Handout 6

[Chapters 1.4 and 1.6 in Brockwell and Davis.]

You will also find below plot of temp and loess fit (window $q = 20$, $span = .25$), plot of the rough $\hat{X}_t = Y_t - \hat{m}_t$ against time, plot of $\hat{X}_t$ against $\hat{X}_{t+1}$ and the plot of the autocorrelations obtained from $\{\hat{X}_t\}$ with a $\pm 1.96/\sqrt{n}$ bars.

Mathematical set up

In order to model the rough $\{X_t\}$, we have to understand its properties. Over the years, methods have been developed to handle the rough. There are three properties of the plot of $\{X_t\}$ against time that make it easy to analyze and model the rough.

a) The series $\{X_t\}$ fluctuates about a constant value μ,

b) The size of the fluctuations about the constant value μ do not change over time.

c) The correlations of lag 1 (i.e., correlation between X_t and X_{t+1}) are the same in different time segments, correlations of lag 2 ((i.e., correlation between X_t and X_{t+2}) are the same in different time segments. In general, correlations of lag j ((i.e., correlation between X_t and X_{t+j}) are the same in different time segments (for any $j = 1, 2, 3.....$).

It is important to keep in mind that not all rough series satisfy these three. If a series does not, then there are methods to handle them. But these three properties provide a convenient framework for analysis of the series of roughs $\{X_t\}$. A series satisfying the three above properties is call a (weak) stationary sequence. Mathematically, we can describe them as

a) $E(X_t) = \mu$, for all t,

b) $Var(X_t)$ is the same for all t,

c) $Corr(X_t, X_{t+j})$ is the same for all t (for any positive integer j). Denote $Corr(X_t, X_{t+j})$ by $\rho(j)$.

[We can combine (b) and (c) into one condition as indicated in the Remark below.]

We will introduce a few more notations. Note that we have, by the definition of correlation

$$\rho(j) = Corr(X_t, X_{t+j}) = \frac{Cov(X_t, X_{t+j})}{\sqrt{Var(X_t)Var(X_{t+j})}},$$

$$Cov(X_t, X_{t+j}) = \rho(j)\sqrt{Var(X_t)Var(X_{t+j})}$$

Since $Var(X_t)$ is the same for all t, we have $Cov(X_t, X_{t+j})$ is proportional to $\rho(j)$, i.e., $Cov(X_t, X_{t+j})$ is the same for all t. Denote by $\gamma(j)$ this common value of $Cov(X_t, X_{t+j})$. Note that $Var(X_t) = Cov(X_t, X_t)$. So we can denote $Var(X_t)$ by $\gamma(0)$. Combining all these we then have

$$\gamma(j) = \rho(j)\gamma(0), \ \rho(j) = \gamma(j)/\gamma(0).$$

Here are a few **important remarks**.

Remarks:

1. We can combine the conditions (b) and (c) above by a single condition:

"$Cov(X_t, X_{t+j})$ is the same for all t (for any NONNEGATIVE integer j)".

2. The quantities $\gamma(j)$ and $\rho(j)$ are called **autocovariance** and **autocorrelation** of order j, respectively. Note that $\rho(0) = 1$. Here is an important fact to keep in mind

$$\rho(-j) = Corr(X_{t+j}, X_t) = Corr(X_t, X_{t+j}) = \rho(j),\ j = 0, 1, ...,$$
$$\gamma(-j) = Cov(X_{t+j}, X_t) = Cov(X_t, X_{t+j}) = \gamma(j),\ j = 0, 1, ...$$

This fact tells us it is enough to investigate the autocorrelation $\rho(j)$ for nonnegative integer j and there is no need to look at $\rho(-j)$.

c) If a sequence $\{X_t\}$ is stationary with mean μ, then the centered sequence $\{X_t^{(c)} = X_t - \mu\}$ is also stationary with mean zero. The two sequences $\{X_t^{(c)}\}$ and $\{X_t\}$ have the same autocovariances and autocorrelations.

Some Examples

Notational issues: In Examples 1-6 below, the notation for white noise is $\{\varepsilon_t\}$. In the text (Brockwell and Davis), the notations for white noise is $\{Z_t\}$.

Example 1. If $\{X_t\}$ are independent and identically distributed (i.i.d.) with mean μ and variance σ^2, then note that

$$\rho(j) = Corr(X_t, X_{t+j}) = \begin{cases} 0 & j = 1, 2, ... \\ 1 & j = 0. \end{cases}$$

"Identically distributed" means that the probability distribution of X_t is the same for all t. Note that, an i.i.d. sequence $\{X_t\}$ is stationary, albeit a rather simple one. If $\{X_t\}$ is i.i.d. with zero mean (i.e. $\mu = 0$) and variance σ^2, then the sequence $\{X_t\}$ is called a WHITE NOISE.

Example 2. If a mean zero sequence $\{X_t\}$ has the structure

$$X_t = \varepsilon_t + \theta\varepsilon_{t-1},$$

where the sequence $\{\varepsilon_t\}$ are i.i.d. with mean zero and variance σ^2, then $\{X_t\}$ is called a MOVING AVERAGE of order 1 (notation: $MA(1)$). In this handout you will find a plot of a simulated $MA(1)$ model (R command: arima.sim).

When $\{X_t\}$ is $MA(1)$, then for all t,

$$E(X_t) = 0,\ Var(X_t) = (1 + \theta^2)\sigma^2.$$

Also note that for any t,

$$\gamma(j) = Cov(X_t, X_{t+j}) = \begin{cases} (1+\theta^2)\sigma^2 & j = 0 \\ \theta\sigma^2 & j = 1 \\ 0 & j \geq 2. \end{cases}$$

Hence $\{X_t\}$ is stationary. Note that

$$\rho(j) = \begin{cases} 1 & j = 0 \\ \theta/(1+\theta^2) & j = 1 \\ 0 & j \geq 2. \end{cases}$$

Here the autocorrelation is nonzero for lag 1 and is zero otherwise.
If $\theta = 0.5$ and $\sigma = 2$, then we have

$$\gamma(0) = (1 + (0.5)^2)(2^2) = 5,\ \gamma(1) = (0.5)(2^2) = 2,\ 0 = \gamma(2) = \gamma(3) = \cdots,$$
$$\rho(0) = 1,\ \rho(1) = 0.4,\ 0 = \rho(2) = \rho(3) = \cdots.$$

Note that for an $MA(1)$ sequence with mean μ, the representation is as follows. The centered sequence (or the mean subtracted sequence) $X_t^{(c)} = X_t - \mu$ which has mean zero and this centered sequence has the form given in (1), i.e.,

$$X_t^{(c)} = \varepsilon_t + \theta\varepsilon_{t-1}, or\ X_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}.$$

Example 3. A mean zero MOVING AVERAGE ($MA(2)$) sequence (with mean zero) of order 2 has the form

$$X_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2},$$

where $\{\varepsilon_t\}$ are i.i.d. with mean zero. An $MA(2)$ sequence with mean μ has the form

$$X_t^{(c)} = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2}, or\ X_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2},$$

where $X_t^{(c)} = X_t - \mu$.
Example 4. If a mean zero sequence $\{X_t\}$ has the structure

$$X_t = \phi X_{t-1} + \varepsilon_t,$$

where, $-1 < \phi < 1$, $\{\varepsilon_t\}$ are i.i.d. with mean zero and variance σ^2. This is called an AUTOREGRESSIVE sequence of order 1 (notation: $AR(1)$). On the last page there is a plot of a simulated AR(1) model with $\phi = 0.5$ (R command: arima.sim).When $|\phi| \geq 1$, the sequence $\{X_t\}$ is not stationary as shown in Example 6 below.

For the $AR(1)$ case, in order to satisfy the condition that $Var(X_t)$ is the same for all t, we must have

$$Var(X_t) = \phi^2 Var(X_{t-1}) + \sigma^2, i.e.,\ Var(X_t) = \sigma^2/(1-\phi^2).$$

In this case it turns out

$$\gamma(j) = \phi^j(1-\phi^2)^{-1}\sigma^2,\ \rho(j) = \phi^j,\ j = 0, 1, 2, ...$$

Note that the autocorrelations here are nonzero (unless $\phi = 0$), but $\rho(j)$ decay rapidly to zero as the lag j becomes large.

An $AR(1)$ sequence with mean μ has the form

$$X_t^{(c)} = \phi X_{t-1}^{(c)} + \varepsilon_t \ \ or\ X_t - \mu = \phi(X_{t-1} - \mu) + \varepsilon_t,$$

where $X_t^{(c)} = X_t = \mu$ and where $\{\varepsilon_t\}$ are i.i.d. with mean zero and variance σ^2, and $-1 < \phi < 1$. Note that this can be rewritten in the form

$$X_t = \beta_0 + \beta_1 X_{t-1} + \varepsilon_t,$$

where $\beta_0 = \mu(1-\phi)$ and $\beta_1 = \phi$. This is clearly a simple linear regression model with X_t as the dependent variable and X_{t-1} as the independent variable. And this justifies the name "autoregressive".

Example 5. A mean zero $AR(2)$ (autoregressive model of order 2) has the form

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t,$$

where $\{\varepsilon_t\}$ are i.i.d. with mean zero and variance σ^2. In order for the sequence to be stationary, there are restrictions on ϕ_1 and ϕ_2 and these issues are discussed in chapter 2 of the text. For an $AR(2)$ sequence, the autocorrelations $\rho(j)$ have complicated formulas, but $\rho(j)$ decay rapidly to zero as lag j becomes large.

An AR(2) sequence with mean μ has the form

$$X_t^{(c)} = \phi_1 X_{t-1}^{(c)} + \phi_2 X_{t-2}^{(c)} + \varepsilon_t, or\ X_t - \mu = \phi_1(X_{t-1} - \mu) + \phi_2(X_{t-2} - \mu) + \varepsilon_t,$$

where $X_t^{(c)} = X_t - \mu$.

Example 6.(Random walk)

A mean zero $AR(1)$ sequence with $\phi = 1$ is called a random walk and it has the form $X_t = X_{t-1} + \varepsilon_t$, where $\{\varepsilon_t\}$ are i.i.d. with mean zero and variance σ^2. For this sequence

$$Cov(X_t, X_{t+j}) = t\sigma^2,\ j \geq 1.$$

Clearly this is not stationary since $Cov(X_t, X_{t+j})$ is not the same for all t. Plot of a generated random walk sequence (with variance $\sigma = 1$ and $n = 100$) is given here. Note that the plot clearly shows that it is not stationary. It does not fluctuate about a constant value and, as a matter of fact, it can be "explosive" (i.e., the magnitude of X_t can drift to infinity for large t).

Computing the autocorrelations from data.

Given a sequence $\{X_t : t = 1, ..., n\}$ (or an estimate $\{\hat{X}_t\}$ as in the temperature data), we can obtain estimates of the autocorrelation functions. Here is the method that is followed in all packages. First the autocovariances are calculated and then the autocorrelations:

$$\hat{\gamma}(j) = \frac{1}{n} \sum_{t=1}^{n-j} (X_t - \bar{X})(X_{t+j} - \bar{X}),\ \hat{\rho}(j) = \hat{\gamma}(j)/\hat{\gamma}(0),\ j = 0, 1, 2....$$

For the temperature data $\{Y_t\}$, we have applied loess with span=0.25 to estimate the trend $\{m_t\}$.The plot of the estimated rough $\{\hat{X}_t = Y_t = \hat{m}_t\}$ looks stationary. Some estimated autocovariances of $\{\hat{X}_t\}$ (the function "acf" in R) are

$$\hat{\gamma}(0) = 0.00717, \hat{\gamma}(1) = 0.00215, \hat{\gamma}(2) = -0.00056, \hat{\gamma}(3) = -0.00107, \hat{\gamma}(4) = -0.00045,$$
$$\hat{\rho}(0) = 1, \hat{\rho}(1) = 0.300, \hat{\rho}(2) = -0.078, \hat{\rho}(3) = -0.149, \hat{\rho}(4) = 0.063.$$

In data analysis, usually these estimated autcorrelations are plotted in order to investigate the dependence lags. Under the assumptions of normality, it can be shown mathematically that the distribution of $\hat{\rho}(j)$ is approximately normally distributed with mean $\rho(j)$ and a variance which has a complicated formula. However, when $\rho(j) = 0$, the variance is approximately equal to $1/n$. Often we may be interested in testing $H_0 : \rho(j) = 0$ against $H_1 : \rho(j) \neq 0.$, and the test statistic is $z^* = \sqrt{n}\hat{\rho}(j)$. So, when the level of significance is $\alpha = 0.05$, we can reject H_0 if $|z^*| > 1.96$, i.e. $|\hat{\rho}(j)| > 1.96/\sqrt{n}$. Alternatively we can check if $\hat{\rho}(j)$ is inside or outside the range $\pm 1.96/\sqrt{n}$. It is a common practice to plot the estimated autocorrelations against lag j and it is called the ACF (autocorrelation function) plot (the function "acf" in R).

Diagnostics.

Diagnostic tools are employed to investigate issues such as if there is trend, if the assumption of equal variance is reasonable, if the sequence is i.i.d., if the assumption of normality is reasonable etc. Graphical tools can often be used. Here are a few diagnostic methods - some are graphical, some are formal testing procedures.

a) Plot of the series against time reveals if there is a trend or if the assumption of equal variance (across time) is reasonable.

b) Checking if the sequence is i.i.d.

Plot of the estimated autocorrelations (also called the ACF plot) along with $\pm 1.96/\sqrt{n}$ bars may reveal if the autocorrelations are close to zero. The ACF plot for the estimated rough $\{\hat{X}_t\}$ for the temperature data indicates that correlations of lags 1, 3 and 6 may not be close to zero. More analysis is needed for this data.

c) Formal tests for checking if the sequence is i.i.d.

Suppose we decide to test, at a given level α, $H_0 : \rho(1) = \cdots = \rho(h) = 0$ against H_1 : at least one of $\rho(1), ..., \rho(h)$ is nonzero.

We will mention two tests: Portmanteau and Ljung-Box. The test statistics are

$$Q = n \sum_{j=1}^{h} \hat{\rho}(j)^2, \; Q_{LB} = n(n+2) \sum_{j=1}^{h} \hat{\rho}(j)^2/(n-j).$$

Under H_0, each of the two the statistics (Q and Q_{LB}) has an approximate chi-square distribution with h degrees of freedom (denoted as χ^2_h).

So we can reject the null hypothesis if $Q > \chi^2_{1-a}$ or $Q_{LB} > \chi^2_{1-\alpha}$, where area to the right of $\chi^2_{1-\alpha}$ is α under the chi-square curve with h degrees of freedom.

For the temperature data, we have calculated these statistics for $h = 10$ and the values are $Q = 33.22$ and $Q_{LB} = 34.55$.

From the chi-square table we find $\chi^2_{0.95} = 18.31$, $\chi^2_{0.99} = 23.21$. Clearly the null hypothesis of $\{X_t\}$ being i.i.d. can be rejected.

d) Test for an increasing or decreasing trend: (Rank test).

Let P equal to the number of times $X_j > X_i, j > i$. If the sequence $\{X_t\}$ are i.i.d., then the distribution of P is approximately normal with mean $\mu_P = n(n-1)/4$ and variance $\sigma^2_P = n(n-1)(2n+5)/72$. So we can calculate the statistic $z^* = (P - \mu_P)/\sigma_P$ and, reject the null hypothesis of no trend (increasing or decreasing) if $|z^*| > z_{1-\alpha/2}$, where area to the right $z_{1-\alpha/2}$ is $\alpha/2$ under the standard normal curve.

e) Normality.

One can examine the histogram of $\{X_t\}$ to check if the assumption of normality is justifiable. Alternatively, one can look at the normal probability plot of $\{X_t\}$. If this graph is linear, then the assumption of normality is reasonable. For the temperature data, the assumption of normality seems to be reasonable if we examine the histogram and the normal probability plot.

Appendix A

Here are some useful R commands:

Loess: loess,

Normal probability plot: qqnorm,

ACF plot: acf,

Histogram: hist,

Simulating AR or MA series: arima.sim.

Appendix B

Autocovariance and autocorrelation calculations.

In order to calculate the autocovariances and autocorrelations, we need some results form probability. Here is an important result.

Fact: Let $V_1, V_2, , , ,$ and $W_1, W_2, ...$ be random variables, and $c_1, c_2, ...$and $d_1, d_2, ...$be constants.

a) The following are true

$$\begin{aligned}
E(c_1V_1 + c_2V_2) &= c_1E(V_1) + c_2E(V_2), \\
Var(c_1V_1 + c_2V_2) &= c_1^2Var(V_1) + c_2^2Var(V_2) + 2c_1c_2Cov(V_1, V_2), \\
Cov(c_1V_1 + c_2V_2, d_1W_1 + d_2W_2) &= c_1d_1Cov(V_1, W_1) + c_1d_2Cov(V_1, W_2) \\
&\quad + c_2d_1Cov(V_2, W_1) + c_2d_2Cov(V_2, W_2).
\end{aligned}$$

b) The following are generalizations of the results in (a)

$$\begin{aligned}
E(\sum c_iV_i) &= \sum c_iE(V_i), \\
Var(\sum c_iV_i) &= \sum c_i^2Var(V_i) + \sum_i \sum_{j \neq i} c_ic_jCov(V_i, V_j) \\
&= \sum_i \sum_j c_ic_jCov(V_i, V_j), \\
Cov(\sum c_iV_i, \sum d_jW_j) &= \sum_i \sum_j c_id_jCov(V_i, W_j).
\end{aligned}$$

c) A consequence of the results in part (b) is that if V_i are mutually uncorrelated (i.e., $Cov(V_i, V_j) = 0$ whenever $i \neq j$), then

$$Var(\sum c_iV_i) = \sum c_i^2Var(V_i).$$

[It is important to keep in mind that, for any random variable V, $Cov(V, V) = Var(V)$.]

Autocovariances and autocorrelations: MA(1)

Since $\{\varepsilon_i\}$ are mutually uncorrelated, we have

$$\begin{aligned}
\gamma(0) &= Var(X_t) = Var(\varepsilon_t + \theta\varepsilon_{t-1}) = Var(\varepsilon_t) + \theta^2Var(\varepsilon_{t-1}) \\
&= \sigma^2 + \theta^2\sigma^2 = (1 + \theta^2)\sigma^2, \\
\gamma(1) &= Cov(X_t, X_{t+1}) = Cov(\varepsilon_t + \theta\varepsilon_{t-1}, \varepsilon_{t+1} + \theta\varepsilon_t) \\
&= Cov(\varepsilon_t, \varepsilon_{t+1}) + \theta Cov(\varepsilon_t, \varepsilon_t) + \theta Cov(\varepsilon_{t-1}, \varepsilon_{t+1}) + \theta^2Cov(\varepsilon_{t-1}, \varepsilon_t) \\
&= 0 + \theta\sigma^2 + 0 + 0 = \theta\sigma^2, \\
\gamma(2) &= Cov(X_t, X_{t+2}) = Cov(\varepsilon_t + \theta\varepsilon_{t-1}, \varepsilon_{t+2} + \theta\varepsilon_{t+1}) = 0.
\end{aligned}$$

Note that $\gamma(2) = 0$ since there is no common ε-term in X_t and X_{t+2}. Incidentally, the same argument applies for $\gamma(3), \gamma(4), ...$ Since $\rho(j) = \gamma(j)/\gamma(0)$, we have

$$\rho(0) = 1, \ \rho(1) = [\theta\sigma^2]/[(1+\theta^2)\sigma^2] = \theta/(1+\theta^2),$$
$$\rho(2) = 0/\gamma(0) = 0, \rho(3) = 0, ...$$

Autocovariances and autocorrelations: AR(1)

The formulas for autocovariances and auto correlations for AR(1) are simple. But there are no simple formulas for AR(2) or higher order autoregressive sequences,

In order to carry out the calculations note that $X_{t-1}, X_{t-2}, ...$ are uncorrelated with ε_t. Now

$$\gamma(0) = Var(X_t) = Var(\phi X_{t-1} + \varepsilon_t) = \phi^2 Var(X_t) + Var(\varepsilon_t) = \phi^2\gamma(0) + \sigma^2.$$

The last identity tells us $\gamma(0)$ must be equal to $(1-\phi^2)^{-1}\sigma^2$.Since X_{t-1} is uncorrelated with ε_t, we have

$$\gamma(1) = Cov(X_t, X_{t-1}) = Cov(\phi X_{t-1} + \varepsilon_t, X_{t-1}) = \phi Cov(X_{t-1}, X_{t-1}) = \phi\gamma(0).$$

Since X_{t-2} is uncorrelated with ε_t, we conclude that

$$\gamma(2) = Cov(X_t, X_{t-2}) = Cov(\phi X_{t-1} + \varepsilon_t, X_{t-2}) = \phi Cov(X_{t-1}, X_{t-2}) = \phi\gamma(1) = \phi^2\gamma(0).$$

A repetition of the same argument will show $\gamma(3) = \phi^3\gamma(0)$, $\gamma(4) = \phi^4\gamma(0), ...$ Hence we can conclude that $\gamma(j) = \phi^j\gamma(0)$, $j = 0, 1, 2, , , , ,$. Consequently, $\rho(j) = \gamma(j)/\gamma(0) = \phi^j, j = 0, 1, 2, ...$

Figure 1: Temperature series: 1850 - 2012

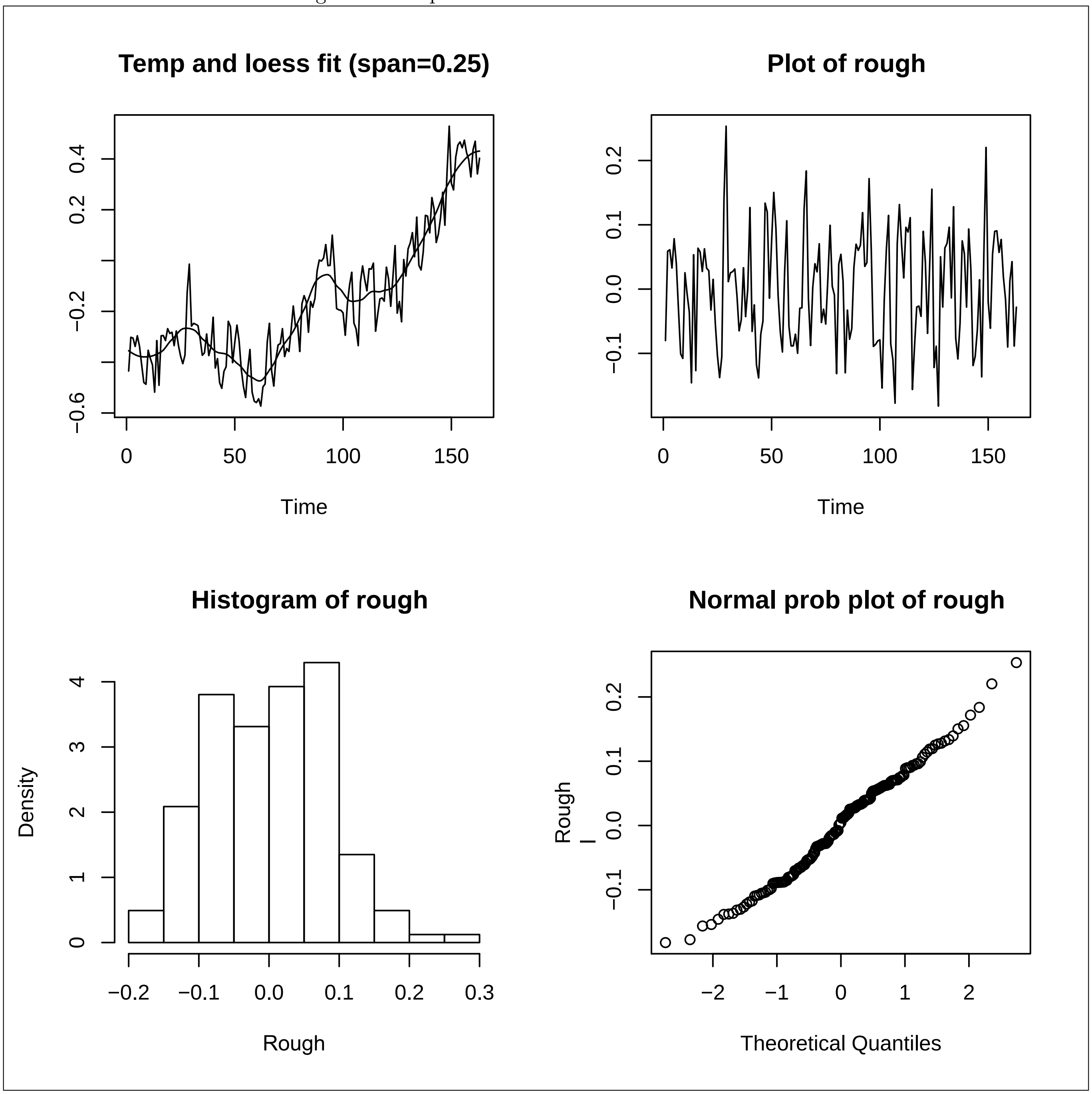

Figure 2: Temperature series: 1850 - 2012

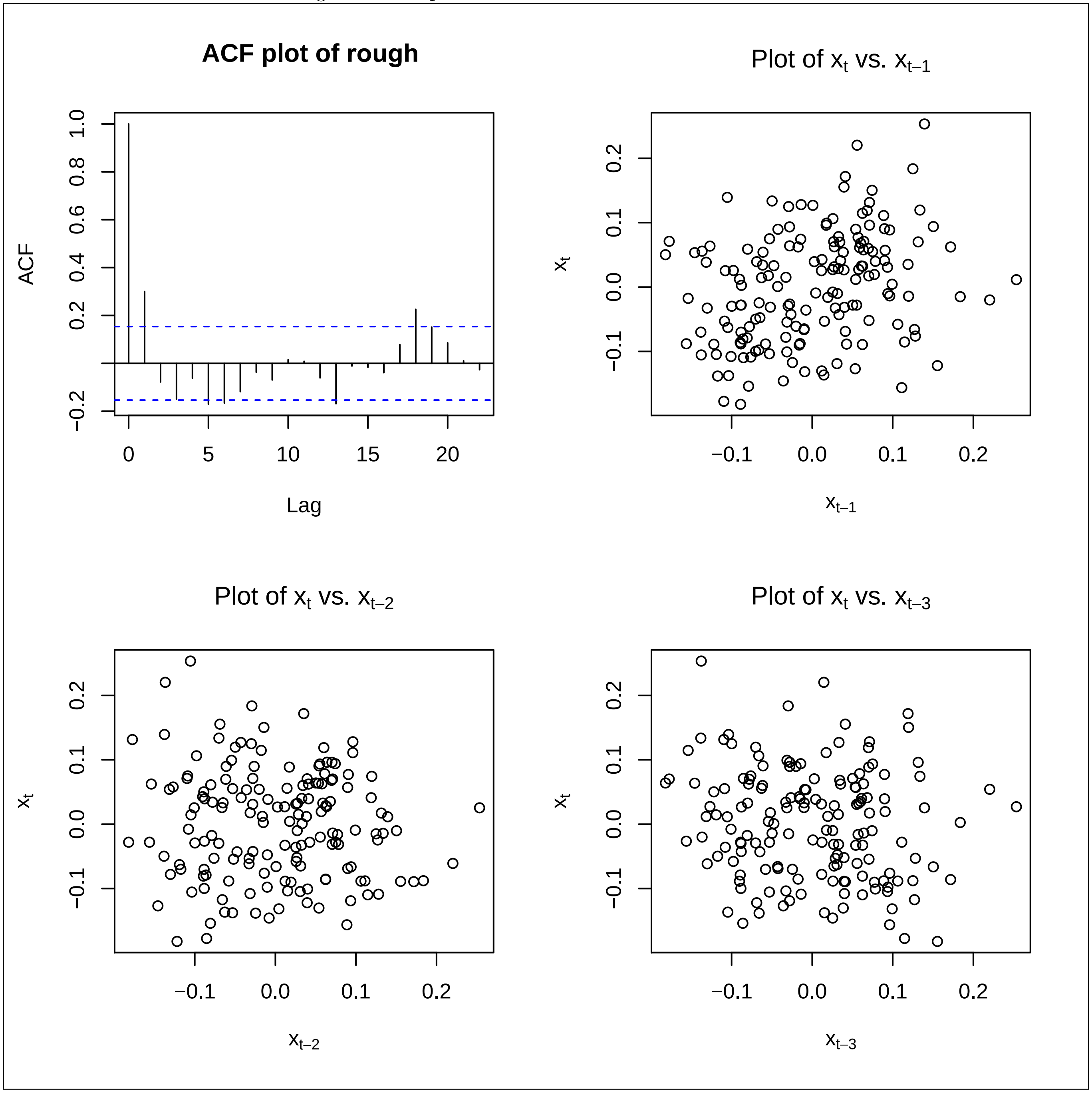

Figure 3: Simulated series