MPCS58020 Assignment 3

Yecheng Li

Derive expressions for the following:

The covariance and correlation of an AR(1) process

ANS:
Expression of an AR(1) process: $x_t = \varphi x_{t-1} + w_t$.

For the **covariance** between observations h time periods apart, we have:

$$x_{t-h}x_t = \varphi x_{t-h}x_{t-1} + x_{t-h}w_t$$

$$\Rightarrow E(x_{t-h}x_t) = E(\varphi x_{t-h}x_{t-1}) + E(x_{t-h}w_t)$$

$\Rightarrow$ Covariance is: $\gamma_h = \varphi\gamma_{h-1} = \varphi^h\gamma_0 = \varphi^h Var(x_t) = \varphi^h \frac{\sigma_w{}^2}{1-\varphi^2}, h \geq 0$

For the **correlation** between observations h time periods apart, we have:

$$\rho_h = \frac{\gamma_h}{Var(x_t)} = \frac{\varphi^h Var(x_t)}{Var(x_t)} = \varphi^h, h \geq 0$$

The covariance and correlation of an MA(1) process

ANS:
Expression of a MA(1) process: $x_t = w_t + \theta w_{\square-1}$

For the **covariance** between observations h time periods apart, we have:

$$x_{t-h}x_t = \theta x_{t-h}w_{t-1} + x_{t-h}w_t$$

$$\Rightarrow E(x_{t-h}x_t) = E(\theta x_{t-h}\omega_{t-1}) + E(x_{t-h}w_t)$$

$\Rightarrow$ Covariance $= E[(w_t + \theta w_{t-1})(w_{t-h} + \theta w_{t-h-1})] = E[w_t w_{t-h} + \theta w_{t-1}w_{t-h} + \theta w_t w_{t-h-1} + \theta^2 w_{t-1}w_{t-h-1}]$

As for any $i \neq j$, we have $E\left(w_i w_j\right) = 0$ by definition of independence, we have:

- When h = 0, $\gamma_h = (1+\theta^2)\sigma_w{}^2$
- When h = 1, $\gamma_h = \theta\sigma_\square{}^2$
- When h > 1, $\gamma_h = 0$

For the **correlation** between observations h time periods apart, we have:

$$\rho_h = \frac{\gamma_h}{Var(x_t)}$$

As $Var(x_t) = (1+\theta^2)\sigma_w{}^2$, we have:

- When h = 1, $\rho_h = \frac{\theta}{1+\theta^2}$
- When h > 1, $\rho_h = 0$

Show that the principal components in a PCA transformation maximize the variance in each coordinate direction (given orthogonality constraint).

ANS:

The central idea of PCA is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. That is, our goal of maximizing the total variance is equivalent to maximizing the sum of variances in each dimension. This goal is achieved by using greedy algorithm: first choose one axis (eigenvector) that maximizes the variance, then another one under the constraint that it has to be orthogonal to the first eigenvector, and so on. So we only need to show why this greedy algorithm gives a global maximum.

- For any two orthogonal vectors u and v with total variance λ_u and λ_v, the total variance in this plane is $u'\Sigma u + v'\Sigma v = \Sigma\lambda_i {u_i}^2 + \Sigma\lambda_i {v_i}^2 = \Sigma\lambda_i({u_i}^2 + {v_i}^2)$. Thus, it is a linear combination of eigenvalues λ_i with coefficients that are all positive, less or equal to 1 (because it is equal to the length of a unit vector's projection on a plane to the power of 2) and sum to 2. As both u and v are unit vectors, it is obvious that the maximum is reached at $\lambda_u+\lambda_v$. Therefore, we know that this greedy procedure gives a global maximum.

Shumway Problem 3.10

Please see Q3_10.r (Q3_10.m has some problem in it)

After run the **fit** command, we get $x_t = 11.45 + 0.4286x_{t-1} + 0.4418x_{t-1} + w_t$

After prediction, we have the following four intervals:

- x_{n+1}^n: [76.45756, 98.74217]
- x_{n+2}^n: [74.64094, 98.88604]
- x_{n+3}^n: [73.35405, 101.32022]
- x_{n+4}^n: [72.33052, 102.09648]

Shumway Problem 3.18

Please see Q3_18.m

Shumway Problem 7.12

Please see Q7_12.m

Shumway Problem 3.20

$$x_t = 0.9x_{t-1} + w_t - 0.9w_{t-1}$$

⇨ $x_t - 0.9x_{t-1} = w_t - 0.9w_{t-1}$

⇨ $(1-0.9B)x_t = (1-0.9B)w_t$ (re-write the expression)

⇨ $x_t = w_x$ (apply $\phi(B)^{-1} = (1-0.5B)^{-1}$ operator)

Therefore, there is parameter redundancy, and the model is actually not ARMA(1,1)

Shumway Problem 3.21

Please see Q3_21.R. It is observed that for the 10 repetitions, the fitted coefficients are close to the true parameters

```
> ans = matrix(0, 10, 3)
> for(i in 1:10){
+   sim = arima.sim(list(order=c(1, 0, 1), ar=0.9, ma=0.5, sd=1), n=200)
+   fit = arima(sim, order=c(1,0,1), include.mean=F)
+   ans[i, 1:2]=fit$coef # AR & MA coefficients
+   ans[i, 3]=sqrt(fit$sigma2) # sigma
+ }
> ans
           [,1]      [,2]      [,3]
 [1,] 0.9018659 0.5691605 0.9802404
 [2,] 0.9260357 0.5559993 0.9669847
 [3,] 0.9275372 0.3699781 1.0046454
 [4,] 0.8313370 0.5572480 0.9052310
 [5,] 0.9305736 0.5026613 1.0383352
 [6,] 0.8716577 0.5104406 0.9921176
 [7,] 0.8447907 0.5648052 0.9908837
 [8,] 0.9095341 0.5138434 0.9858549
 [9,] 0.9585693 0.4515495 1.0150235
[10,] 0.8816878 0.4513629 0.9641110
```

Shumway Problem 3.22

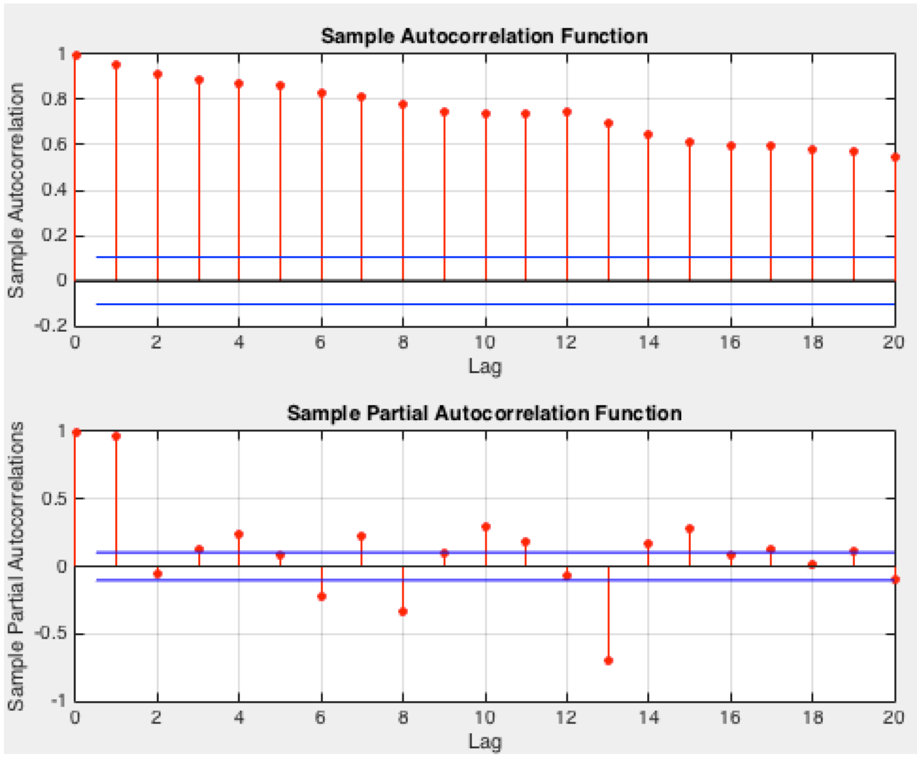
Please see Q3_22.m

Shumway Problem 3.36

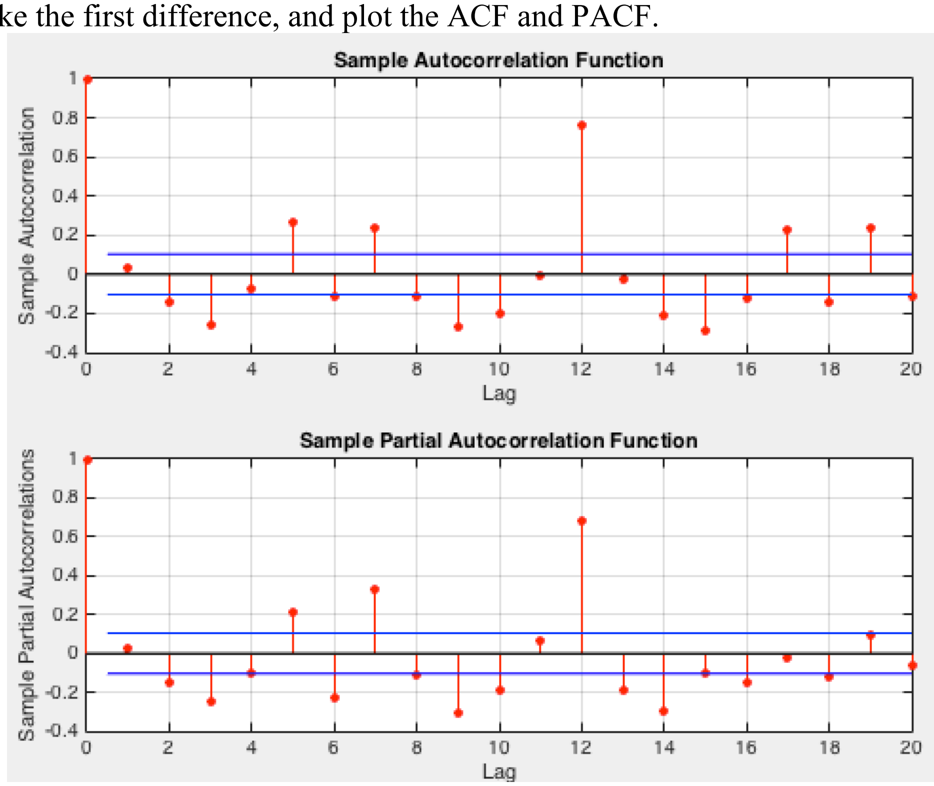
Please see Q3_36.r and Q3_36.m

Shumway Problem 3.37

Plot the ACF and PACF of the original series. It is observed that there is slow decay in the ACF and the peak at lag h=1 and h = 13 in the PACF.



Then take the first difference, and plot the ACF and PACF.



Let s=12, we see peaks at 1s, 2s, 3s, 4s with relatively slow decay, thus we take a seasonal difference of the first difference to get series Y. Below is the ACF/PACF of Y.

We see one peak at 1s on ACF, and 3 peaks at 1s, 2s, 3s on PACF. This suggests models of seasonal MA(1), seasonal AR(3), or seasonal ARMA(3,1) Inspecting ACF and PACF within season lag, PACF cuts off after lag 2, ACF tails off. This suggests that the model is AR(2).

Choose the model to be ARMA(2, 1, 0)×$(0, 1, 1)_{12}$, the coefficients we have are:

	AR1	AR2	SMA1
Non-seasonal	0.1351	0.2462	-0.6953
Seasonal	0.0513	0.0515	0.0381

Use `arima.for(data, 12, 2, 1, 0, 0, 1, 1, 12)` to get the forecasts:

Predicted: `1979 676.4664 685.1172 653.2388 585.6939 553.8813 664.4072 647.0657 611.0828 594.6414 569.3997 587.5801 581.1833`

Seasonal: `1979 21.20465 32.07710 43.70167 53.66329 62.85364 71.12881 78.73590 85.75096 92.28663 98.41329 104.19488 109.67935`