MATH4425 (T1A) – Tutorial 7

Kazovskaia Anastasiia

March, 21

Important information

- T1A: **Thursday 19:00 - 19:50** (Rm 1033, LSK Bldg)
- Office hours: **Wednesday 14:00 - 14:50** (Math support center, 3rd floor, Lift 3)
- Any questions to be addressed to **akazovskaia@connect.ust.hk**

1 Forecasting. Minimum Mean Square Error Forecasts for ARMA/ARIMA models

Given a sequence of data $Z_1, Z_2, \cdots, Z_n$ from ARMA or ARIMA model, you can forecast $Z_{n+1}, \cdots, Z_{n+l}$ and give their forecasting intervals.

Here, we usually consider $\boldsymbol{a_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_a^2)}$.

1.1 Best linear predictor

We already know that **best predictor** (in terms of MSE) is a *conditional expectation* $\mathbb{E}(Z_{n+l} \mid Z_n, Z_{n-1}, \dots)$.

1.2 Formulas of computing forecasts

Forecasts can be calculated recursively as follows

$$\hat{Z}_n(l) = \alpha_1 \hat{Z}_n(l-1) + \alpha_2 \hat{Z}_n(l-2) + \cdots + \alpha_m \hat{Z}_n(l-m) +$$

$$\hat{a}_n(l) - \theta_1 \hat{a}_n(l-1) - \theta_2 \hat{a}_n(l-2) - \cdots - \theta_q \hat{a}_n(l-q),$$

where α_j are either the **initial** AR coefficients of $\mathbf{ARMA}(\boldsymbol{p, q})$ model (then, $m = p$) or the AR coefficients obtained from $\mathbf{ARMA}(\boldsymbol{p+d, q})$ representation model of $\mathbf{ARIMA}(\boldsymbol{p, d, q})$ model (then, $m = p + d$) and

$$\hat{Z}_n(j) = \begin{cases} \mathbb{E}(Z_{n+j} \mid Z_n, Z_{n-1}, \dots) & \text{if } j = 1, 2, \dots, l \\ Z_{n+j} & \text{if } j = 0, -1, \dots \end{cases}$$

$$\hat{a}_n(j) = \begin{cases} 0 & \text{if } j = 1, 2, \dots, l \\ a_{n+j} & \text{if } j = 0, -1, \dots \end{cases}$$

Actually, $\hat{a}_n(j)$ for $j \leq 0$ can be calculated as follows:

$$\hat{a}_n(j) = a_{n+(j-1)+1} = e_{n+j-1}(1) = Z_{n+j} - \hat{Z}_{n+j-1}(1)$$

1.3 Forecasting error

Forecasting error can be calculated as follows

$$e_n(l) = Z_{n+l} - \hat{Z}_n(l) = \sum_{j=0}^{l-1} \psi_j a_{n+l-j},$$

where ψ_j either correspond to $\mathbf{MA}(\infty)$ representation of $\mathbf{ARMA}(\boldsymbol{p},\boldsymbol{q})$ model or can be calculated **recursively** for $\mathbf{ARIMA}(\boldsymbol{p},\boldsymbol{d},\boldsymbol{q})$ model:

$$\psi_j = \sum_{k=0}^{j-1} \pi_{j-k}\psi_k \quad \forall j = 1, 2, \ldots, l-1$$

where π_j correspond to $\mathbf{AR}(\infty)$ representation of $\mathbf{ARIMA}(\boldsymbol{p},\boldsymbol{d},\boldsymbol{q})$ model.

1.4 Forecasting variance

First, note that $e_n(l) \sim \mathcal{N}(0, \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2)$. So,

$$\mathrm{var}[e_n(l)] = \mathbb{E}e_n^2(l) = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 \nearrow \text{ as } l \uparrow$$

1.5 Forecast interval (FI)

Forecast interval can be estimated as follows:

$$\left[\hat{Z}_n(l) - \mathcal{N}_{\frac{\alpha}{2}} \sigma_a \sqrt{\sum_{j=0}^{l-1} \psi_j^2}\ ,\ \hat{Z}_n(l) + N_{\frac{\alpha}{2}} \sigma_a \sqrt{\sum_{j=0}^{l-1} \psi_j^2}\right],$$

where $\mathcal{N}_{\frac{\alpha}{2}}$ is the $\frac{\alpha}{2}$-quantile of the standard normal distribution, i.e. $P(\mathcal{N}(0,1) > \mathcal{N}_{\frac{\alpha}{2}}) = \alpha/2$.

1.6 Updating forecasts for ARMA and ARIMA

If Z_{n+1} turns out to be available, how to forecast Z_{n+l}?

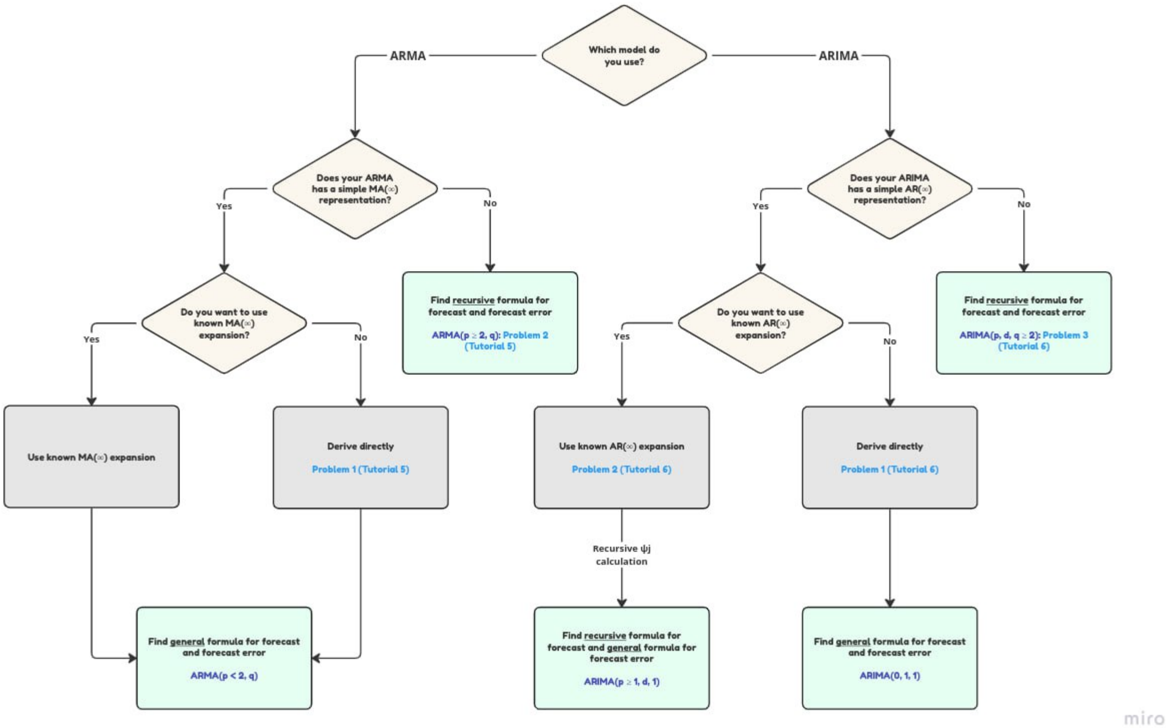
Method 1

We can forecast value of Z_{n+l} as $\hat{Z}_{n+1}(l-1)$.

Method 2

We can update the forecast $\hat{Z}_{n+1}(l-1) = \hat{Z}_n(l) + \psi_{l-1}[Z_{n+1} - \hat{Z}_n(1)]$

2 Guide to ARMA and ARIMA forecasting



3 Problems

Problem 1 (Updated Problem 3 from Tutorial 5)

Consider a model

$$Z_t - 1.2Z_{t-1} + 0.6Z_{t-2} = 26 + a_t,$$

given $\sigma_a^2 = 1$.

Suppose that we have the observations from this model:

$$Z_{76} = 60.4$$

$$Z_{77} = 58.9$$

$$Z_{78} = 64.7$$

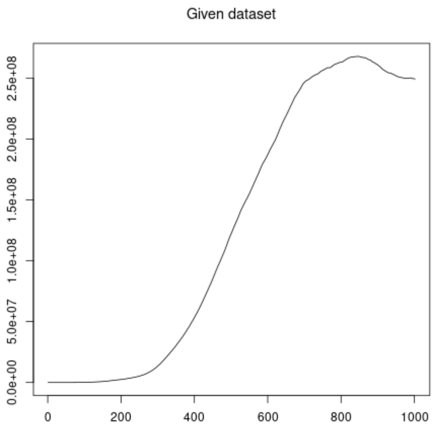
$$Z_{79} = 70.4$$

$$Z_{80} = 62.6$$

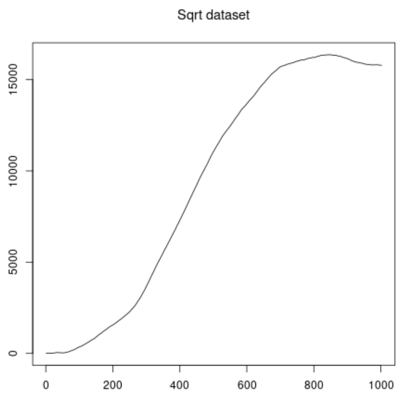
1) Forecast $Z_{81}, Z_{82}, Z_{83}, Z_{84}$
2) Find the 99% forecast interval ($\mathcal{N}_{\frac{0.01}{2}} = 2.576$)
3) Suppose that the observation at $t = 81$ turns out to be $Z_{81} = 62.2$. Find the updated forecasts for Z_{82}, Z_{83}, Z_{84}

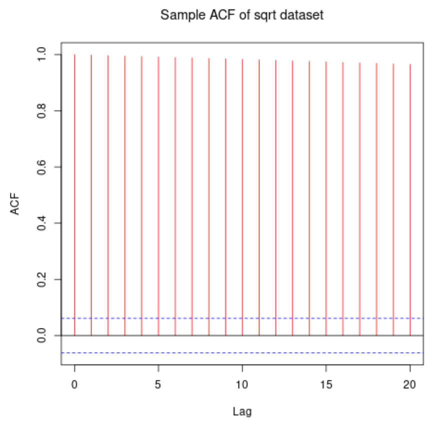
Problem 2

Let's consider the following example:

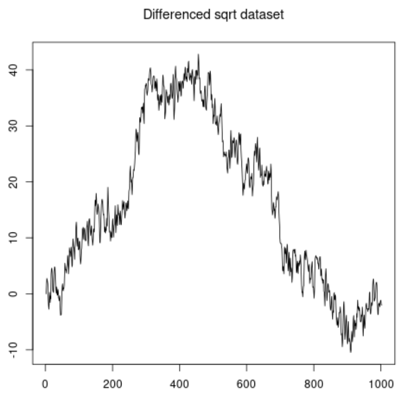


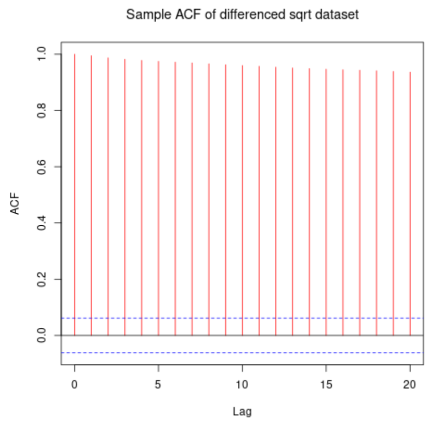
Given time series is increasing very fast but still not «exploding». Let's apply square root transformation:



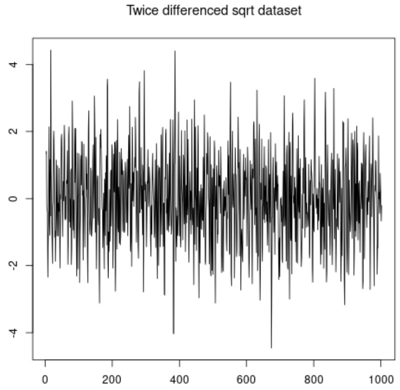


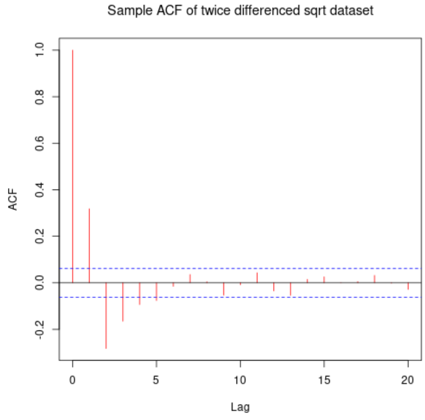
Obviously, the time series is not stationary. Let's difference it:



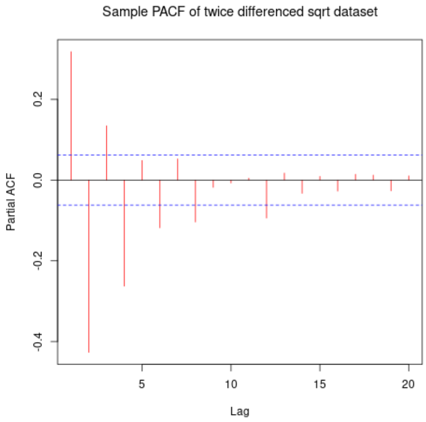


Seems, the time series is still not stationary. Let's difference it once again:





Now the time series looks stationary. We can take a look at sample PACF to pick the model:



I suggest to try to fit ARMA(1/2/3, 2/3) models (i.e. estimate the parameters using conditional least squares, for example). But there are always plenty of other options to try, of course.

Here is the (filtered) output of fitting via `arima` from `tseries` package:

```
Model:
ARMA(1,2)

Coefficient(s):
           Estimate  Std. Error  t value Pr(>|t|)
ar1        0.568160    0.051353   11.064   <2e-16 ***
ma1       -0.007969    0.041124   -0.194    0.846
ma2       -0.668155    0.029482  -22.663   <2e-16 ***
intercept -0.001835    0.010730   -0.171    0.864
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 1.091,  Conditional Sum-of-Squares = 1087.84,  AIC = 2933.07
```

```
Model:
ARMA(1,3)

Coefficient(s):
           Estimate  Std. Error  t value Pr(>|t|)
ar1        0.480941    0.102282    4.702 2.57e-06 ***
ma1        0.094448    0.106171    0.890    0.374
ma2       -0.631486    0.046363  -13.621  < 2e-16 ***
ma3       -0.072377    0.055147   -1.312    0.189
intercept -0.001646    0.012933   -0.127    0.899
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 1.09,  Conditional Sum-of-Squares = 1085.49,  AIC = 2933.92


Model:
ARMA(2,2)

Coefficient(s):
           Estimate  Std. Error  t value Pr(>|t|)
ar1        0.573464    0.057540    9.966   <2e-16 ***
ar2       -0.064403    0.045784   -1.407    0.160
ma1        0.003906    0.049660    0.079    0.937
ma2       -0.625252    0.047426  -13.184   <2e-16 ***
intercept -0.002070    0.012525   -0.165    0.869
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 1.089,  Conditional Sum-of-Squares = 1085.69,  AIC = 2933.1


Model:
ARMA(2,3)

Coefficient(s):
            Estimate  Std. Error  t value Pr(>|t|)
ar1        1.1168702   0.2451489    4.556 5.22e-06 ***
ar2       -0.3520229   0.1400247   -2.514   0.0119 *
ma1       -0.5478094   0.2478145   -2.211   0.0271 *
ma2       -0.6330182   0.0373184  -16.963  < 2e-16 ***
ma3        0.3608807   0.1676641    2.152   0.0314 *
intercept  0.0003612   0.0059592    0.061   0.9517
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 1.088,  Conditional Sum-of-Squares = 1083.68,  AIC = 2934.23
```

```
Model:
ARMA(3,2)

Coefficient(s):
           Estimate  Std. Error  t value Pr(>|t|)
ar1        0.618185    0.071619    8.632   <2e-16 ***
ar2       -0.084857    0.056066   -1.514    0.130
ar3        0.028796    0.046809    0.615    0.538
ma1       -0.040681    0.064440   -0.631    0.528
ma2       -0.631771    0.042539  -14.851   <2e-16 ***
intercept -0.001389    0.010845   -0.128    0.898
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 1.089,  Conditional Sum-of-Squares = 1084.7,  AIC = 2935.19


Model:
ARMA(3,3)

Coefficient(s):
            Estimate  Std. Error  t value Pr(>|t|)
ar1        1.0971711   0.3077812    3.565 0.000364 ***
ar2       -0.3605254   0.1826583   -1.974 0.048408 *
ar3        0.0118757   0.0464066    0.256 0.798024
ma1       -0.5165756   0.3075977   -1.679 0.093077 .
ma2       -0.6246548   0.0386421  -16.165  < 2e-16 ***
ma3        0.3332951   0.2023241    1.647 0.099490 .
intercept  0.0007086   0.0063889    0.111 0.911681
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 1.088,  Conditional Sum-of-Squares = 1083.55,  AIC = 2936.1
```

All models look reasonable, though AIC of ARMA(1, 2) (`AIC = 2933.07`) and ARMA(2, 2) (`AIC = 2933.1`) are a bit smaller (true parameters are `ar=c(0.5), ma=c(0.1, -0.7)`, by the way).

Finally, the task. The last few observations of $X_t = (1-B)^2\sqrt{Z_t}$ (Z_t is the given dataset) are:

$$X_{998} = 0.09$$

$$X_{999} = -0.67$$

$$X_{1000} = -0.21$$

Considering the above observations are coming from ARMA(2, 2) with $\phi_1 = 0.573, \phi_2 = -0.064$ and $\theta_1 = -0.004, \theta_2 = 0.625$ (using our notation for MA coefficients) with $\sigma_a^2 = 1.089$

1) Forecast $X_{1001}, X_{1002}, X_{1003}$

2) Find the 95% forecast interval ($\mathcal{N}_{\frac{0.05}{2}} = 1.96$)

4 Solutions

Solution 1

Here we have the following **stationary** AR(2) model with drift:

$$Z_t = 26 + 1.2Z_{t-1} - 0.6Z_{t-2} + a_t$$

Notice that here $(1 - \phi_1 - \phi_2)\mu = 26$, so

$$\mu = \frac{26}{1 - 1.2 + 0.6} = 65$$

1) First, introduce the notation $\dot{Z}_t := Z_t - \mu$. Then

$$\hat{Z}_{80}(l) = \mathbb{E}(Z_{80+l} \mid Z_{80}, Z_{79}, \dots) = \mathbb{E}(\dot{Z}_{80+l} + \mu \mid Z_{80}, Z_{79}, \dots) =$$

$$\mathbb{E}(\dot{Z}_{80+l} \mid Z_{80}, Z_{791}, \dots) + \mu = \hat{\dot{Z}}_{80}(l) + 65$$

We can forecast for the AR(2) model **without** drift. Let's find **recursive** formula for $\hat{\dot{Z}}_{80}(l)$ (in this case, it is a better option because the MA(∞) expansion is complicated):

$$\hat{\dot{Z}}_{80}(l) = \mathbb{E}(\dot{Z}_{80+l} \mid Z_{80}, Z_{79}, \dots) = \mathbb{E}(1.2\dot{Z}_{n+l-1} - 0.6\dot{Z}_{n+l-2} + a_{80+l} \mid Z_{80}, Z_{79}, \dots) =$$

$$1.2\hat{\dot{Z}}_{80}(l-1) - 0.6\hat{\dot{Z}}_{80}(l-2) + 0 = 1.2\hat{\dot{Z}}_{80}(l-1) - 0.6\hat{\dot{Z}}_{80}(l-2)$$

For $l = 1, 2, 3, 4$ we will thus only have to calculate $\hat{\dot{Z}}_{80}(l)$ recursively having

$$\hat{\dot{Z}}_{80}(1) = 1.2\hat{\dot{Z}}_{80}(0) - 0.6\hat{\dot{Z}}_{80}(-1) = 1.2\dot{Z}_{80} - 0.6\dot{Z}_{79} =$$

$$1.2(62.6 - 65) - 0.6(70.4 - 65) = -6.12$$

$$\hat{\dot{Z}}_{80}(2) = 1.2\hat{\dot{Z}}_{80}(1) - 0.6\hat{\dot{Z}}_{80}(0) = 1.2\hat{\dot{Z}}_{80}(1) - 0.6\dot{Z}_{80} =$$

$$1.2(-6.12) - 0.6(62.6 - 65) \approx -5.90$$

as initial conditions.

For example, for $l = 3$ we have

$$\hat{\dot{Z}}_{80}(3) = 1.2\hat{\dot{Z}}_{80}(2) - 0.6\hat{\dot{Z}}_{80}(1) =$$

$$1.2(-5.90) - 0.6(-6.12) \approx -3.41$$

Finally, for $l = 1, 2, 3, 4$ the forecast for the initial AR(2) model with drift can be obtained. For example:

$$\hat{Z}_{80}(1) = \hat{\dot{Z}}_{80}(1) + 65 = -6.12 + 65 = 58.88$$

$$\hat{Z}_{80}(2) = \hat{\dot{Z}}_{80}(2) + 65 = -5.90 + 65 = 59.10$$

$$\hat{Z}_{80}(3) = \hat{\dot{Z}}_{80}(3) + 65 = -3.41 + 65 = 61.59$$

2) For $l = 1, 2, 3, 4$ we have

$$e_{80}(l) = Z_{80+l} - \hat{Z}_{80}(l) =$$

$$[\underbrace{(1 - 1.2 + 0.6)65}_{26} + 1.2Z_{80+l-1} - 0.6Z_{80+l-2} + a_{80+l}] - [1.2(\hat{Z}_{80}(l-1) - 65) - 0.6(\hat{Z}_{80}(l-2) - 65) + 65] =$$

$$1.2(Z_{80+l-1} - \hat{Z}_{80}(l-1)) - 0.6(Z_{80+l-2} - \hat{Z}_{80}(l-2)) + a_{80+l} =$$

$$1.2e_{80}(l-1) - 0.6e_{80}(l-2) + a_{80+l}$$

with initial conditions

$$e_{80}(1) = Z_{81} - \hat{Z}_{80}(1) =$$

$$[\underbrace{(1-1.2+0.6)65}_{26} + 1.2Z_{80} - 0.6Z_{79} + a_{81}] - [1.2(Z_{80} - 65) - 0.6(Z_{79} - 65) + 65] = a_{81}$$

$$e_{80}(2) = Z_{82} - \hat{Z}_{80}(2) =$$

$$[\underbrace{(1-1.2+0.6)65}_{26} + 1.2Z_{81} - 0.6Z_{80} + a_{82}] - [1.2(\hat{Z}_{80}(1) - 65) - 0.6(Z_{80} - 65) + 65] =$$

$$1.2e_{80}(1) + a_{82} = 1.2a_{81} + a_{82}$$

The next $e_{80}(l)$ can be obtained using **recursive** formula. For example, for $l = 3$ we have

$$e_{80}(3) = 1.2e_{80}(2) - 0.6e_{80}(1) + a_{83} = 1.2(1.2a_{81} + a_{82}) - 0.6a_{81} + a_{83} =$$

$$0.84a_{81} + 1.2a_{82} + a_{83}$$

Notice, that we have just obtained $\psi_0 = 1, \psi_1 = 1.2, \psi_2 = 0.84$.

For $l = 1, 2$

$$\mathrm{var}(e_{80}(1)) = \mathrm{var}(a_{81}) = \sigma_a^2 = 1$$

$$\left[58.88 - 2.58 \times \sqrt{1}\ ,\ 32.88 + 2.58 \times \sqrt{1}\right],$$

that is,

$$\left[56.30\ ,\ 61.46\right]$$

and

$$\mathrm{var}(e_{80}(2)) = \mathrm{var}(1.2a_{81} + a_{82}) = \sigma_a^2[1.2^2 + 1] = 2.44$$

$$\left[59.10 - 2.58 \times \sqrt{2.44}\ ,\ 59.10 + 2.58 \times \sqrt{2.44}\right],$$

that is,

$$\left[55.07\ ,\ 63.13\right]$$

3) We already know how to update the forecasts:

$$\hat{Z}_{81}(l-1) = \hat{Z}_{80}(l) + \psi_{l-1}[Z_{81} - \hat{Z}_{80}(1)]$$

Thus, for $l = 2, 3$ we get updated forecasts:

$$\hat{Z}_{82} = \hat{Z}_{81}(1) = \hat{Z}_{80}(2) + \psi_1[Z_{81} - \hat{Z}_{80}(1)] =$$

$$59.10 + 1.2[62.20 - 58.88] \approx 63.08$$

$$\hat{Z}_{83} = \hat{Z}_{81}(2) = \hat{Z}_{80}(3) + \psi_2[Z_{81} - \hat{Z}_{80}(1)] =$$

$$61.59 + 0.84[62.20 - 58.88] \approx 64.38$$

4.1 Solution 2

Here we have the following **stationary** and **invertible** ARMA(1, 2) model without drift:

$$X_t = 0.573X_{t-1} - 0.064X_{t-2} + a_t - 0.004a_{t-1} - 0.625a_{t-2}$$

1) Let's find **general** formula (although it is a bit complicated) for $\hat{X}_{1000}(l)$. First, let's find MA(∞) representation of X_t:

$$X_t = (1 - 0.573B + 0.064B^2)^{-1}(1 - 0.004B - 0.625B^2)a_t = \sum_{i=0}^{\infty} \psi_i B^i a_t \Leftrightarrow$$

$$(1 - 0.004B - 0.625B^2)a_t = \sum_{i=0}^{\infty} \psi_i(1 - 0.573B + 0.064B^2)B^i a_t =$$

$$\psi_0 a_t + (\psi_1 - 0.573\psi_0)a_{t-1} + \sum_{i=2}^{\infty}(\psi_i - 0.573\psi_{i-1} + 0.064\psi_{i-2})B^i a_t$$

Thus, we get

$$\psi_0 = 1$$
$$\psi_1 - 0.573\psi_0 = -0.004$$
$$\psi_2 - 0.573\psi_1 + 0.064\psi_0 = -0.625$$
$$\psi_i - 0.573\psi_{i-1} + 0.064\psi_{i-2} = 0 \ \ \forall i > 2$$

Let's calculate a few first ψ_i:

$$\psi_0 = 1$$
$$\psi_1 = 0.569$$
$$\psi_2 \approx -1.015$$

Finally, the **general** formula for $\hat{X}_{1000}(l)$:

$$\hat{X}_{1000}(l) = \psi_l a_n + \psi_{l+1}a_{n-1} + \psi_{l+2}a_{n-2} + \cdots$$

However, this formula is not convenient for forecasting. It is still much easier to use **recursive** formula (though, ψ_i will come in handy for $e_{1000}(l)$ calculation):

$$\hat{X}_{1000}(l) = 0.573\hat{X}_{1000}(l-1) - 0.064\hat{X}_{1000}(l-2) + \hat{a}_{1000}(l) - 0.004\hat{a}_{1000}(l-1) - 0.625\hat{a}_{1000}(l-2)$$

First let's calculate noise predictions:

1) Since there's no way to calculate *very past* noises (we don't have enough past data), we assume

$$a_{100+j} = 0 \quad \forall j \leq -3,$$

2) a) For $j = -2$ we have

$$\hat{a}_{1000}(-2) = a_{998} = X_{998} - \hat{X}_{997}(1) =$$

$$X_{998} - (0.573X_{997} - 0.064X_{996} + 0 - 0 - 0) = X_{998} = 0.09,$$

since X_{1000+j} for $j \leq -3$ is a linear combination of past noises which are assumed to be 0 (the model is stationary $\Rightarrow$ MA(∞) representation exists)

b) For $j = -1$ we have

$$\hat{a}_{1000}(-1) = a_{999} = X_{999} - \hat{X}_{998}(1) =$$

$$X_{999} - (0.573X_{998} - 0.064X_{997} + 0 - 0.004a_{998} - 0) \approx -0.721,$$

c) For $j = 0$ we have

$$\hat{a}_{1000}(0) = a_{1000} = X_{1000} - \hat{X}_{999}(1) =$$

$$X_{1000} - (0.573X_{999} - 0.064X_{998} + 0 - 0.004a_{999} - 0.625a_{998}) \approx 0.233$$

For $l = 1$ we have

$$\hat{X}_{1001} = \hat{X}_{1000}(1) = 0.573X_{1000} - 0.064X_{999} + 0 - 0.004\hat{a}_{1000}(0) - 0.625\hat{a}_{999} \approx 0.372$$

For $l = 2$ we have

$$\hat{X}_{1002} = \hat{X}_{1000}(2) = 0.573\hat{X}_{1001} - 0.064X_{1000} + 0 - 0 - 0.625\hat{a}_{1000}(0) \approx 0.081$$

2) For any l we have

$$e_{1000}(l) = X_{1000+l} - \hat{X}_{1000}(l) = \sum_{j=0}^{l-1} \psi_j a_{1000+l-j}$$

For $l = 1, 2$, for example,

$$\mathrm{var}(e_{1000}(1)) = \mathrm{var}(a_{1001}) = \sigma_a^2 = 1.089$$

$$\left[0.372 - 1.96 \times \sqrt{1.089}\ ,\ 0.372 + 1.96 \times \sqrt{1.089}\right],$$

that is,

$$\left[-1.725\ ,\ 2.365\right]$$

and

$$\mathrm{var}(e_{1000}(2)) = \mathrm{var}(0.569a_{1001} + a_{1002}) = \sigma_a^2[0.569^2 + 1] \approx 1.381$$

$$\left[0.081 - 1.96 \times \sqrt{1.381}\ ,\ 0.081 + 1.96 \times \sqrt{1.381}\right],$$

that is,

$$\left[-2.222\ ,\ 2.384\right]$$