NTNU

Department of Mathematical Sciences

Examination paper for **TMA4285 Time series models**

Academic contact during examination: Håkon Tjelmeland
Phone: 4822 1896

Examination date: December 7th 2013
Examination time (from–to): 09:00–13:00
Permitted examination support material: C:

- Calculator CITIZEN SR-270X, CITIZEN SR-270X College or HP30S.
- Statistiske tabeller og formler, Tapir forlag.
- K. Rottman: Matematisk formelsamling.
- One yellow, stamped A5 sheet with own handwritten formulas and notes.

Other information:
Note that all answers should be justified.
In your solution you can use English and/or Norwegian.

Language: English
Number of pages: 5
Number pages enclosed: 0

Checked by:

Date Signature

Problem 1

Figure 1 shows two observed time series, ts1 and ts2, with corresponding estimated autocorrelation function (acf) and partial autocorrelation function (pacf).

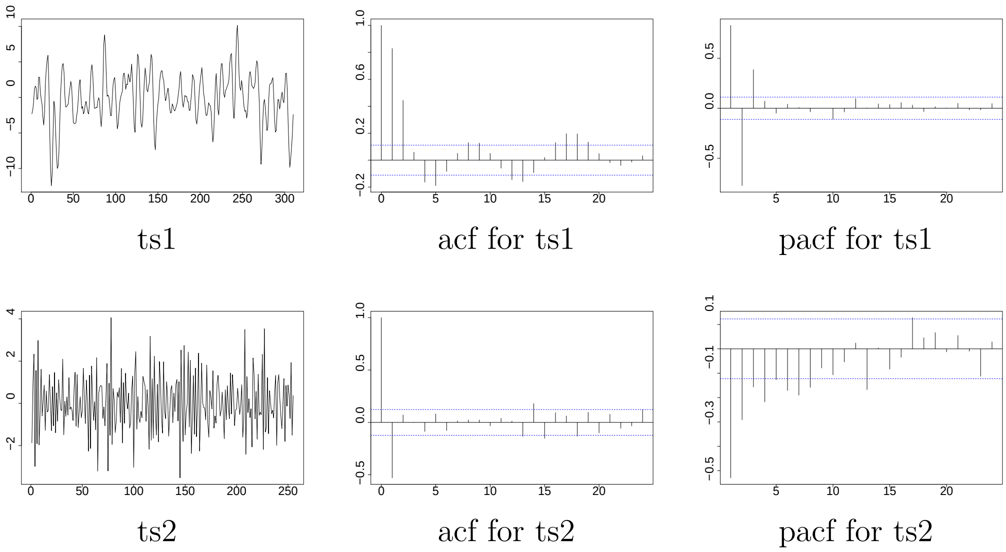


Figure 1: Time series ts1 and ts2 with estimated acf and pacf.

a) Assuming you want to fit ARIMA(p,d,q) models to these data sets, discuss for each of the two observed time series what values you would have started with for p, d and q. Remember to give reasons for your choices.

For a third time series, ts3, we have fitted an ARMA(1,1) model, an ARMA(1,2) model and an ARMA(2,1) model. The R output for fitted ARMA(1,1) model is

```
Call:
arima(x = x, order = c(1, 0, 1))

Coefficients:
         ar1     ma1  intercept
      0.9841  0.2162    -1.4961
s.e.  0.0091  0.0568     3.4168

sigma^2 estimated as 0.9445:  log likelihood = -558.09,  aic = 1124.18
```

The R output for the fitted ARMA(1,2) model is

```
Call:
arima(x = x, order = c(1, 0, 2))

Coefficients:
         ar1     ma1      ma2  intercept
      0.9869  0.2514  -0.1283    -1.8741
s.e.  0.0082  0.0574   0.0684     3.7526

sigma^2 estimated as 0.9353:  log likelihood = -556.21,  aic = 1122.41
```

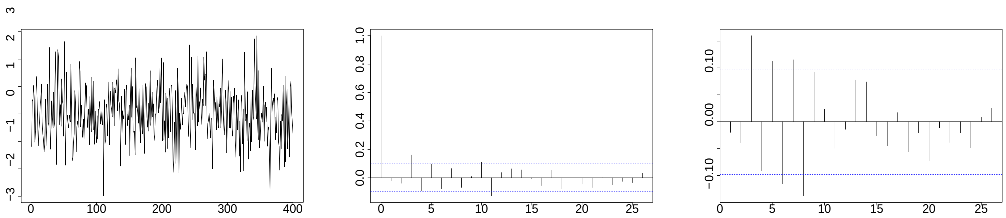
The R output for the fitted ARMA(2,1) model is

```
Call:
arima(x = x, order = c(2, 0, 1))

Coefficients:
         ar1     ar2     ma1  intercept
      0.3440  0.6345  0.9049    -1.6786
s.e.  0.0572  0.0573  0.0307     3.6834

sigma^2 estimated as 0.8779:  log likelihood = -543.76,  aic = 1097.52
```

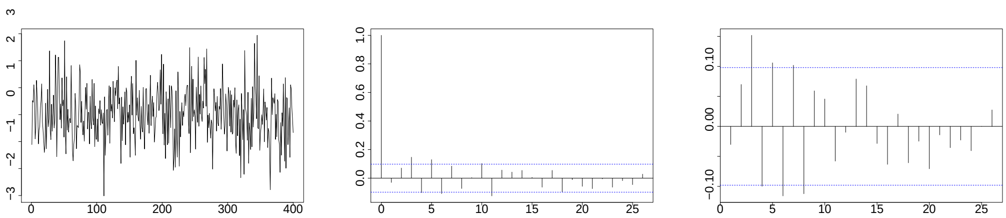
Plots of estimated residuals for the three models and corresponding estimated autocorrelation function (acf) and partial autocorrelation function (pacf) are shown in Figure 2.

b) For each of the three fitted models, discuss briefly whether or not the estimated model is a reasonable model for the observed time series.

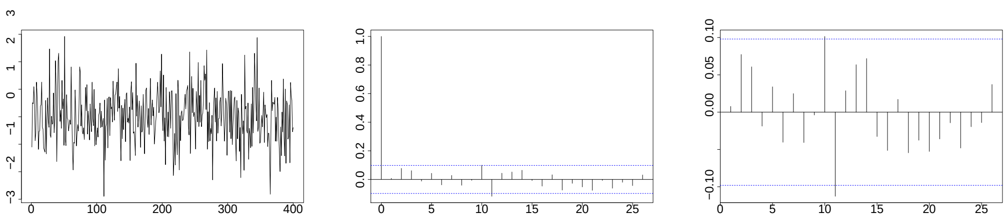
Which of the three fitted models would you use for the observed time series?



ARMA(1,1) residuals | acf for ARMA(1,1) residuals | pacf for ARMA(1,1) residuals



ARMA(1,2) residuals | acf for ARMA(1,2) residuals | pacf for ARMA(1,2) residuals



ARMA(2,1) residuals | acf for ARMA(2,1) residuals | pacf for ARMA(2,1) residuals

Figure 2: Estimated residuals with corresponding estimated acf and pacf for fitted ARMA(1,1), ARMA(1,2) and ARMA(2,1) models.

Problem 2

Let $\{z_t\}$ be a time series process.

a) What properties must $\{z_t\}$ have to be second-order stationary?

What properties must $\{z_t\}$ have to be covariance stationary (second-order weakly stationary)?

Can a time series process $\{z_t\}$ be covariance stationary, but not second-order stationary? Can a time series process $\{z_t\}$ be second-order stationary, but not covariance stationary? Explain your answers!

In the remaining part of this problem we will assume that $\{z_t\}$ follows the zero-mean ARMA(2,1) model

$$(1 - \varphi_1 B - \varphi_2 B^2) z_t = (1 - \theta_1 B) a_t,$$

where φ_1, φ_2 and θ are (real) parameters, B denotes the *backshift operator*, i.e. $B^k z_t = z_{t-k}$, and a_t is a white noise process with zero mean and variance $\mathrm{Var}[a_t] = \sigma_a^2$.

b) For what values of φ_1, φ_2 and θ_1 is the model for $\{z_t\}$ invertible?

Is the model covariance stationary when $\varphi_1 = 1.7$, $\varphi_2 = -0.72$ and $\theta_1 = 0.75$?

c) Assuming the above model to be covariance stationary, show that the auto-covariance function γ_k must fulfil the second order homogeneous difference equation

$$\gamma_k - \varphi_1 \gamma_{k-1} - \varphi_2 \gamma_{k-2} = 0 \quad \text{for} \quad k = 2, 3, \ldots.$$

Develop a set of equations that define initial conditions for this homogeneous difference equation.

d) Assuming the model to be invertible, find the coefficients $\psi_k, k = 1, 2, \ldots$ in the $AR(\infty)$ representation of the above model,

$$z_t = \sum_{k=1}^{\infty} \psi_k z_{t-k} + a_t.$$

Assume that we have observed $\{z_t\}_{t=-\infty}^{n}$ and that we want to use these observations to forecast future values of the process. Moreover, still assume that $\{z_t\}$ follows the zero-mean ARMA(2,1) model specified above and that the values of the model parameters φ_1, φ_2, θ_1 and σ_a^2 are known.

e) Find an expressions for the one-step ahead forecast $\widehat{z}_n(1)$.

When $\varphi_1 = 1.7$, $\varphi_2 = -0.72$ and $\theta_1 = 0.75$, develop the eventual forecast function, i.e. an expression of the k-step forecast $\widehat{z}_n(k), k \geq 2$ as a function of $\widehat{z}_n(1)$ and observed data.

Problem 3 Consider the state-space model

$$\begin{aligned} x_t &= \Phi x_{t-1} + \omega_t \quad \text{for } t = 1, 2, \ldots, \\ y_t &= A_t x_t + v_t \quad \text{for } t = 0, 1, 2, \ldots, \end{aligned}$$

where $\omega_1, \omega_2, \ldots$ and $v_0, v_1, \ldots$ are independent and $\omega_t \sim \mathrm{N}(0, Q)$ and $v_t \sim \mathrm{N}(0, R)$. Moreover, assume x_0 to be independent of all ω_t's and v_t's and $x_0 \sim \mathrm{N}(\mu_0, \Sigma_0)$.

In the following we assume $y_0, \ldots, y_n$ to be observed and that we want to forecast x_{n+k} for $k = 1, 2, \ldots$ with associated uncertainties. You can assume that the filtering problem is already solved, i.e. that we have available $\mathrm{E}[x_n | y_0, \ldots, y_n]$ and $\mathrm{Cov}[x_n | y_0, \ldots, y_n]$.

a) Develop recursive formulas that can be used to compute the k-steps ahead forecast $\mathrm{E}[x_{n+k} | y_0, \ldots, y_n]$ and the associated $\mathrm{Cov}[x_{n+k} | y_0, \ldots, y_n]$.

Explain why x_{n+k} given the observations $y_0, \ldots, y_n$ is Gaussian distributed, and use this to construct a 95% prediction interval for x_{n+k}.