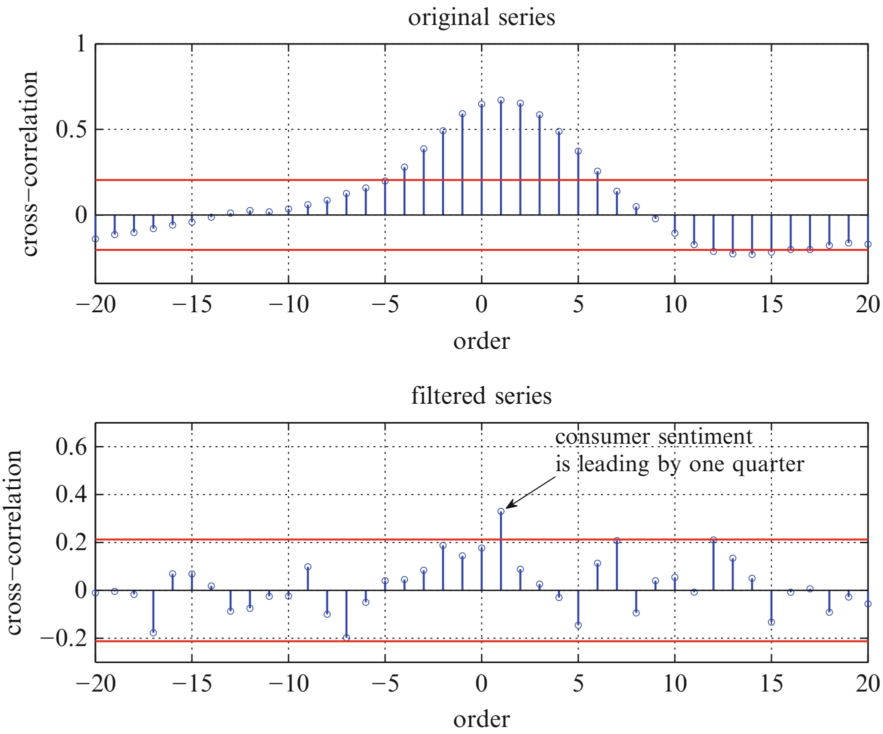

Fig. 11.3 Cross-correlations between real growth of GDP and the consumer sentiment index

cross-correlations are plotted in the upper panel of Fig. 11.3. It shows several correlations outside the conventional confidence interval. The use of this confidence interval is, however, misleading as the distribution of the raw cross-correlations depends on the autocorrelations of each series. Thus, instead we filter both time series by an AR(8) model and investigate the cross-correlations of the residuals.[6] The order of the AR model was chosen deliberately high to account for all autocorrelations. The cross-correlations of the filtered data are displayed in the lower panel of Fig. 11.3. As it turns out, only the cross-correlation which is significantly different from zero is for $h = 1$. Thus the Consumer Sentiment Index is leading the growth rate in GDP. In other words, an unexpected higher consumer sentiment is reflected in a positive change in the GDP growth rate of next quarter.[7]

[6]With quarterly data it is wise to set to order as a multiple of four to account for possible seasonal movements. As it turns out $p = 8$ is more than enough to obtain white noise residuals.

[7]During the interpretation of the cross-correlations be aware of the ordering of the variables because $\rho_{12}(1) = \rho_{21}(-1) \neq \rho_{21}(1)$.