HSB
Hochschule Bremen
City University of Applied Sciences

INSTITUTE OF WATERACOUSTICS, SONAR ENGINEERING AND SIGNAL THEORY

Stochastic Signals and Systems

Contents

5 Spectrum Estimation

5.1 Estimation of Moment Functions

5.1.1 Ergodicity

In the following the relationship between statistical averages and time averages is considered. Suppose we would like to determine the mean of a stationary stochastic process (X_t). For this purpose, we observe a large number of samples $X_t(\xi_l)$, $l = 1,\ldots,L$ and use their ensemble average

$$\hat{\mu}_X = \frac{1}{L}\sum_{l=1}^{L} X_t(\xi_l)$$

as estimate for $\mu_X = \mathrm{E}(X_t)$. However, if we have access only to a single sample $x_t = X_t(\xi)$ for each $t = 1,\ldots,N$ then we can ask whether the time average

$$\overline{x} = \frac{1}{N}\sum_{t=1}^{N} x_t$$

can be used as estimate for μ_X.

Definition:
Let (X_t) be a stationary stochastic process with mean μ_X and covariance function $c_{XX}(\tau)$. Then (X_t) is said to be

1) mean square ergodic in the mean if

$$\lim_{N\to\infty} \mathrm{E}\left(\left(\frac{1}{N}\sum_{t=1}^{N} X_t - \mu_X\right)^2\right) = 0.$$

2) mean square ergodic in the covariance function if

$$\lim_{N\to\infty} \mathrm{E}\left(\left(\frac{1}{N}\sum_{t=1}^{N}(X_{t+\tau}-\mu_X)(X_t-\mu_X)-c_{XX}(\tau)\right)^2\right)=0 \quad \forall\,\tau.$$

<u>*Theorem*</u>:

A stationary process (X_t) is mean square ergodic

1) in the mean if its covariance function satisfies

$$\sum_{\tau=-\infty}^{\infty}\left|c_{XX}(\tau)\right|<\infty,$$

2) in the covariance function if its covariance function and its fourth order cumulant function satisfy

$$\sum_{\tau=-\infty}^{\infty}\left|c_{XX}(\tau)\right|<\infty \quad \text{and} \quad \sum_{\tau=-\infty}^{\infty}\left|\kappa_{XXXX}(\tau+m,\tau,n)\right|<\infty.$$

Exercise 5.1-1:
Cumulants and cumulant functions

Corollary:

A normally distributed stationary stochastic process is mean square ergodic in the mean and in the covariance function if its covariance function is absolute summable.

Theorem:

A stationary ARMA(p,q)-Process possesses a absolute summable covariance function and is therefore mean square ergodic in the mean.

Corollary:

A stationary ARMA(p,q)-Process is mean square ergodic in the mean and in the covariance function if its white noise input process is normally distributed.

5.1.2 Estimation of the Mean

Let (X_t) be a stationary stochastic process with mean μ_X and covariance function $c_{XX}(\tau)$. For given observations $x_1,\dots,x_N$ we propose to estimate μ_X by the time average

$$\bar{x} = \frac{1}{N}\sum_{t=1}^{N} x_t$$

which due to

$$\mathrm{E}\left(\bar{X}\right) = \mathrm{E}\left(\frac{1}{N}\sum_{t=1}^{N} X_t\right) = \frac{1}{N}\sum_{t=1}^{N}\underbrace{\mathrm{E}\left(X_t\right)}_{\mu_X} = \frac{1}{N}N\mu_X = \mu_X$$

provides unbiased estimates of μ_X. The variance of the estimator $\bar{X}$ can be derived as follows.

$$\mathrm{Var}\left(\bar{X}\right) = \mathrm{E}\left(\left(\bar{X} - \mu_X\right)^2\right) = \mathrm{E}\left(\left(\frac{1}{N}\sum_{t=1}^{N}(X_t - \mu_X)\right)^2\right)$$

$$= \frac{1}{N^2}\sum_{n=1}^{N}\sum_{m=1}^{N}\mathrm{E}(X_n - \mu_X)(X_m - \mu_X)$$

$$= \frac{1}{N^2}\sum_{n=1}^{N}\sum_{m=1}^{N}c_{XX}(n-m) = \frac{1}{N^2}\sum_{\tau=-(N-1)}^{N-1}\left(N - |\tau|\right)c_{XX}(\tau)$$

$$= \frac{1}{N}\sum_{\tau=-(N-1)}^{N-1}\left(1 - \frac{|\tau|}{N}\right)c_{XX}(\tau)$$

This result is exact for all values of N. However, if

$$\sum_{\tau=-\infty}^{\infty} c_{XX}(\tau) < \infty$$

one can show, that

$$\lim_{N\to\infty} \sum_{\tau=-(N-1)}^{N-1} \left(1-\frac{|\tau|}{N}\right) c_{XX}(\tau) = \sum_{\tau=-\infty}^{\infty} c_{XX}(\tau) = C_{XX}(0),$$

i.e. the Caesaro sum converges to the unweighted sum and that consequently,

$$\operatorname{Var}\left(\bar{X}\right) = \frac{1}{N} \sum_{\tau=-(N-1)}^{N-1} \left(1-\frac{|\tau|}{N}\right) c_{XX}(\tau) \xrightarrow[N\to\infty]{} 0.$$

Thus, $\bar{X}$ is a mean square consistent estimator of μ_X.

For large N the useful approximation

$$\operatorname{Var}\left(\bar{X}\right) \approx \frac{1}{N} C_{XX}(0)$$

can be applied. This may give larger/smaller values then the usual expression $\mathrm{Var}(\bar{X}) \approx \sigma_X^2 / N$ which applies in case of uncorrelated observations.

For example, if (X_t) is a stationary AR(1)-Process, i.e.

$$X_t + aX_{t-1} = Z_t, \quad |a| < 1$$

with

$$C_{xx}(\Omega) = \frac{\sigma_Z^2}{|1 + a e^{-j\Omega}|^2} \quad \text{and} \quad c_{xx}(\tau) = \frac{\sigma_Z^2 (-a)^{|\tau|}}{1 - a^2}$$

we have

$$\mathrm{Var}(\bar{X}) \approx \frac{\sigma_X^2}{N} \cdot \frac{1-a}{1+a} = \frac{\sigma_X^2}{\tilde{N}} \quad \text{with} \quad \tilde{N} = N \cdot \frac{1+a}{1-a},$$

where $\tilde{N}$ denotes the equivalent number of uncorrelated observations which would provide the same accuracy.

Theorem:

If (X_t) is a general linear process of the form

$$X_t = \mu_X + \sum_{\tau=-\infty}^{\infty} h_\tau Z_{t-\tau},$$

where (Z_t) is a sequence of independently and identically distributed random variables with

$$\mathrm{E}(Z_t) = 0,\ \mathrm{E}(Z_t^2) < \infty \ \text{ and } \ \sum_{\tau=-\infty}^{\infty} |h_\tau| < \infty$$

then for $N \to \infty$ we have

$$\sqrt{N}\left(\bar{X} - \mu_X\right) \overset{as.}{\sim} \mathcal{N}\left(0, C_{XX}(0)\right)$$

with

$$C_{XX}(0) = \sum_{\tau=-\infty}^{\infty} c_{XX}(\tau).$$

5.1.3 Estimation of the Covariance Function

Let $x_1,\ldots,x_N$ be N consecutive observations of a stationary stochastic process (X_t) with mean μ_X and covariance function $c_{XX}(\tau)$. Then $c_{XX}(\tau)$ can be estimated by

1) the sample covariance function

a) $$\hat{c}_{XX}(\tau) = \frac{1}{N} \sum_{t=1}^{N-|\tau|} \left(x_{t+|\tau|} - \bar{x}\right)\left(x_t - \bar{x}\right), \quad |\tau| < N$$

if the mean is unknown

b) $$\hat{c}^{\mu}_{XX}(\tau) = \frac{1}{N} \sum_{t=1}^{N-|\tau|} (x_{t+|\tau|} - \mu_X)(x_t - \mu_X), \quad |\tau| < N$$

if the mean is known

2) the modified sample covariance function

a) $$\tilde{c}_{XX}(\tau) = \frac{1}{N-|\tau|}\sum_{t=1}^{N-|\tau|}\left(x_{t+|\tau|}-\bar{x}\right)\left(x_t-\bar{x}\right), \quad |\tau|<N$$

if the mean is unknown

b) $$\tilde{c}^{\mu}_{XX}(\tau) = \frac{1}{N-|\tau|}\sum_{t=1}^{N-|\tau|}(x_{t+|\tau|}-\mu_X)(x_t-\mu_X), \quad |\tau|<N$$

if the mean is known

The mean values of the estimators $\tilde{c}^{\mu}_{XX}(\tau)$ and $\hat{c}^{\mu}_{XX}(\tau)$, which assume that the mean value μ_X is given, can be determined by

$$\mathrm{E}\left(\tilde{c}_{XX}^{\mu}(\tau)\right) = \frac{1}{N-|\tau|}\sum_{t=1}^{N-|\tau|}\mathrm{E}\left((X_{t+|\tau|}-\mu_X)(X_t-\mu_X)\right)$$
$$= \frac{1}{N-|\tau|}\sum_{t=1}^{N-|\tau|} c_{XX}(\tau) = c_{XX}(\tau),$$

$$\mathrm{E}\left(\hat{c}_{XX}^{\mu}(\tau)\right) = \mathrm{E}\left(\frac{N-|\tau|}{N}\tilde{c}_{XX}^{\mu}(\tau)\right) = \left(1-\frac{|\tau|}{N}\right)c_{XX}(\tau)$$
$$= c_{XX}(\tau) + O(1/N).$$

If the covariance function $c_{XX}(\tau)$ is absolutely summable, the mean values of the estimators $\hat{c}_{XX}(\tau)$ and $\tilde{c}_{XX}(\tau)$ can be expressed by

$$\begin{aligned}
\mathrm{E}\left(\tilde{c}_{XX}(\tau)\right) &= \frac{1}{N-|\tau|}\sum_{t=1}^{N-|\tau|}\mathrm{E}\left((X_{t+|\tau|}-\bar{X})(X_t-\bar{X})\right) \\
&= \frac{1}{N-|\tau|}\sum_{t=1}^{N-|\tau|}\mathrm{E}\Big(\left((X_{t+|\tau|}-\mu_X)-(\bar{X}-\mu_X)\right) \\
&\qquad\qquad \times\left((X_t-\mu_X)-(\bar{X}-\mu_X)\right)\Big) \\
&= \frac{1}{N-|\tau|}\sum_{t=1}^{N-|\tau|}\Bigg(c_{XX}(\tau)+\mathrm{Var}(\bar{X}) \\
&\qquad\qquad -\frac{1}{N}\sum_{n=1}^{N}\left(c_{XX}(t+|\tau|-n)+c_{XX}(n-t)\right)\Bigg) \\
&= c_{XX}(\tau)+O\left(1/N\right)
\end{aligned}$$

and

$$\mathrm{E}\left(\hat{c}_{XX}(\tau)\right) = \mathrm{E}\left(\frac{N-|\tau|}{N}\tilde{c}_{XX}(\tau)\right)$$
$$= \left(1-\frac{|\tau|}{N}\right)\left(c_{XX}(\tau)+O\left(\frac{1}{N}\right)\right)$$
$$= c_{XX}(\tau)+O\left(1/N\right),$$

respectively.

Hence, we can conclude that $\tilde{c}^{\mu}_{XX}(\tau)$ is an unbiased estimator for $c_{XX}(\tau)$ and that $\hat{c}^{\mu}_{XX}(\tau)$, $\tilde{c}_{XX}(\tau)$, $\hat{c}_{XX}(\tau)$ are only asymptotically unbiased estimators for $c_{XX}(\tau)$.

We are now going to consider the variance, covariance and mean square error of the covariance function estimators.

The second order moments of the unbiased estimator $\tilde{c}^{\mu}_{XX}(\tau)$ can be expressed by

$$(N-k)(N-l)\,\mathrm{E}\left(\tilde{c}^{\mu}_{XX}(k)\tilde{c}^{\mu}_{XX}(l)\right)=$$

$$=\sum_{n=1}^{N-k}\sum_{m=1}^{N-l}\mathrm{E}\left((X_{n+k}-\mu_X)(X_n-\mu_X)(X_{m+l}-\mu_X)(X_m-\mu_X)\right)$$

$$=\sum_{n=1}^{N-k}\sum_{m=1}^{N-l}\left(c_{XX}(k)c_{XX}(l)+c_{XX}(n-m+k-l)c_{XX}(n-m)+\right.$$

$$\left.c_{XX}(n-m+k)c_{XX}(n-m-l)+\kappa_{XXXX}(n-m+k,n-m,l)\right)$$

with $0 \le k < N$ and $0 \le l < N$. Exploiting

$$\operatorname{Cov}\left(\tilde{c}^{\mu}_{xx}(k), \tilde{c}^{\mu}_{xx}(l)\right) = \operatorname{E}\left(\tilde{c}^{\mu}_{xx}(k), \tilde{c}^{\mu}_{xx}(l)\right) - c_{xx}(k)\, c_{xx}(l)$$

the covariances of $\tilde{c}^{\mu}_{xx}(\tau)$ are given by

$$(N-k)(N-l)\operatorname{Cov}\left(\tilde{c}^{\mu}_{xx}(k), \tilde{c}^{\mu}_{xx}(l)\right) =$$

$$= \sum_{n=1}^{N-k} \sum_{m=1}^{N-l} \Big(c_{xx}(n-m+k-l)\, c_{xx}(n-m)$$

$$+ c_{xx}(n-m+k)\, c_{xx}(n-m-l) + \kappa_{xxxx}(n-m+k, n-m, l) \Big).$$

After changing the variables from n and m to $t = n - m$ and n the summand depends only on t. A careful examination of the limits of n provides the result

$$(N-k)(N-l)\operatorname{Cov}\left(\tilde{c}^{\mu}_{XX}(k),\tilde{c}^{\mu}_{XX}(l)\right)=$$

$$=\sum_{t=-(N-l)+1}^{N-k-1}\phi(t)\big(c_{XX}(t+k-l)c_{XX}(t)+c_{XX}(t+k)c_{XX}(t-l)$$

$$+\kappa_{XXXX}(t+k,t,l)\big),$$

where the function $\phi(t)$ is defined by

$$\phi(t)=\begin{cases} N-k-t & t\geq 0 \\ N-k & -(l-k)\leq t\leq 0 \\ N-l-t & -(N-l)\leq t\leq -(l-k) \end{cases}.$$

Setting now $0\leq \tau=k=l<N$, we obtain

$$(N-\tau)^2\operatorname{Cov}\left(\tilde{c}^{\mu}_{XX}(\tau),\tilde{c}^{\mu}_{XX}(\tau)\right)=(N-\tau)^2\operatorname{Var}\left(\tilde{c}^{\mu}_{XX}(\tau)\right)=$$

$$= \sum_{t=-(N-\tau)+1}^{N-\tau-1} (N-\tau-|t|)\times$$

$$\times\left(c_{XX}^2(t)+c_{XX}(t+\tau)c_{XX}(t-\tau)+\kappa_{XXXX}(t+\tau,t,\tau)\right).$$

Consequently, if the covariance function and the fourth order cumulant function of (X_t) are absolute summable the following orders of convergence hold for $N \to \infty$.

$$\operatorname{Cov}\left(\tilde{c}_{XX}^{\mu}(k),\tilde{c}_{XX}^{\mu}(l)\right)=O(1/N),\quad \operatorname{Var}\left(\tilde{c}_{XX}^{\mu}(\tau)\right)=O(1/N)$$

and

$$\operatorname{MSE}\left(\tilde{c}_{XX}^{\mu}(\tau)\right)=\underbrace{\operatorname{Var}\left(\tilde{c}_{XX}^{\mu}(\tau)\right)}_{=O(1/N)}+\underbrace{b^2\left(\tilde{c}_{XX}^{\mu}(\tau)\right)}_{=0}=O(1/N).$$

All the above results are easily modified for the biased estimator $\hat{c}^{\mu}_{XX}(\tau)$. Since

$$\hat{c}^{\mu}_{XX}(\tau) = \frac{(N-\tau)}{N}\tilde{c}^{\mu}_{XX}(\tau) \quad 0 \le \tau < N$$

the covariance, variance and mean square error can be simply derive as follows.

$$\operatorname{Cov}\left(\hat{c}^{\mu}_{XX}(k), \hat{c}^{\mu}_{XX}(l)\right) = \operatorname{Cov}\left(\frac{(N-k)}{N}\tilde{c}^{\mu}_{XX}(k), \frac{(N-l)}{N}\tilde{c}^{\mu}_{XX}(l)\right)$$

$$= \underbrace{\frac{(N-k)(N-l)}{N^2}}_{=O(1)} \underbrace{\operatorname{Cov}\left(\tilde{c}^{\mu}_{XX}(k), \tilde{c}^{\mu}_{XX}(l)\right)}_{=O(1/N)} = O\left(\frac{1}{N}\right),$$

$$\operatorname{Var}\left(\hat{c}^{\mu}_{XX}(\tau)\right) = \operatorname{Var}\left(\frac{N-\tau}{N}\hat{c}^{\mu}_{XX}(\tau)\right) = \underbrace{\left(\frac{N-\tau}{N}\right)^2}_{=O(1)} \underbrace{\operatorname{Var}\left(\hat{c}^{\mu}_{XX}(\tau)\right)}_{=O(1/N)}$$

$$= O\left(\frac{1}{N}\right)$$

and

$$\operatorname{MSE}\left(\hat{c}^{\mu}_{XX}(\tau)\right) = \underbrace{\operatorname{Var}\left(\hat{c}^{\mu}_{XX}(\tau)\right)}_{=O(1/N)} + \underbrace{b^2\left(\hat{c}^{\mu}_{XX}(\tau)\right)}_{=O\left(1/N^2\right)} = O\left(\frac{1}{N}\right).$$

Although the evaluation of the estimators $\tilde{c}_{XX}(\tau)$ and $\hat{c}_{XX}(\tau)$ is more difficult, cumbersome calculations demonstrate that similar asymptotic properties can be deduced.

Theorem:

If (X_t) is a general linear process of the form

$$X_t = \mu_X + \sum_{\tau=-\infty}^{\infty} h_\tau Z_{t-\tau},$$

where (Z_t) is a sequence of independently and identically distributed random variables with

$$\mathrm{E}(Z_t) = 0,\ \mathrm{E}(Z_t^2) < \infty,\ \mathrm{E}(Z_t^4) < \infty \text{ and } \sum_{\tau=-\infty}^{\infty} |h_\tau| < \infty$$

then for $N \to \infty$ we have

$$\sqrt{N}\left(\hat{c}_{XX}^{\mu}(0) - c_{XX}(0), \ldots, \hat{c}_{XX}^{\mu}(n) - c_{XX}(n)\right)^T \overset{as.}{\sim} \mathcal{N}_{n+1}(\mathbf{0}, \mathbf{\Sigma})$$

with

$$\mathbf{\Sigma} = \left(N \cdot \mathrm{Cov}\left(\hat{c}_{XX}^{\mu}(k), \hat{c}_{XX}^{\mu}(l)\right)\right)_{k=0,\ldots,n;\, l=0,\ldots,n}.$$

5.1.4 Estimation of the Cross-Covariance Function

Let $(x_1,y_1),\ldots,(x_N,y_N)$ be N successive realizations of a bi-variate stochastic process (X_t, Y_t) with mean $\boldsymbol{\mu} = (\mu_X, \mu_Y)^T$, covariance functions $c_{XX}(\tau), c_{YY}(\tau)$ and cross-covariance function $c_{XY}(\tau)$. Then $c_{XY}(\tau)$ can be estimated by

1) the sample cross-covariance function

a) $$\hat{c}_{XY}(\tau) = \begin{cases} \dfrac{1}{N}\displaystyle\sum_{t=1}^{N-\tau}(x_{t+\tau}-\bar{x})(y_t-\bar{y}) & 0 \le \tau < N \\ \dfrac{1}{N}\displaystyle\sum_{t=-\tau+1}^{N}(x_{t+\tau}-\bar{x})(y_t-\bar{y}) & -N < \tau < 0 \end{cases}$$

with $$\bar{x} = \frac{1}{N}\sum_{t=1}^{N} x_t, \quad \bar{y} = \frac{1}{N}\sum_{t=1}^{N} y_t$$

b) $$\hat{c}^{\mu}_{XY}(\tau) = \begin{cases} \dfrac{1}{N}\sum_{t=1}^{N-\tau}(x_{t+\tau}-\mu_X)(y_t-\mu_Y) & 0 \le \tau < N \\ \dfrac{1}{N}\sum_{t=-\tau+1}^{N}(x_{t+\tau}-\mu_X)(y_t-\mu_Y) & -N < \tau < 0 \end{cases}$$

2) the modified sample covariance function

a) $$\tilde{c}_{XY}(\tau) = \begin{cases} \dfrac{1}{N-\tau}\sum_{t=1}^{N-\tau}(x_{t+\tau}-\bar{x})(y_t-\bar{y}) & 0 \le \tau < N \\ \dfrac{1}{N+\tau}\sum_{t=-\tau+1}^{N}(x_{t+\tau}-\bar{x})(y_t-\bar{y}) & -N < \tau < 0 \end{cases}$$

with $$\bar{x} = \frac{1}{N}\sum_{t=1}^{N} x_t, \quad \bar{y} = \frac{1}{N}\sum_{t=1}^{N} y_t$$

b) $$\tilde{c}^{\mu}_{XY}(\tau) = \begin{cases} \dfrac{1}{N-\tau}\displaystyle\sum_{t=1}^{N-\tau}\left(x_{t+\tau}-\mu_X\right)\left(y_t-\mu_Y\right) & 0 \le \tau < N \\ \dfrac{1}{N+\tau}\displaystyle\sum_{t=-\tau+1}^{N}\left(x_{t+\tau}-\mu_X\right)\left(y_t-\mu_Y\right) & -N < \tau < 0 \end{cases}$$

The mean values of the estimators $\tilde{c}^{\mu}_{XY}(\tau)$ and $\hat{c}^{\mu}_{XY}(\tau)$, which assume that the mean values μ_X, μ_Y are known, are given exemplarily for $0 \le \tau < N$ by

$$\begin{aligned} \mathrm{E}\left(\tilde{c}^{\mu}_{XY}(\tau)\right) &= \frac{1}{N-\tau}\sum_{t=1}^{N-\tau}\mathrm{E}\left((X_{t+\tau}-\mu_X)(Y_t-\mu_Y)\right) \\ &= \frac{1}{N-\tau}\sum_{t=1}^{N-\tau}c_{XY}(\tau) = c_{XY}(\tau) \end{aligned}$$

and

$$\mathrm{E}\left(\hat{c}_{XY}^{\mu}(\tau)\right) = \mathrm{E}\left(\frac{N-\tau}{N}\tilde{c}_{XY}^{\mu}(\tau)\right)$$

$$= \left(1-\frac{\tau}{N}\right)c_{XY}(\tau) = c_{XY}(\tau) + O\left(\frac{1}{N}\right).$$

If the cross-covariance function $c_{XY}(\tau)$ is absolutely summable, the mean values of the estimators $\tilde{c}_{XY}(\tau)$, $\hat{c}_{XY}(\tau)$ for $0 \le \tau < N$ can be expressed by

$$\mathrm{E}\left(\tilde{c}_{XY}(\tau)\right) = \frac{1}{N-\tau}\sum_{t=1}^{N-\tau}\mathrm{E}\left((X_{t+\tau}-\bar{X})(Y_t-\bar{Y})\right)$$

$$= \frac{1}{N-\tau}\sum_{t=1}^{N-\tau}\mathrm{E}\left(\left((X_{t+\tau}-\mu_X)-(\bar{X}-\mu_X)\right)\left((Y_t-\mu_Y)-(\bar{Y}-\mu_Y)\right)\right)$$

$$= \frac{1}{N-\tau}\sum_{t=1}^{N-\tau}\left(c_{XY}(\tau)+\frac{1}{N^2}\sum_{k=1}^{N}\sum_{l=1}^{N}c_{XY}(k-l)\right.$$

$$\left.+\frac{1}{N}\sum_{n=1}^{N}\left(c_{XY}(t+\tau-n)+c_{XY}(n-t)\right)\right)$$

$$= c_{XY}(\tau)+\frac{1}{N}\sum_{n=-N+1}^{N-1}\left(1-\frac{|n|}{N}\right)c_{XY}(n)$$

$$+\frac{1}{(N-\tau)N}\sum_{t=1}^{N-\tau}\sum_{n=1}^{N}\left(c_{XY}(t+\tau-n)+c_{XY}(n-t)\right)$$

$$= c_{XY}(\tau)+O\left(1/N\right)$$

and

$$\mathrm{E}\left(\hat{c}_{XY}(\tau)\right)=\mathrm{E}\left(\frac{N-\tau}{N}\tilde{c}_{XY}(\tau)\right)$$
$$=\left(1-\frac{\tau}{N}\right)\left(c_{XY}(\tau)+O\left(\frac{1}{N}\right)\right)=c_{XY}(\tau)+O\left(\frac{1}{N}\right),$$

respectively.

As for the covariance function estimators, we can conclude for the cross-covariance estimators that $\tilde{c}^{\mu}_{XY}(\tau)$ provides unbiased and $\hat{c}^{\mu}_{XY}(\tau)$, $\tilde{c}_{XY}(\tau)$, $\hat{c}_{XY}(\tau)$ only asymptotically unbiased estimates of $c_{XY}(\tau)$.

The variance, covariance and mean square error of the proposed cross-covariance function estimators will be investigated in the following.

We first consider the second order moments of the un-biased estimator $\tilde{c}^{\mu}_{XY}(\tau)$ which are given by

$$(N-k)(N-l)\,\mathrm{E}\left(\tilde{c}^{\mu}_{XY}(k)\tilde{c}^{\mu}_{XY}(l)\right)=$$

$$=\sum_{n=1}^{N-k}\sum_{m=1}^{N-l}\mathrm{E}\left((X_{n+k}-\mu_X)(Y_n-\mu_Y)(X_{m+l}-\mu_X)(Y_m-\mu_Y)\right)$$

$$=\sum_{n=1}^{N-k}\sum_{m=1}^{N-l}\Big(c_{XY}(k)c_{XY}(l)+c_{XX}(n-m+k-l)c_{YY}(n-m)$$

$$+c_{XY}(n-m+k)c_{XY}(m-n+l)+\kappa_{XYXY}(n-m+k,n-m,l)\Big)$$

for $0\le k<N$ and $0\le l<N$. Hence, utilization of

$$\mathrm{Cov}\left(\tilde{c}^{\mu}_{XY}(k),\tilde{c}^{\mu}_{XY}(l)\right)=\mathrm{E}\left(\tilde{c}^{\mu}_{XY}(k),\tilde{c}^{\mu}_{XY}(l)\right)-c_{XY}(k)c_{XY}(l)$$

provides the covariances

$$(N-k)(N-l)\mathrm{Cov}\left(\tilde{c}^{\mu}_{XY}(k),\tilde{c}^{\mu}_{XY}(l)\right)=$$

$$=\sum_{n=1}^{N-k}\sum_{m=1}^{N-l}\Big(c_{XX}(n-m+k-l)c_{YY}(n-m)$$

$$+c_{XY}(n-m+k)c_{XY}(m-n+l)+\kappa_{XYXY}(n-m+k,n-m,l)\Big)$$

$$=\sum_{t=-(N-l)+1}^{N-k-1}\phi(t)\Big(c_{XX}(t+k-l)c_{YY}(t)+c_{XY}(t+k)c_{XY}(t-l)$$

$$+\kappa_{XYXY}(t+k,t,l)\Big)$$

with

$$\phi(t)=\begin{cases} N-k-t & t\geq 0 \\ N-k & -(l-k)\leq t\leq 0 \\ N-l-t & -(N-l)\leq t\leq -(l-k) \end{cases},$$

where the double sum is converted into a single sum by exploiting the arguments already applied for deriving the covariances of the covariance function estimators.

Setting now $0 \leq \tau = k = l < N$, we obtain

$$(N-\tau)^2 \mathrm{Cov}\left(\tilde{c}^{\mu}_{XY}(\tau), \tilde{c}^{\mu}_{XY}(\tau)\right) = (N-\tau)^2 \mathrm{Var}\left(\tilde{c}^{\mu}_{XY}(\tau)\right) =$$

$$= \sum_{t=-(N-\tau)+1}^{N-\tau-1} (N-\tau-|t|)\big(c_{XX}(t)c_{YY}(t) + c_{XY}(t+\tau)c_{XY}(t-\tau)$$

$$+\kappa_{XYXY}(t+\tau, t, \tau)\big).$$

Consequently, if the covariance functions, cross-covariance function and the fourth order cross-cumulant function of (X_t, Y_t) are absolute summable the following orders of convergence hold for $N \to \infty$.

$$\mathrm{Cov}\left(\tilde{c}^{\mu}_{XY}(k),\tilde{c}^{\mu}_{XY}(l)\right)=O\left(\frac{1}{N}\right),\quad \mathrm{Var}\left(\tilde{c}^{\mu}_{XY}(\tau)\right)=O\left(\frac{1}{N}\right)$$

and

$$\mathrm{MSE}\left(\tilde{c}^{\mu}_{XY}(\tau)\right)=\underbrace{\mathrm{Var}\left(\tilde{c}^{\mu}_{XY}(\tau)\right)}_{=O(1/N)}+\underbrace{b^{2}\left(\tilde{c}^{\mu}_{XY}(\tau)\right)}_{=0}=O\left(\frac{1}{N}\right).$$

Since

$$\hat{c}^{\mu}_{XY}(\tau)=\frac{(N-\tau)}{N}\tilde{c}^{\mu}_{XY}(\tau)\quad 0\leq\tau<N$$

the asymptotic behavior of the covariance, variance and mean square error of the biased estimator $\hat{c}^{\mu}_{XY}(\tau)$ can be easily deduced from the previous results. Thus

$$\operatorname{Cov}\left(\hat{c}^{\mu}_{XY}(k),\hat{c}^{\mu}_{XY}(l)\right)=\underbrace{\frac{(N-k)(N-l)}{N^2}}_{=O(1)}\underbrace{\operatorname{Cov}\left(\tilde{c}^{\mu}_{XY}(k),\tilde{c}^{\mu}_{XY}(l)\right)}_{=O(1/N)}=O\left(\frac{1}{N}\right),$$

$$\operatorname{Var}\left(\hat{c}^{\mu}_{XY}(\tau)\right)=\underbrace{\left((N-\tau)/N\right)^2}_{=O(1)}\underbrace{\operatorname{Var}\left(\hat{c}^{\mu}_{XY}(\tau)\right)}_{=O(1/N)}=O\left(\frac{1}{N}\right)$$

and

$$\operatorname{MSE}\left(\hat{c}^{\mu}_{XY}(\tau)\right)=\underbrace{\operatorname{Var}\left(\hat{c}^{\mu}_{XY}(\tau)\right)}_{=O(1/N)}+\underbrace{b^2\left(\hat{c}^{\mu}_{XY}(\tau)\right)}_{=O(1/N^2)}=O\left(\frac{1}{N}\right).$$

The evaluation of the estimators $\tilde{c}_{XY}(\tau)$ and $\hat{c}_{XY}(\tau)$ is more difficult. However, again laborious calculations show that similar asymptotic properties can be stated.

5.2 Nonparametric Spectrum Estimation

5.2.1 Finite Discrete-Time Fourier Transform

Let (X_t) be a stationary stochastic process. The samples taken at the time instances $t = 0,1,...,T-1$ are supposed to be modelled by the random Variables $X_0,\ldots,X_{T-1}$.

Hence,

$$X^T(\Omega) = \sum_{t=0}^{T-1} X_t \, e^{-j\Omega t}$$

denotes the finite discrete-time Fourier transform (finite DTFT) of the model. It is called discrete Fourier transform (DFT) if only the discrete frequencies

$$\Omega_k = 2\pi k/T\,,\;\; k = 0,1,\ldots,T-1,$$

are considered. The DFT

$$X^T(2\pi k/T) = \sum_{t=0}^{T-1} X_t \, e^{-j(2\pi k/T)t}, \quad k = 0,1,\ldots,T-1$$

can be efficiently determined by the so-called fast Fourier transform (FFT). The corresponding inverse finite DTFT and inverse DFT are given by

$$X_t = \frac{1}{2\pi} \int_{-\pi}^{\pi} X^T(\Omega) e^{j\Omega t} d\Omega$$

and

$$X_t = \frac{1}{T} \sum_{k=0}^{T-1} X^T(2\pi k/T) e^{j(2\pi k/T)t} \quad \text{with} \quad X_t = X_{t+T}$$

respectively.

Let w_t denote a window of bounded variation with

$$w_t = 0 \quad \forall t \notin [0,1] \quad \text{and} \quad \int_0^1 w_t^2 \, dt = 1.$$

The windowed DTFT is defined by

$$X_w^T(\Omega) = \sum_{t=0}^{T-1} w_{t/T} X_t e^{-j\Omega t}$$

and windowed DFT accordingly by

$$X_w^T(2\pi k / T) = \sum_{t=0}^{T-1} w_{t/T} X_t e^{-j(2\pi k/T)t}.$$

Now, we suppose that the set $\{0,1,\ldots,T-1\}$ can be partitioned into L disjoint sets of length $T' = T/L$. Hence, the finite DTFT of the resulting L consecutive data pieces can be expressed by

$$X_w^{T'}(\Omega, l) = \sum_{t=0}^{T'-1} w_{t/T'} X_{(l-1)T'+t}\, e^{-j\Omega t}, \quad l = 1, \ldots, L.$$

Circular-Symmetric Complex Normal Distribution

A complex valued random vector $\mathbf{X} = (X_1, \ldots, X_n)^T$ is said to be circular-symmetric complex normally distributed with mean vector $\boldsymbol{\mu}_{\mathbf{X}}$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{XX}}$, i.e.

$$\mathbf{X} \sim \mathcal{CN}_n(\boldsymbol{\mu}_{\mathbf{X}}, \boldsymbol{\Sigma}_{\mathbf{XX}}),$$

if the corresponding real valued random vector

$$\begin{pmatrix} \mathrm{Re}(\mathbf{X}) \\ \mathrm{Im}(\mathbf{X}) \end{pmatrix} = \left(\mathrm{Re}(X_1), \ldots, \mathrm{Re}(X_n), \mathrm{Im}(X_1), \ldots, \mathrm{Im}(X_n)\right)^T$$

is distributed as

$$\begin{pmatrix} \mathrm{Re}(\mathbf{X}) \\ \mathrm{Im}(\mathbf{X}) \end{pmatrix} \sim \mathcal{N}_{2n}\left(\begin{pmatrix} \mathrm{Re}(\boldsymbol{\mu}_{\mathbf{X}}) \\ \mathrm{Im}(\boldsymbol{\mu}_{\mathbf{X}}) \end{pmatrix}, \frac{1}{2} \begin{pmatrix} \mathrm{Re}(\boldsymbol{\Sigma}_{\mathbf{XX}}) & -\mathrm{Im}(\boldsymbol{\Sigma}_{\mathbf{XX}}) \\ \mathrm{Im}(\boldsymbol{\Sigma}_{\mathbf{XX}}) & \mathrm{Re}(\boldsymbol{\Sigma}_{\mathbf{XX}}) \end{pmatrix} \right),$$

where

$$\mathrm{E}(\mathbf{X}) = \mathrm{E}\big(\mathrm{Re}(\mathbf{X}) + j\,\mathrm{Im}(\mathbf{X})\big) = \mathrm{Re}(\boldsymbol{\mu}_{\mathbf{X}}) + j\,\mathrm{Im}(\boldsymbol{\mu}_{\mathbf{X}}) = \boldsymbol{\mu}_{\mathbf{X}}$$

$$\mathrm{E}\left((\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^H\right) = \boldsymbol{\Sigma}_{\mathbf{XX}} \quad \text{(covariance matrix)}$$

and

$$\mathrm{E}\left((\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T\right) = \mathbf{0}. \quad \text{(pseudo covariance matrix)}$$

Suppose $\boldsymbol{\Sigma}_{\mathbf{XX}}$ is regular, then the density function of the circular-symmetric complex normally distributed random vector $\mathbf{X}$ is given by

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\pi^n \det(\boldsymbol{\Sigma}_{\mathbf{XX}})} \exp\left(-(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^H \boldsymbol{\Sigma}_{\mathbf{XX}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})\right)$$

For $n = 1$, i.e. $\mathbf{X} = X \sim \mathcal{CN}(\mu_X, \sigma_X^2)$ with

$$\mathrm{E}(X) = \mathrm{E}\left(\mathrm{Re}(X) + j\,\mathrm{Im}(X)\right) = \mathrm{Re}(\mu_X) + j\,\mathrm{Im}(\mu_X) = \mu_X$$

$$\mathrm{E}\left((X - \mu_X)(X - \mu_X)^*\right) = \mathrm{E}\left(\left|X - \mu_X\right|^2\right) = \sigma_X^2$$

$$\mathrm{E}\left((X - \mu_X)(X - \mu_X)\right) = \mathrm{E}\left((X - \mu_X)^2\right) = 0,$$

$\mathrm{Re}(X)$ and $\mathrm{Im}(X)$ are independent and normally distributed random variables with

$$\mathrm{Re}(X) \sim \mathcal{N}\left(\mathrm{Re}(\mu_X), \sigma_X^2/2\right), \quad \mathrm{Im}(X) \sim \mathcal{N}\left(\mathrm{Im}(\mu_X), \sigma_X^2/2\right).$$

Exercise 5.2-1:

Verify the density function of a circular-symmetric complex normally distributed random vector

Exercise 5.2-2:

Show the independence of Re(X) and Im(X)

Distributional Properties of the finite DTFT

Assumption:

(X_t) is a discrete time strictly stationary stochastic process whose cumulant functions satisfy

$$\sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{\infty} \left| \kappa_{X \ldots X}(\tau_1, \ldots, \tau_{k-1}) \right| < \infty$$

for all $k = 2,3,\ldots$

Theorem: (properties of DTFT)

Suppose (X_t) satisfies the assumption above.

1) $X^T(\Omega)$ $(0 < \Omega < \pi)$ is asymptotically complex normally distributed with mean zero and variance $TC_{XX}(\Omega)$.

$X^T(0)$ is asymptotically normally distributed with mean $T\mu_X$ and variance $TC_{XX}(0)$.

$X^T(\pi)$ is asymptotically normally distributed with mean zero and variance $TC_{XX}(\pi)$.

2) For $0 \le \Omega_m \le \pi$, $m = 1,\ldots,M$ and $|\Omega_m - \Omega_n| \ge 2\pi/T$, $m \ne n$ the $X^T(\Omega_1),\ldots,X^T(\Omega_M)$ are asymptotically independent random variables.

3) For L successive data pieces of length T', i.e. $T = LT'$, the $X^{T'}(\Omega,1),\ldots,X^{T'}(\Omega,L)$ are asymptotically independent random variables.

<u>Theorem:</u> (properties of windowed DTFT)

Suppose (X_t) satisfies the assumption above and the window w_t is of bounded variation with $\int_0^1 w_t^2\,dt = 1$.

1) $X_w^T(\Omega)$ $(0 < \Omega < \pi)$ is asymptotically complex normally distributed with mean zero and variance $TC_{XX}(\Omega)$.

   $X_w^T(0)$ is asymptotically normally distributed with mean $T\mu_X \int_0^1 w_t\,dt$ and variance $TC_{XX}(0)$.

   $X_w^T(\pi)$ is asymptotically normally distributed with mean zero and variance $TC_{XX}(\pi)$.

2) For $0 \le \Omega_m \le \pi$, $m = 1, \dots, M$ and $|\Omega_m - \Omega_n| \ge 2\pi/T$, $m \ne n$ the $X_w^T(\Omega_1), \dots, X_w^T(\Omega_M)$ are asymptotically independent random variables.

3) For L successive data pieces of length T', i.e. $T = LT'$, the $X_w^{T'}(\Omega, 1), \dots, X_w^{T'}(\Omega, L)$ are asymptotically independent random variables.

Remark:
In case that (X_t) is a Gauss process its (windowed) finite discrete-time Fourier transform is exactly normally distributed with mean and variance asymptotically given by the results stated in preceding theorems.

5.2.2 Periodogram

Let (X_t) be a zero mean stationary stochastic process with an existing but unknown power spectral density function $C_{XX}(\Omega)$ which shall be estimated.

Now, if (X_t) can be observed at the time instances $t = 0$, $1,\ldots,T-1$ the distributional properties of the finite DTFT suggest the periodogram defined by

$$I_{XX}^{T}(\Omega) = \frac{1}{T}\left|X^{T}(\Omega)\right|^{2} = \frac{1}{T}\left|\sum_{t=0}^{T-1} X_t\, e^{-j\Omega t}\right|^{2}$$

as an suitable estimator for $C_{XX}(\Omega)$.

Exercise 5.2-3:
Periodogram and Fourier transformed sample covariance function

Theorem: *(moment properties of the periodogram)*

Let (X_t) be a zero mean stationary stochastic process.

1) If the covariance function of (X_t) satisfies

$$\sum_{\tau=-\infty}^{\infty} \left| c_{XX}(\tau) \right| < \infty,$$

i.e. $C_{XX}(\Omega)$ is continuous, then

$$\mathrm{E} I_{XX}^{T}(\Omega) = \frac{1}{2\pi T} \int_{-\pi}^{\pi} \left(\frac{\sin\left((\Omega-\Lambda)T/2\right)}{\sin\left((\Omega-\Lambda)/2\right)} \right)^{2} C_{XX}(\Lambda)\, d\Lambda,$$

$$\mathrm{E} I_{XX}^{T}(\Omega) \xrightarrow[T\to\infty]{} C_{XX}(\Omega).$$

2) If the covariance function of (X_t) satisfies

$$\sum_{\tau=-\infty}^{\infty} |\tau| \left| c_{XX}(\tau) \right| < \infty,$$

i.e. $C_{XX}(\Omega)$ is continuous differentiable, then
$\mathrm{E} I_{XX}^T(\Omega) = C_{XX}(\Omega) + O(1/T)$.

3) If the cumulant functions of (X_t) satisfy

$$\sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{\infty} \left(1+|\tau_n|\right) \left|\kappa_{X\ldots X}(\tau_1,\ldots,\tau_{k-1})\right| < \infty$$

for $n = 1,\ldots,k-1$ when $k = 2,3,\ldots$, then

$$\operatorname{Cov}\left(I_{XX}^T(\Omega), I_{XX}^T(\Lambda)\right) = \left(\left(\frac{\sin((\Omega+\Lambda)T/2)}{T\sin((\Omega+\Lambda)/2)}\right)^2 + \left(\frac{\sin((\Omega-\Lambda)T/2)}{T\sin((\Omega-\Lambda)/2)}\right)^2\right) C_{XX}^2(\Omega) + O\left(\frac{1}{T}\right).$$

Exercise 5.2-4:
Proof of the Theorem if (X_t) is a Gauss process

Corollary:

If $\Omega = 2\pi n/T$ and $\Lambda = 2\pi m/T$ with $n, m = 0,1,...,T-1$ the moments stated in the previous theorem can be simplified to

$$\mathrm{Cov}\left(I_{XX}^T(\Omega), I_{XX}^T(\Lambda)\right) = O(1/T) \qquad \Omega \neq \Lambda$$

$$\mathrm{Var}\left(I_{XX}^T(\Omega)\right) = \begin{cases} C_{XX}^2(\Omega) + O(1/T) & \Omega \neq 0, \pi \\ 2C_{XX}^2(\Omega) + O(1/T) & \Omega = 0, \pi \end{cases}.$$

Assuming that the time interval $[0, T)$ is divided in L disjoint pieces of length $T' = T/L$ the periodogram of the resulting L subsequent data pieces can be expressed by

$$I_{XX}^{T'}(\Omega, l) = \frac{1}{T'}\left|X^{T'}(\Omega, l)\right|^2 = \frac{1}{T'}\left|\sum_{t=0}^{T'-1} X_{(l-1)T'+t}\, e^{-j\Omega t}\right|^2, \quad l = 1, \ldots, L.$$

Theorem: (distributional properties of the periodogram)

Suppose (X_t) is a zero mean stochastic process that satisfies the assumption stated on p. 43.

1) $I_{XX}^T(\Omega)$ $(0 < \Omega < \pi)$ is up to the factor $C_{XX}(\Omega)/2$ asymptotically chi-square distributed with two degrees of freedom.

   $I_{XX}^T(0)$ is up to the factor $C_{XX}(0)$ asymptotically chi-square distributed with one degree of freedom.

   $I_{XX}^T(\pi)$ is up to the factor $C_{XX}(\pi)$ asymptotically chi-square distributed with one degree of freedom.

2) For $0 \le \Omega_m \le \pi$, $m = 1,\ldots,M$ and $\left|\Omega_m - \Omega_n\right| \ge 2\pi/T$, $m \ne n$ the $I_{XX}^T(\Omega_1),\ldots,I_{XX}^T(\Omega_M)$ are asymptotically independent random variables.

3) For L successive data pieces of length T', i.e. $T = LT'$, the $I_{XX}^{T'}(\Omega,1),\ldots,I_{XX}^{T'}(\Omega,L)$ are asymptotically independent random variables.

Remark:
Similar results as those stated in the previous theorems can also be derived for the windowed periodogram

$$I_{X_w X_w}^T(\Omega) = \frac{1}{T}\left|X_w^T(\Omega)\right|^2 = \frac{1}{T}\left|\sum_{t=0}^{T-1} w_{t/T} X_t \, e^{-j\Omega t}\right|^2 .$$

5.2.3 Smoothing and Averaging of Periodograms

Although the periodogram is asymptotically unbiased it is due to its variance behavior an inconsistent and consequently an inadequate estimator for $C_{XX}(\Omega)$.

However, the distributional properties of the periodogram immediately suggest the following two improvements.

Smoothing of the Periodogram

$$\tilde{C}_{XX}(\Omega)=\frac{1}{M}\sum_{m=1}^{M} I_{XX}^{T}(\Omega_m)=\frac{1}{MT}\sum_{m=1}^{M}\left|X^{T}(\Omega_m)\right|^2,\ \ \Omega_m=\frac{2\pi k_m}{T}\in N_{\Omega},$$

where N_{Ω} is a neighborhood of Ω, i.e. a set of frequencies Ω_m which are located around Ω symmetrically.

For $0 < \Omega < \pi$ the mean and variance of the smoothed periodogram are given by

$$E\left(\tilde{C}_{XX}(\Omega)\right) = \frac{1}{M}\sum_{m=1}^{M} E\left(I_{XX}^{T}(\Omega_m)\right) = \frac{1}{M}\sum_{m=1}^{M} C_{XX}(\Omega_m) + O(1/T)$$

and

$$\mathrm{Var}\left(\tilde{C}_{XX}(\Omega)\right) = \frac{1}{M^2}\sum_{m}^{M}\sum_{n=1}^{M} \mathrm{Cov}\left(I_{XX}^{T}(\Omega_m), I_{XX}^{T}(\Omega_n)\right)$$
$$= \frac{1}{M^2}\sum_{m=1}^{M} C_{XX}^{2}(\Omega_m) + O(1/T).$$

Suppose $C_{XX}(\Omega)$ is sufficiently smooth over N_Ω we can write for large T approximately

$$E\left(\tilde{C}_{XX}(\Omega)\right) \approx C_{XX}(\Omega) \quad \text{and} \quad \mathrm{Var}\left(\tilde{C}_{XX}(\Omega)\right) \approx C_{XX}^{2}(\Omega)/M.$$

Averaging over Periodograms

$$\hat{C}_{XX}(\Omega) = \frac{1}{L}\sum_{l=1}^{L} I_{XX}^{T'}(\Omega,l) = \frac{1}{LT}\sum_{l=1}^{L}\left|X^{T'}(\Omega,l)\right|^2, \quad T = LT',$$

where L denotes the number of consecutive data pieces of length T' and $I_{XX}^{T'}(\Omega,l)$, $l = 1,\ldots,L$ represent the corresponding periodograms.

For $0 < \Omega < \pi$ the mean and variance of the averaged periodograms are given by

$$\mathrm{E}\left(\hat{C}_{XX}(\Omega)\right) = \frac{1}{L}\sum_{l=1}^{L}\mathrm{E}\left(I_{XX}^{T'}(\Omega,l)\right) = \frac{1}{L}\sum_{l=1}^{L}\left(C_{XX}(\Omega) + O(1/T')\right)$$

$$= C_{XX}(\Omega) + O(1/T') = C_{XX}(\Omega) + O(L/T)$$

and

$$\operatorname{Var}\left(\hat{C}_{XX}(\Omega)\right) = \frac{1}{L^2}\sum_{l=1}^{L}\sum_{k=1}^{L}\operatorname{Cov}\left(I_{XX}^{T'}(\Omega,l), I_{XX}^{T'}(\Omega,k)\right)$$

$$= \frac{1}{L^2}\sum_{l=1}^{L} C_{XX}^2(\Omega) + O(1/T') = C_{XX}^2(\Omega)/L + O(L/T).$$

For a suitably chosen number of data pieces L we can write for large T approximately

$$\mathrm{E}\left(\hat{C}_{XX}(\Omega)\right) \approx C_{XX}(\Omega) \quad \text{and} \quad \operatorname{Var}\left(\hat{C}_{XX}(\Omega)\right) \approx C_{XX}^2(\Omega)/L.$$

Remark:
Smoothing of a periodogram and averaging of periodograms do not provide consistent estimators. Nevertheless, they allow to control the variances of the estimators in a desired manner by properly selecting M and L.

5.2.4 Consistent Spectrum Estimation

Moreover, the estimators proposed in the previous Chapter might become consistent if the number M used for smoothing or L used for averaging is growing suitably to infinity as the number of observations T tends to infinity.

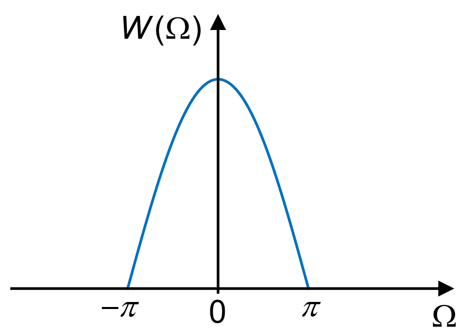
Smoothing of the Periodogram

Let $W(\Omega)$ denote a real valued and even spectral window of bounded variation and finite support $(-\pi, \pi)$ with

$$\int_{-\infty}^{\infty} W(\Omega)\, d\Omega / 2\pi = 1$$

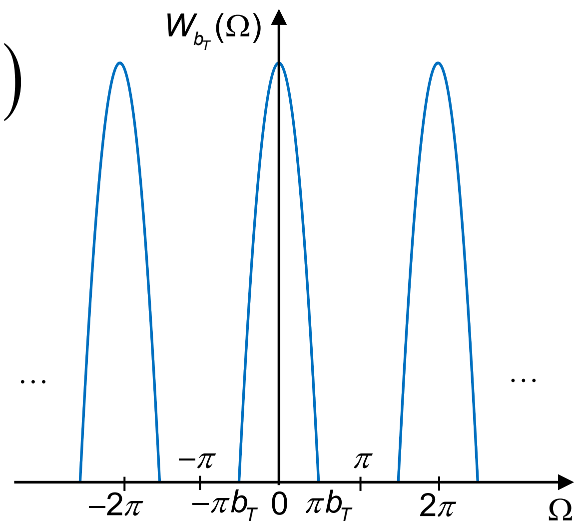
and

$$\int_{-\infty}^{\infty} |W(\Omega)|\, d\Omega < \infty .$$



Then, we can define the 2π-periodic spectral window

$$W_{b_T}(\Omega) = \frac{1}{b_T} \sum_{k=-\infty}^{\infty} W\left((\Omega - 2\pi k)/b_T\right)$$



with scalable bandwidth and

$$\int_{-\pi}^{\pi} W_{b_T}(\Omega)\, d\Omega / 2\pi = 1$$

<u>*Assumption:*</u>

The bandwidth b_T obeys

$$b_T \xrightarrow[T\to\infty]{} 0 \quad \text{such that} \quad b_T T \xrightarrow[T\to\infty]{} \infty$$

applies for the time-bandwidth product.

Example:

Let $b_T \propto 1/\sqrt{T}$, i.e. $b_T \xrightarrow[T\to\infty]{} 0$, then

$$b_T T \propto T/\sqrt{T} = \sqrt{T} \xrightarrow[T\to\infty]{} \infty .$$

Now, we are going to estimate $C_{XX}(\Omega)$ by

$$\tilde{C}_{XX}^{T}(\Omega) = \frac{1}{T}\sum_{m=1}^{T-1} W_{b_T}(\Omega - 2\pi m/T)\, I_{XX}^{T}(2\pi m/T),$$

where T is assumed to be large enough such that

$$\frac{1}{T}\sum_{m=1}^{T-1} W_{b_T}(2\pi m/T) \approx \frac{1}{T}\sum_{m=0}^{T-1} W_{b_T}(2\pi m/T) \approx \int_0^{2\pi} W_{b_T}(\Omega)\frac{d\Omega}{2\pi} = 1.$$

Since

$$W_{b_T}(2\pi m/T) = 0 \text{ if } m \in \{\lceil b_T T/2) \rceil, \ldots, \lfloor T(1-b_T/2) \rfloor\}$$

$\tilde{C}_{XX}^{T}(\Omega)$ can be understood as a weighted averaging of the periodogram over the

$$M = T - \left(\lfloor T(1-b_T/2) \rfloor - \lceil b_T T/2) \rceil + 1\right) \approx b_T T$$

frequencies that are closest to Ω, where asymptotically

$$b_T \xrightarrow[T\to\infty]{} 0 \quad \text{and} \quad M \approx b_T T \xrightarrow[T\to\infty]{} \infty.$$

<u>Theorem:</u> (mean and covariance properties of $\tilde{C}_{XX}^{T}(\Omega)$)

Let (X_t) be a zero mean stationary stochastic process.

1) If the covariance function of (X_t) satisfies

$$\sum_{\tau=-\infty}^{\infty} |\tau| \, |c_{XX}(\tau)| < \infty,$$

i.e. $C_{XX}(\Omega)$ is continuous differentiable, then

$$\mathrm{E}\tilde{C}_{XX}^{T}(\Omega) = \frac{1}{T}\sum_{m=1}^{T-1} W_{b_T}(\Omega - 2\pi m/T)\, C_{XX}(2\pi m/T) + O(1/T)$$

$$= \int_0^{2\pi} W_{b_T}(\Omega - \Lambda)\, C_{XX}(\Lambda)\, d\Lambda + O\big(1/(b_T T)\big),$$

holds and $\tilde{C}_{XX}^{T}(\Omega)$ is asymptotically unbiased, i.e.

$$\mathrm{E}\tilde{C}_{XX}^{T}(\Omega) \xrightarrow[b_T \to 0,\ b_T T \to \infty]{T \to \infty} C_{XX}(\Omega).$$

2) If the cumulant functions of (X_t) satisfy

$$\sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{\infty} \big(1 + |\tau_n|\big) \left|\kappa_{X\ldots X}(\tau_1, \ldots, \tau_{k-1})\right| < \infty$$

for $n = 1, \ldots, k-1$ when $k = 2, 3, \ldots$, then

$$b_T T \ \mathrm{Cov}\left(\tilde{C}_{XX}^{T}(\Omega), \tilde{C}_{XX}^{T}(\Lambda)\right) =$$

$$= \left(\eta(\Omega-\Lambda) + \eta(\Omega+\Lambda)\right) C_{XX}^{2}(\Omega) \int_{0}^{2\pi} W_{b_T}^{2}(\Omega) \frac{d\Omega}{2\pi} + O\left(\frac{1}{b_T T}\right)$$

with

$$\eta(\Omega) = \begin{cases} 1 & \Omega = 2k\pi, \quad k \in \mathbb{Z} \\ 0 & \text{elsewhere} \end{cases}.$$

<u>Corollary:</u> *(mean square consistency of* $\tilde{C}_{XX}^{T}(\Omega)$*)*

$$\mathrm{E}\left(\tilde{C}_{XX}^{T}(\Omega) - C_{XX}(\Omega)\right)^{2} =$$

$$= \mathrm{Var}\left(\tilde{C}_{XX}^{T}(\Omega)\right) + \left(\mathrm{E}\tilde{C}_{XX}^{T}(\Omega) - C_{XX}(\Omega)\right)^{2} \xrightarrow[b_T \to 0,\ b_T T \to \infty]{T \to \infty} 0$$

Averaging over Periodograms

$$\hat{C}_{XX}^{T}(\Omega) = \frac{1}{L(T)} \sum_{l=1}^{L(T)} I_{XX}^{T'(T)}(\Omega, l) \quad \text{with} \quad T = L(T) \cdot T'(T),$$

where the number of consecutive data pieces $L(T)$ and the length of each data piece $T'(T)$ are functions that monotonically increase with the number of observations T.

For $0 < \Omega < \pi$ the mean and variance of periodograms averaged in this way are given by

$$\mathrm{E}\left(\hat{C}_{XX}^{T}(\Omega)\right) = \frac{1}{L(T)} \sum_{l=1}^{L(T)} \mathrm{E}\left(I_{XX}^{T'(T)}(\Omega, l)\right)$$

$$= C_{XX}(\Omega) + O\left(\frac{1}{T'(T)}\right) = C_{XX}(\Omega) + O\left(\frac{L(T)}{T}\right)$$

and

$$\operatorname{Var}\left(\hat{C}_{XX}^{T}(\Omega)\right)=\frac{1}{L^2(T)}\sum_{l=1}^{L(T)}\sum_{k=1}^{L(T)}\operatorname{Cov}\left(I_{XX}^{T'(T)}(\Omega,l),I_{XX}^{T'(T)}(\Omega,k)\right)$$

$$=\frac{1}{L^2(T)}\sum_{l=1}^{L(T)}C_{XX}^2(\Omega)+O\left(\frac{1}{T'(T)}\right)$$

$$=\frac{C_{XX}^2(\Omega)}{L(T)}+O\left(\frac{L(T)}{T}\right).$$

<u>*Example:*</u>

Let $T'(T)\propto T^{1-\alpha}$ with $0<\alpha<1$, then

$$L(T)=T/T'(T)\propto T/T^{1-\alpha}=T^{\alpha}\xrightarrow[T\to\infty]{}\infty$$

and

$$L(T)/T \propto T^{\alpha}/T = 1/T^{1-\alpha} \xrightarrow[T\to\infty]{} 0.$$

<u>*Theorem:*</u> *(mean and covariance properties of $\hat{C}_{XX}^{T}(\Omega)$)*
Let (X_t) be a zero mean stationary stochastic process.

1) If the covariance function of (X_t) satisfies

$$\sum_{\tau=-\infty}^{\infty} |\tau| \, |c_{XX}(\tau)| < \infty,$$

then $\hat{C}_{XX}^{T}(\Omega)$ is asymptotically unbiased, i.e.

$$\mathrm{E}\left(\hat{C}_{XX}^{T}(\Omega)\right) \xrightarrow[L(T)/T\to 0]{T\to\infty} C_{XX}(\Omega).$$

2) If the cumulant functions of (X_t) satisfy

$$\sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{\infty} \left(1+\left|\tau_n\right|\right)\left|\kappa_{X\ldots X}(\tau_1,\ldots,\tau_{k-1})\right| < \infty$$

for $n = 1,\ldots,k-1$ when $k = 2,3,\ldots$, then

$$\mathrm{Var}\left(\hat{C}_{XX}^{T}(\Omega)\right) \xrightarrow[L(T)\to\infty,\; L(T)/T\to 0]{T\to\infty} 0.$$

<u>*Corollary:*</u> *(mean square consistency of $\hat{C}_{XX}^{T}(\Omega)$)*

$$\mathrm{E}\left(\hat{C}_{XX}^{T}(\Omega)-C_{XX}(\Omega)\right)^2 = \mathrm{Var}\left(\hat{C}_{XX}^{T}(\Omega)\right)+$$

$$+\left(\mathrm{E}\hat{C}_{XX}^{T}(\Omega)-C_{XX}(\Omega)\right)^2 \xrightarrow[L(T)\to\infty,\; L(T)/T\to 0]{T\to\infty} 0$$

5.3 Parametric Spectrum Estimation

5.3.1 Parametric Models

Auto-Regressive (AR)-Process

(X_t) is called p-th order auto-regressive process (denoted by AR(p)) if it satisfies the difference equation

$$X_t + \sum_{n=1}^{p} a_n X_{t-n} = Z_t,$$

where $a_1, a_2, \ldots, a_p$ are constant coefficients and (Z_t) is white noise, i.e.

$$\mathrm{E}(Z_t) = 0 \quad \text{and} \quad c_{ZZ}(\tau) = \sigma_Z^2\, \delta_t.$$

Thus, using the transfer function of an AR(p)-Filter and the power spectral density of white noise, i.e.

$$H(\Omega) = \frac{1}{1 + \sum_{n=1}^{p} a_n e^{-j\Omega n}} \quad \text{and} \quad C_{ZZ}(\Omega) = \sigma_Z^2,$$

the power spectral density of an AR(p)-Process can be parameterized by its coefficients $a_1, a_2, \ldots, a_p$ and the white noise variance σ_Z^2 as follows.

$$C_{XX}(\Omega) = \left|H(\Omega)\right|^2 C_{ZZ}(\Omega) = \frac{\sigma_Z^2}{\left|1 + \sum_{n=1}^{p} a_n e^{-j\Omega n}\right|^2}$$

Moving-Average (MA)-Process

A process (X_t) that obeys an equation of the form

$$X_t = Z_t + \sum_{n=1}^{q} b_n Z_{t-n}$$

is called moving-average process of order q (denoted by MA(q)), where $b_1, b_2, \ldots, b_q$ and (Z_t) denote constant coefficients and white noise, respectively.

The transfer function of a MA(q)-Filter is given by

$$H(\Omega) = 1 + \sum_{n=1}^{q} b_n e^{-j\Omega n}.$$

Hence, the power spectral density of a MA(q)-Process can be parameterized by its coefficients $b_1, b_2, \ldots, b_q$ and the white noise variance σ_Z^2 as follows.

$$C_{XX}(\Omega) = |H(\Omega)|^2 C_{ZZ}(\Omega) = \sigma_Z^2 \left| 1 + \sum_{n=1}^{q} b_n e^{-j\Omega n} \right|^2$$

Auto-Regressive-Moving-Average (ARMA)-Process

We say that (X_t) is an auto-regressive-moving-average process of order (p,q) (denoted by ARMA(p,q)) if it can be represented in the form

$$X_t + \sum_{n=1}^{p} a_n X_{t-n} = Z_t + \sum_{n=1}^{q} b_n Z_{t-n}.$$

With the transfer function of an ARMA(*p*,*q*)-Filter

$$H(\Omega) = \frac{1 + \sum_{n=1}^{q} b_n e^{-j\Omega n}}{1 + \sum_{n=1}^{p} a_n e^{-j\Omega n}},$$

the power spectral density of an ARMA(*p*,*q*)-Process can be parameterized by its coefficients $a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q$ and the white noise variance σ_Z^2 as follows.

$$C_{XX}(\Omega) = \left|H(\Omega)\right|^2 C_{ZZ}(\Omega) = \sigma_Z^2 \left| \frac{1 + \sum_{n=1}^{q} b_n e^{-j\Omega n}}{1 + \sum_{n=1}^{p} a_n e^{-j\Omega n}} \right|^2$$

AR-Process + Noise

A process (Y_t) defined by

$$Y_t = \underbrace{X_t}_{\text{random signal}} + \underbrace{V_t}_{\text{noise}} \quad \text{with} \quad \underbrace{X_t + \sum_{n=1}^{p} a_n X_{t-n} = Z_t}_{\text{AR}(p)\text{-Process}}$$

is known as a special case of a more general random signal plus noise model, where (Z_t) and (V_t) denote independently distributed white noise processes.

After some manipulations the equation stated above can be reformulated via

$$Y_t - V_t = X_t = Z_t - \sum_{n=1}^{p} a_n X_{t-n} = Z_t - \sum_{n=1}^{p} a_n (Y_{t-n} - V_{t-n})$$

into the difference equation

$$Y_t + \sum_{n=1}^{p} a_n Y_{t-n} = V_t + \sum_{n=1}^{p} a_n V_{t-n} + Z_t.$$

The covariance function and spectral density function of (Y_t) are given by

$$c_{YY}(\tau) = \mathrm{E}\left(Y_{t+\tau} Y_t\right) = \mathrm{E}\left((X_{t+\tau} + V_{t+\tau})(X_t + V_t)\right) = c_{XX}(\tau) + c_{VV}(\tau)$$

and

$$C_{YY}(\Omega) = C_{XX}(\Omega) + C_{VV}(\Omega) = \left|H(\Omega)\right|^2 C_{ZZ}(\Omega) + C_{VV}(\Omega)$$

$$= \sigma_Z^2 \Big/ \left|1 + \sum\nolimits_{n=1}^{p} a_n e^{-j\Omega n}\right|^2 + \sigma_V^2,$$

respectively.

Oscillation + Noise

A process (Y_t) defined by

$$Y_t = \underbrace{X_t}_{\text{totally predictable signal}} + \underbrace{Z_t}_{\text{white noise}} \quad \text{with} \quad \underbrace{X_t = A\cos(\Omega_0 t) + B\sin(\Omega_0 t)}_{\text{Oscillation}},$$

where the random amplitudes satisfy

$$\mathrm{E}(A) = \mathrm{E}(B) = 0,\ \mathrm{E}(A^2) = \mathrm{E}(B^2) = \sigma^2 \ \text{ and } \ \mathrm{Cov}(A,B) = 0$$

and (Z_t) denotes white noise independent of *A* and *B*, is known to be wide sense stationary with covariance function, cf. Exercise 2.3-1, given by

$$c_{YY}(\tau) = c_{XX}(\tau) + c_{ZZ}(\tau) = \sigma^2\cos(\Omega_0\tau) + \sigma_Z^2\delta_\tau .$$

Hence, the spectral density function of (Y_t) can be expressed as

$$C_{YY}(\Omega) = \pi\sigma^2\big(\eta(\Omega-\Omega_0)+\eta(\Omega+\Omega_0)\big)+\sigma_Z^2$$

with

$$\eta(\Omega) = \sum_{n=-\infty}^{\infty}\delta(\Omega-2\pi n).$$

Pisarenko Model

A generalization of the Oscillation + Noise model provides the Pisarenko Model which is defined by

$$Y_t = \underbrace{X_t}_{\text{totally predictable signal}} + \underbrace{Z_t}_{\text{white noise}} \quad \text{with} \quad \underbrace{X_t = \sum_{m=1}^{M} A_m \cos(\Omega_m t + \phi_m)}_{\text{Oscillations}},$$

where A_m and ϕ_m are independent random amplitudes and phases with

$$\mathrm{E}(A_m) = 0,\; \mathrm{E}(A_m^2) = \sigma_m^2 \quad \text{and} \quad \phi_m \sim \mathcal{R}(-\pi, \pi)$$

and (Z_t) denotes white noise which is independent of A_m and ϕ_m. Finally, with the covariance function given by

$$c_{YY}(\tau) = c_{XX}(\tau) + c_{ZZ}(\tau) = \frac{1}{2}\sum_{m=1}^{M} \sigma_m^2 \cos(\Omega_m \tau) + \sigma_Z^2 \delta_\tau$$

the spectral density function of (Y_t) can be expressed as

$$C_{YY}(\Omega) = \frac{\pi}{2}\sum_{m=1}^{M} \sigma_m^2 \left(\eta(\Omega - \Omega_m) + \eta(\Omega + \Omega_m)\right) + \sigma_Z^2.$$

Exercise 5.3-1:

Show that the Oscillation + Noise model can be interpreted as unstably rationally filtered white noise

Exponential Model of Bloomfield

In case that nonparametric spectral density estimates or prior knowledge about the physics provide information about the frequency behavior of a process, a direct modeling in the frequency domain might be advantageous.

An example of such an approach gives Bloomfield's exponential model

$$C_{YY}(\Omega) = \sigma^2 \exp\left(\sum_{m=1}^{M} C_m \cos(\Omega m) \right),$$

which is appropriate for problems where the spectral density function shows a ripple behavior. Applying the logarithm to both sides of the last equation leads to the linear model given by

$$\ln\left(C_{YY}(\Omega)\right) = \ln(\sigma^2) + \sum_{m=1}^{M} C_m \cos(\Omega m).$$

Remark:
The inverse Fourier transform of the logarithm of the power spectral density function, i.e.

$$c_{XX}^{cep}(\tau) = \mathcal{F}^{-1}\left\{\ln\left(C_{YY}(\Omega)\right)\right\}$$

called Cepstrum, is often used in audio signal analysis.

5.3.2 Consistent Parameter Estimators

<u>*Auto-Regressive (AR)-Process*</u>

Let (X_t) be an AR(p)-Process that can be represented by the difference equation

$$X_t + a_1 X_{t-1} + \ldots + a_p X_{t-p} = X_t + \sum_{n=1}^{p} a_n X_{t-n} = Z_t,$$

where $a_1, a_2, \ldots, a_p$ are constants and (Z_t) is white noise.

Multiplying both sides of the difference equation by X_{t-m} from the right and taking expectations

$$\mathrm{E}\left(\left(X_t + \sum_{n=1}^{p} a_n X_{t-n}\right) X_{t-m}\right) = \mathrm{E}\left(Z_t X_{t-m}\right)$$

we obtain

$$\mathrm{E}\left(X_t X_{t-m}\right)+\sum_{n=1}^{p} a_n \mathrm{E}\left(X_{t-n} X_{t-m}\right)=\mathrm{E}\left(Z_t X_{t-m}\right)$$

$$c_{XX}(m)+\sum_{n=1}^{p} a_n c_{XX}(m-n)=c_{ZX}(m)=\sigma_Z^2\delta_m=\begin{cases}\sigma_Z^2 & m=0\\ 0 & m>0\end{cases}.$$

For m > 0 these equations, known as Yule-Walker equations, can be expressed in matrix notation by

$$\underbrace{\begin{pmatrix} c_{XX}(0) & c_{XX}(1) & \cdots & c_{XX}(p-1)\\ c_{XX}(1) & c_{XX}(0) & \cdots & c_{XX}(p-2)\\ \vdots & \vdots & \ddots & \vdots\\ c_{XX}(p-1) & c_{XX}(p-2) & \cdots & c_{XX}(0)\end{pmatrix}}_{\mathbf{C}_{XX}}\underbrace{\begin{pmatrix} a_1\\ a_2\\ \vdots\\ a_p\end{pmatrix}}_{\mathbf{a}}=-\underbrace{\begin{pmatrix} c_{XX}(1)\\ c_{XX}(2)\\ \vdots\\ c_{XX}(p)\end{pmatrix}}_{\mathbf{c}_{XX}}.$$

If all roots of

$$\alpha(z) = z^p + \sum_{n=1}^{p} a_n z^{p-n}$$

are lying within the unit circle one can show that the coefficient matrix of the equation system $\mathbf{C}_{XX}$, which is a symmetric Toeplitz matrix, is positive definite.

Thus, assuming $c_{XX}(0),\ldots,c_{XX}(p)$ to be known the equation system can be uniquely solved for $a_1,a_2,\ldots,a_p$, e.g. by means of the Levinson-Durbin algorithm, and the variance of the white noise can be subsequently determined by the Yule-Walker equations for $m=0$, i.e.

$$\sigma_Z^2 = c_{XX}(0) + \sum_{n=1}^{p} a_n\, c_{XX}(n).$$

However, $c_{XX}(\tau)$ is typically unknown. On the other hand following Chapter 5.1.3, $c_{XX}(\tau)$ can be consistently estimated using the sample covariance function $\hat{c}_{XX}(\tau)$.

Thus, after replacing $c_{XX}(0),\ldots,c_{XX}(p)$ in the Yule-Walker equations by its estimates $\hat{c}_{XX}(0),\ldots,\hat{c}_{XX}(p)$, one obtains the so-called empirical Yule-Walker equation system

$$\underbrace{\begin{pmatrix} \hat{c}_{XX}(0) & \hat{c}_{XX}(1) & \cdots & \hat{c}_{XX}(p-1) \\ \hat{c}_{XX}(1) & \hat{c}_{XX}(0) & \cdots & \hat{c}_{XX}(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{c}_{XX}(p-1) & \hat{c}_{XX}(p-2) & \cdots & \hat{c}_{XX}(0) \end{pmatrix}}_{\hat{\mathbf{C}}_{XX}} \underbrace{\begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_p \end{pmatrix}}_{\hat{\mathbf{a}}} = -\underbrace{\begin{pmatrix} \hat{c}_{XX}(1) \\ \hat{c}_{XX}(2) \\ \vdots \\ \hat{c}_{XX}(p) \end{pmatrix}}_{\hat{\mathbf{c}}_{XX}}$$

for $m = 1,\ldots,p$ and for $m = 0$ the equation

$$\hat{\sigma}_Z^2 = \hat{c}_{XX}(0) + \sum_{n=1}^{p} \hat{a}_n \hat{c}_{XX}(n).$$

From the consistency property of the sample covariance function, i.e.

$$\hat{c}_{XX}(\tau) \xrightarrow[T \to \infty]{m.s.} c_{XX}(\tau),$$

follows

$$\left(\hat{c}_{XX}(0), \ldots, \hat{c}_{XX}(p)\right)^T \xrightarrow[T \to \infty]{m.s.} \left(c_{XX}(0), \ldots, c_{XX}(p)\right)^T$$

and due to the continuity of the functions

$$\hat{\mathbf{a}} = \mathbf{f}\left(\hat{c}_{XX}(0), \ldots, \hat{c}_{XX}(p)\right) = -\hat{\mathbf{C}}_{XX}^{-1} \hat{\mathbf{c}}_{XX},$$

$$\hat{\sigma}_Z^2 = f\left(\hat{c}_{XX}(0), \ldots, \hat{c}_{XX}(p)\right) = \hat{c}_{XX}(0) + \hat{\mathbf{c}}_{XX}^T \hat{\mathbf{a}}$$

$$= \hat{c}_{XX}(0) - \hat{\mathbf{c}}_{XX}^T \hat{\mathbf{C}}_{XX}^{-1} \hat{\mathbf{c}}_{XX}$$

and

$$\hat{C}_{XX}(\Omega) = F(\Omega \,|\, \hat{\mathbf{a}}, \hat{\sigma}_Z^2) = \hat{\sigma}_Z^2 \Big/ \left| 1 + \sum_{n=1}^{p} \hat{a}_n e^{-j\Omega n} \right|^2,$$

finally

$$\hat{\mathbf{a}} \xrightarrow[T\to\infty]{m.s.} \mathbf{a} = \mathbf{f}\left(c_{XX}(0),\ldots,c_{XX}(p)\right) = -\mathbf{C}_{XX}^{-1}\mathbf{c}_{XX},$$

$$\hat{\sigma}_Z^2 \xrightarrow[T\to\infty]{m.s.} \sigma_Z^2 = f\left(c_{XX}(0),\ldots,c_{XX}(p)\right) = c_{XX}(0) + \mathbf{c}_{XX}^T\mathbf{a}$$

and

$$\hat{C}_{XX}(\Omega) \xrightarrow[T\to\infty]{m.s.} C_{XX}(\Omega) = F(\Omega \,|\, \mathbf{a}, \sigma_Z^2)$$

$$= \sigma_Z^2 \Big/ \left| 1 + \sum_{n=1}^{p} a_n e^{-j\Omega n} \right|^2.$$

Moving-Average (MA)-Process

Now, we suppose that (X_t) is a MA(q)-Process that can be expressed in the form

$$X_t = Z_t + b_1 Z_{t-1} + \ldots + b_q Z_{t-q} = \sum_{n=0}^{q} b_n Z_{t-n} \quad \text{with} \quad b_0 = 1,$$

where $b_1,\ldots,b_q$ are constants and (Z_t) is white noise.

Since (X_t) is a linear combination of uncorrelated random variables its mean and variance are given by

$$\mu_X = \mu_Z \sum_{n=0}^{q} b_n \quad \text{and} \quad \sigma_X^2 = \sigma_Z^2 \sum_{n=0}^{q} b_n^2 .$$

Furthermore, (X_t) is always stationary (irrespective of the values of $b_1,\ldots,b_q$) and has the covariance function

$$c_{XX}(\tau)=\mathrm{E}(X_t X_{t-\tau})=\mathrm{E}\left(\sum_{n=0}^{q} b_n Z_{t-n} \sum_{m=0}^{q} b_m Z_{t-\tau-m}\right)$$

$$=\sum_{n=0}^{q}\sum_{m=0}^{q} b_n b_m \mathrm{E}(Z_{t-n} Z_{t-\tau-m})=\sum_{n=0}^{q}\sum_{m=0}^{q} b_n b_m c_{ZZ}(m+\tau-n)$$

$$=\sum_{n=0}^{q}\sum_{m=0}^{q} b_n b_m \sigma_Z^2 \delta_{m+\tau-n}=\begin{cases}\sigma_Z^2 \sum_{m=0}^{q-|\tau|} b_{m+|\tau|} b_m & |\tau|\le q\\ 0 & |\tau|>q\end{cases},$$

where $\mathrm{E}(Z_t)=0 \Rightarrow \mathrm{E}(X_t)=0$ has been exploited.

For given covariances $c_{XX}(0),\ldots,c_{XX}(q)$ the parameters $b_1, b_2,\ldots, b_q$ and σ_Z^2 can be determined by solving the system of $(q+1)$ non-linear equations

$$c_{XX}(\tau)=\sigma_Z^2\sum_{m=0}^{q-|\tau|} b_{m+|\tau|}b_m=\sum_{m=0}^{q-|\tau|}\gamma_{m+|\tau|}\gamma_m \text{ with } \gamma_n=\begin{cases}\sigma_Z & n=0\\ \sigma_Z b_n & n=1,\ldots,q\end{cases}$$

which usually possesses 2^q solution vectors $(\gamma_0,\ldots,\gamma_q)^T$.

However, under the additional constraint that the MA(q)-filter has to be of minimum phase a unique solution can be derived as follows.

First, the bilateral z-Transform of $c_{XX}(\tau)$ provides

$$C(z)=\sum_{\tau=-q}^{q} c_{XX}(\tau)z^{-\tau}=\sum_{\tau=-q}^{q}\sum_{m=0}^{q-|\tau|}\gamma_{m+|\tau|}\gamma_m z^{-\tau}=\sum_{n=0}^{q}\sum_{m=0}^{q}\gamma_n\gamma_m z^{m-n}$$

$$=\sum_{n=0}^{q}\gamma_n z^{-n}\sum_{m=0}^{q}\gamma_m z^{m}=\Gamma(z)\Gamma(z^{-1})$$

with

$$z^q\Gamma(z) = z^q \sum_{n=0}^{q} \gamma_n z^{-n} = \sigma_Z z^q \sum_{n=0}^{q} b_n z^{-n} = \sigma_Z z^q b(z^{-1})$$

$$= \sigma_Z \left(z^q + b_1 z^{q-1} + \cdots + b_{q-1} z^1 + b_q \right) = \sigma_Z \beta(z),$$

where $\beta(z)$ is known as the characteristic polynomial of the associated MA(q)-filter.

To satisfy the minimum phase (invertibility) constraint the roots of $\beta(z)$ and accordingly of $z^q\Gamma(z)$ must not lie outside the unit circle.

Moreover, the roots of $z^q C(z) = z^q \Gamma(z)\Gamma(z^{-1}) = 0$ always occur in pairs, i.e. if z' is a root then $1/z'$ is also a root.

Hence, after selecting from the $2q$ roots of the polynomial $z^q C(z)$ those q roots that satisfy $|z_n| \le 1$ for $n = 1, \ldots, q$, the coefficients $b_1, \ldots, b_q$ and the white noise variance σ_Z^2 can be determined by means of

$$\beta(z) = \frac{z^q}{\sigma_Z}\Gamma(z) = \frac{z^q}{\sigma_Z}\sum_{n=0}^{q}\gamma_n z^{-n} = z^q\left(1 + \sum_{n=1}^{q}\frac{\gamma_n}{\gamma_0}z^{-n}\right)$$

$$= z^q\prod_{n=1}^{q}(1 - z_n z^{-1}) = \prod_{n=1}^{q}(z - z_n) = z^q + b_1 z^{q-1} + \cdots + b_{q-1}z^1 + b_q$$

and

$$\sigma_Z^2 = c_{XX}(0) \Big/ \left(1 + \sum_{n=1}^{q} b_n^2\right)$$

respectively.

Generally, $c_{XX}(\tau)$ is unknown. Thus, replacing $c_{XX}(0)$,..., $c_{XX}(q)$ in $z^q C(z)$ by its estimates $\hat{c}_{XX}(0),\dots,\hat{c}_{XX}(p)$, we obtain the empirical polynomial

$$z^q \hat{C}(z) = z^q \sum_{\tau=-q}^{q} \hat{c}_{XX}(\tau) z^{-\tau} = \sum_{\tau=-q}^{q} \hat{c}_{XX}(\tau) z^{q-\tau},$$

whose $2q$ roots usually have to be calculated numerically. If from the $2q$ roots those q roots are selected that satisfy $|\hat{z}_n| \le 1$ for $n=1,\dots,q$ the estimates $\hat{b}_1,\dots,\hat{b}_q$ and $\hat{\sigma}_Z^2$ can be derived using

$$\hat{\beta}(z) = \prod_{n=1}^{q} (z - \hat{z}_n) = z^q + \hat{b}_1 z^{q-1} + \dots + \hat{b}_{q-1} z^1 + \hat{b}_q$$

and

$$\hat{\sigma}_Z^2 = \hat{c}_{XX}(0) \Big/ \left(1 + \sum_{n=1}^{q} \hat{b}_n^2 \right).$$

The power spectral density can then be estimated by

$$\hat{C}_{XX}(\Omega) = \hat{\sigma}_Z^2 \left| 1 + \sum_{n=1}^{q} \hat{b}_n e^{-j\Omega n} \right|^2 .$$

Finally, from the consistency property of $\hat{c}_{XX}(\tau)$, i.e.

$$\hat{c}_{XX}(\tau) \xrightarrow[T\to\infty]{m.s.} c_{XX}(\tau),$$

and the inherent continuous functional relations follows

$$(\hat{z}_1, \ldots, \hat{z}_q)^T \xrightarrow[T\to\infty]{m.s.} (z_1, \ldots, z_q)^T$$

and consequently

$$\left(\hat{b}_1, \ldots, \hat{b}_q, \hat{\sigma}_Z^2\right)^T \xrightarrow[T\to\infty]{m.s.} \left(b_1, \ldots, b_q, \sigma_Z^2\right)^T$$

as well as

$$\hat{C}_{XX}(\Omega) \xrightarrow[T\to\infty]{m.s.} C_{XX}(\Omega) = \sigma_Z^2 \left| 1 + \sum_{n=1}^{q} b_n e^{-j\Omega n} \right|^2 .$$

Auto-Regressive-Moving-Average (ARMA)-Process

In the following (X_t) is considered to be an ARMA(p,q)-Process that satisfies the difference equation

$$X_t + \sum_{n=1}^{p} a_n X_{t-n} = \sum_{n=0}^{q} b_n Z_{t-n} \quad \text{with} \quad b_0 = 1.$$

Multiplying both sides of the equation by X_{t-m} from the right and taking expectations

$$\mathrm{E}\left(\left(X_t + \sum_{n=1}^{p} a_n X_{t-n}\right) X_{t-m}\right) = \mathrm{E}\left(\left(\sum_{n=0}^{q} b_n Z_{t-n}\right) X_{t-m}\right)$$

$$\mathrm{E}(X_t X_{t-m}) + \sum_{n=1}^{p} a_n \mathrm{E}(X_{t-n} X_{t-m}) = \sum_{n=0}^{q} b_n \mathrm{E}(Z_{t-n} X_{t-m})$$

we obtain

$$c_{XX}(m)+\sum_{n=1}^{p} a_n\, c_{XX}(m-n)=0 \quad m>q,$$

where $\mathrm{E}(Z_t)=0 \Rightarrow \mathrm{E}(X_t)=0$ has been exploited. This set of equations is sometimes called modified Yule-Walker equations.

For $m=q+1,q+2,\ldots,q+p$ the modified Yule-Walker equations can be expressed by

$$\underbrace{\begin{pmatrix} c_{XX}(q) & c_{XX}(q-1) & \cdots & c_{XX}(q-p+1) \\ c_{XX}(q+1) & c_{XX}(q) & \cdots & c_{XX}(q-p+2) \\ \vdots & \vdots & \ddots & \vdots \\ c_{XX}(q+p-1) & c_{XX}(q+p-2) & \cdots & c_{XX}(q) \end{pmatrix}}_{\mathbf{C}_{XX}^{\text{mod}}} \underbrace{\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}}_{\mathbf{a}} = -\underbrace{\begin{pmatrix} c_{XX}(q+1) \\ c_{XX}(q+2) \\ \vdots \\ c_{XX}(q+p) \end{pmatrix}}_{\mathbf{c}_{XX}^{\text{mod}}}.$$

The coefficient matrix of the equation system is obviously again a Toeplitz matrix but it is not anymore symmetric.

However, one can show that the coefficient matrix $\mathbf{C}_{XX}^{\text{mod}}$ is regular if all roots of $\alpha(z)=z^p+\sum_{n=1}^{p} a_n z^{p-n}$ are lying within the unit circle.

Hence, assuming $c_{XX}(0),\ldots,c_{XX}(q+p)$ to be known the equation system can be uniquely solved for $a_1,a_2,\ldots,a_p$.

Subsequently, the parameters $a_1,a_2,\ldots,a_p$ and the covariances $c_{XX}(0),\ldots,c_{XX}(q+p)$ allow the calculation of the covariance function $c_{YY}(\tau)$ of the MA(q)-process introduced by

$$Y_t = b(B)Z_t = a(B)X_t = \sum_{n=0}^{p} a_n X_{t-n} \quad \text{with} \quad a_0 = 1$$

as follows.

$$c_{YY}(\tau) = \mathrm{E}(Y_t Y_{t-\tau}) = \mathrm{E}\left(\sum\nolimits_{n=0}^{p} a_n X_{t-n} \sum\nolimits_{m=0}^{p} a_m X_{t-\tau-m}\right)$$

$$= \sum_{n=0}^{p}\sum_{m=0}^{p} a_n a_m \mathrm{E}(X_{t-n} X_{t-\tau-m}) = \sum_{n=0}^{p}\sum_{m=0}^{p} a_n a_m c_{XX}(\tau + m - n)$$

$$= \sum_{k=-p}^{p}\sum_{m=0}^{p-|k|} a_{m+|k|} a_m c_{XX}(\tau - k)$$

Thus, after determining the covariances $c_{YY}(0),\ldots,c_{YY}(q)$ the parameters $b_1, b_2, \ldots, b_q$ and σ_Z^2 can be determined by solving the equation system

$$c_{YY}(\tau) = \sigma_Z^2 \sum_{m=0}^{q-|\tau|} b_{m+|\tau|} b_m = \sum_{m=0}^{q-|\tau|} \gamma_{m+|\tau|} \gamma_m \quad \text{with } \gamma_n = \begin{cases} \sigma_Z & n = 0 \\ \sigma_Z b_n & n = 1, \ldots, q \end{cases}$$

where a unique solution can be derived by following the approach mentioned in conjunction with a MA(q)-Process.

Since $c_{XX}(\tau)$ is typically unknown we replace it here again by its consistent estimate $\hat{c}_{XX}(\tau)$. This leads us to the empirical modified Yule-Walker equation system

$$\underbrace{\begin{pmatrix} \hat{c}_{XX}(q) & \hat{c}_{XX}(q-1) & \cdots & \hat{c}_{XX}(q-p+1) \\ \hat{c}_{XX}(q+1) & \hat{c}_{XX}(q) & \cdots & \hat{c}_{XX}(q-p+2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{c}_{XX}(q+p-1) & \hat{c}_{XX}(q+p-2) & \cdots & \hat{c}_{XX}(q) \end{pmatrix}}_{\hat{\mathbf{C}}_{XX}^{\text{mod}}} \underbrace{\begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_p \end{pmatrix}}_{\hat{\mathbf{a}}} = -\underbrace{\begin{pmatrix} \hat{c}_{XX}(q+1) \\ \hat{c}_{XX}(q+2) \\ \vdots \\ \hat{c}_{XX}(q+p) \end{pmatrix}}_{\hat{\mathbf{c}}_{XX}^{\text{mod}}},$$

which for a regular $\hat{\mathbf{C}}_{XX}^{\text{mod}}$ provides with

$$\hat{\mathbf{a}} = -\left(\hat{\mathbf{C}}_{XX}^{\text{mod}}\right)^{-1} \hat{\mathbf{c}}_{XX}^{\text{mod}}$$

consistent estimates of $a_1,\ldots,a_p$. These estimates can be used along with $\hat{c}_{XX}(\tau)$ to estimate $c_{YY}(\tau)$ by means of

$$\hat{c}_{YY}(\tau) = \sum_{k=-p}^{p} \sum_{m=0}^{p-|k|} \hat{a}_{m+|k|}\hat{a}_m \hat{c}_{XX}(\tau - k)$$

consistently. Replacing $c_{YY}(\tau)$ in $z^q C(z)$ by $\hat{c}_{YY}(\tau)$ yields the empirical polynomial

$$z^q \hat{C}(z) = z^q \sum_{\tau=-q}^{q} \hat{c}_{YY}(\tau) z^{-\tau} = \sum_{\tau=-q}^{q} \hat{c}_{YY}(\tau) z^{q-\tau}.$$

From its $2q$ roots, which are typically to be determined numerically, we then have to select those q roots $\hat{z}_n$ ($n=1,\ldots,q$) that do not lie outside the unit circle.

Since the roots of $z^q\hat{C}(z)$ are consistent estimates of the roots of $z^q C(z)$, the coefficients $b_1,\ldots,b_q$ and the white noise variance σ_Z^2 can be consistently estimated via

$$\hat{\beta}(z)=\prod_{n=1}^{q}(z-\hat{z}_n)=z^q+\hat{b}_1 z^{q-1}+\cdots+\hat{b}_{q-1}z^1+\hat{b}_q$$

and

$$\hat{\sigma}_Z^2=\hat{c}_{XX}(0)\Big/\left(1+\sum_{n=1}^{q}\hat{b}_n^2\right).$$

Finally, the consistent power spectral density estimate of the ARMA(p,q)-Process is given by

$$\hat{C}_{XX}(\Omega)=\hat{\sigma}_Z^2\frac{\left|1+\sum_{n=1}^{q}\hat{b}_n e^{-j\Omega n}\right|^2}{\left|1+\sum_{n=1}^{p}\hat{a}_n e^{-j\Omega n}\right|^2}.$$

5.3.3 Asymptotically Efficient Parameter Estimators

Asymptotic Efficiency

The Cramer-Rao inequality

$$\mathbf{a}^T \mathbf{C}_{\hat{\boldsymbol{\Theta}}\hat{\boldsymbol{\Theta}}} \mathbf{a} \geq \mathbf{a}^T \mathcal{I}(\boldsymbol{\theta})^{-1} \mathbf{a} \quad \forall \mathbf{a} \in \mathbb{R}^p,$$

provides a lower bound to the covariance matrix $\mathbf{C}_{\hat{\boldsymbol{\Theta}}\hat{\boldsymbol{\Theta}}}$ of any unbiased estimator of $\boldsymbol{\theta}$, cf. Chapter 3.5.

An estimator for which the inequality takes the equality sign, i.e. whose $\mathbf{C}_{\hat{\boldsymbol{\Theta}}\hat{\boldsymbol{\Theta}}}$ coincides with the inverse of the Fisher information matrix $\mathcal{I}(\boldsymbol{\theta})$ is called efficient.

However, estimators used in practice are often neither unbiased nor mean square consistent, i.e. the Cramer-Rao inequality can not be applied.

On the other hand, one can often observe that the deviation of the covariance matrix from the inverse Fisher information matrix decreases for large sample sizes N.

In these cases, estimators are typically evaluated using their limiting distribution. If the limiting distribution has the property

$$\lim_{N\to\infty} \sqrt{N}\,(\hat{\boldsymbol{\Theta}}_N - \boldsymbol{\theta}) \sim \mathcal{N}_p\left(\mathbf{0}, \boldsymbol{\Gamma}(\boldsymbol{\theta})^{-1}\right),$$

where

$$\boldsymbol{\Gamma}(\boldsymbol{\theta})^{-1} = \lim_{N\to\infty} N\,\mathbf{C}_{\hat{\boldsymbol{\Theta}}_N\hat{\boldsymbol{\Theta}}_N} \quad \text{and} \quad \boldsymbol{\Gamma}(\boldsymbol{\theta}) = \lim_{N\to\infty} \frac{1}{N}\mathcal{I}_N(\boldsymbol{\theta})$$

with

$$\mathcal{I}_N(\boldsymbol{\theta}) = \mathrm{E}\left(\frac{\partial \ln\left(f_{\mathbf{X}}(\mathbf{X}|\boldsymbol{\theta})\right)}{\partial \theta_i} \cdot \frac{\partial \ln\left(f_{\mathbf{X}}(\mathbf{X}|\boldsymbol{\theta})\right)}{\partial \theta_j}\right)_{i,j=1,\ldots,p}$$

and $\mathbf{X} = (X_1,\ldots,X_N)^T$, the estimator is said to be asymptotically efficient.

Let (X_t) be a stationary Gaussian process with zero mean and power spectral density $C_{XX}(\Omega) = C_{XX}(\Omega\,|\,\boldsymbol{\theta})$ parameterized by $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_p)^T$.

Hence, the probability density function of the random vector $\mathbf{X} = (X_1,\ldots,X_N)^T$ modeling N consecutive observations of the process is given by

$$f_{\mathbf{X}}(\mathbf{x}\,|\,\boldsymbol{\theta}) = \frac{1}{(2\pi)^{N/2}\sqrt{\det \mathbf{C}_N(\boldsymbol{\theta})}} \exp\left\{-\frac{1}{2}\mathbf{x}^T\mathbf{C}_N^{-1}(\boldsymbol{\theta})\mathbf{x}\right\}.$$

Employing the rules

$$\mathrm{E}(\mathbf{X}^T\mathbf{A}\mathbf{X}) = \mathrm{tr}\left(\mathbf{A}\mathrm{E}(\mathbf{X}\mathbf{X}^T)\right) = \mathrm{tr}\left(\mathbf{A}\mathbf{C}_N(\boldsymbol{\theta})\right)$$

and

$$\mathrm{E}(\mathbf{X}^T\mathbf{A}\mathbf{X}\mathbf{X}^T\mathbf{B}\mathbf{X}) = \mathrm{tr}\left(\mathbf{A}\mathbf{C}_N(\boldsymbol{\theta})\right)\mathrm{tr}\left(\mathbf{B}\mathbf{C}_N(\boldsymbol{\theta})\right) + 2\,\mathrm{tr}\left(\mathbf{A}\mathbf{C}_N(\boldsymbol{\theta})\mathbf{B}\mathbf{C}_N(\boldsymbol{\theta})\right)$$

as well as the identities

$$\frac{\partial \mathbf{C}_N^{-1}(\boldsymbol{\theta})}{\partial \theta_i} = -\mathbf{C}_N^{-1}\frac{\partial \mathbf{C}_N(\boldsymbol{\theta})}{\partial \theta_i}\mathbf{C}_N^{-1}$$

and

$$\frac{\partial \ln\left(\det \mathbf{C}_N(\boldsymbol{\theta})\right)}{\partial \theta_i} = \mathrm{tr}\left(\mathbf{C}_N^{-1}(\boldsymbol{\theta})\frac{\partial \mathbf{C}_N(\boldsymbol{\theta})}{\partial \theta_i}\right)$$

the Fisher information matrix for a zero mean Gaussian random vector can be expressed by

$$\mathcal{I}_N(\boldsymbol{\theta}) = \frac{1}{2}\left(\operatorname{tr}\left(\mathbf{C}_N^{-1}(\boldsymbol{\theta}) \frac{\partial \mathbf{C}_N(\boldsymbol{\theta})}{\partial \theta_i} \mathbf{C}_N^{-1}(\boldsymbol{\theta}) \frac{\partial \mathbf{C}_N(\boldsymbol{\theta})}{\partial \theta_j} \right)\right)_{i,j=1,\dots,p} .$$

If one parameterizes the covariance matrix according to

$$\begin{aligned} \mathbf{C}_N(\boldsymbol{\theta}) &= \left(c_{XX}(i-j) \right)_{i,j=1,\dots,N} \\ &= \left(\int_{-\pi}^{\pi} C_{XX}(\Omega \mid \boldsymbol{\theta}) e^{j\Omega(i-j)} \frac{d\Omega}{2\pi} \right)_{i,j=1,\dots,N} \\ &= \overline{\mathbf{C}}_N\left(C_{XX}(\Omega \mid \boldsymbol{\theta}) \right) \end{aligned}$$

via the power spectral density, the Fisher information matrix becomes

$$\mathcal{I}_N(\boldsymbol{\theta}) = \frac{1}{2}\left(\operatorname{tr}\left\{ \bar{\mathbf{C}}_N^{-1}(C_{XX}(\Omega\,|\,\boldsymbol{\theta}))\bar{\mathbf{C}}_N\left(\frac{\partial C_{XX}(\Omega\,|\,\boldsymbol{\theta})}{\partial\theta_i}\right)\cdot \bar{\mathbf{C}}_N^{-1}(C_{XX}(\Omega\,|\,\boldsymbol{\theta}))\bar{\mathbf{C}}_N\left(\frac{\partial C_{XX}(\Omega\,|\,\boldsymbol{\theta})}{\partial\theta_j}\right)\right\}\right)_{i,j=1,\ldots,p}.$$

Furthermore, under certain regularity conditions on can show that

$$\lim_{N\to\infty}\frac{1}{N}\operatorname{tr}\left\{\bar{\mathbf{C}}_N^{-1}(C_{XX}(\Omega\,|\,\boldsymbol{\theta}))\bar{\mathbf{C}}_N\big(a_i(\Omega\,|\,\boldsymbol{\theta})C_{XX}(\Omega\,|\,\boldsymbol{\theta})\big)\cdot \bar{\mathbf{C}}_N^{-1}(C_{XX}(\Omega\,|\,\boldsymbol{\theta}))\bar{\mathbf{C}}_N\big(a_j(\Omega\,|\,\boldsymbol{\theta})C_{XX}(\Omega\,|\,\boldsymbol{\theta})\big)\right\} = \cdots$$

$$\cdots = \int_{-\pi}^{\pi} a_i(\Omega\,|\,\boldsymbol{\theta})\, a_j(\Omega\,|\,\boldsymbol{\theta}) \frac{d\Omega}{2\pi}$$

holds. Exploiting this result by choosing

$$a_{i,j}(\Omega\,|\,\boldsymbol{\theta}) = \frac{\partial \ln C_{XX}(\Omega\,|\,\boldsymbol{\theta})}{\partial \theta_{i,j}} = \frac{1}{C_{XX}(\Omega\,|\,\boldsymbol{\theta})} \frac{\partial C_{XX}(\Omega\,|\,\boldsymbol{\theta})}{\partial \theta_{i,j}}$$

for $i,j = 1,\ldots,p$, we obtain

$$\lim_{N\to\infty} \frac{1}{N} \mathcal{I}_N(\boldsymbol{\theta}) = \frac{1}{2} \left(\int_{-\pi}^{\pi} \frac{\partial \ln C_{XX}(\Omega\,|\,\boldsymbol{\theta})}{\partial \theta_i} \frac{\partial \ln C_{XX}(\Omega\,|\,\boldsymbol{\theta})}{\partial \theta_j} \frac{d\Omega}{2\pi} \right)_{i,j=1,\ldots,p}$$

$$= \frac{1}{2} \int_{-\pi}^{\pi} \nabla_{\boldsymbol{\theta}} \ln C_{XX}(\Omega\,|\,\boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}}^T \ln C_{XX}(\Omega\,|\,\boldsymbol{\theta}) \frac{d\Omega}{2\pi} = \boldsymbol{\Gamma}(\boldsymbol{\theta})$$

Maximum Likelihood Estimator

Whittle Estimator

Simple Approach for Designing Asymptotically Efficient Estimators

References to Chapter 5

[1] T.W. Anderson, *The Analysis of Time Series,* Wiley, 1971

[2] D.R. Brillinger, *Time Series: Data Analysis and Theory*, Holden Day, 1981

[3] G.E.P. Box and G.M. Jenkins, *Time Series Analysis: Forecasting and Control,* Holden Day, 1976

[4] K.O. Dzhaparidze, *Spectrum Analysis of Stationary Time Series: Parameter Estimation and Hypothesis Testing*, Springer, 1985

[5] E.J. Hannan, *Multiple Time Series,* Wiley, 1970

[6] G.M. Jenkins and D.G. Watts, *Spectral Analysis and its Applications,* Holden Day, 1978

[7] M.B. Priestley, *Spectral Analysis and Time Series*, Academic Press, 1984

[8] P. Stoica and R.L. Moses, *Spectral Analysis of Signals*, Prentice Hall, 2005