Homework 3 solutions

Joe Neeman

September 22, 2010

1. Recall that the best linear predictor of Y given Z is and linear function of Z (say, $P(Y|Z) = aZ + b$) that satisfies (by the projection theorem) $\mathbb{E}P(Y|Z) = \mathbb{E}Y$ and $\mathbb{E}Z(Y - P(Y|Z)) = 0$. Therefore, $a\mathbb{E}Z + b = \mathbb{E}Y$ and

$$\begin{aligned}0 &= \mathbb{E}Z(Y - aZ - b)\\ &= \mathbb{E}YZ - a\mathbb{E}Z^2 - b\mathbb{E}Z\\ &= \mathbb{E}YZ - a\mathbb{E}Z^2 - (\mathbb{E}Y - a\mathbb{E}Z)\mathbb{E}Z\\ &= \mathbb{E}YZ - \mathbb{E}Y\mathbb{E}Z + a(\mathbb{E}Z)^2 - a\mathbb{E}Z^2\\ &= \operatorname{Cov}(Y, Z) - a\operatorname{Var}(Z).\end{aligned}$$

So $a = \operatorname{Cov}(Y, Z)/\operatorname{Var}(Z)$ and $b = \mathbb{E}Y - \operatorname{Cov}(Y, Z)\mathbb{E}Z/\operatorname{Var}(Z)$ and so

$$P(Y|Z) = \frac{\operatorname{Cov}(Y, Z)}{\operatorname{Var}(Z)}Z + \mathbb{E}Y - \frac{\operatorname{Cov}(Y, Z)}{\operatorname{Var}(Z)}\mathbb{E}Z.$$

Thus,

$$\begin{aligned}P(\alpha_1 Y_1 + \alpha_2 Y_2|Z) &= \frac{\operatorname{Cov}(\alpha_1 Y_1 + \alpha_2 Y_2, Z)}{\operatorname{Var}(Z)}Z + \mathbb{E}Y - \frac{\operatorname{Cov}(\alpha_1 Y_1 + \alpha_2 Y_2, Z)}{\operatorname{Var}(Z)}\mathbb{E}Z\\ &= \frac{\alpha_1\operatorname{Cov}(Y_1, Z) + \alpha_2\operatorname{Cov}(Y_2, Z)}{\operatorname{Var}(Z)}Z + \mathbb{E}Y - \frac{\alpha_1\operatorname{Cov}(Y_1, Z) + \alpha_2\operatorname{Cov}(Y_2, Z)}{\operatorname{Var}(Z)}\mathbb{E}Z\\ &= \alpha_1 P(Y_1|Z) + \alpha_2 P(Y_2|Z).\end{aligned}$$

2. We can assume that the mean of the process is zero and that $W_t \sim WN(0, 1)$. Then the best linear predictor of X_2 given X_1 and X_3 is a random variable of the form $Y = aX_1 + bX_3$ which satisfies $\mathbb{E}X_1(X_2 - Y) = 0$ and $\mathbb{E}X_3(X_2 - Y) = 0$. Let's write our AR(1) process as $X_t = \phi X_{t-1} + W_t$ for $\phi \neq 1$. Then the covariance function is $\gamma(h) = \phi^h/(1 - \phi^2)$ and so we can solve our conditions for the best linear predictor:

$$\begin{aligned}0 &= \mathbb{E}X_1(X_2 - aX_1 - bX_3)\\ &= \frac{\phi}{1 - \phi^2} - \frac{a}{1 - \phi^2} - \frac{\phi^2 b}{1 - \phi^2}\end{aligned}$$

and

$$\begin{aligned} 0 &= \mathbb{E}X_3(X_2 - aX_1 - bX_3) \\ &= \frac{\phi}{1-\phi^2} - \frac{a\phi^2}{1-\phi^2} - \frac{b}{1-\phi^2}. \end{aligned}$$

The first of these gives $a = \phi - \phi^2 b$, which we substitute into the second to obtain $0 = \phi - \phi^3 + \phi^4 b - b$, which gives $b = (\phi - \phi^3)/(1-\phi^4) = \phi/(1+\phi^2)$. Substituting back and solving for a gives $a = \phi/(1+\phi^2)$ also, and so

$$P(X_2|X_1, X_3) = \frac{\phi}{1+\phi^2}(X_1 + X_3).$$

3. For an AR(1) process of the form $X_t = \phi X_{t-1} + W_t$, the correlation function is $\rho(h) = \phi^h$. Let $X_{t+n}^t = \sum_{s=1}^t \alpha_{t-s+1} X_s$ be the best linear predictor of X_{t+n} given $X_1, \dots, X_t$. Then $\alpha_1, \dots, \alpha_t$ satisfy the equations

$$\begin{aligned} 0 &= \mathbb{E}X_s(X_{t+n} - X_{t+n}^t) \\ &= \mathbb{E}X_s X_{t+n} - \sum_{\tau=1}^t \alpha_{t-\tau+1}\mathbb{E}X_s X_\tau \end{aligned}$$

for $s = 1, 2, \dots, t$. Dividing both sides by $\gamma(0)$ and rearranging, we have

$$\rho(t+n-s) = \sum_{\tau=1}^t \alpha_{t-\tau+1}\rho(s-\tau)$$

and so

$$\phi^{t+n-s} = \sum_{\tau=1}^t \alpha_{t-\tau+1}\phi^{|s-\tau|}$$

for every $s = 1, \dots, t$. In matrix form,

$$\begin{pmatrix} 1 & \phi & \phi^2 & \dots & \phi^{t-1} \\ \phi & 1 & \phi & \dots & \phi^{t-2} \\ \phi^2 & \phi & 1 & \dots & \phi^{t-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi^{t-1} & \phi^{t-2} & \phi^{t-3} & \dots & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_t \end{pmatrix} = \begin{pmatrix} \phi^n \\ \vdots \\ \phi^{t+n-1} \end{pmatrix}$$

If you stare at this matrix for long enough, it becomes clear that one solution is $\alpha_1 = \phi^n$ and $\alpha_s = 0$ for $s > 1$. That is, $X_{t+n}^t = \phi^n X_t$ solves the prediction equations and so it is the best linear predictor of X_{t+n} given $X_1, \dots, X_t$.

To compute its mean-squared error,

$$\begin{aligned} \mathbb{E}(X_{t+n} - X_{t+n}^t)^2 &= \mathbb{E}(X_{t+n} - \phi^n X_t)^2 \\ &= \gamma(0) - 2\phi^n\gamma(n) + \phi^{2n}\gamma(0). \end{aligned}$$

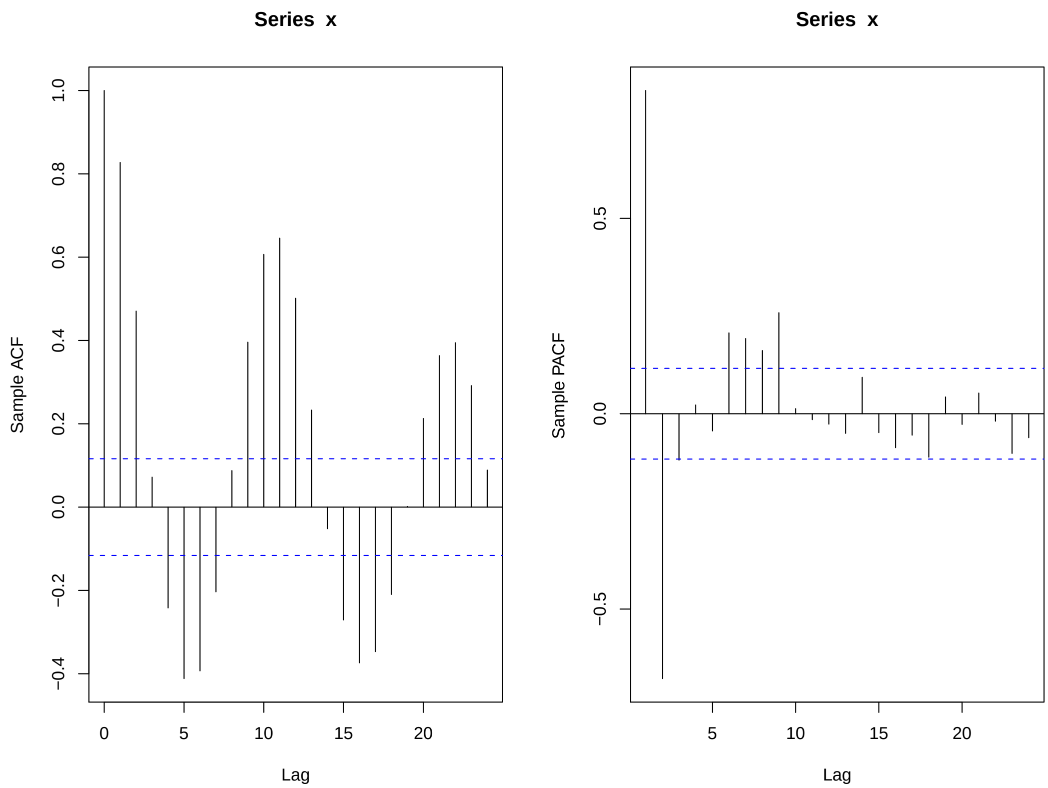

Figure 1: Sample ACF and sample PACF for the sunspot data.

But $\gamma(h) = \sigma_w^2 \phi^h/(1-\phi^2)$ for an AR(1) process, and so this is just

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

(a) The sample ACF and sample PACF are given in Figure 1. The R code that computes and plots them is:

```
z <- read.table("sunspot.dat")$V1
x <- sqrt(z)
postscript(file="stat_153_solutions3_4a.eps")
par(mfcol=c(1,2))
a <- acf(x, ylab="Sample ACF")
pa <- acf(x, type="partial", ylab="Sample PACF")
dev.off()
```

4. An MA(1) model would have a correlation function that was zero for lags of 2 or more. Similarly, an MA(2) model would have a correlation function that was zero for lags of 3 or more. Neither of these correponds to the sample ACF shown in Figure 1. An AR(1) model, on the other hand, would show a PACF that was zero for lags of 2 or more and an AR(2) model would have a PACF that was zero for lags of 3 or more. This last model looks the most likely, because the PACF is fairly large for the first two lags and then it drops off fairly substantially.

5. We estimated the parameters with the command `ar.yw(x, order = 2)`, which gave us the estimate $X_t = 1.388X_{t-1} - 0.678X_{t-2} + W_t + 6.351$, where $W_t \sim WN(0, 2.082)$.

6. We used the `predict` function to predict the next four values. After we squared the results, the prediction intervals were about $[20.3, 47.8]$, $[9.7, 52.4]$, $[6.8, 59.3]$ and $[7.6, 67.4]$.

7. The plots of the original time series, our predictions and their prediction intervals are in Figure 2.

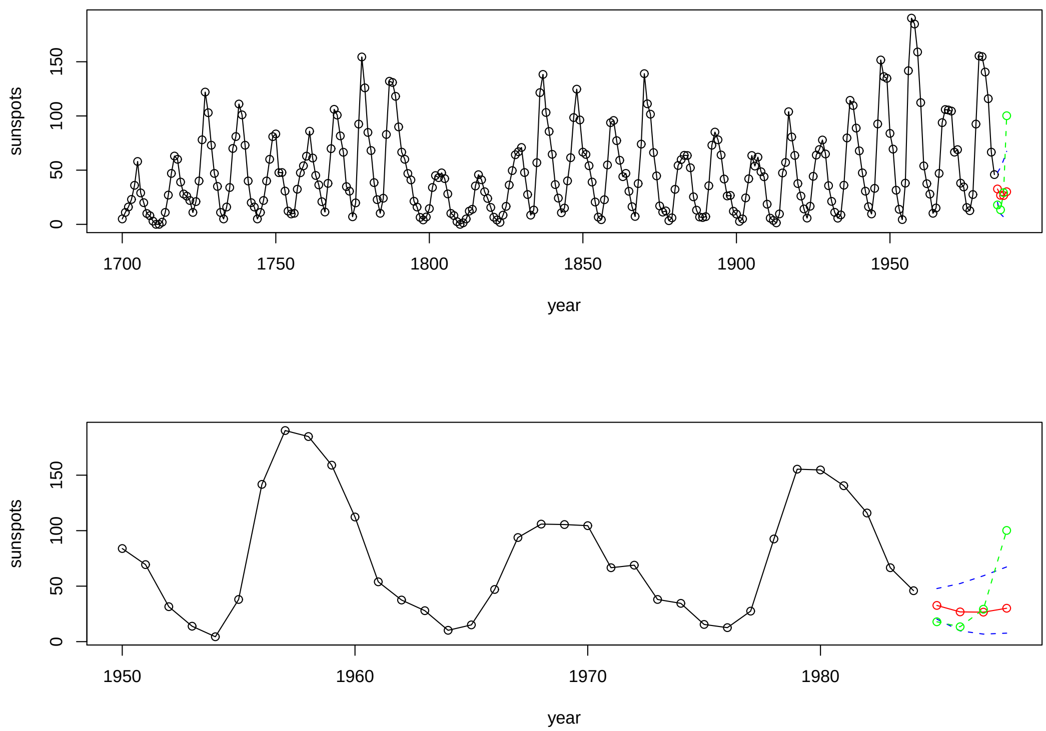

Figure 2: In black, the original time series. Our predictions for 1985–1988 are in red and the prediction intervals are in blue. The actual values for 1985-1988 are in green. The second graph zooms in on the years 1950-1988.