Time Series Analysis

Solutions to problems in Chapter 3

IMM

Solution 3.1

Question 1.

In the light of plotting Y_t vs. x_t the following simple linear regression model seems reasonable

$$Y_t = \alpha + \beta x_t + \epsilon_t$$

where $\{\epsilon\}$ is assumed to be a sequence of mutually uncorrelated normal distributed random variables with mean value 0 and variance σ^2.

The observations can be written on matrix form:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \\ Y_7 \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \\ 1 & x_5 \\ 1 & x_6 \\ 1 & x_7 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \end{bmatrix}$$

or

$$Y = \mathbf{x}\theta + \epsilon$$

We find

$$\mathbf{x}^\top \mathbf{x} = \begin{bmatrix} 7 & 21.5 \\ 21.5 & 72.75 \end{bmatrix} \quad , \quad \mathbf{x}^\top \mathbf{Y} = \begin{bmatrix} 16 \\ 41.5 \end{bmatrix}$$

I.e. from Theorem 3.1 (3.35) we get

$$\hat{\theta} = \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = (\mathbf{x}^\top \mathbf{x})^{-\mathbf{1}} \mathbf{x}^\top \mathbf{Y} = \begin{bmatrix} 1.5479 & -0.4574 \\ -0.4574 & 0.1489 \end{bmatrix} \begin{bmatrix} 16 \\ 41.5 \end{bmatrix} = \underline{\underline{\begin{bmatrix} 5.784 \\ -1.139 \end{bmatrix}}}$$

An unbiased estimate of σ^2 is from Theorem 3.4 (3.44)

$$\begin{aligned}\hat{\sigma^2} &= \frac{(\mathbf{Y}-\mathbf{x}\hat{\theta})^\top(\mathbf{Y}-\mathbf{x}\hat{\theta})}{n-p} \\ &= [(-0.228)^2+(-0.728)^2+(0.203)^2+(0.772)^2+(-0.006)^2+ \\ &\quad (0.064)^2+(-0.076)^2]/(7-2) \\ &= \underline{\underline{(0.496)^2}}\end{aligned}$$

where n is the number of observations and p is the number of parameters.
Question 2.

$$\hat{Y}_{8|7} = \mathbf{x_8}^\top\hat{\theta} = \begin{bmatrix} 1 & 0.5 \end{bmatrix}\begin{bmatrix} 5.784 \\ -1.139 \end{bmatrix} = \underline{\underline{5.21}}$$

The prediction error is $e_8 = Y_8 - \hat{Y}_{8|7}$. An estimate of the prediction error variance is according to Theorem 3.10 (3.60)

$$\begin{aligned}\hat{V}[e_8] &= \hat{\sigma}^2[1+\mathbf{x_8}^\top(\mathbf{x}^\top\mathbf{x})^{-\mathbf{1}}\mathbf{x_8}] \\ &= 0.496^2\left[1+\begin{bmatrix} 1 & 0.5 \end{bmatrix}\begin{bmatrix} 1.5479 & -0.4574 \\ -0.4574 & 0.1489 \end{bmatrix}\begin{bmatrix} 1 \\ 0.5 \end{bmatrix}\right] \\ &= \underline{\underline{0.724^2}}\end{aligned}$$

which leads to the following 90% confidence interval for Y_8

$$\hat{Y}_{8|7} \pm t_{0.05}(n-p)\cdot\hat{V}[e_8]^{\frac{1}{2}} = \underline{\underline{[3.75, 6.67]}}$$

Question 3.

A plot of Y_t versus t reveals that a (global) linear trend would be a reasonable description of the variations in Y_t (a larger data set could give rise to considering local linear trend models).

I.e. the model is

$$Y_{n+j} = \mathbf{f}^\top(j)\theta + \epsilon_{n+j}$$

where $\mathbf{f}(j) = [1\ j]^\top$ and $\theta = [\theta_0\ \theta_1]^\top$. $\{\epsilon_t\}$ is again assumed to be a sequence of mutually uncorrelated normal distributed random variables with mean value 0 and variance σ^2.

$$\begin{aligned} \mathbf{F}_7 &= \mathbf{x}_7^\top \mathbf{x}_7 = \sum_{j=0}^{6} \mathbf{f}(-j)\mathbf{f}^\top(-j) = \begin{bmatrix} 7 & -21 \\ -21 & 91 \end{bmatrix} \\ \mathbf{h}_7 &= \mathbf{x}_7^\top \mathbf{Y} = \sum_{j=0}^{6} \mathbf{f}(-j)\mathbf{Y}_{n-j} = \begin{bmatrix} 16 \\ -32.5 \end{bmatrix} \end{aligned}$$

I.e. from (3.89)

$$\hat{\theta}_7 = \mathbf{F}_7^{-1}\mathbf{h}_7 = \begin{bmatrix} 0.46429 & 0.10714 \\ 0.10724 & 0.03571 \end{bmatrix} \begin{bmatrix} 16 \\ -32.5 \end{bmatrix} = \begin{bmatrix} 3.947 \\ 0.554 \end{bmatrix}$$

An unbiased estimate of σ^2 is given by (3.44)

$$\begin{aligned} \hat{\sigma}^2 &= \frac{\sum_{j=0}^{6}(Y_{7-j} - \mathbf{f}^\top(-j)\hat{\theta}_7)^2}{7-2} \\ &= [(0.053)^2 + (-0.393)^2 + (0.661)^2 + (-0.285)^2 + (0.269)^2 + \\ &\quad (-0.677)^2 + (0.377)^2]/5 \\ &= \underline{\underline{(0.5193)^2}} \end{aligned}$$

And the variance of the prediction error is obtained from (3.91)

$$\begin{aligned} \hat{V}[e_8] &= \hat{\sigma}^2[1 + \mathbf{f}^\top(1)\mathbf{F}_7^{-1}\mathbf{f}(1)] \\ &= 0.5193^2 \left[1 + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 0.46429 & 0.10714 \\ 0.10714 & 0.03571 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right] \\ &= \underline{\underline{0.680}}^2 \end{aligned}$$

According to (3.92) the 90% confidence interval is

$$\hat{\mathbf{Y}}_{8|7} \pm t_{0.05}(7-2) \cdot \hat{V}[e_8]^{\frac{1}{2}} = \underline{\underline{[3.131, 5.871]}}$$

Solution 3.2

Question 1.

The unweighted least square estimator ($\hat{\beta}^*$) is given by

$$\hat{\beta}^* = (X^T X)^{-1} X^T Y \ .$$

The expected value of the estimator is

$$\begin{aligned} \mathrm{E}[\hat{\beta}^*] =& \mathrm{E}[(X^T X)^{-1} X^T Y] \\ =& \mathrm{E}[(X^T X)^{-1} X^T (X\beta + \epsilon)] \\ =& \beta + (X^T X)^{-1} \mathrm{E}[\epsilon] = \beta \end{aligned}$$

I.e. the estimate is unbiased.
The variance of the estimator can be calculated as

$$V[\hat{\beta}^*] = \mathrm{E}[(\hat{\beta}^* - \mathrm{E}[\hat{\beta}^*])(\hat{\beta}^* - \mathrm{E}[\hat{\beta}^*])^T]$$

where,

$$\begin{aligned} \hat{\beta}^* - \mathrm{E}[\hat{\beta}^*] =& \hat{\beta}^* - \beta = (X^T X)^{-1} X^T Y - \beta \\ =& (X^T X)^{-1} X^T (X\beta + \epsilon) - \beta = (X^T X)^{-1} X^T \epsilon \end{aligned}$$

i.e

$$\begin{aligned} V[\hat{\beta}^*] =& (X^T X)^{-1} X^T V[\epsilon] X (X^T X)^{-1} \\ =& (X^T X)^{-1} \begin{bmatrix} X_1 & \cdots & X_N \end{bmatrix} \begin{bmatrix} \frac{\sigma^2}{X_1^2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\sigma^2}{X_N^2} \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix} (X^T X)^{-1} \\ =& \sigma^2 N (X^T X)^{-2} = \underline{\underline{\frac{N\sigma}{\left(\sum X_t^2\right)^2}}} \end{aligned}$$

Question 2.

The weighted least square estimator $\hat{\beta}$ is given by

$$\hat{\beta} = (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}Y$$

We have

$$\begin{aligned}\hat{\beta} - \beta =& (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}(X\beta + \epsilon) - \beta \\ =& (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}\epsilon\end{aligned}$$

As $\mathrm{E}[\hat{\beta}] = \beta$ the expression for the variance becomes

$$\begin{aligned}V[\hat{\beta}] =& (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}V[\epsilon]X(X^T\Sigma^{-1}X)^{-1} \\ =& \sigma^2(X^T\Sigma^{-1}X)^{-1} = \underline{\underline{\frac{\sigma^2}{\sum X_t^4}}}\end{aligned}$$

Question 3.

$$\begin{aligned}\frac{\left(\sum^N X_t^2\right)^2}{N} =& \frac{X_1^2X_1^2 + X_1^2X_2^2 + X_1^2X_3^2 + ... + X_N^2X_N^2}{N} \\ =& \frac{X_1^2X_1^2 + X_2^2X_2^2 + 2X_1^2X_2^2 + 2X_1^2X_3^2 + ...}{N} \\ \leq& \frac{X_1^2X_1^2 + X_2^2X_2^2 + X_1^2X_1^2 + X_2^2X_2^2 + X_1^2X_1^2 + X_3^2X_3^2 + ...}{N}\end{aligned}$$

(as $(X_i^2 + X_j^2)^2 \geq 0 \Leftrightarrow X_i^4 + X_j^4 \geq 2X_i^2X_j^2$). The numerator of the above fraction contains N terms with X_i^4. I.e. the fraction is equal to $\sum^N X_t^4$ and

$$\underline{\underline{V[\hat{\beta}^*] \geq V[\hat{\beta}]}}$$

Question 4.

The variance of ϵ is now set to

$$V[\epsilon] = \sigma^2 \Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & 0 & \cdots & 0 & 0 \\ \rho & 1 & \rho & \cdots & 0 & 0 \\ 0 & \rho & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \rho \\ 0 & 0 & 0 & \cdots & \rho & 1 \end{bmatrix}$$

The unweighted least square estimator

$$\hat{\beta}^* = (X^T X)^{-1} X^T Y \ ,$$

is still unbiased, as

$$\mathrm{E}[\hat{\beta}^*] = \beta + (X^T X)^{-1} X^T \mathrm{E}[\epsilon] = \beta$$

The variance of the estimator is

$$\begin{aligned} V[\hat{\beta}^*] =& (X^T X)^{-1} X^T V[\epsilon] X (X^T X)^{-1} \\ =& \underline{\underline{\frac{\sigma^2 \left(\sum_{t=1}^{N} X_t^2 + 2\rho \sum_{t=1}^{N-1} X_t X_{t+1} \right)}{\left(\sum_{t=1}^{N} X_t^2 \right)^2}}} \end{aligned}$$

Solution 3.3

Question 1.:

The observations $Y_1, Y_2, \cdots, Y_N$ can be described by the model

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} 1 & \cos\omega & \sin\omega \\ 1 & \cos 2\omega & \sin 2\omega \\ \vdots & \vdots & \vdots \\ 1 & \cos N\omega & \sin N\omega \end{bmatrix} \begin{bmatrix} \mu \\ \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix}$$

Or

$$\mathbf{Y} = \mathbf{x}\theta + \epsilon$$

The stochastic vector ϵ has $E[\epsilon] = 0$ and $V[\epsilon] = \sigma^2 I$ (I is the identity matrix).

The least square estimate of $\theta = [\mu \ \alpha \ \beta]^\top$ is calculated as the solution to the normal equation (Theorem 3.1)

$$\hat{\theta} = (\mathbf{x}^\top \mathbf{x})^{-\mathbf{1}} \mathbf{x}^\top \mathbf{Y} \ ,$$

where

$$\mathbf{x}^\top \mathbf{x} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \cos\omega & \cos 2\omega & \cdots & \cos N\omega \\ \sin\omega & \sin 2\omega & \cdots & \sin N\omega \end{bmatrix} \begin{bmatrix} 1 & \cos\omega & \sin\omega \\ 1 & \cos 2\omega & \sin 2\omega \\ \vdots & \vdots & \vdots \\ 1 & \cos N\omega & \sin N\omega \end{bmatrix}$$

$$= \begin{bmatrix} N & \sum_{t=1}^N \cos t\omega_i & \sum_{t=1}^N \sin t\omega \\ \sum_{t=1}^N \cos t\omega & \sum_{t=1}^N \cos^2 t\omega & \sum_{t=1}^N \cos t\omega \sin t\omega_i \\ \sum_{t=1}^N \sin t\omega & \sum_{t=1}^N \cos t\omega \sin t\omega & \sum_{t=1}^N \sin^2 t\omega \end{bmatrix}$$

and

$$\mathbf{x}^\top \mathbf{Y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \cos\omega & \cos 2\omega & \cdots & \cos N\omega \\ \sin\omega & \sin 2\omega & \cdots & \sin N\omega \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{t=1}^N Y_t \\ \sum_{t=1}^N (\cos t\omega) Y_t \\ \sum_{t=1}^N (\sin t\omega) Y_t \end{bmatrix}$$

Question 2.:

The fundamental frequency (where a period in oscillations covers exactly all observations) is $\omega_1 = 2\pi/N$. We set the frequency to be equal a multiplum of the fundamental frequency, i.e. $\omega = \omega_i = 2\pi i/N \quad (i \in \mathbb{Z})$.

Aplying the hints and the fact that

$$\sum_{t=1}^N \cos(t\omega_i) = \sum_{t=1}^N \cos\left(\frac{2\pi i}{N} t\right) = 0 \text{ and}$$
$$\sum_{t=1}^N \sin(t\omega_i) = \sum_{t=1}^N \cos\left(\frac{2\pi i}{N} t\right) = 0$$

the following is obtained.

$$\mathbf{x}^\top\mathbf{x} = \begin{bmatrix} N & 0 & 0 \\ 0 & \frac{N}{2} & 0 \\ 0 & 0 & \frac{N}{2} \end{bmatrix}$$

and

$$\mathbf{x}^\top\mathbf{Y} = \begin{bmatrix} \sum_{t=1}^N Y_t \\ \sum_{t=1}^N (\cos t\omega_i) Y_t \\ \sum_{t=1}^N (\sin t\omega_i) Y_t \end{bmatrix}$$

I.e.

$$\hat{\theta} = \begin{bmatrix} \hat{\mu} \\ \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} \frac{1}{N}\sum_{t=1}^N Y_t \\ \frac{2}{N}\sum_{t=1}^N Y_t\cos(\omega_i t) \\ \frac{2}{N}\sum_{t=1}^N Y_t\sin(\omega_i t) \end{bmatrix}$$

Question 3.:

$$\begin{aligned} I(\omega_i) &= \sum_{t=1}^N (\hat{\alpha}\cos\omega_i t + \hat{\beta}\sin\omega_i t)^2 \\ &= \sum_{t=1}^N \left[\hat{\alpha}^2\cos^2\omega_i t + \hat{\beta}^2\sin^2\omega_i t + \hat{\alpha}\hat{\beta}\cos\omega_i t\sin\omega_i t\right] \\ &= \frac{\hat{\alpha}^2 N}{2} + \frac{\hat{\beta}^2 N}{2} + 0 \\ &= \left[\hat{\alpha}^2 + \hat{\beta}^2\right]\frac{N}{2} \end{aligned}$$

Solution 3.4

Question 1.

In figure 1 the velocity is plotted as a function of observations. By inspecting

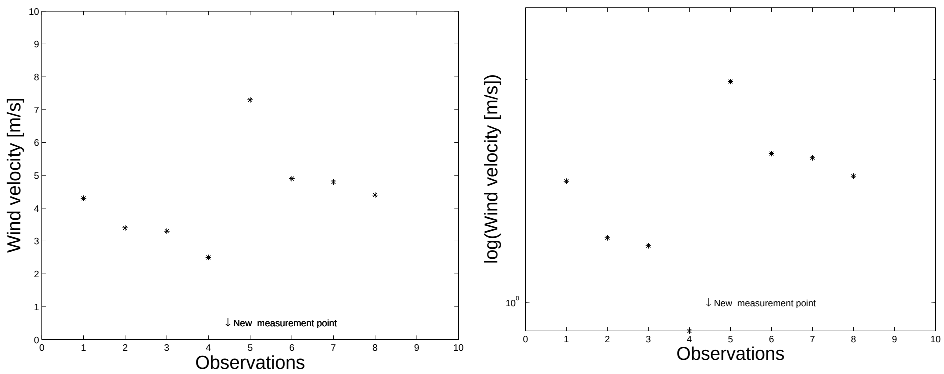

Figure 1: Left: Plot of wind velocity vs. observations. Right: Plot of log wind velocity vs. observations.

the graph it is assumed that $\log Y_t$ depends linear on t in each area (due to the few observations it is impossible to determine if the dependency is more linear in the logarithmic case). Furthermore the slope is believed to be the same in the two areas. It is obvious that having measurement from a longer period of time the model will not could be described by a linear model, but for the restricted area, which is considered, it is reasonable to use a linear model with constant parameters.

The mixed model will be suitable for describing the variations in the observed wind speed

$$\log Y_t = \theta_1 + \theta_2 t + \theta_3 \rho_t + \epsilon_t$$

where

$$\rho_t = \begin{cases} 0 & \text{for} \quad t \leq 4 \\ 1 & \text{for} \quad t \leq 5 \end{cases}$$

It is assumed that $V[\epsilon] = \sigma^2$ (i.e. constant).

Question 2.

Using the model from question 2 the observations can be written as

$$\begin{bmatrix} \log Y_1 \\ \vdots \\ \log Y_4 \\ \log Y_5 \\ \vdots \\ \log Y_8 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 4 & 0 \\ 1 & 5 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 8 & 1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_4 \\ \epsilon_5 \\ \vdots \\ \epsilon_8 \end{bmatrix}$$

or

$$\mathbf{Y}^{\star} = \mathbf{x}\theta + \epsilon$$

We find

$$\mathbf{x}^{\top}\mathbf{x} = \begin{bmatrix} 8 & 36 & 4 \\ 36 & 204 & 26 \\ 4 & 26 & 4 \end{bmatrix}, \ \mathbf{x}^{\top}\mathbf{Y}^{\star} = \begin{bmatrix} 4.9696 \\ 23.2280 \\ 2.8782 \end{bmatrix}$$

which by equation (3.35) leads to

$$\hat{\theta} = (\mathbf{x}^{\top}\mathbf{x})^{-1}\mathbf{x}^{\top}\mathbf{Y}^{\star} = \begin{bmatrix} 0.875 & -0.25 & 0.75 \\ -0.25 & 0.1 & -0.4 \\ 0.75 & -0.4 & 2.1 \end{bmatrix} \begin{bmatrix} 4.9696 \\ 23.228 \\ 2.8782 \end{bmatrix} = \underline{\underline{\begin{bmatrix} 0.700 \\ -0.0709 \\ 0.480 \end{bmatrix}}}$$

By using (3.44) we get the unbiased estimate of σ^2

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{8}(\log Y_t - \mathbf{x}_t^{\top}\hat{\theta})^2}{8-3} = \underline{\underline{0.0419^2}}$$

The difference in wind speed at the two measurement locations is determined by θ_3. Let $Y_t^{\downarrow}$ and $Y_t^{\uparrow}$ be the wind speed at the old and new measurement locations, respectively.

$$\log Y_t^{\uparrow} - \log Y_t^{\downarrow} = \theta_3$$

or

$$\frac{Y_t^{\uparrow}}{Y_t^{\downarrow}} = 10^{\hat{\theta_3}} = \underline{\underline{3.02}}$$

The estimated model indicates that the wind speed at the top of the building is approx. 3 times higher than the wind speed at the normal measurement location (2 m. above ground level).
We were able to calculate the relative difference in wind speed, as the model

were made for $\log Y$ in stead of Y. The relative value often gives more information than the absolute value. The same calculations as above with a model for Y would result in an absolute difference in wind speed on approx. 4.9.

Question 3.

The predicted wind speed in one hour ($t = 9$) at the old measurement location is given as

$$\begin{aligned}\log \hat{Y}_9 &= \mathbf{x}^\top \hat{\theta} = \begin{bmatrix} 1 & 9 & 0 \end{bmatrix} \begin{bmatrix} 0.700 \\ -0.0709 \\ 0.480 \end{bmatrix} = 0.0619 \quad \Rightarrow \\ \hat{Y}_9 &= \underline{\underline{1.15 \text{ m/s}}}\end{aligned}$$

Solution 3.5

Question 1.

The model with local constant mean:

$$Y_{n+j} = \theta_0 + \epsilon_{N+j} \qquad \text{(forgetting factor: } 0 < \lambda \leq 1\text{)}$$

I.e. $f^\top(j) = 1$. The prediction of $Y_{N+\ell}$ given the observations $Y_1, Y_2, ..., Y_N$ is given by (3.101)

$$\hat{Y}_{N+\ell|N} = f^\top(\ell)\hat{\theta}_N = \hat{\theta}_N \ ,$$

where $\hat{\theta}_N$ can be estimated by (3.99)

$$\hat{\theta}_N = F_N^{-1} h_N \ .$$

From equation (1.100) it follows that

$$F_N = \sum_{j=0}^{N-1} \lambda^j f(-j) f^\top(-j) = \sum_{j=0}^{N-1} \lambda^j = \frac{1-\lambda^N}{1-\lambda}$$

$$h_N = \sum_{j=0}^{N-1} \lambda^j f(-j) Y_{N-j} = \sum_{j=0}^{N-1} \lambda^j Y_{N-j}$$

I.e.

$$\hat{Y}_{N+\ell|N} = f^\top(\ell)\hat{\theta}_N = \frac{1-\lambda}{1-\lambda^N} \sum_{j=0}^{N-1} \lambda^j Y_{N-j}$$

Question 2.

Using L'Hospital's rule we obtain the following

$$\frac{1-\lambda}{1-\lambda^N} \rightarrow \frac{1}{N} \text{ for } \lambda \rightarrow 1 \ ,$$

and the prediction equation for $\lambda \rightarrow 1$ becomes

$$\hat{Y}_{N+\ell|N} \rightarrow \frac{1}{N} \sum_{j=0}^{N-1} Y_{N-j} = \bar{Y}$$

which is identical to (3.64).

Question 3.

The steady state value of F is according to (3.105)

$$\lim_{n\to\infty} F_{N+1} = F = \sum_{j\geq 0} \lambda^j = \frac{1}{1-\lambda}$$

From theorem 3.14 we get the updating of the parameters in the locally constant trend model at steady state

$$\hat{\theta}_{N+1} = \hat{\theta}_n + (1-\lambda)\left[Y_{N+1} - \hat{Y}_{N+1|N}\right]$$

Since $\hat{\theta}_N = \hat{Y}_{N+1|N}$ the updating of the on-step prediction is

$$\begin{aligned}\hat{Y}_{N+2|N+1} &= \hat{Y}_{N+1|N} + (1-\lambda)\left[Y_{N+1} - \hat{Y}_{N+1|N}\right]\\ &= (1-\lambda)Y_{N+1} + \lambda\hat{Y}_{N+1|N}\end{aligned}$$

which is identical to (3.74).

Solution 3.6

Question 1.

We consider a linear trend model of the form

$$Y_{N+j} = \mathbf{f}^\top(j)\theta + \epsilon_{N+j} = \mathbf{x_N}\theta + \epsilon$$

where $\mathbf{f}(j) = [1\ j]^\top$, $\theta = [\theta_0\ \theta_1]^\top$ and $\mathbf{x} = (\mathbf{f}^T(-N+1), ...\mathbf{f}^T(1), \mathbf{f}^T(0)$. V$[\epsilon] = \sigma^2\mathbf{\Sigma}$ where $\mathbf{\Sigma} = \text{diag}[1/\lambda^{N-1},...,1/\lambda,1]$ and λ is known to be $\lambda = 0.9$. Using (3.100)

$$\begin{aligned}
\mathbf{F}_7 &= \mathbf{x}_7^\top\mathbf{\Sigma}^{-1}\mathbf{x}_7 = \sum_{j=0}^{6}\lambda^j\mathbf{f}(-j)\mathbf{f}^\top(-j) = \begin{bmatrix} 5.22 & -13.47 \\ -13-47 & 55.09 \end{bmatrix} \\
\mathbf{h}_7 &= \mathbf{x}_7^\top\mathbf{\Sigma}^{-1}\mathbf{Y} = \sum_{j=0}^{6}\lambda^j\mathbf{f}(-j)\mathbf{Y}_{n-j} = \begin{bmatrix} 13.13 \\ -22.66 \end{bmatrix}
\end{aligned}$$

From (3.99)

$$\hat{\theta}_7 = \mathbf{F}_7^{-1}\mathbf{h}_7 = \begin{bmatrix} 3.949 \\ 0.554 \end{bmatrix},$$

and the prediction of Y_8 giving the 7 previous observation is

$$\hat{Y}_{8|7} = \mathbf{f}^\top(1)\hat{\theta}_7 = \begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} 3.949 \\ 0.554 \end{bmatrix} = \underline{\underline{4.50}}$$

An unbiased estimate of σ^2 is given by (3.44)

$$\hat{\sigma_7}^2 = \frac{\sum_{j=0}^{6}\lambda^j(Y_{7-j} - \mathbf{f}^\top(-j)\hat{\theta}_j)^2}{7-2},$$

where θ is not a constant as previously but depend on j, i.e.

$$\theta_i = \mathbf{F}_i^{-1}\mathbf{h}_i$$

The residual is thus found recursively and the variance of the residual is estimated to

$$\hat{\sigma_7}^2 = \underline{\underline{(0.235)^2}}$$

The variance of the prediction error is obtained from (3.102)

$$\begin{aligned}\hat{V}[e_7] &= \hat{\sigma}^2[1 + \mathbf{f}^\top(1)\mathbf{F}_7^{-1}\mathbf{f}(1)] \\ &= 0.235^2 \left[1 + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 0.520 & 0.127 \\ 0.127 & 0.049 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right] \\ &= 0.317^2\end{aligned}$$

According to (3.92) the 90% confidence interval is

$$\hat{\mathbf{Y}}_{8|7} \pm t_{0.05}(7-2) \cdot \hat{V}[e_7]^{\frac{1}{2}} = \underline{\underline{[3.864, 5.142]}}$$

Question 2.

We now assume that $y_8 = 5.0$ is given. Using theorem 3.13 we get

$$\mathbf{F}_8 = \mathbf{F}_7 + \lambda^8 \mathbf{f}(-8)\mathbf{f}\top(-8) = \begin{bmatrix} 5.64 & -16.92 \\ -16.92 & 82.64 \end{bmatrix}$$

$$\mathbf{h}_8 = \lambda \mathbf{L}^{-1}\mathbf{h}_7 + \mathbf{f}(0)Y_8 = \begin{bmatrix} 16.32 \\ -32.21 \end{bmatrix},$$

where $\mathbf{L} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. The updated parameter estimate is

$$\hat{\theta}_8 = \mathbf{F}_8^{-1}\mathbf{h}_8 = \begin{bmatrix} 4.45 \\ 0.52 \end{bmatrix},$$

and the predicted value of Y_9 given y_8 is

$$\hat{Y}_{9|8} = \mathbf{f}^\top(1)\hat{\theta}_8 = 4.97.$$

As in *Question 1* $\hat{\sigma^2}$ is estimated by recursive calculation of the residual

($\sigma^2 = 0.229^2$). The variance of the prediction error is

$$\begin{aligned}\hat{V}[e_8] &= \hat{\sigma}^2[1 + \mathbf{f}^\top(1)\mathbf{F}_8^{-1}\mathbf{f}(1)] \\ &= 0.229^2 \left[1 + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 0.478 & 0.102 \\ 0.102 & 0.035 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right] \\ &= 0.300^2\end{aligned}$$

Thus the 90% confidence interval is

$$\hat{\mathbf{Y}}_{9|8} \pm t_{0.05}(8-2) \cdot \hat{V}[e_8]^{\frac{1}{2}} = \underline{\underline{[4.762, 5.930]}}$$