Chapter 4

Stationary TS Models

A time series is a sequence of random variables $\{X_t\}_{t=1,2,\ldots}$, hence it is natural to ask about distributions of these r.vs. There may be an infinite number of r.vs, so we consider multivariate distributions of random vectors, i.e. of finite subsets of the sequence $\{X_t\}_{t=1,2,\ldots}$.

Definition 4.1. *We define a* ***time series model*** *for the observed data* $\{x_t\}$ *to be a specification of all the joint distributions of the random vectors* $\boldsymbol{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$, $n = 1, 2, \ldots,$ *of which* $\{x_t\}$ *are possible realizations, that is all the probabilities*

$$P(X_1 \leq x_1, \ldots, X_n \leq x_n), \quad -\infty < x_1, \ldots, x_n < \infty, \quad n = 1, 2, \ldots.$$

□

Such a specification is rather impractical. Instead, we consider first and second-order moments of the joint distributions, i.e.,

$$\mathrm{E}(X_t) \quad \text{and} \quad \mathrm{E}(X_{t+\tau}X_t) \quad \text{for} \quad t = 1, 2, \ldots, \quad \tau = 0, 1, 2, \ldots.$$

and examine properties of the TS which depend on these. They are called **second-order properties**.

Remark 4.1. For the multivariate normal joint distributions, the first and second order moments completely determine the distributions. Hence for a TS having all the joint distributions normal the second-order properties of TS give its complete characterization (model).

4.1 Weak Stationarity and Autocorrelation

For an n dimensional random vector $\boldsymbol{X}$ we can calculate the variance-covariance matrix. However, a TS usually involves a very large (infinite in theory) number of r.vs. Then, there is a very large number of pairs of the variables and so we define so called **autocovariance** as an extension of the variance-covariance matrix . It is usually denoted by Greek letter γ and we write

$$\gamma_{(X_{t+\tau},X_t)} = \operatorname{cov}(X_{t+\tau}, X_t), \quad \text{for all indexes } t \text{ and lags } \tau. \tag{4.1}$$

Definition 4.2. *A time series $\{X_t\}$ is called **weakly stationary** or just stationary if*

1. $\operatorname{E} X_t = \mu_{X_t} = \mu < \infty$*, that is the expectation of X_t is finite and does not depend on t, and*
2. $\gamma_{(X_{t+\tau},X_t)} = \gamma_\tau$*, that is for each τ the autocovariance of r.vs $(X_{t+\tau}, X_t)$ does not depend on t (it is constant for a given lag τ).* □

Remark 4.2. If $\{X_t\}$ is a weakly stationary TS then the autocovariance $\gamma_{(X_{t+\tau},X_t)}$ may be viewed as a function of one variable τ. It is called **autocovariance function (ACVF)** and we often write $\gamma_X(\tau)$ or just $\gamma(\tau)$ when it is clear which TS it refers to.

Note that

$$\gamma(0) = \operatorname{var}(X_t),$$

that is, the variance is constant for all t.

Remark 4.3. Similarly, we define so called **autocorrelation function (ACF)** as

$$\rho_X(\tau) = \frac{\gamma_X(\tau)}{\gamma_X(0)} = \operatorname{corr}(X_{t+\tau}, X_t) \quad \text{for all } \ t, \tau. \tag{4.2}$$

Example 4.1. i.i.d. noise
$\{X_t\}$ is a sequence of r.vs which have no trend or seasonality and are independently, identically distributed (i.i.d.). Then the joint c.d.f. can be written as

$$\begin{aligned} F(x_1, \ldots, x_n) &= P(X_1 \leq x_1, \ldots, X_n \leq x_n) \\ &= P(X_1 \leq x_1) \ldots P(X_n \leq x_n) \\ &= F(x_1) \ldots F(x_n). \end{aligned}$$

So the conditional distribution of $X_{n+\tau}$ given values of $(X_1, \dots, X_n)$ is

$$P(X_{n+\tau} \le x | X_1 = x_1, \dots, X_n = x_n) = P(X_{n+\tau} \le x).$$

It means that knowledge of the past has no value for predicting future.

$\{X_t\}$ has zero trend, hence $\mathrm{E}\, X_t = 0$. If $\mathrm{E}(X_t^2) = \sigma^2 < \infty$, then it has finite variance and by the independence we get

$$\gamma(\tau) = \begin{cases} \sigma^2, & \text{if } \tau = 0, \\ 0, & \text{if } \tau \ne 0. \end{cases}$$

This meets the requirements of Definition 4.2. Hence i.i.d. noise with finite second moment is a weakly stationary process, usually denoted by

$$\{X_t\} \sim IID(0, \sigma^2).$$

□

Example 4.2. White noise
A sequence $\{X_t\}$ of uncorrelated r.vs, each with zero mean and variance σ^2 is called **white noise**. It is denoted by

$$\{X_t\} \sim WN(0, \sigma^2).$$

The name 'white' comes from the analogy with white light and indicates that all possible periodic oscillations are present with equal strength.

A particularly useful white noise process is the Gaussian white noise series of iid r.vs, which we denote by

$$\{X_t\} \underset{iid}{\sim} \mathcal{N}(0, \sigma^2).$$

White noise meets the requirements of the definition of weak stationarity. □

Remark 4.4. Note that every $IID(0, \sigma^2)$ series is $WN(0, \sigma^2)$, but not conversely. In general uncorrelation does not imply independence. Gaussian white noise is however an IID process.

Example 4.3. MA(1) process
The series defined as the following combination of two neighbouring White Noise variables

$$X_t = Z_t + \theta Z_{t-1}, \quad t = 0, \pm 1, \pm 2, \dots, \tag{4.3}$$

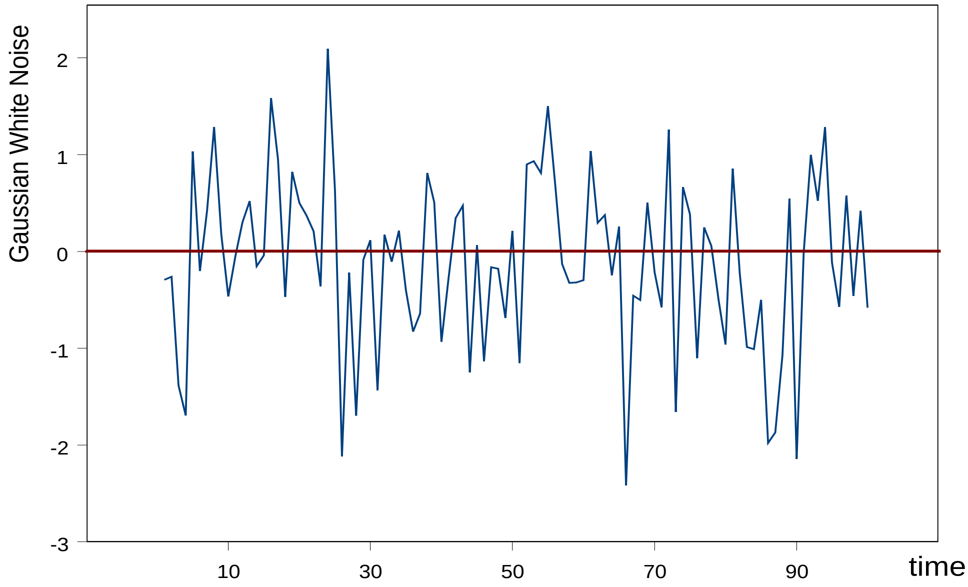

Figure 4.1: Simulated Gaussian White Noise Time Series

where

$$\{Z_t\} \sim WN(0, \sigma^2),$$

and θ is a constant, is called **first order moving average**, which we denote by **MA(1)**.

Is the MA(1) a weakly stationary series?

From equation 4.3 we obtain

$$\mathrm{E}(X_t) = \mathrm{E}(Z_t + \theta Z_{t-1}) = \mathrm{E}(Z_t) + \theta\,\mathrm{E}(Z_{t-1}) = 0.$$

Now, we need to check if the autocovariance function does not depend on time, i.e., it depends only on lag τ.

$$\begin{aligned}
\mathrm{cov}(X_t, X_{t+\tau}) &= \mathrm{cov}(Z_t + \theta Z_{t-1}, Z_{t+\tau} + \theta Z_{t-1+\tau}) \\
&= \mathrm{E}[(Z_t + \theta Z_{t-1})(Z_{t+\tau} + \theta Z_{t-1+\tau})] \\
&\quad - \underbrace{\mathrm{E}(Z_t + \theta Z_{t-1})}_{=0}\underbrace{\mathrm{E}(Z_{t+\tau} + \theta Z_{t-1+\tau})}_{=0} \\
&= \mathrm{E}(Z_t Z_{t+\tau}) + \theta\,\mathrm{E}(Z_t Z_{t-1+\tau}) + \theta\,\mathrm{E}(Z_{t-1} Z_{t+\tau}) + \theta^2\,\mathrm{E}(Z_{t-1} Z_{t-1+\tau}).
\end{aligned}$$

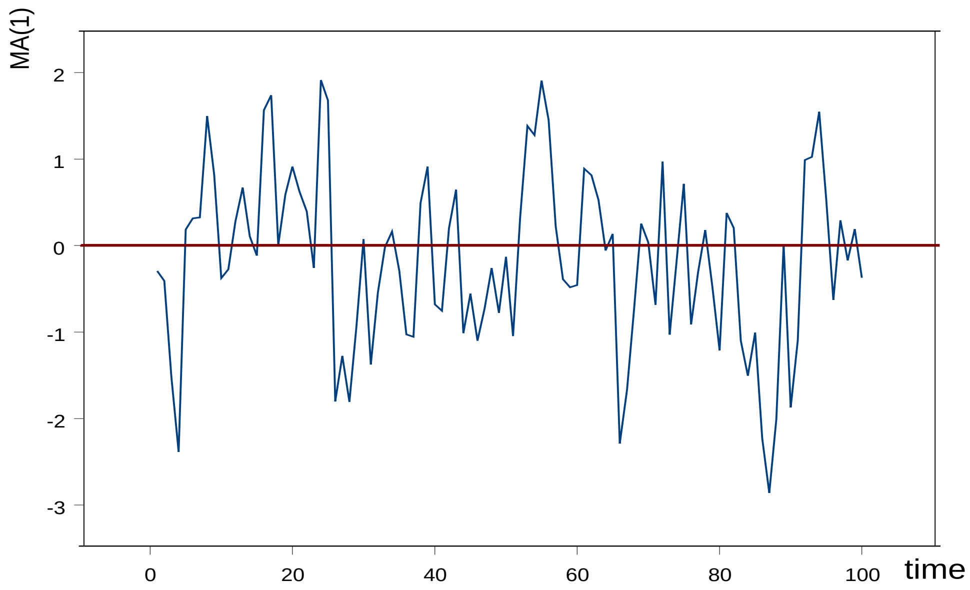

Figure 4.2: Simulated MA(1) Time Series

Now, taking various values of the lag τ we obtain

$$\operatorname{cov}(X_t, X_{t+\tau}) = \begin{cases} \mathrm{E}(Z_t^2) + \theta^2\,\mathrm{E}(Z_{t-1}^2) = (1+\theta^2)\sigma^2, & \text{if } \tau = 0, \\ \theta\,\mathrm{E}(Z_t^2) = \theta\sigma^2, & \text{if } \tau = \pm 1, \\ 0, & \text{if } |\tau| > 1. \end{cases} \tag{4.4}$$

Hence, the covariance does not depend on t and we can write the autocovariance function of lag τ as

$$\gamma_X(\tau) = \operatorname{cov}(X_t, X_{t+\tau}) \quad \text{for any } t.$$

So, the conclusion is that MA(1) is a weakly stationary process. Also, from (4.4) we obtain the form of the autocorrelation function

$$\rho_X(\tau) = \begin{cases} 1, & \text{if } \tau = 0, \\ \frac{\theta}{1+\theta^2} & \text{if } \tau = \pm 1, \\ 0, & \text{if } |\tau| > 1. \end{cases} \tag{4.5}$$

Figure 4.2 shows the MA(1) process obtained from the simulated white noise taking $\theta = 0.5$.

□

4.1.1 Sample Autocorrelation Function

The autocorrelation function is a helpful tool in assessing the degree of dependence and in recognizing what kind of model the TS follows. When we try to fit a

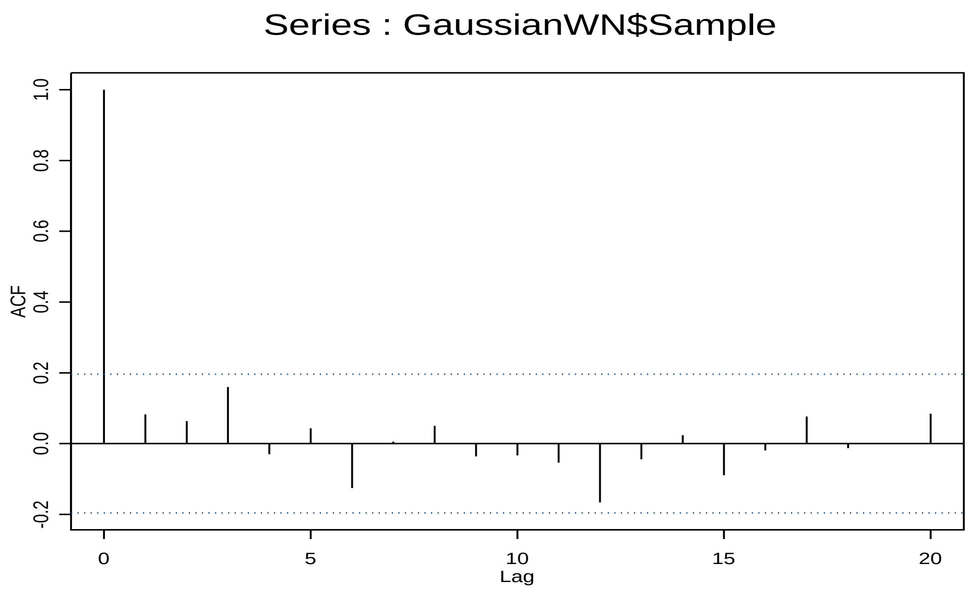

Figure 4.3: Correlogram of the Simulated Gaussian White Noise Time Series

model to an observed TS we use so called sample autocorrelation function based on the data. It is defined analogously to the ACF for a TS $\{X_t\}$.

Definition 4.3. *Let $x_1, \ldots, x_n$ be observations of a TS. The* ***sample autocovariance function*** *is defined as*

$$\widehat{\gamma}(\tau) = \frac{1}{n} \sum_{t=1}^{n-|\tau|} (x_t - \bar{x})(x_{t+|\tau|} - \bar{x}), \quad -n < \tau < n \tag{4.6}$$

where

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

The ***sample autocorrelation function*** *is defined as*

$$\widehat{\rho}(\tau) = \frac{\widehat{\gamma}(\tau)}{\widehat{\gamma}(0)}, \quad -n < \tau < n. \tag{4.7}$$

Remark 4.5. For lag $\tau \geq 0$ the sample autocovariance function is approximately equal to the sample covariance of the $n - \tau$ pairs $(x_1, x_{1+\tau}), \ldots, (x_{n-\tau}, x_n)$. Note that, in (4.6), we divide the sum by n, not by $n - \tau$ and also we use the overall mean $\bar{x}$ for both x_t and $x_{t+\tau}$.

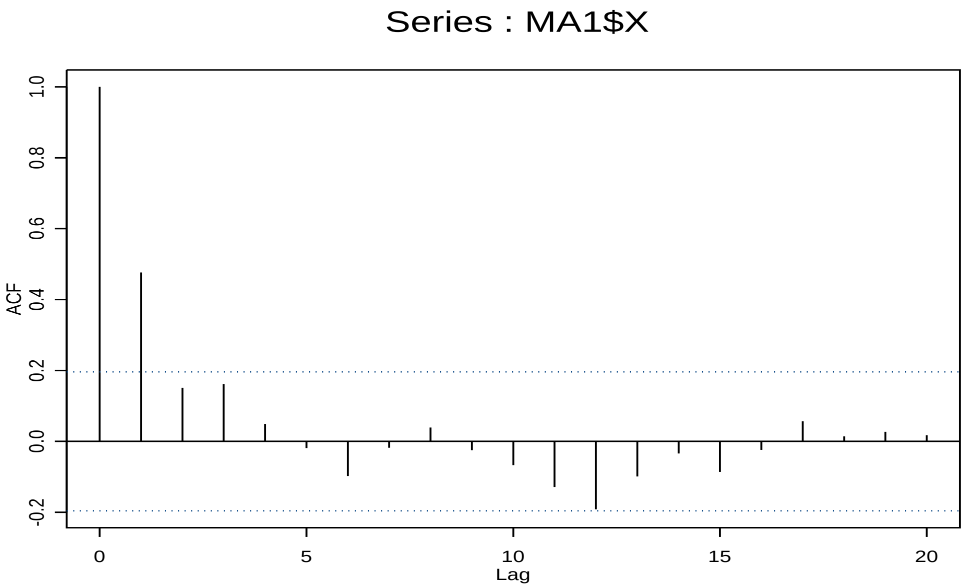

Figure 4.4: Correlogram of the Simulated MA(1) Time Series

A graph of sample autocorrelation function is called **correlogram**.

Figures 4.3 and 4.4, respectively, show the correlogram of the Gaussian white noise time series given in Figure 4.1 and the correlogram of the MA(1) TS with $\theta = 0.5$ calculated from the white noise. As expected, there is no significant correlation for lag $\tau \geq 1$ for the white noise, but there is one for the MA(1) for lag $\tau = 1$.

The role of ACF in prediction

Suppose that $\{X_t\}$ is a stationary Gaussian TS and we have observed X_n. We would like to predict to predict $X_{n+\tau}$ with high precision. The Mean Square Error is a good measure of precision of the prediction,

$$MSE = \mathrm{E}[X_{n+\tau} - f(X_{n+\tau}|X_n)]^2$$

and is minimized when the function f is the conditional expectation of $X_{n+\tau}$ given X_n,

$$f(X_{n+\tau}|X_n) = \mathrm{E}(X_{n+\tau}|X_n).$$

For a Gaussian stationary TS we have, see (3.33),

$$\mathrm{E}(X_{n+\tau}|X_n = x_n) = \mu_{n+\tau} + \rho(\tau)\sigma_{n+\tau}\sigma_n^{-1}(x_n - \mu_n) = \mu + \rho(\tau)(x_n - \mu).$$

Then, see (3.34),

$$MSE = var(X_{n+\tau}|X_n = x_n) = \sigma^2(1 - \rho(\tau)).$$

It shows that as $\rho \longrightarrow 1$ the value of precision measure $MSE \longrightarrow 0$. The higher is the correlation at lag τ the more precise is prediction of $X_{n+\tau}$ based on the observed X_n. Similar conclusions can be drawn for prediction of $X_{n+\tau}$ based on the observed $X_n, X_{n-1}, \ldots$. We will come back to this problem later.

4.1.2 Properties of ACVF and ACF

First we examine some basic properties of the Autocovariance function (ACVF).

Proposition 4.1. *The ACVF of a stationary TS is a function $\gamma(\cdot)$ such that*

1. $\gamma(0) \geq 0$,
2. $|\gamma(\tau)| \leq \gamma(0)$ *for all* τ,
3. $\gamma(\cdot)$ *is even, i.e.,*
$$\gamma(\tau) = \gamma(-\tau), \quad \textit{for all } \tau.$$

Proof. 1. Obvious, as $\gamma(0) = \mathrm{var}(X_t) \geq 0$.

2. From the definition of correlation (3.30) and stationarity of the TS we have
$$|\gamma(\tau)| = |\rho(\tau)|\sigma^2,$$
where $\sigma^2 = \mathrm{var}(X_t)$. Also, $|\rho(\tau)| \leq 1$. Hence
$$|\gamma(\tau)| = |\rho(\tau)|\sigma^2 \leq \sigma^2 = \gamma(0).$$

3. Here we have
$$\gamma(\tau) = \mathrm{cov}(X_{t+\tau}, X_t) = \mathrm{cov}(X_t, X_{t+\tau}) = \gamma(-\tau).$$

□

Another important property of the ACVF is given by the following theorem.

Theorem 4.1. *A real-valued function defined on the integers is the autocovariance function of a stationary TS if and only if it is even and nonnegative definite.*

Proof. We say that a real-valued function κ defined on integers is nonnegative definite if

$$\sum_{i,j=1}^{n} a_i \kappa(i-j) a_j \geq 0 \tag{4.8}$$

for all positive integers n and real-valued vectors $\boldsymbol{a} = (a_1, \ldots, a_n)^{\mathrm{T}}$.

It is easy to show that an ACVF is nonnegative definite and this is what we do below. Take vector r.v. $\boldsymbol{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ whose variance-covariance matrix $\boldsymbol{V}$ is given by

$$\boldsymbol{V} = \begin{pmatrix} \gamma(0) & \gamma(1-2) & \ldots & \gamma(1-n) \\ \gamma(2-1) & \gamma(0) & \ldots & \gamma(2-n) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \ldots & \gamma(0) \end{pmatrix}.$$

Then, denoting $\boldsymbol{Z} = (X_1 - \mathrm{E}\, X_1, \ldots, X_n - \mathrm{E}\, X_n)^{\mathrm{T}}$, we can write

$$\begin{aligned} 0 \leq \operatorname{var}(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{X}) &= \mathrm{E}[(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{Z})(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{Z})^{\mathrm{T}}] \\ &= \mathrm{E}[\boldsymbol{a}^{\mathrm{T}}\boldsymbol{Z}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{a}] \\ &= \boldsymbol{a}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{a} = \sum_{i,j=1}^{n} a_i \gamma(i-j) a_j. \end{aligned}$$

Hence $\gamma(\tau)$ is a nonnegative definite function. □