Avd. Matematisk statistik

KTH Matematik

TENTAMEN I SF2945 TIME SERIES ANALYSIS/TIDSSERIEANALYS FREDAGEN DEN 4 JUNI 2010 KL 08.00–13.00.

Examinator: Timo Koski, tel. 7907134

Tillåtna hjälpmedel: Formulas and survey, Time series analysis. Handheld calculator.

Införda beteckningar skall förklaras och definieras. Resonemang och uträkningar skall vara så utförliga och väl motiverade att de är lätta att följa.

Varje korrekt lösning ger 10 poäng. Gränsen för godkänt är 25 poäng. De som erhåller 23 eller 24 poäng på tentamen kommer att erbjudas möjlighet att komplettera till betyget E. Den som är aktuell för komplettering skall till examinator anmäla önskan att få en sådan inom en vecka från publicering av tentamensresultatet.

Lösningarna får givetvis skrivas på svenska.

Resultatet skall vara klart senast torsdag den 25 juni 2010 och blir tillgängligt via ”Mina sidor”.

Lösningarna får givetvis skrivas på svenska.

\- -

Quantiles of the normal distribution
(Normalfördelningens kvantiler)

$P(X > \lambda_\alpha) = \alpha$ where $X \sim N(0,1)$

α	λ_α	α	λ_α
0.10	1.2816	0.001	3.0902
0.05	1.6449	0.0005	3.2905
0.025	1.9600	0.0001	3.7190
0.010	2.3263	0.00005	3.8906
0.005	2.5758	0.00001	4.2649

Problem 1

The process $\{X_t \mid t = 0, \pm 1, \pm 2, \ldots\}$ satisfies the equation

$$X_t - \frac{9}{20} X_{t-1} + \frac{1}{20} X_{t-2} = Z_t,$$

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

(a) Show that $\{X_t \mid t = 0, \pm 1, \pm 2, \ldots\}$ is a stationary and causal AR(2) -process. (3 p)

(b) In view of (a) we can write the autocovariance function (ACVF) of the process $\{X_t \mid t = 0, \pm 1, \pm 2, \ldots\}$ as

$$\gamma_X(h) = \mathrm{Cov}\left(X_{t+h}, X_t\right), \quad |h| = 0, \pm 1, \pm 2, \ldots,$$

Show by a detailed argument that $\gamma_X(h)$ satisfies for $h \geq 2$ the following homogeneous linear difference equation with constant coefficients

$$\gamma_X(h) - \frac{9}{20} \gamma_X(h-1) + \frac{1}{20} \gamma_X(h-2) = 0. \tag{1}$$

(3 p)

(c) Use (1) to show that

$$\lim_{h \to +\infty} \gamma_X(h) = 0.$$

Note that you do not need to determine explictly the initial conditions $\gamma_X(1)$ and $\gamma_X(0)$ for this. (4 p)

Problem 2

The MA(2) - process $\{X_t \mid t = 0, \pm 1, \pm 2, \ldots\}$ satisfies

$$X_t = Z_t + \frac{1}{2} Z_{t-1} + \frac{1}{4} Z_{t-2},$$

where $Z_t \sim \mathrm{WN}(0, 1)$ is a Gaussian white noise.

(a) Find the probability distribution of the column vector $\left(X_{t+2}, X_{t+1}, X_t\right)^T$. Explain carefully your reasoning and show your calculations. (4 p)

(b) Find the probability distribution of the column vector $\left(X_{t+12}, X_{t+11}, X_{t+10}\right)^T$. Show your reasoning. (1 p)

(c) Find a constant a such that

$$P\left(X_{t+2} > a \mid X_{t+1} = 1.0\right) = 0.90.$$

(5 p)

Problem 3

$(X_t)_{t=-\infty}^{\infty}$ is an $MA(1)$-process

$$X_t = Z_t + Z_{t-1},$$

where $(Z_t)_{n=-\infty}^{\infty}$ is an I.I.D. process with the distribution given by the probability density

$$f_Z(z) = \begin{cases} 0 & z < -1 \\ e^{-(z+1)} & z \geq -1. \end{cases}$$

Then it holds that probability density of X_t is

$$f_{X_t}(x) = \begin{cases} 0 & x < -2 \\ (x+2)e^{-(x+2)} & x \geq -2. \end{cases}$$

You need not show this.

(a) Check that the conditional density of Z_t given that $X_t = c$ is

$$f_{Z_t|X_t=c}(z) = \begin{cases} \frac{1}{c+2} & -1 \leq z < c+1 \\ 0 & \text{elsewhere.} \end{cases}$$

(4 p)

(b) Show that

$$E\left[Z_t \mid X_t = c\right] = \frac{c}{2}.$$

(2 p)

(c) Show that

$$E\left[X_{t+1} \mid X_t = c\right] = \frac{c}{2}.$$

(4 p)

Problem 4

Let $\{Y_t\}_{t=-\infty}^{\infty}$ be a linear process with

$$Y_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j}, \quad t \in \{0, \pm 1, \pm 2, \ldots, \},$$

where $\{X_t\}_{t=-\infty}^{\infty}$ is a stationary time series with mean zero and spectral density $f_X(\lambda)$, and $\sum_{j=-\infty}^{\infty} \mid \psi_j \mid < \infty$. We know that the autocovariance function (ACVF) of $\{Y_t\}_{t=-\infty}^{\infty}$ is for $h \geq 0$

$$\gamma_Y(h) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_X(h + k - j),$$

where $\gamma_X(\cdot)$ is the ACVF of $\{X_t\}_{t=-\infty}^{\infty}$. Show that the spectral density $f_Y(\lambda)$ of $\{Y_t\}_{t=-\infty}^{\infty}$ is

$$f_Y(\lambda) = \mid \Psi\left(e^{-i\lambda}\right) \mid^2 f_X(\lambda),$$

where $\Psi\left(e^{-i\lambda}\right) = \sum_{j=-\infty}^{\infty} \psi_j e^{-ij\lambda}$. (10 p)

Problem 5

We have an AR(1) process

$$X_n = \phi X_{n-1} + Z_n, n = 0, 1, 2, \ldots, \tag{2}$$

where $\mid \phi \mid < 1$, $\{Z_n\} \sim \mathrm{WN}(0, \sigma^2)$. We assume that $\mathrm{Var}(X_0) = \sigma_0^2$. The true state X_n is not observed directly, but we observe X_n with added white measurement noise $\{V_n\} \sim \mathrm{WN}(0, \sigma_V^2)$, as Y_n in

$$Y_n = cX_n + V_n, n = 0, 1, 2, \ldots, \tag{3}$$

The white noises $\{Z_n\}$ and $\{V_n\}$ are independent processes.

(a) Find the predictor $\widehat{X}_{n+1} = E\left[X_{n+1} \mid Y_0, Y_1, \ldots, Y_n\right]$ by the Kalman recursions. *Hint: You may apply in a straightforward manner the collection of formulas.* (5 p)

(b) What is the Riccati equation for this predictor and what quantity does it give ? What is the computational advantage of the Riccati equation in recursive prediction ? (4 p)

(c) Give the innovations form of the predictor $\widehat{X}_{n+1} = E\left[X_{n+1} \mid Y_0, Y_1, \ldots, Y_n\right]$. (1 p)

Avd. Matematisk statistik

KTH Matematik

LÖSNINGAR TILL TENTAMEN
SF2945 TIME SERIES ANALYSIS/TIDSSERIEANALYS ONSDAGEN DEN 16 DECEMBER 2009 KL 08.00–13.00.

Problem 1

(a) We note that the linear stochastic difference equation with constant coefficients

$$X_t - \frac{9}{20}X_{t-1} + \frac{1}{20}X_{t-2} = Z_t,$$

can be written as

$$\phi(B)X_t = Z_t,$$

where

$$\phi(B) = 1 - \frac{9}{20}B + \frac{1}{20}B^2.$$

In order to answer the question about ARMA(2) we consider zeros of the polynomial $\phi(z)$ for $z \in \mathbf{C}$,

$$\phi(z) = 1 - \frac{9}{20}z + \frac{1}{20}z^2.$$

We get that

$$\phi(\xi_i) = 0 \Leftrightarrow \xi_1 = 4, \xi_2 = 5 \Leftrightarrow \phi(z) = \left(1 - \frac{z}{4}\right)\left(1 - \frac{z}{5}\right).$$

Thus we see that all roots of $\phi(z) = 0$ are satisfy $\mid z \mid \neq 1$. Hence $\{X_t \mid t = 0, \pm 1, \pm 2, \ldots\}$ is stationary. The roots of $\phi(z) = 0$ satisfy $\mid z \mid > 1$. By definition this means that $\{X_t \mid t = 0, \pm 1, \pm 2, \ldots\}$ is a causal process.

(b) Since $\{X_t \mid t = 0, \pm 1, \pm 2, \ldots\}$ is stationary

$$E[X_t] = \mu.$$

Hence

$$E[X_t] - \frac{9}{20}E[X_{t-1}] + \frac{1}{20}E[X_{t-2}] = E[Z_t]$$
$$\Leftrightarrow$$
$$\mu\left[1 - \frac{9}{20} + \frac{1}{20}\right] = 0,$$

which means that

$$\mu = 0.$$

Hence

$$\gamma_X(h) = \mathrm{Cov}(X_{t+h}, X_t) = E[X_{t+h}X_t].$$

We multiply by X_{t-h} and take the expectations:

$$E\left[X_t X_{t-h}\right] - \frac{9}{20} E\left[X_{t-1} X_{t-h}\right] + \frac{1}{20} E\left[X_{t-2} X_{t-h}\right] = E\left[Z_t X_{t-h}\right]$$

Since the process is causal, we have that X_{t-h} is for $h \geq 2$ noncorrelated with Z_t and

$$E\left[X_{t-h} Z_t\right] = E\left[X_{t-h}\right] \cdot E\left[Z_t\right] = 0 \cdot 0.$$

Thus we have from the above, e.g., that

$$E\left[X_{t-1} X_{t-h}\right] = \mathrm{Cov}\left(X_{t-1}, X_{t-h}\right) = \gamma_X(h-1)$$

therefore we get (1), i.e.,

$$\gamma_X(h) - \frac{9}{20}\gamma_X(h-1) + \frac{1}{20}\gamma_X(h-2) = 0.$$

(c) We shall find the general solution to (1). We make the standard Ansatz

$$z_h = z^h.$$

We inser this in (1) and get

$$z^h - \frac{9}{20} z^{h-1} + \frac{1}{20} z^{h-2} = 0 \Leftrightarrow 1 - \frac{9}{20} z^{-1} + \frac{1}{20} z^{-2} = 0 \Leftrightarrow z^2 - \frac{9}{20} z + \frac{1}{20} = 0 \Leftrightarrow$$

$$z_1 = \frac{1}{4}, \quad z_2 = \frac{1}{5}.$$

Then

$$\gamma_X(h) = c_1 \left(\frac{1}{4}\right)^h + c_2 \left(\frac{1}{5}\right)^h$$

where c_1 and c_2 are two constants to be determined by the initial conditions. But since $|\frac{1}{4}| < 1$ and $|\frac{1}{5}| < 1$, we get directly that

$$\lim_{h \to +\infty} \gamma_X(h) = 0.$$

Problem 2

(a) Since MA(2) - process $\{X_t \mid t = 0, \pm 1, \pm 2, \ldots\}$ satisfies

$$X_t = Z_t + \frac{1}{2} Z_{t-1} + \frac{1}{4} Z_{t-2},$$

it is stationary. As $Z_t \sim \mathrm{WN}(0,1)$ is a Gaussian white noise, the process $\{X_t \mid t = 0, \pm 1, \pm 2, \ldots\}$ is Gaussian. Therefore $\{X_t \mid t = 0, \pm 1, \pm 2, \ldots\}$ n, the probability distribution of the column vector $\left(X_{t+2}, X_{t+1}, X_t\right)^T$ is multivariate normal.

We have

$$E\left[X_t\right] = E\left[Z_t\right] + \frac{1}{2} E\left[Z_{t-1}\right] + \frac{1}{4} E\left[Z_{t-2}\right] = 0,$$

and therefore the mean vector of $(X_{t+2}, X_{t+1}, X_t)^T$ is the 3×1 zero vector. Thus we have

$$\begin{pmatrix} X_{t+2} \\ X_{t+1} \\ X_t \end{pmatrix} \sim N_3 \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(1) & \gamma(0) & \gamma(1) \\ \gamma(2) & \gamma(1) & \gamma(0) \end{pmatrix} \right)$$

where the entries of the covariance matrix are obtained from the ACVF. A straightforward computation using

$$E\left[Z_{t+h} + \frac{1}{2}Z_{t+h-1} + \frac{1}{4}Z_{t+h-2}\right]\left[Z_t + \frac{1}{2}Z_{t-1} + \frac{1}{4}Z_{t-2}\right]$$

for $h = 0, 1, 2$ gives

$$\gamma(0) = 1 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{4}\right)^2 = \frac{21}{16}$$

$$\gamma(1) = \frac{1}{2} + \frac{1}{2}\frac{1}{4} = \frac{5}{8}$$

$$\gamma(2) = \frac{1}{4}.$$

C.f., Brockwell & Davis Example 3.2.2 on p. 89.

ANSWER (a) : $\underline{\begin{pmatrix} X_{t+2} \\ X_{t+1} \\ X_t \end{pmatrix} \sim N_3 \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{21}{16} & \frac{5}{8} & \frac{1}{4} \\ \frac{5}{8} & \frac{21}{16} & \frac{5}{8} \\ \frac{1}{4} & \frac{5}{8} & \frac{21}{16} \end{pmatrix} \right)}$

(b) Because the stationary process $\{X_t \mid t = 0, \pm1, \pm2, \ldots\}$ is Gaussian, it is also strictly stationary. This means amongst other things that the distribution of $(X_{t+2}, X_{t+1}, X_{t+10})^T$ is not changed by a common shift of the time indices. As $(X_{t+12}, X_{t+11}, X_{t+10})^T$ contains these random variables shifted by ten time units, we get the answer.

ANSWER (b): $\underline{\begin{pmatrix} X_{t+12} \\ X_{t+11} \\ X_{t+10} \end{pmatrix} \sim \begin{pmatrix} X_{t+2} \\ X_{t+1} \\ X_t \end{pmatrix}}.$

(c) By strict stationarity we need to find a constant a such that

$$P\left(X_{t+1} > a \mid X_t = 1.0\right) = 0.95,$$

we need to find probability distribution of $X_{t+1} \mid X_t = 1.0$. As the process $\{X_t \mid t = 0, \pm1, \pm2, \ldots\}$ is Gaussian, the joint distribution of $(X_{t+1}, X_t)^T$ is a bivariate normal distribution. We have by **Collection of Formulas**,

$$X_{t+1} \mid X_t = 1.0 \sim N\left(\rho\frac{\sigma_{X_{t+1}}}{\sigma_{X_t}}, \sigma^2_{X_{t+1}}\left(1 - \rho^2\right)\right)$$

where, as means are zero, and from part (a) $\sigma_{X_{t+1}} = \sigma_{X_t} = \sqrt{\gamma(0)} = \sqrt{\frac{21}{16}}$ and the coefficient of correlation is by part (a)

$$\rho = \frac{\gamma(1)}{\gamma(0)} = \frac{\frac{5}{8}}{\frac{21}{16}} = \frac{10}{21}.$$

Thereby

$$X_{t+1} \mid X_t = 1 =$$

$$\sim N\left(\frac{10}{21}, \frac{21}{16}\left(1 - \left(\frac{10}{21}\right)^2\right)\right) = N\left(\frac{10}{21}, 1.0149\right).$$

Then

$$P\left(X_{t+1} > a \mid X_t = 1.0\right) = P\left(\frac{X_{t+1} - \frac{10}{21}}{\sqrt{1.0149}} > \frac{a - \frac{10}{21}}{\sqrt{1.0149}} \mid X_t = 1.0\right)$$

and since now $\frac{X_{t+1} - \frac{10}{21}}{\sqrt{1.02}} \mid X_t = 1.0 \sim N(0,1)$, we obtain

$$= P\left(\xi > \frac{a - \frac{10}{21}}{\sqrt{1.02}}\right),$$

where $\xi \sim N(0,1)$. Hence we want to find a such that

$$P\left(\xi > \frac{a - \frac{10}{21}}{\sqrt{1.02}}\right) = 0.90.$$

By the table on $N(0,1)$ we have that $\frac{a - \frac{10}{21}}{\sqrt{1.02}} = \lambda_{0.90}$, where by symmetry $\lambda_{0.90} = -\lambda_{0.10}$ and

$$P\left(\xi > -\lambda_{0.10}\right) = 0.90,$$

and $\lambda_{0.10} = 1.2816$ is given in the Quantiles of the normal distribution recapitulated in the ingress.

In other words,

$$a = -\lambda_{0.10}\sqrt{1.02} + \frac{10}{21} = -1.2816\sqrt{1.02} + \frac{10}{21} = -0.8182$$

ANSWER: $\underline{a = -0.8182.}$

Problem 3

(a) By definition of conditional density we get

$$f_{Z_t|X_t=c}(z) = \frac{f_{X_t|Z_t=z}(c) f_{Z_t}(z)}{f_{X_t}(c)}$$

But if $X_t = c$ and $Z_t = z$, then $Z_{t-1} = c - z$ and therefore

$$f_{X_t|Z_t=z}(c) = f_{Z_{t-1}}(c - z)$$

We have thus with $Z_{t-1} = x$ that $c - z = x \geq -1$, i.e., $c + 1 \geq z$. For these c and z

$$\frac{f_{X_t|Z_t=z}(c) f_{Z_t}(z)}{f_{X_t}(c)} = \frac{f_{Z_{t-1}}(c-z) f_{Z_t}(z)}{f_{X_t}(c)}$$

$$= \frac{e^{-(c-z+1)} e^{-(z+1)}}{(c+2)e^{-(c+2)}}$$

$$= \frac{e^{-(c+2)}}{(c+2)e^{-(c+2)}} = \frac{1}{c+2}.$$

The sought density $= 0$ as soon as $z < -1$.

(b) Show that

$$E\left[Z_t \mid X_t = c\right] = \int_{-1}^{c+1} z f_{Z_t|X_t=c}(z)dz =$$

$$= \frac{1}{c+2}\int_{-1}^{c+1} zdz = \frac{1}{2(c+2)}\left[z^2\right]_{-1}^{c+1} = \frac{1}{2(c+2)}\left((c+1)^2 - (-1)^2\right)$$

$$= \frac{1}{2(c+2)}\left(c^2 + 2c + 1 - 1\right) = \frac{c}{2(c+2)}(c+2) = \frac{c}{2}.$$

(c)

$$E\left[X_{t+1} \mid X_t = c\right] = E\left[Z_{t+1} + Z_t \mid X_t = c\right] = \frac{c}{2}.$$

$$= E\left[Z_{t+1} \mid X_t = c\right] + E\left[Z_t \mid X_t = c\right]$$

$$= E\left[Z_{t+1}\right] + \frac{c}{2},$$

since Z_{t+1} is independent of X_t and by part (b). Here

$$E\left[Z_{t+1}\right] = \int_{-1}^{\infty} ze^{-(z+1)}dz = e^{-1}\int_{-1}^{\infty} ze^{-z}dz$$

$$= e^{-1}\left\{\left[-ze^{-z}\right]_{-1}^{\infty} + \int_{-1}^{\infty} e^{-z}dz\right\}$$

$$= e^{-1}\left\{-e^1 + \left[-e^{-z}\right]_{-1}^{\infty}\right\}$$

$$= e^{-1}\left\{-e^1 + e^1\right\} = 0.$$

Hence

$$E\left[Z_t \mid X_t = c\right] = \frac{c}{2}.$$

Problem 4

The result was earlier at KTH known under the name *superfomula*. The required proof is found on pp. 129−130 in Brockwell & Davis.

Problem 5

(a) This assignment asks one to simply identify the timevariant matrices in the Formulas and survey in the scalar and timeinvariant case as follows:

$$F_t \leftrightarrow \phi$$

$$Q_t \leftrightarrow \sigma^2$$

$$G_t \leftrightarrow c$$

$$R_t \leftrightarrow \sigma_V^2.$$

$$\Omega_t \leftrightarrow E\left[e_{n+1}^2\right].$$

A matrix transpose like F_t^T boils down to the scalar itself, and

$$F_t \Omega_t F_t^T \leftrightarrow \phi^2 E\left[e_{n+1}^2\right]$$

$$F_t \Omega_t G_t^T \leftrightarrow \phi c E\left[e_n^2\right].$$

$$\triangle_t = G_t \Omega_t G_t^T + R_t \leftrightarrow \triangle_n = c^2 E\left[e_n^2\right] + \sigma_V^2$$

Then the desired predictor $\widehat{X}_{n+1} = E\left[X_{n+1} \mid Y_0, Y_1, \ldots, Y_n\right]$ is computed recursively by the Kalman recursions obtained by the identifications above from the Formulas and survey,

$$\widehat{X}_{n+1} = \phi \widehat{X}_n + \frac{\theta_n}{\triangle_n} \varepsilon_n, \tag{4}$$

where

$$\varepsilon_n = Y_n - c\widehat{X}_n, \tag{5}$$

and

$$e_{n+1} = X_{n+1} - \widehat{X}_{n+1},$$

and

$$E\left[e_{n+1}^2\right] = \phi^2 E\left[e_n^2\right] + \sigma^2 - \frac{\theta_n^2}{\triangle_n}, \tag{6}$$

where

$$E\left[e_0^2\right] = \sigma_0^2, \tag{7}$$

and

$$\theta_n = \phi c E\left[e_n^2\right], \tag{8}$$

and

$$\triangle_n = \sigma_V^2 + c^2 E\left[e_n^2\right]. \tag{9}$$

(b) The Riccati equation is (6), and it gives a nonlinear difference equation for the variance $E\left[e_{n+1}^2\right]$ of the prediction error

$$e_{n+1} = X_{n+1} - \widehat{X}_{n+1}.$$

A computational advantage is that the Riccati equation does not depend on the observed time series, and can therefore be precomputed (when values of parameters are known) prior to any prediction.

(c) The innovation variables are

$$\varepsilon_n = \left[Y_n - c\widehat{X}_n\right].$$

Here ε_n is the part of Y_n which is not exhausted by $c\widehat{X}_n$. Then the prediction has an innovations representation

$$\widehat{X}_{n+1} = \phi \widehat{X}_n + K(n)\varepsilon_n$$

$$Y_n = c\widehat{X}_n + \varepsilon_n,$$

where

$$K(n) = \frac{\theta_n}{\triangle_n}$$

is the Kalman gain.