Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4285 Time Series Models, Solution, December 2017

Problem 1

a)

$$\rho_k = \begin{cases} 1, k = 0 \\ \frac{3}{4}, k = \pm 1 \\ \frac{2}{4}, k = \pm 2 \\ \frac{1}{4}, k = \pm 3 \\ 0, |k| \geq 4 \end{cases}$$

b) (i) The pacf cuts of after lags $k > 1$ and the acf tails off. This suggest an AR(1) model.

(ii) The acf of the non-differenced series decays slowly which suggest that the process is non-stationary. The acf of the differenced series is significant at one multiple of $s = 4$ but cuts off at higher multiple while the pacf has tailing off behaviour at 4, 8, 12, ... This suggest that the seasonal part of the model is MA(1). Both the acf and pacf is zero at lag $k = 1$ and around each multiple of s. This suggest that the regular part of the model has no AR or MA part. Overall the model is thus $\text{ARIMA}(0,1,0) \times (0,0,1)_4$.

(iii) The acf cuts off for lags $k > 3$ while the pacf tails off. This suggests a MA(3) model. Note that if the MA-polynomial has complex roots this would translate to cycles in the pacf. These cycles happen to have a period of 4. The peaks in the pacf at multiples of 4 could perhaps suggest a seasonal MA(1) part but this is excluded by the fact that the acf at lag 4 is non-significant.

c) (i) $\hat{\phi}_1 = -.9$. (ii) $\Theta_1 = -1$. (iii) This looks like the acf of the model in point a. Thus, reasonable estimates are $\theta_1 = \cdots = \theta_3 = -1$.

Problem 2

a) The partial autocorrelation at lag k can be defined as

$$\phi_{kk} = \text{corr}(Z_t - \hat{Z}_t, Z_{t+k} - \hat{Z}_{t+k})$$

where $\hat{Z}_t$ and $\hat{Z}_{t+k}$ are minimum mean square linear predictors of Z_t and Z_{t+k} based on the intermediate observations $Z_{t+1}, \ldots, Z_{t+k-1}$.

Alternatively, ϕ_{kk} can be defined as the last coefficient in the regression of Z_{t+k} on $Z_{t+k-1}, \ldots, Z_t$ or, for Gaussian processes, as the correlation conditional on the intermediate observations.

For an AR(2) process we have $\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}$ for $k > 0$ such that

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = \frac{\begin{vmatrix} 1 & \rho_1 & \phi_1 + \phi_2\rho_1 \\ \rho_1 & 1 & \phi_1\rho_1 + \phi_2 \\ \rho_2 & \rho_1 & \phi_1\rho_2 + \phi_2\rho_1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}.$$

The third column of the matrix in the numerator is thus a linear combination of columns 1 and 2. The determinant and ϕ_{33} are thus both zero.

Problem 3

a) The autoregressive polynomial for the model $1 - B$ has a unit root and the process is thus not stationary. It is also not invertible since the root $B_1 = 1/2$ of the moving average polynomial $1 - 2B$ is inside the unit circle.

The moving average $a'_t - \frac{1}{2}a'_{t-1}$ will have the same autocovariance function if we let $\sigma_a^{2'} = 1$. Hence, the process can be represented by the invertible model

$$(1 - B)Z_t = (1 - \theta'_1 B)a'_t.$$

where $\theta'_1 = \frac{1}{2}$.

To write the model in pure autoregressive form

$$\pi(B)Z_t = a'_t$$

we must have

$$\begin{aligned} \pi(B)(1 - \theta'_1 B) &= 1 - B \\ 1 - \pi_1 B - \pi_2 B^2 - \ldots & \\ -\theta'_1 B + \theta_1\pi_1 B^2 + \ldots &= 1 - B \end{aligned}$$

Equating coefficients we find that $\pi_j = (\frac{1}{2})^j$.

b) The 1-step ahead forecast becomes

$$\begin{aligned}\hat{Z}_5(1) &= \pi_1 Z_t + \pi_2 Z_4 + \dots \\ &= .5 \cdot .32 + .25 \cdot 1.75 + .125 \cdot .12 + .0625 \cdot .72 \\ &= .6575.\end{aligned}$$

The 1-step ahead forecast variance $\operatorname{Var} e_5(1) = \sigma_a^{2\prime} = 1$.

For lead times $l > 1$ the forecast function satisifies

$$\begin{aligned}\hat{Z}_5(l) &= E(Z_{5+l} | Z_5, Z_4, \dots) \\ &= E(Z_{5+l-1} + a'_{5+l} - \theta'_1 a'_{5+l-1} | Z_5, Z_4, \dots) \\ &= \hat{Z}_5(l-1) \\ &= \hat{Z}_5(1) = .6575.\end{aligned}$$

c) To write the model in pure moving average form

$$Z_t = \psi(B) a'_t$$

we must have

$$(1 - B)\psi(B) = 1 - \theta'_1 B$$

which leads to $\psi_i = 1 - \theta'_1 = \frac{1}{2}$ for all i. A non-stationary model can not be repesented in pure moving average form, however, and the ψ_i's are indeed not square summable.

The variance of the l-step ahead forecast error can still be computed in the usual way, however, see the lecture summary p. 24 or Wei, ch. 5. For $l > 1$ we obtain

$$\begin{aligned}\operatorname{Var} e_5(l) &= \sigma_a^{2\prime} \sum_{j=0}^{l-1} \psi_j^2 \\ &= 1 + \frac{1}{4}(l-1) \\ &= \frac{3}{4} + \frac{1}{4} l.\end{aligned}$$

d) We have

$$Z_t = (1 - \theta_1 B)\xi_t \tag{1}$$

and

$$(1 - B)\xi_t = a_t. \tag{2}$$

Applying $(1 - B)$ to both sides of (1) and using (2) yields

$$(1 - B)Z_t = (1 - B)(1 - \theta_1 B)\xi_t = (1 - \theta_1 B)a_t \tag{3}$$

which shows that Z_t is an ARIMA(0,1,1) process.

A state space representation can be obtained by first writing (2) in vector AR(1) form as

$$\underbrace{\begin{bmatrix}\xi_t\\ \xi_{t-1}\end{bmatrix}}_{Y_t} = \underbrace{\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}\xi_{t-1}\\ \xi_{t-2}\end{bmatrix}}_{Y_{t-1}} + \underbrace{\begin{bmatrix}1\\ 0\end{bmatrix}}_{G} a_t. \tag{4}$$

The observed values depends on the the unobserved states through

$$Z_t = \underbrace{\begin{bmatrix}1 & -2\end{bmatrix}}_{H}\begin{bmatrix}\xi_t\\ \xi_{t-1}\end{bmatrix} \tag{5}$$

e) It follows that

$$\begin{aligned}
\hat{Y}_{6|5} &= E(Y_6|Z_1,\ldots,Z_5)\\
&= E(AY_5 + Ga_t|Z_1,\ldots,Z_5)\\
&= AE(Y_5|Z_1,\ldots,Z_5)\\
&= A\hat{Y}_{5|5}\\
&= \begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}\begin{bmatrix}-.668\\ -.494\end{bmatrix}\\
&= \begin{bmatrix}-.668\\ -.668\end{bmatrix}
\end{aligned}$$

Similarly,

$$\begin{aligned}
V_{6|5} &= \mathrm{Var}(Y_6|Z_1,\ldots,Z_5)\\
&= \mathrm{Var}(AY_5 + Ga_t|Z_1,\ldots,Z_5)\\
&= A\,\mathrm{Var}(Y_5|Z_1,\ldots,Z_5)A^T + G\sigma_a^2G^T\\
&= \begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}\begin{bmatrix}.7507 & .3754\\ .3754 & .1877\end{bmatrix}\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix} + \begin{bmatrix}0.25 & 0\\ 0 & 0\end{bmatrix}\\
&= \begin{bmatrix}1.007 & .7507\\ .7507 & .7507\end{bmatrix}
\end{aligned}$$

From this, the forecast of Z_6 is given by

$$\begin{aligned}\hat{Z}_5(1) &= E(Z_6|Z_1,\dots,Z_5)\\ &= E(HY_6|Z_1,\dots,Z_5)\\ &= HE(Y_6|Z_1,\dots,Z_5)\\ &= H\hat{Y}_{6|5}\\ &= [1 \quad -2]\begin{bmatrix}-.668\\ -.668\end{bmatrix}\\ &= 0.668.\end{aligned}$$

Similarly, the forecast error variance is given by

$$\begin{aligned}&\operatorname{Var}(Z_6|Z_1,\dots,Z_5)\\ &\qquad= \operatorname{Var}(HY_6|Z_1,\dots,Z_5)\\ &\qquad= H\operatorname{Var}(Y_6|Z_1,\dots,Z_5)H^T\\ &\qquad= HV_{6|5}H^T\\ &\qquad= [1 \quad -2]\begin{bmatrix}1.007 & .7507\\ .7507 & .7507\end{bmatrix}\begin{bmatrix}1\\ -2\end{bmatrix}\\ &\qquad= 1.007\end{aligned}$$

The actual forecast is very similar to the infinite history forecast in point b. This is expected since the forecast only depends stongly on the last few observations (the AR(∞) coefficients decays by a factor of $\theta_1' = 1/2$ for every time step).

The variance of the finite history forecast error is slightly larger than the inifinite history forecast error variance in point b (computed as if all the past is known) which again is expected since the finite history forecast is based on less information.

f) First consider ξ_0. To represent that we only have vague knowledge of this quantitiy it would be reasonable to assume that $E(\xi_0) = 0$ and that the variance is large, say $\operatorname{Var}(\xi_0) = 10^6$. Now, from the assumption $\xi_1 = \xi_0 + a_1$ it follows that

$$Y_{1|0} = E\begin{bmatrix}\xi_0 + a_1\\ \xi_0\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$$

and

$$V_{1|0} = \operatorname{Var}\begin{bmatrix}\xi_0 + a_1\\ \xi_0\end{bmatrix} = \begin{bmatrix}1000000.25 & 10^6\\ 10^6 & 10^6\end{bmatrix}$$

This makes ξ_1 and ξ_0 strongly correlated reflecting the fact that these two quantities are both highly uncertain but that they must have similar values (they are subsequent values in a random walk).

An alternative would perhaps be to assume that ξ_0 and ξ_1 are independent with large variances but this would lead to different results (much more uncertainty after having conditioned on Z_1) and would not utilize what we know about the process a priori.