KTH Mathematics

Examination in SF2943/SF2945 Time Series Analysis, August 25, 2014, 08:00–13:00.

Examiner: Filip Lindskog, tel. 790 7217, e-mail: lindskog@kth.se

Allowed aids: Pocket calculator, "Formulas and survey, Time series analysis" by Jan Grandell, without notes.

Any notation introduced must be explained and defined. Arguments and computations must be detailed so that they are easy to follow.

The formulas $\sum_{k=1}^{n} k = n(n+1)/2$ and $\sum_{k=1}^{n} k^2 = n(n+1)(2n+1)/6$ may be useful.

- - - - - - - - - - - - - - - - - - - -

Problem 1

Consider a causal AR(1) process $\{X_t\}$. Due to imperfect measurement equipment, the prediction of X_{t+2} is based on noisy observations Y_s for $s \leq t$. Let

$$\begin{aligned} X_t &= \phi X_{t-1} + Z_t, \quad \{Z_t\} \sim \mathrm{WN}(0, \sigma_z^2), \\ Y_t &= X_t + W_t, \quad \{W_t\} \sim \mathrm{IID}(0, \sigma_w^2), \end{aligned}$$

where the noise sequences $\{Z_t\}$ and $\{W_t\}$ are independent.

(a) Determine the best linear predictor (minimizing the mean squared prediction error) of X_{t+2} based on Y_t, Y_{t-1}. (5 p)

(b) Explain why the expression for the predictor is reasonable/expected in the case when σ_z is much larger than σ_w and in the case when σ_w is much larger than σ_z. (5 p)

Problem 2

Consider the following estimates, for lags $0, 1, \ldots, 7$, of the ACVF and PACF based on a sample of size 1000 from a stationary time series:

ACVF	0.620	-0.200	0.005	-0.013	0.005	-0.012	-0.028	0.017
PACF	1	-0.323	-0.107	-0.060	-0.021	-0.028	-0.070	-0.017

Suggest an AR(p) or MA(q) model and estimate its parameters. (10 p)

Problem 3

Consider a causal AR(2) process $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t$, $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$. Determine the best linear predictor of X_{t+1} based on X_t, X_{t-1}, X_{t-2}. (10 p)

Problem 4

Consider the ARMA(1, 1) process $X_t - 0.5X_{t-1} = Z_t + 1.5Z_{t-1}$, $\{Z_t\} \sim \mathrm{WN}(0, 1)$.

(a) Express $\{X_t\}$ as an infinite order moving average process. (5 p)

(b) Compute the variance of $\{X_t\}$. (5 p)

Problem 5

Consider a time series expressed as a sum of a linear trend and $\mathrm{IID}(0, \sigma^2)$ noise $\{X_t\}$. Without removing the trend, for a given lag, the sample autocorrelation based on a sufficiently large sample will be close to one. Make this claim plausible. You may use that, with probability one, $n^{-3} \sum_{t=1}^{n} tX_t \approx 0$ for n sufficiently large. (10 p)

Problem 1

(a) Due to zero means, $E[X_t] = E[Y_t] = 0$ for all t, we look for the best linear predictor of the form $\widehat{X}_{t+2} = a_1 Y_t + a_2 Y_{t-1}$. Notice that

$$\begin{pmatrix} \mathrm{Cov}(Y_t, Y_t) & \mathrm{Cov}(Y_t, Y_{t-1}) \\ \mathrm{Cov}(Y_{t-1}, Y_t) & \mathrm{Cov}(Y_{t-1}, Y_{t-1}) \end{pmatrix} = \begin{pmatrix} \gamma(0) + \sigma_w^2 & \gamma(1) \\ \gamma(1) & \gamma(0) + \sigma_w^2 \end{pmatrix},$$

$\mathrm{Cov}(Y_t, Y_{t+2}) = \gamma(2)$ and $\mathrm{Cov}(Y_{t-1}, Y_{t+2}) = \gamma(3)$, where γ is the ACVF of $\{X_t\}$.

$$\begin{aligned} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} &= \begin{pmatrix} \gamma(0) + \sigma_w^2 & \gamma(1) \\ \gamma(1) & \gamma(0) + \sigma_w^2 \end{pmatrix}^{-1} \begin{pmatrix} \gamma(2) \\ \gamma(3) \end{pmatrix} \\ &= \frac{1}{(\gamma(0) + \sigma_w^2)^2 - \gamma(1)^2} \begin{pmatrix} \gamma(0) + \sigma_w^2 & -\gamma(1) \\ -\gamma(1) & \gamma(0) + \sigma_w^2 \end{pmatrix} \begin{pmatrix} \gamma(2) \\ \gamma(3) \end{pmatrix}. \end{aligned}$$

With $\gamma(h) = c\phi^h$, where $c = \sigma_z^2/(1-\phi^2)$, we get

$$\begin{aligned} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} &= \frac{1}{c^2(1-\phi^2) + \sigma_w^2(2c + \sigma_w^2)} \begin{pmatrix} c^2\phi^2(1-\phi^2) + \sigma_w^2 c\phi^2 \\ \sigma_w^2 c\phi^3 \end{pmatrix} \\ &= \frac{\phi^2\sigma_z^2}{\sigma_z^4 + \sigma_w^2(2\sigma_z^2 + \sigma_w^2(1-\phi^2))} \begin{pmatrix} \sigma_z^2 + \sigma_w^2 \\ \sigma_w^2\phi \end{pmatrix}. \end{aligned}$$

(b) In particular, for $\sigma_w << \sigma_z$ we have

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \approx \begin{pmatrix} \phi^2 \\ 0 \end{pmatrix}$$

(the usual two-step predictor for an AR(1)) and for $\sigma_w >> \sigma_z$ we have

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \approx \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

(the signal drowns in the noise and the best predictor is simply the zero mean).

Problem 2

If the time series data were iid data, then the sample ACF values $\widehat{\gamma}(h)/\widehat{\gamma}(0)$, $h \geq 1$, would be approximately iid and with absolute values exceeding $1.96/\sqrt{1000} \approx 0.062$ with a probability of approximately 0.05. We observe an exceedance for $h = 1$ whereas the other sample ACF values are, in absolute values, much smaller than 0.062. The observed sample ACF values suggest an MA(1) process.
The sample PACF values do become much smaller (in absolute values) from some lag onwards. Therefore the sample PACF values do not suggest an AR(p) process of reasonably small order p.
We may estimate the parameters θ, σ of the suggested MA(1) process $X_t = Z_t + \theta Z_{t-1}$, $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$ by equating $\widehat{\gamma}(h)$, $h = 0, 1$, with there theoretical counterparts and solving the equation system. Here, $\gamma(0) = (1+\theta^2)\sigma^2$ and $\gamma(1) = \theta\sigma^2$, which yields

$$\frac{\theta}{1+\theta^2} = -\frac{0.2}{0.62}.$$

Note that $0.62/0.2 = 3.1$ and that the equation $\theta^2 + 3.1\theta + 1 = 0$ has the solutions

$$\theta = -1.55 \pm \sqrt{1.55^2 - 1},$$

i.e. $\theta = -2.734272$ and $\theta = -0.3657281$. The latter solution is preferable since it corresponds to an invertible MA(1) process. Now, $\widehat{\gamma}(1) = \theta\sigma^2$, with $\theta = -0.3657281$, yields $\sigma = 0.739496$. Summing up, we get the parameter estimates $\widehat{\theta} \approx -0.37$ and $\widehat{\sigma} \approx 0.74$. (The sample was generated from an MA(1) process with $\theta = -0.4$ and $\sigma = 0.7$.)

Problem 3

F&S p. 12: For a stationary zero-mean time series, $P(X_{t+1} \mid X_t, \dots, X_{t-n+1}) = a_1 X_t + \dots + a_n X_{t-n+1}$, where $a = (a_1, \dots, a_n)'$ solves $\Gamma_n a = \gamma_n$, i.e.

$$\begin{pmatrix} \gamma(0) & \dots & \gamma(n-1) \\ \vdots & \ddots & \vdots \\ \gamma(n-1) & \dots & \gamma(0) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \gamma(1) \\ \vdots \\ \gamma(n) \end{pmatrix}.$$

F&S p. 19: For a causal zero-mean AR(p) process, $\Gamma_p \phi = \gamma_p$, where $\phi = (\phi_1, \dots, \phi_p)'$ (the Yule-Walker equations). In particular, for $n \geq p$ the coefficients a_k of the best linear 1-step predictor satisfy $a_k = \phi_k$ for $k \leq p$ and $a_k = 0$ for $k > p$. In particular, for the AR(2) process, $P(X_{t+1} \mid X_t, X_{t-1}, X_{t-2}) = \phi_1 X_t + \phi_2 X_{t-1}$.

Problem 4

Here $\phi(z) = 1 - \phi z$ and $\theta(z) = 1 + \theta z$ have no common zeros and $\phi(z) \neq 0$ for $|z| \leq 1$. Therefore (Def. 6.5 and Thm 6.1 on p. 9 in F&S) $\{X_t\}$ is causal (hence also stationary) with representation

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad \sum_{j=0}^{\infty} |\psi_j| < \infty,$$

where the coefficients ψ_j are given by

$$\begin{aligned} \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)} = \frac{1 + \theta z}{1 - \phi z} &= (1 + \theta z) \sum_{j=0}^{\infty} (\phi z)^j \\ &= 1 + \sum_{j=1}^{\infty} (\phi + \theta)\phi^{j-1} z^j. \end{aligned}$$

From F&S p. 8 we know that the causal ARMA$(1,1)$ process has the ACVF

$$\gamma(h) = \sum_{j=0}^{\infty} \psi_j \psi_{j+h} \sigma^2.$$

In particular,

$$\begin{aligned} \gamma(0) = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2 &= \sigma^2 \Big(1 + (\phi + \theta)^2 \sum_{j=1}^{\infty} \phi^{2(j-1)}\Big) = \sigma^2 \Big(1 + (\phi + \theta)^2 \sum_{j=0}^{\infty} \phi^{2j}\Big) \\ &= \sigma^2 \Big(1 + \frac{(\phi + \theta)^2}{1 - \phi^2}\Big). \end{aligned}$$

Inserting $\phi = 0.5$, $\theta = 1.5$ and $\sigma = 1$ gives

$$\mathrm{Var}(X_t) = 1 + \frac{2^2}{3/4} = 1 + \frac{16}{3} = 19/3 \approx 6.33.$$

Problem 5

The sample ACVF based on a sample of size n from the time series $Y_t = X_t + ct$ is given by

$$\begin{aligned}
\widehat{\gamma}_Y(h) &= \frac{1}{n}\sum_{t=1}^{n-h}(Y_t - \overline{Y}_n)(Y_{t+h} - \overline{Y}_n) \\
&= \frac{1}{n}\sum_{t=1}^{n-h}(X_t + ct - \overline{Y}_n)(X_{t+h} + c(t+h) - \overline{Y}_n) \\
&= \frac{1}{n}\sum_{t=1}^{n-h} X_t X_{t+h} + \frac{1}{n}\sum_{t=1}^{n-h} X_t(c(t+h) - \overline{Y}_n) + \frac{1}{n}\sum_{t=1}^{n-h} X_{t+h}(ct - \overline{Y}_n) \\
&\quad + \frac{1}{n}\sum_{t=1}^{n-h}(ct - \overline{Y}_n)(c(t+h) - \overline{Y}_n).
\end{aligned}$$

Due to $E[X_t] = 0$,

$$\frac{1}{n}\sum_{t=1}^{n-h} X_t X_{t+h} \approx \frac{1}{n}\sum_{t=1}^{n-h}(X_t - \overline{X}_n)(X_{t+h} - \overline{X}_n) = \widehat{\gamma}_X(h),$$

$$\overline{Y}_n = \overline{X}_n + \frac{1}{n}\sum_{t=1}^{n} ct \approx \frac{cn}{2}.$$

Set

$$A_n = \frac{1}{n}\sum_{t=1}^{n-h} X_t(c(t+h) - \overline{Y}_n) + \frac{1}{n}\sum_{t=1}^{n-h} X_{t+h}(ct - \overline{Y}_n)$$

and notice that we may use that $A_n/n^2 \approx 0$ for n sufficiently large. Moreover,

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n-h}(ct - \overline{Y}_n)(c(t+h) - \overline{Y}_n) &\approx \frac{1}{n}\sum_{t=1}^{n-h}(ct - cn/2)(c(t+h) - cn/2) \\
&= \frac{1}{n}\sum_{t=1}^{n}(ct - cn/2)^2 + O(n).
\end{aligned}$$

Finally,

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}(ct - cn/2)^2 &= \frac{1}{n}\Big(c^2\frac{n(n+1)(2n+1)}{6} - 2c(cn/2)\frac{n(n+1)}{2} + n(cn/2)^2\Big) \\
&\approx n^2c^2\Big(\frac{2}{6} - \frac{1}{2} + \frac{1}{4}\Big) = n^2c^2/12.
\end{aligned}$$

In particular, for n sufficiently large,

$$\widehat{\rho}_Y(h) \approx \frac{\widehat{\gamma}_X(h) + n^2(A_n/n^2) + n^2c^2/12}{\widehat{\gamma}_X(0) + n^2(A_n/n^2) + n^2c^2/12} \approx 1.$$