Lab 3

Pstat 174/274

Model Identification

For each of the following time series processes:

(a) $X_t = 0.7X_{t-1} - 0.1X_{t-2} + Z_t$
(b) $X_t = 0.1Z_{t-2} - 0.7Z_{t-1} + Z_t$
(c) $X_t = 0.5X_{t-1} + Z_t + 0.4Z_{t-1}$
(d) $X_t = 0.75X_{t-1} - 0.5625X_{t-2} + Z_t + 1.25Z_{t-1}$

Compute the following:

1. State whether it is an $MA(q)$, $AR(p)$, or $ARMA(p, q)$ and determine p and/or q.

   (a) AR(2)
   (b) MA(2)
   (c) ARMA(1,1)
   (d) ARMA(2,1)

2. Express the processes in terms of the back shift operator, B.

   (a) $(1 - 0.7B + 0.1B^2)X_t = Z_t$
   (b) $X_t = (1 - 0.7B + 0.1B^2)Z_t$
   (c) $(1 - 0.5B)X_t = (1 + 0.4B)Z_t$
   (d) $(1 - 0.75B + 0.5625B^2)X_t = (1 + 1.25B)Z_t$

3. Determine whether each process is causal and/or invertible.

   Causality [p.85 BD]: An ARMA(p,q) process X_t is causal, or a causal function of Z_t if:

$$\phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p \neq 0 \text{ for all } |z| \leq 1.$$

   Invertibility [p.86 BD]: An ARMA(p,q) process X_t is invertible if:

$$\theta(z) = 1 + \theta_1 z + \ldots + \theta_p z^p \neq 0 \text{ for all } |z| \leq 1.$$

   (a) The autoregressive (AR) polynomial for this process has the factorization $\phi(z) = 1 - 0.7z + 0.1z^2 = (1 - 0.5z)(1 - 0.2z)$ and is therefore zero at $z = 2$ and $z = 5$. Since these zeroes lie outside de unit circle, we conclude that X_t is a causal $AR(2)$.

```
polyroot(c(1,-0.7,0.1))
```

```
## [1] 2+0i 5-0i
```

   (b) The moving average (MA) polynomial is given by $\theta(z) = 1 - 0.7z + 0.1z^2$ with corresponding zero at $z = 2$ and $z = 5$, which is greater than 1 in absolute value so the process X_t is invertible.

```
polyroot(c(1,-0.7,0.1))
```

```
## [1] 2+0i 5-0i
```

   (c) The AR and MA polynomial are given by $\phi(z) = 1 - 0.5z$ and $\theta(z) = 1 + 0.4z$, respectively. The AR root is $z = 2$ so the model is causual; and the MA root is $z = -2.5$, which lies outside the unit circle, so the process is invertible.

```
polyroot(c(1,-0.5))
```

```
## [1] 2+0i
```

```r
polyroot(c(1,0.4))
```

```
## [1] -2.5+0i
```

(d) $(1-0.75B+0.5625B^2)X_t = (1+1.25)Z_t$. The AR polynomial is $\phi(z) = 1-0.75z+0.5625z^2$ has zeroes at $z = 2(1 \pm i\sqrt{3})/3$. which lie outside the unit circle. The process is therefore causal. On the other hand, the MA polynomial $\theta(z) = 1+1.25z$ has zero at $z = -0.8$, and hence the X_t is not invertible.

```r
polyroot(c(1,-0.75,0.5625))
```

```
## [1] 0.666667+1.154701i 0.666667-1.154701i
```

```r
polyroot(c(1,1.25))
```

```
## [1] -0.8+0i
```

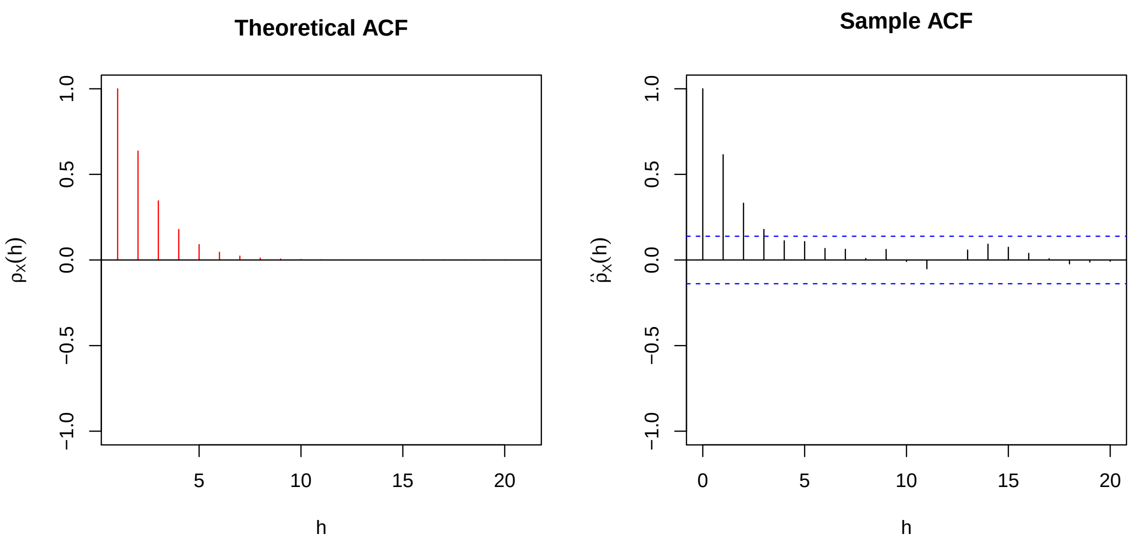
4. Simulate 200 observations from each model and compare the sample ACF and PACF with the model ACF and PACF (white noise variance is 1).

(a) AR(2)

```r
# Simulate model
set.seed(1)
ar2 <- arima.sim(model = list(ar = c(0.7,-0.1),sd = 1),n = 200)

# Theoretical ACF
?ARMAacf
theo_acf <- ARMAacf(ar = c(0.7,-0.1),lag.max = 20, pacf = FALSE)
op <- par(mfrow = c(1,2))
# Theoretical ACF
plot(theo_acf,type = "h",ylim = c(-1,1),
         main = "Theoretical ACF",
         col = "red",
         ylab = expression(rho[X](h)),
         xlab = "h")
abline(h = 0) # Add horizontal line

# Sample ACF
acf(ar2,lag.max = 20,
        main = "Sample ACF",
        ylim = c(-1,1),
        xlab = "h",
        ylab = expression(hat(rho)[X](h)))
```



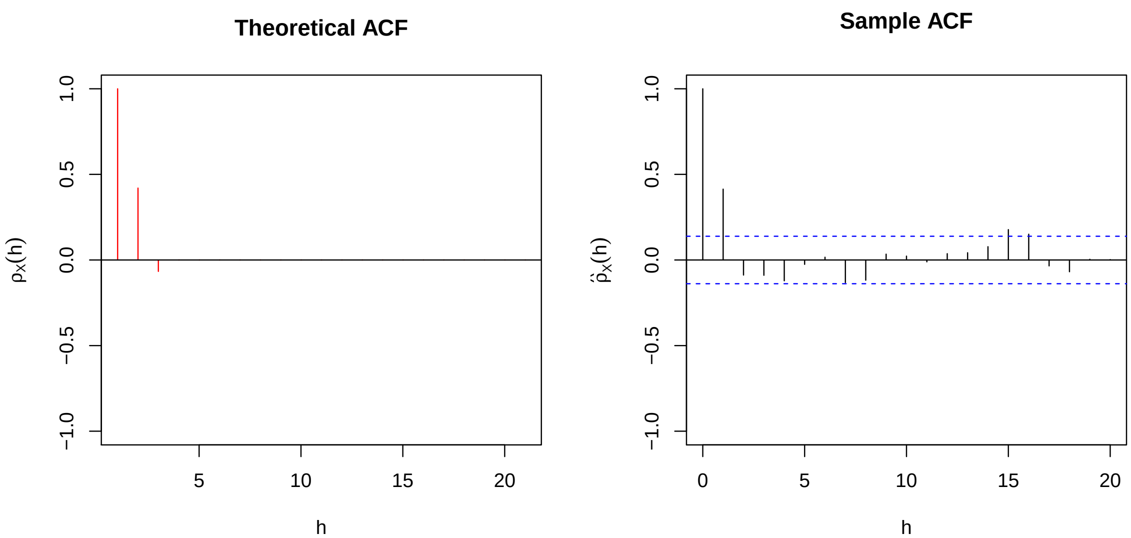

```r
par(op)
```

(b) MA(2)

```r
# Simulate model
set.seed(2)
ma1 <- arima.sim(model = list(ma = c(0.7,-0.1),sd = 1),n = 200)

# Theoretical ACF
theo_acf <- ARMAacf(ma = c(0.7,-0.1),lag.max = 20, pacf = FALSE)
# Plot
op <- par(mfrow = c(1,2))
plot(theo_acf,type = "h",ylim = c(-1,1),
         main = "Theoretical ACF",
         col = "red",
         ylab = expression(rho[X](h)),
         xlab = "h")
abline(h = 0) # Add horizontal line

# Sample ACF
acf(ma1,lag.max = 20,
        main = "Sample ACF",
        ylim = c(-1,1),
        xlab = "h",
        ylab = expression(hat(rho)[X](h)))
```



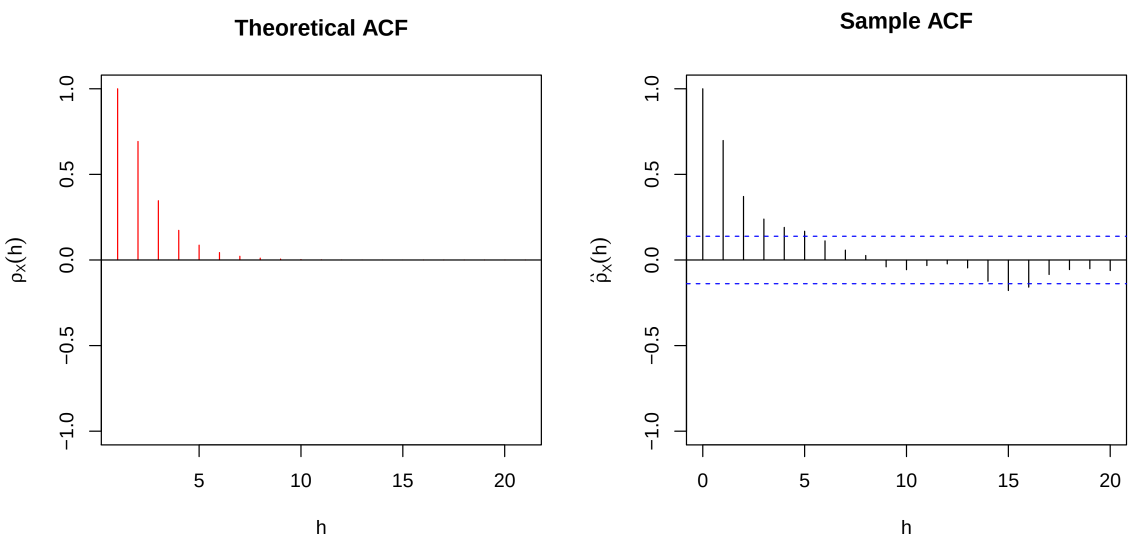

```r
par(op)
```

(c) ARMA(1,1)

```r
# Simulate model
set.seed(3)
arma11 <- arima.sim(model = list(ar = c(0.5),ma = c(0.4),sd = 1),n = 200)

# Theoretical ACF
theo_acf <- ARMAacf(ar = 0.5, ma = 0.4,lag.max = 20, pacf = FALSE)
# Plot
op <- par(mfrow = c(1,2))
plot(theo_acf,type = "h",ylim = c(-1,1),
         main = "Theoretical ACF",
         col = "red",
         ylab = expression(rho[X](h)),
         xlab = "h")
abline(h = 0) # Add horizontal line

# Sample ACF
acf(arma11,lag.max = 20,
        main = "Sample ACF",
        ylim = c(-1,1),
        xlab = "h",
        ylab = expression(hat(rho)[X](h)))
```



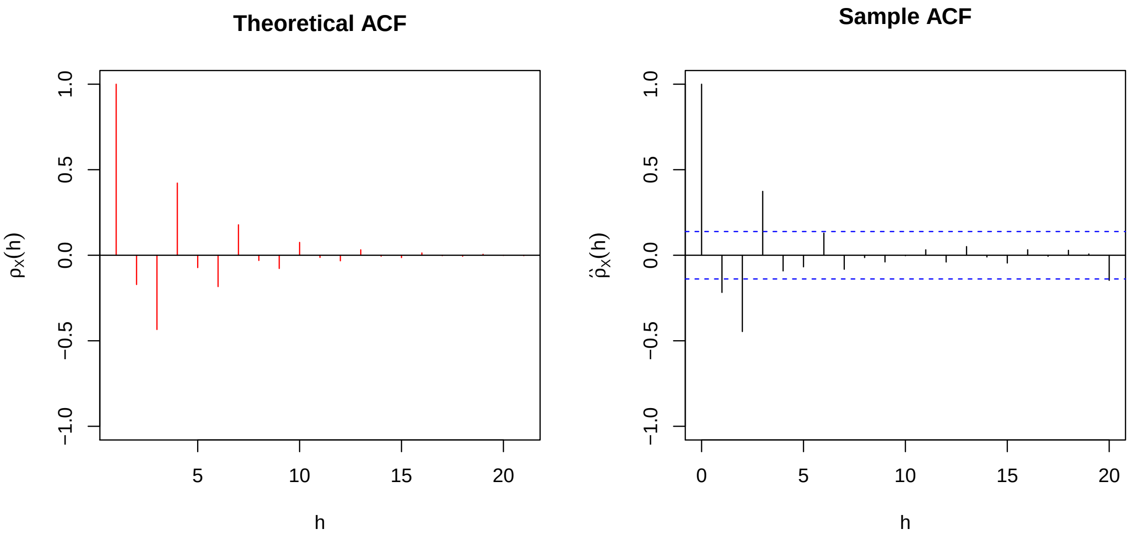

```r
par(op)
```

(d) ARMA(2,1)

```r
# Simulate model
set.seed(4)
arma21 <- arima.sim(model = list(ar = c(-0.75,-0.5625),ma = 1.25,sd = 1),n = 200)

# Theoretical ACF
theo_acf <- ARMAacf(ar = c(-0.75,-0.5625), ma = 1.25,lag.max = 20, pacf = FALSE)
# Plot
op <- par(mfrow = c(1,2))
plot(theo_acf,type = "h",ylim = c(-1,1),
         main = "Theoretical ACF",
         col = "red",
         ylab = expression(rho[X](h)),
         xlab = "h")
abline(h = 0) # Add horizontal line

# Sample ACF
acf(arma21,lag.max = 20,
        main = "Sample ACF",
        ylim = c(-1,1),
        xlab = "h",
        ylab = expression(hat(rho)[X](h)))
```




```r
par(op)
```

Model Estimation

Assume we have n observations $x_1, \ldots, x_n$ from a (causual and invertible) $ARMA(p,q)$ process in which, *initially*, the order parameters, p and q, are known. Our goal is to estimate the parameters, $\phi_1, \ldots, \phi_p$ and $\theta_1, \ldots, \theta_q$.

Autoregressive Model

Consider an autorregressive model of order p, i.e., $AR(p)$:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t, \quad Z_t \sim WN(0, \sigma_Z^2)$$

In order to estimate $\phi_1, \ldots, \phi_p$ and σ_Z^2 we use the Yule-Walker Equations:

$$\begin{aligned} \rho_X(h) &= \phi_1 \rho_X(h-1) + \cdots + \phi_p \rho_X(h-p), \text{ for } h = 1, 2, \ldots, p, \\ \sigma_X^2 &= \frac{\sigma_Z^2}{1 - \phi_1 \rho_X(1) - \cdots - \phi_p \rho_X(p)} \end{aligned}$$

In matrix notation the **Yule-Walker** equations are:

$$\underbrace{\begin{bmatrix} \rho_X(1) \\ \rho_X(2) \\ \vdots \\ \rho_X(p-1) \\ \rho_X(p) \end{bmatrix}}_{\boldsymbol{\rho}_p} = \underbrace{\begin{bmatrix} \rho_X(0) & \rho_X(1) & \ldots & \rho_X(p-1) \\ \rho_X(1) & \rho_X(2) & \ldots & \rho_X(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_X(p-2) & \rho_X(p-1) & \ldots & \rho_X(1) \\ \rho_X(p-1) & \rho_X(p-2) & \ldots & \rho_X(0) \end{bmatrix}}_{\boldsymbol{R}_p} \underbrace{\begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{p-1} \\ \phi_p \end{bmatrix}}_{\boldsymbol{\phi}}$$

hence, estimates of $\phi_1, \ldots, \phi_p$ can be obtained from solving the previous equation as:

$$\boldsymbol{\phi} = \boldsymbol{R}_p^{-1} \boldsymbol{\rho}_p$$

and the estimator for the noise variance σ_Z^2 is given by:

$$\sigma_Z^2 = \gamma_X(0)[1 - \boldsymbol{\rho}_p' \boldsymbol{R}_p^{-1} \boldsymbol{\rho}_p]$$

Using the method of moments in the above equations, we replace $\rho_X(h)$ by its sample estimators $\hat{\rho}_X(h)$ (the sample autocorrelation) and obtain the estimators for $\hat{\boldsymbol{\phi}} = (\hat{\phi}_1, \ldots, \hat{\phi}_p)$ and $\hat{\sigma}_Z^2$.

Example

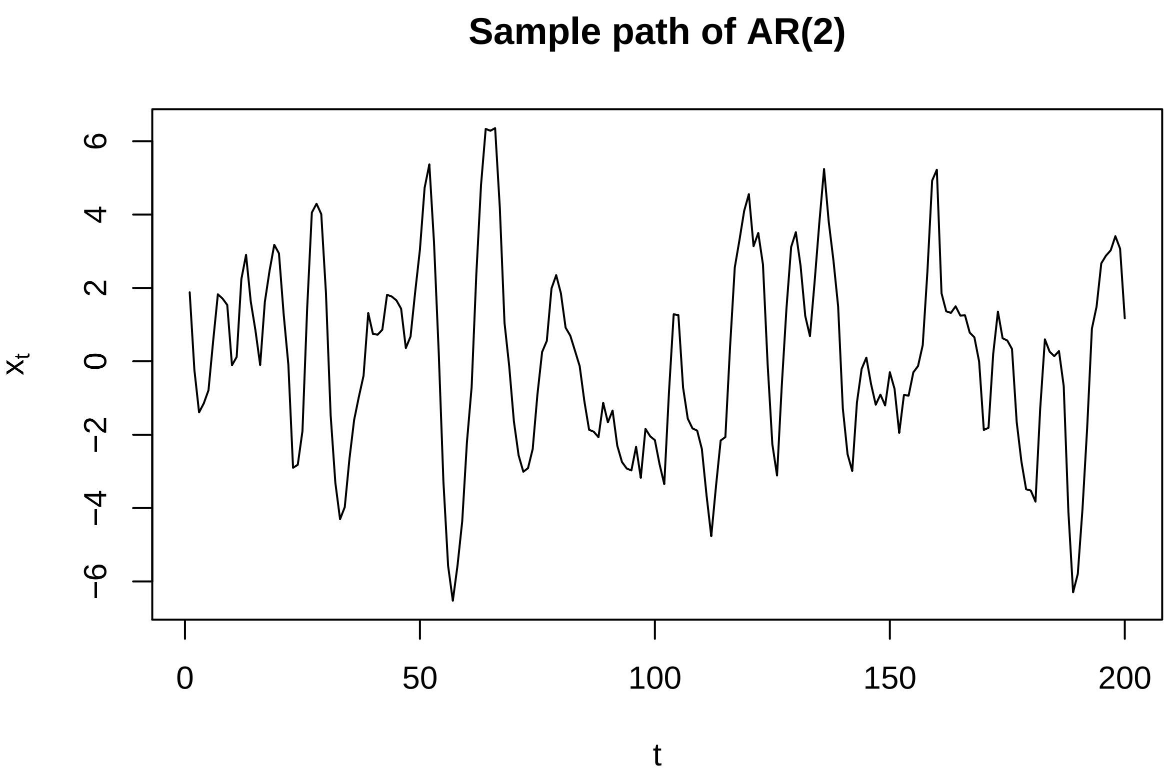
Consider the AR(2) process below:

$$X_t = 1.5X_{t-1} - 0.75X_{t-2} + Z_t \text{ with } Z_t \overset{\text{iid}}{\sim} N(0,1)$$

Suppose that $n = 200$ realization of such process are observed,

```
# Simulate Model
set.seed(1)
ar2 <- arima.sim(model = list(ar = c(1.5,-0.75), sd = 1),n = 200)

# Plot
plot(ar2, main  = "Sample path of AR(2)", xlab = "t", ylab = expression(x[t]))
```



We now manually construct the Yule-Walker estimates. To do so, we use $p = 2$ (as per the ACF/PACF above) estimate the sample autocorrelation function $\rho_X(h)$ and construct the matrix $\hat{R}^p$ utilizing the `toeplitz` function.

```
# Estimation with Yul-Walker eqns
acv_ar <- acf(ar2,type = "covariance",main = "Sample ACF",plot = F)
?toeplitz
Rho <- toeplitz(acv_ar$acf[c(1,2)]/acv_ar$acf[1])
rho <- acv_ar$acf[c(2,3)]/acv_ar$acf[1]
```

```r
phi_hat <- solve(Rho) %*% rho
phi_hat
```

```
##            [,1]
## [1,]  1.4189241
## [2,] -0.6492239
```

```r
# Estimate of noise variance
sigma_z <- acv_ar$acf[1]*(1-t(rho)%*%solve(Rho)%*%rho)
sigma_z
```

```
##          [,1]
## [1,] 1.004926
```

We now compare the latter estimates with the ones obtained using the pre-installed function `ar.yw`:

```r
yw <- ar.yw(ar2,order = 2)
yw$x.mean # mean estimate
```

```
## [1] 0.06918922
```

```r
yw$ar # Parameter estimates
```

```
## [1]  1.4189241 -0.6492239
```

```r
yw$var.pred # Error variance
```

```
## [1] 1.020229
```