Division of Mathematical Statistics

KTH Matematik

EXAM IN SF2943 TIME SERIES ANALYSIS
FRIDAY JUNE 4 2021, 14:00–19:00

Examiner: Joakim Andén, 08-790 80 52.

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

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

Problem 1

Suppose that the first few values of a theoretical acf are given by

$$\rho(1) = 0.341, \quad \rho(2) = 0.268, \quad \rho(3) = 0.0488, \quad \rho(4) = 0.0341.$$

Does this acf belong to a causal AR(2) model? Motivate your answer. (10 p)

Problem 2

a) You would like to analyze a time series with $n = 250$ samples that you suspect is white noise. To test this, you subtract the sample mean to obtain the values $x_1, x_2, \ldots, x_n$ and calculate

$$\hat{\gamma}(0) = 0.140, \quad \hat{\gamma}(1) = -0.014, \quad \hat{\gamma}(2) = -0.014, \quad \hat{\gamma}(3) = 0.005.$$

Show that the null hypothesis that the data arose from a white noise model cannot be rejected at the confidence level 95%. (3 p)

b) The next step is to determine whether the data result from an ARCH(1) process. With this in mind, you form the sequence

$$y_t = x_t^2 - \hat{\gamma}(0), \quad t = 1, 2, \ldots, n,$$

and calculate

$$\frac{1}{n}\sum_{t=1}^{n} y_t^2 = 0.052 \quad \text{and} \quad \frac{1}{n}\sum_{t=1}^{n-1} y_t y_{t+1} = 0.011.$$

Estimate the parameters α_0 and α_1 of an ARCH(1) model for $x_1, x_2, \ldots, x_n$:

$$\begin{cases} X_t = \sqrt{h_t} e_t, & \{e_t\} \sim \text{IID N}(0, 1), \\ h_t = \alpha_0 + \alpha_1 X_{t-1}^2. \end{cases}$$

You may assume that $\hat{\gamma}(0)$ is an accurate estimate of $\text{Var}(X_t)$ and that $V_t = X_t^2 - h_t$ is a white noise process. Test the hypothesis that $\alpha_1 = 0$ (that is, the model is a trivial ARCH process) at confidence level 95%. (7 p)

Problem 3

a) Let $\mathrm{P}(U|\mathbf{W})$ be the best linear predictor of U given $\mathbf{W} = [W_1, W_2, \dots, W_n]^{\mathrm{T}}$ for a random variable U and a random vector $\mathbf{W}$ with finite variances. Prove linearity of this operator, that is show that

$$\mathrm{P}(\alpha_1 U_1 + \alpha_2 U_2 | \mathbf{W}) = \alpha_1 \mathrm{P}(U_1|\mathbf{W}) + \alpha_2 \mathrm{P}(U_2|\mathbf{W}),$$

for $\alpha_1, \alpha_2 \in \mathbb{R}$, where U_1, U_2 are random variables and $\mathbf{W}$ is a random vector. Assume that U_1, U_2, and $\mathbf{W}$ have zero mean and finite variance and that the covariance matrix of $\mathbf{W}$ is invertible. (2 p)

b) Given a stationary process $\{X_t\}$ such that $\Gamma_n = [\gamma(i-j)]_{i,j=1}^{n}$ is non-singular for all $n > 0$, show that the variance of the linear prediction error given $X_1, X_2, \dots, X_{h-1}$ is the same for both X_0 and X_h. In other words, show that

$$\mathrm{Var}\left(X_0 - \mathrm{P}(X_0|X_1, X_2, \dots, X_{h-1})\right) = \mathrm{Var}\left(X_h - \mathrm{P}(X_h|X_1, X_2, \dots, X_{h-1})\right)$$

for any $h > 0$. *Aid: What is the relation between the linear prediction coefficients for X_0 and X_h given $X_1, X_2, \dots, X_{h-1}$?* (8 p)

Good luck!

Division of Mathematical Statistics

KTH Matematik

SOLUTIONS TO EXAM IN SF2943 TIME SERIES ANALYSIS
FRIDAY JUNE 4, 2021, 14.00–19.00.

Problem 1

We have $\rho(1) = 0.341$ and $\rho(2) = 0.268$. We can then use the Yule–Walker equations for $p = 2$ to fit ϕ_1 and ϕ_2. We thus have

$$\begin{bmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \rho(2) \end{bmatrix},$$

since $\rho(0) = 1$. Inverting the matrix, we get

$$\begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \frac{1}{1-\rho(1)^2} \begin{bmatrix} 1 & -\rho(1) \\ -\rho(1) & 1 \end{bmatrix} \begin{bmatrix} \rho(1) \\ \rho(2) \end{bmatrix}.$$

Plugging in the values for $\rho(1)$ and $\rho(2)$ gives $\phi_1 \approx 0.2825$ and $\phi_2 \approx 0.1717$.
If the acf indeed corresponds to an AR(2) process, it should satisfy the Yule–Walker equations for $p = 3$ but with $\phi_3 = 0$. We thus have

$$\begin{bmatrix} 1 & \rho(1) & \rho(2) \\ \rho(1) & 1 & \rho(1) \\ \rho(2) & \rho(1) & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ 0 \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \rho(3) \end{bmatrix}.$$

In particular, the third row gives the equation

$$\phi_1\rho(1) + \phi_2\rho(2) = \rho(3).$$

The left-hand side equals 0.142, which is not equal to $\rho(3) = 0.0488$. As a result, the acf could not have been generated by an AR(2) model.

Problem 2

a) The number of samples $n = 250$ is high enough that we can assume that the sample acf is approximately normally distributed. Under the hypothesis of white noise, $\rho(h) = 0$ for $|h| > 0$, which means that the distribution of $\hat{\rho}(h)$ is approximately $\mathrm{N}(0, 1/\sqrt{n})$. We would therefore reject this null hypothesis if $|\rho(h)| > \lambda_{0.025}/\sqrt{n}$, where λ_α is the α-quantile of the normal distribution. Since $\lambda_{0.025} \approx 1.96$, we have the critical threshold $\lambda_{0.025}/\sqrt{n} \approx 0.124$.
From the given sample acvf, we get

$$\hat{\rho}(1) = -\frac{0.014}{0.140} = -0.10, \quad \hat{\rho}(2) = -\frac{0.014}{0.140} = -0.10, \quad \hat{\rho}(3) = \frac{0.005}{0.140} \approx 0.036.$$

Since all are within the interval $[-0.124, 0.124]$, we cannot reject the null hypothesis that the data is generated from white noise.

b) We use the fact that $V_t = X_t^2 - h_t$ is a white noise process. Since $h_t = X_t^2 - V_t$, we get

$$\begin{aligned} X_t^2 - V_t &= \alpha_0 + \alpha_1 X_{t-1}^2 \\ X_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 + V_t \\ X_t^2 - \frac{\alpha_0}{1-\alpha_1} &= \alpha_1 \left(X_{t-1}^2 - \frac{\alpha_0}{1-\alpha_1} \right) + V_t. \end{aligned}$$

Consequently, $\{X_t^2 - \alpha_0/(1-\alpha_1)\}$ is an AR(1) process. (Equivalently, $\{X_t^2\}$ is an AR(1) process with mean $\alpha_0/(1-\alpha_1)$.)
To fit the parameters, we first use the fact that $\mathbb{E}[X_t^2] = \mathrm{Var}[X_t]$. According to the problem formulation, we may assume that $\hat{\gamma}(0)$ is an accurate estimate of $\mathrm{Var}[X_t]$. Given the above derivation, we know that

$$\mathrm{Var}[X_t] = \frac{\alpha_0}{1-\alpha_1},$$

so we have the constraint $\alpha_0 = \hat{\gamma}(0)(1-\alpha_1)$.
We introduce the notation $Y_t = X_t^2 - \alpha_0/(1-\alpha_1)$. To estimate α_1, we use the fact that $\{Y_t\}$ is an AR(1) process. From the data, we have that its sample acvf $\hat{\gamma}_Y$ satisfies $\hat{\gamma}_Y(0) = 0.052$ and $\hat{\gamma}_Y(1) = 0.011$. Its sample acf at lag one is therefore $\hat{\rho}_Y(1) = 0.011/0.052 \approx 0.2115$. Since it is an AR(1) process, its acf ρ_Y satisfies $\rho_Y(1) = \alpha_1$, so an estimate of α_1 is $\hat{\rho}_Y(1) \approx 0.2115$. Plugging this into the formula for α_0, we obtain the estimate 0.110.
Under the null hypothesis that $\alpha_0 = 0$, we have the same situtation as before, and $\hat{\rho}_Y(1)$ would be approximately distributed as $\mathrm{N}(0, 1/\sqrt{n})$. We obtain the same threshold as before, $\lambda_{0.025}/\sqrt{n} \approx 0.124$. Since our $\hat{\rho}_Y(1) \approx 0.2115$ is above this threshold, we reject the null hypothesis.

Problem 3

a) Since all random variables are of mean zero, we have

$$\mathrm{P}(\alpha_1 U_1 + \alpha_2 U_2 | \mathbf{W}) = \boldsymbol{\phi}^{\mathrm{T}} \mathbf{W},$$

where $\boldsymbol{\phi} = \Gamma^{-1}\boldsymbol{\gamma}$ for $\Gamma = \mathrm{Cov}(\mathbf{W}, \mathbf{W})$ and $\boldsymbol{\gamma} = \mathrm{Cov}(\alpha_1 U_1 + \alpha_2 U_2, \mathbf{W})$. In addition, we have

$$\begin{aligned} \mathrm{P}(U_1 | \mathbf{W}) &= \boldsymbol{\phi}_1^{\mathrm{T}} \mathbf{W}, \\ \mathrm{P}(U_2 | \mathbf{W}) &= \boldsymbol{\phi}_2^{\mathrm{T}} \mathbf{W}, \end{aligned}$$

where $\boldsymbol{\phi}_1 = \Gamma^{-1}\boldsymbol{\gamma}_1$ and $\boldsymbol{\phi}_2 = \Gamma^{-1}\boldsymbol{\gamma}_2$ with $\boldsymbol{\gamma}_1 = \mathrm{Cov}(U_1, \mathbf{W})$ and $\boldsymbol{\gamma}_2 = \mathrm{Cov}(U_2, \mathbf{W})$. We note that

$$\begin{aligned} \alpha_1 \boldsymbol{\phi}_1 + \alpha_2 \boldsymbol{\phi}_2 &= \alpha_1 \Gamma^{-1} \boldsymbol{\gamma}_1 + \alpha_2 \Gamma^{-1} \boldsymbol{\gamma}_2 \\ &= \Gamma^{-1}(\alpha_1 \mathrm{Cov}(U_1, \mathbf{W}) + \alpha_2 \mathrm{Cov}(U_2, \mathbf{W})) \\ &= \Gamma^{-1} \mathrm{Cov}(\alpha_1 U_1 + \alpha_2 U_2, \mathbf{W}) = \Gamma^{-1} \boldsymbol{\gamma}. \end{aligned}$$

So $\boldsymbol{\phi} = \alpha_1 \boldsymbol{\phi}_1 + \alpha_2 \boldsymbol{\phi}_2$. Consequently,

$$\begin{aligned} \mathrm{P}(\alpha_1 U_1 + \alpha_2 U_2 | \mathbf{W}) &= \boldsymbol{\phi}^{\mathrm{T}} \mathbf{W} \\ &= (\alpha_1 \boldsymbol{\phi}_1 + \alpha_2 \boldsymbol{\phi}_2)^{\mathrm{T}} \mathbf{W} \\ &= \alpha_1 \boldsymbol{\phi}_1^{\mathrm{T}} \mathbf{W} + \alpha_2 \boldsymbol{\phi}_2^{\mathrm{T}} \mathbf{W} \\ &= \alpha_1 \mathrm{P}(U_1 | \mathbf{W}) + \alpha_2 \mathrm{P}(U_2 | \mathbf{W}). \end{aligned}$$

b) Let $\mathbf{X} = [X_1, X_2, \ldots, X_{h-1}]^{\mathrm{T}}$, $\Gamma = [\gamma(i-j)]_{i,j=1}^{h-1}$, and $\boldsymbol{\gamma} = [\gamma(1), \gamma(2), \ldots, \gamma(h-1)]^{\mathrm{T}}$. Then

$$\mathrm{P}(X_0 | X_1, X_2, \ldots, X_{h-1}) = \mathrm{P}(X_0|\mathbf{X}) = \mathbf{a}^{\mathrm{T}}\mathbf{X},$$

where $\mathbf{a} = \Gamma^{-1}\boldsymbol{\gamma}$. Similarly, we have

$$\mathrm{P}(X_h | X_1, X_2, \ldots, X_{h-1}) = \mathrm{P}(X_h|\mathbf{X}) = \tilde{\mathbf{a}}^{\mathrm{T}}\mathbf{X},$$

where $\tilde{\mathbf{a}} = \Gamma^{-1}\tilde{\boldsymbol{\gamma}}$ for $\tilde{\boldsymbol{\gamma}} = [\gamma(h-1), \gamma(h-2), \ldots, \gamma(1)]^{\mathrm{T}}$.
To simplify notations, we introduce the reversion matrix

$$J = \begin{bmatrix} 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & \ldots & 1 & 0 \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ 0 & 1 & \vdots & 0 & 0 \\ 1 & 0 & \vdots & 0 & 0 \end{bmatrix}.$$

This matrix has several useful properties, such as $J^2 = J$ and $J^{\mathrm{T}} = J$.
We thus have $\tilde{\boldsymbol{\gamma}} = J\boldsymbol{\gamma}$. Furthermore, we can verify that $J\Gamma J = \Gamma$ (left-hand multiplication reverses the rows and right-hand multiplication reverses the columns, so the result is the same). Using this, we have that

$$\tilde{\mathbf{a}} = \Gamma^{-1}\tilde{\boldsymbol{\gamma}} = \Gamma^{-1}J\boldsymbol{\gamma} = J(J\Gamma J)^{-1}\boldsymbol{\gamma} = J\Gamma^{-1}\boldsymbol{\gamma} = J\mathbf{a}.$$

In other words, $\tilde{\mathbf{a}}$ is the reverse of $\mathbf{a}$.
For the variance, we now have

$$\mathrm{Var}(X_0 - \mathrm{P}(X_0|\mathbf{X})) = \mathrm{Var}(X_0 - \mathbf{a}^{\mathrm{T}}\mathbf{X}) = \mathrm{Var}(X_0) + \mathrm{Var}(\mathbf{a}^{\mathrm{T}}\mathbf{X}) - 2\mathrm{Cov}(X_0, \mathbf{a}^{\mathrm{T}}\mathbf{X}).$$

Let us go through these terms one by one. First, since $\{X_t\}$ is stationary, we have $\mathrm{Var}(X_0) = \mathrm{Var}(X_h)$. Second, we have

$$\begin{aligned}\mathrm{Var}(\mathbf{a}^{\mathrm{T}}\mathbf{X}) &= \mathrm{Var}((J\tilde{\mathbf{a}})^{\mathrm{T}}\mathbf{X}) = \mathrm{Var}(\tilde{\mathbf{a}}^{\mathrm{T}}J\mathbf{X}) = \tilde{\mathbf{a}}^{\mathrm{T}}J\mathrm{Cov}(\mathbf{X},\mathbf{X})J\tilde{\mathbf{a}} = \tilde{\mathbf{a}}^{\mathrm{T}}J\Gamma J\tilde{\mathbf{a}} \\ &= \tilde{\mathbf{a}}^{\mathrm{T}}\Gamma\tilde{\mathbf{a}} = \tilde{\mathbf{a}}^{\mathrm{T}}\mathrm{Cov}(\mathbf{X},\mathbf{X})\tilde{\mathbf{a}} = \mathrm{Var}(\tilde{\mathbf{a}}^{\mathrm{T}}\mathbf{X}).\end{aligned}$$

Finally, we obtain

$$\begin{aligned}\mathrm{Cov}(X_0, \mathbf{a}^{\mathrm{T}}\mathbf{X}) &= \mathrm{Cov}(X_0, (J\tilde{\mathbf{a}})^{\mathrm{T}}\mathbf{X}) = \tilde{\mathbf{a}}^{\mathrm{T}}J\mathrm{Cov}(X_0,\mathbf{X}) = \tilde{\mathbf{a}}^{\mathrm{T}}J\boldsymbol{\gamma} = \tilde{\mathbf{a}}^{\mathrm{T}}\tilde{\boldsymbol{\gamma}} \\ &= \tilde{\mathbf{a}}^{\mathrm{T}}\mathrm{Cov}(X_h,\mathbf{X}) = \mathrm{Cov}(X_h, \tilde{\mathbf{a}}^{\mathrm{T}}\mathbf{X}).\end{aligned}$$

Putting all of this together, we get

$$\begin{aligned}\mathrm{Var}(X_0 - \mathrm{P}(X_0|\mathbf{X})) &= \mathrm{Var}(X_h) + \mathrm{Var}(\tilde{\mathbf{a}}^{\mathrm{T}}\mathbf{X}) - 2\mathrm{Cov}(X_h, \tilde{\mathbf{a}}^{\mathrm{T}}\mathbf{X}) \\ &= \mathrm{Var}(X_h - \mathrm{P}(X_h|\mathbf{X}).\end{aligned}$$