Tutorial 8 - STAT 485/685

Trevor Thomson

Department of Statistics & Actuarial Science
Simon Fraser University, BC, Canada

November 9, 2020

Today's Plan

1. Chapter 4 Review
   - General Linear Processes
   - Moving Average Processes
   - Autoregressive Processes
   - Mixed Autoregressive and Moving Average Processes
   - Example: Assignment 4 Question 3
2. Chapter 5 Review
   - Stationarity Through Differencing
   - ARIMA Models
   - Constant Terms in ARIMA Models
   - Other Transformations
   - Example: Assignment 5 Question 2
3. Chapter 6 Review
   - Properties of the Sample Autocorrelation Function
   - The Partial and Extended Autocorrelation Function
   - Example: Assignment 6 Question 1

Definition: $\{Y_t : t \in \mathcal{I}\}$ is a *general linear process* if it can be written as a (weighted) linear combination of present and past white noise terms:

$$\begin{aligned} Y_t &= e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \psi_3 e_{t-3} + \cdots \\ &= \sum_{j=0}^{\infty} \psi_j e_{t-j}, \end{aligned}$$

where $\psi_0 \equiv 1$.

- Also referred to as ψ-weight representation of $\{Y_t : t \in \mathcal{I}\}$.

In order for the infinite sum above to be *convergent*, we assume that

$$\sum_{j=1}^{\infty} \psi_j^2 < \infty.$$

What is $E(Y_t)$, $Var(Y_t)$, $Cov(Y_t, Y_{t-k})$, and $Corr(Y_t, Y_{t-k})$, for $k \geq 0$?

$$E(Y_t) = E\left(\sum_{j=0}^{\infty} \psi_j e_{t-j}\right) = \sum_{j=0}^{\infty} \psi_j \underbrace{E(e_{t-j})}_{0} = 0.$$

$$Var(Y_t) = Var\left(\sum_{j=0}^{\infty} \psi_j e_{t-j}\right) = \sum_{j=0}^{\infty} \psi_j^2 \underbrace{Var(e_{t-j})}_{\sigma_e^2} = \sigma_e^2 \underbrace{\sum_{j=0}^{\infty} \psi_j^2}_{<\infty} < \infty.$$

For $k \geq 0$:

$$\begin{aligned}
Cov(Y_t, Y_{t-k}) &= Cov\left(\sum_{j=0}^{\infty} \psi_j e_{t-j}, \sum_{\ell=0}^{\infty} \psi_\ell e_{(t-k)-\ell}\right) \\
&= \sum_{j=0}^{\infty} \sum_{\ell=0}^{\infty} \psi_j \psi_\ell Cov(e_{t-j}, e_{t-(k+\ell)}) \\
&= \sum_{\ell=0}^{\infty} \psi_{k+\ell} \psi_\ell \underbrace{Cov(e_{t-(k+\ell)}, e_{t-(k+\ell)})}_{\sigma_e^2} + \sum_{\substack{j=0 \\ j \neq k+\ell}}^{\infty} \sum_{\ell=0}^{\infty} \psi_j \psi_\ell \underbrace{Cov(e_{t-j}, e_{t-(k+\ell)})}_{0} \\
&= \sigma_e^2 \sum_{\ell=0}^{\infty} \psi_{k+\ell} \psi_\ell.
\end{aligned}$$

For $k \geq 0$:

$$
\begin{aligned}
Corr(Y_t, Y_{t-k}) &= \frac{Cov(Y_t, Y_{t-k})}{\sqrt{Var(Y_t)}\sqrt{Var(Y_{t-k})}} \\
&= \frac{\sigma_e^2 \sum_{\ell=0}^{\infty} \psi_{k+\ell}\psi_\ell}{\sqrt{\sigma_e^2 \sum_{j=0}^{\infty} \psi_j^2} \times \sqrt{\sigma_e^2 \sum_{j=0}^{\infty} \psi_j^2}} \\
&= \frac{\sum_{\ell=0}^{\infty} \psi_{k+\ell}\psi_\ell}{\sum_{j=0}^{\infty} \psi_j^2}.
\end{aligned}
$$

Remark: $E(Y_t)$ and $Cov(Y_t, Y_{t-k})$ do not depend on time t

$\Rightarrow \{Y_t : t \in \mathcal{I}\}$ is stationary.

Definition: $\{Y_t : t \in \mathcal{I}\}$ is a moving average of order q if q of the ψ_j's of the general linear process are non-zero

$$\begin{aligned} Y_t &= \sum_{j=0}^{\infty} \psi_j e_{t-j} \\ &= \sum_{j=0}^{q} \psi_j e_{t-j}, \quad \text{with } \psi_0 = 1, \\ &= e_t + \psi_1 e_{t-1} + \cdots + \psi_q e_{t-q}. \end{aligned}$$

If so, we say that $\{Y_t : t \in \mathcal{I}\}$ is an $MA(q)$ process.

People often express $MA(q)$ processes by slightly changing the notation $\Rightarrow$ letting $\theta_j = -\psi_j$:

$$\begin{aligned} Y_t &= -\sum_{j=0}^{q} \theta_j e_{t-j}, \quad \text{with } \theta_0 = -1, \\ &= e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}. \end{aligned}$$

What is $E(Y_t)$, $Var(Y_t)$, $Cov(Y_t, Y_{t-k})$, and $Corr(Y_t, Y_{t-k})$, for $k \geq 0$?

By expressing the $MA(q)$ process in terms of the general linear process (from earlier)

$$Y_t = \sum_{j=0}^{\infty} \psi_j e_{t-j} = -\sum_{j=0}^{q} \theta_j e_{t-j},$$

where

$$\psi_\ell = \begin{cases} -\theta_\ell & \text{if } \ell \leq q \\ 0 & \text{if } \ell > q \end{cases}.$$

$\Rightarrow$ We can use results from general linear processes to help us!

$$E(Y_t) = E\left(\sum_{j=0}^{\infty} \psi_j e_{t-j}\right) = E\left(-\sum_{j=0}^{q} \theta_j e_{t-j}\right) = -\sum_{j=0}^{q} \theta_j \underbrace{E(e_{t-j})}_{0} = 0.$$

$$\begin{aligned} Var(Y_t) &= \sigma_e^2 \sum_{j=0}^{\infty} \psi_j^2 \qquad \text{(from earlier)} \\ &= \sigma_e^2 \sum_{j=0}^{q} (-\theta_j)^2 \\ &= \sigma_e^2 \sum_{j=0}^{q} \theta_j^2. \end{aligned}$$

$$
\begin{aligned}
Cov(Y_t, Y_{t-k}) &= \sigma_e^2 \sum_{\ell=0}^{\infty} \psi_{k+\ell}\psi_\ell \qquad \text{(from earlier)} \\
&= \sigma_e^2 \sum_{\ell=0}^{q-k} \psi_{k+\ell}\psi_\ell
\end{aligned}
$$

We consider cases:

- **Case 1**: $q - k < 0 \Rightarrow k > q \Rightarrow Cov(Y_t, Y_{t-k}) = 0$
- **Case 2**: $q - k \geq 0 \Rightarrow k \leq q$

$$
\begin{aligned}
Cov(Y_t, Y_{t-k}) &= \sigma_e^2 \sum_{\ell=0}^{q-k} (-\theta_{k+\ell})(-\theta_\ell) \\
&= \sigma_e^2 \sum_{\ell=0}^{q-k} \theta_{k+\ell}\theta_\ell
\end{aligned}
$$

Therefore,

$$
Cov(Y_t, Y_{t-k}) = \begin{cases} \sigma_e^2 \sum_{\ell=0}^{q-k} \theta_{k+\ell}\theta_\ell & \text{if } k = 0, 1, \cdots, q \\ 0 & \text{if } k > q \end{cases}.
$$

$$Corr(Y_t, Y_{t-k}) = \frac{Cov(Y_t, Y_{t-k})}{\sqrt{Var(Y_t)}\sqrt{Var(Y_{t-k})}}.$$

We consider cases:

- **Case 1**: $q - k < 0 \Rightarrow k > q \Rightarrow Corr(Y_t, Y_{t-k}) = 0$
- **Case 2**: $q - k \geq 0 \Rightarrow k \leq q$

$$\begin{aligned} Corr(Y_t, Y_{t-k}) &= \frac{\sigma_e^2 \sum\limits_{\ell=0}^{q-k} \theta_{k+\ell}\theta_\ell}{\sqrt{\sigma_e^2 \sum\limits_{j=0}^{q} \theta_j^2} \times \sqrt{\sigma_e^2 \sum\limits_{j=0}^{q} \theta_j^2}} \\ &= \frac{\sum\limits_{\ell=0}^{q-k} \theta_{k+\ell}\theta_\ell}{\sum\limits_{j=0}^{q} \theta_j^2}. \end{aligned}$$

Therefore,

$$Corr(Y_t, Y_{t-k}) = \begin{cases} \dfrac{\sum\limits_{\ell=0}^{q-k} \theta_{k+\ell}\theta_\ell}{\sum\limits_{j=0}^{q} \theta_j^2} & \text{if } k = 0, 1, \cdots, q \\ 0 & \text{if } k > q \end{cases}.$$

Remark: The ACF *cuts off* after lag q.

- **Definition**: $\{Y_t : t \in \mathcal{I}\}$ is an autoregressive process of order p if the time series $\{Y_t : t \in \mathcal{I}\}$ satisfies the following equation

$$\begin{aligned} Y_t &= \sum_{j=1}^{p} \phi_j Y_{t-j} + e_t \\ &= \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t. \end{aligned}$$

If so, we say that $\{Y_t : t \in \mathcal{I}\}$ is an $AR(p)$ process.

Rather than writing an $AR(p)$ process in terms of Y_t, we can rearrange for e_t:

$$e_t = \phi(B) Y_t,$$

where

$$\begin{aligned} \phi(B) &= 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p, \\ B^k Y_t &= Y_{t-k}. \end{aligned}$$

$\Rightarrow$ B is the *backshift operator*
$\Rightarrow$ $\phi(B)$ is the *autoregressive process characteristic polynomial*.

- **Result**: An $AR(p)$ process is stationary if the roots of $\phi(B)$ lie outside the “unit circle” in $\mathbb{R}^p$-space.

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t$$

$$e_t = \underbrace{[1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p]}_{\phi(B)} Y_t$$

$\Rightarrow$ Setting $\phi(B) = 0 \cdots$
$\Rightarrow \cdots$ Turns out we need

$$\begin{aligned} \phi_1 + \phi_2 + \cdots + \phi_p &< 1, \\ |\phi_p| &< 1 \end{aligned}$$

Examples:

$p = 1$: Stationarity condition: $|\phi_1| < 1$
$p = 2$: Stationarity conditions:

$$\phi_1 + \phi_2 < 1 \qquad \phi_2 - \phi_1 < 1 \qquad |\phi_2| < 1$$

To obtain γ_k and ρ_k, solve the Yule-Walker equations

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 + \phi_3 \rho_2 + \cdots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \phi_3 \rho_1 + \cdots + \phi_p \rho_{p-2} \\ &\vdots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \phi_3 \rho_{p-3} + \cdots + \phi_p. \end{aligned}$$

Definition: $\{Y_t : t \in \mathcal{I}\}$ is a mixed autoregressive moving average process of orders p and q, respectively, if the time series $\{Y_t : t \in \mathcal{I}\}$ satisfies the following equation

$$
\begin{aligned}
Y_t &= [\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t] + [-\theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}] \\
&= \sum_{j=1}^{p} \phi_j Y_{t-j} - \sum_{j=0}^{q} \theta_j e_{t-j},
\end{aligned}
$$

with $\theta_0 \equiv -1$. If so, we say that $\{Y_t : t \in \mathcal{I}\}$ is an $ARMA(p, q)$ process.

The $ARMA(p, q)$ characteristic polynomial is

$$
\begin{aligned}
&\theta(B)e_t = \phi(B)Y_t, \\
&\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \quad (MA(q) \text{ characteristic polynomial}) \\
&\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \quad (AR(p) \text{ characteristic polynomial})
\end{aligned}
$$

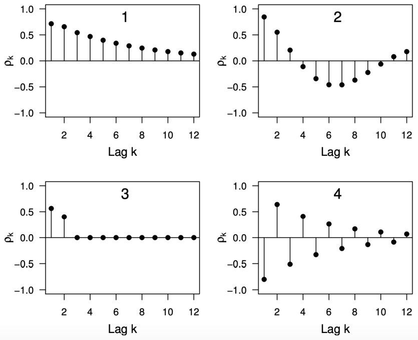
It is assumed that $\theta(B)$ and $\phi(B)$ have no common factors.

- **Result**: An $ARMA(p, q)$ process is stationary if the roots of $\phi(B)$ lie outside the “unit circle” in $\mathbb{R}^p$-space.

$\Rightarrow$ Solve the Yule-Walker equations to derive γ_k and ρ_k (see Appendix C, page 85).

Example: Assignment 4 Question 3

For each of the models below, choose one of the four plots, 1-4, that best represents what you think ρ_k will look like. Explain your reasoning.



(a) $AR(2)$, with $\phi_1 = 1.6$ and $\phi_2 = -0.8$

(b) $MA(2)$, with $\theta_1 = -0.7$ and $\theta_2 = -0.99$

(c) $AR(2)$, with $\phi_1 = 0.5$ and $\phi_2 = 0.3$

(d) $AR(1)$, with $\phi = -0.8$

Recall that for

- an $AR(1)$ process,

$$\rho_k = \phi^k.$$

- an $AR(2)$ process,

$$\begin{aligned}
\rho_1 &= \frac{\phi_1}{1-\phi_2}, \\
\rho_2 &= \frac{\phi_2(1-\phi_2)+\phi_1^2}{1-\phi_2}, \\
\rho_k &= \phi_1\rho_{k-1} + \phi_2\rho_{k-2}, \quad \text{for } k > 2
\end{aligned}$$

- an $MA(2)$ process,

$$\rho_k = \begin{cases} 1 & \text{if } k = 0 \\ \frac{-\theta_1+\theta_1\theta_2}{1+\theta_1^2+\theta_2^2} & \text{if } k = 1 \\ \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} & \text{if } k = 2 \\ 0 & \text{otherwise} \end{cases}.$$

$\Rightarrow$ See `Week10_Tutorial.R`

Question: If a time series $\{Y_t : t \in \mathcal{I}\}$ is not stationary, can we find a stationary time series $\{W_t : t \in \mathcal{I}\}$, such that W_t is derived from $\{Y_t : t \in \mathcal{I}\}$?

- **Approach 1**: Define $W_t = \nabla^d Y_t = \nabla(\nabla^{d-1} Y_t)$, where $d = 1$ or $d = 2$.
- **Approach 2**: Define $W_t = f(Y_t)$, for some function $f(.)$.

Question: If a time series $\{Y_t : t \in \mathcal{I}\}$ is not stationary, can we find a stationary time series $\{W_t : t \in \mathcal{I}\}$, such that W_t is derived from $\{Y_t : t \in \mathcal{I}\}$?

- **Approach 1**: Define $W_t = \nabla^d Y_t = \nabla(\nabla^{d-1} Y_t)$, where $d = 1$ or $d = 2$.
- **Approach 2**: Define $W_t = f(Y_t)$, for some function $f(.)$.

For **Approach 1**, we showed that if

- $Y_t = \beta_0 + \beta_1 t + X_t$, where $\{X_t : t \in \mathcal{I}\}$ is a zero-mean stationary series with autocovariance function γ_k, and β_0 and β_1 are non-zero constants,

  $\Rightarrow \{W_t : t \in \mathcal{I}\}$ is stationary, where $W_t = \nabla Y_t$.

- $Y_t = Y_{t-1} + e_t$, where $\{e_t : t \in \mathcal{I}\}$ is white noise,

  $\Rightarrow \{W_t : t \in \mathcal{I}\}$ is stationary, where $W_t = \nabla Y_t$.

Definition: $\{Y_t : t \in \mathcal{I}\}$ is an integrated autoregressive moving average model if the dth difference $W_t = \nabla^d Y_t$ is a stationary $ARMA(p,q)$. That is, we can construct $\{W_t : t \in \mathcal{I}\}$, where

$$\begin{aligned} W_t &= [\phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} + e_t] + [-\theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}] \\ &= \sum_{j=1}^{p} \phi_j W_{t-j} - \sum_{j=0}^{q} \theta_j e_{t-j}. \end{aligned}$$

If so, we say that $\{Y_t : t \in \mathcal{I}\}$ is an $ARIMA(p,d,q)$ process.

Note: For practical purposes, we only allow for $d \in \{0,1,2\}$.

We can then apply the models from Chapter 4 with $\{W_t : t \in \mathcal{I}\}$.

$\Rightarrow$ Use the fact that $W_t = \nabla^d Y_t$ to then apply the model to $\{Y_t : t \in \mathcal{I}\}$.

Special Cases:

- $ARIMA(0,d,q) \Rightarrow IMA(d,q)$
- $ARIMA(p,d,0) \Rightarrow ARI(p,d)$
- $ARIMA(p,0,q) \Rightarrow ARMA(p,q)$

Constant Terms in ARIMA Models

If $\{W_t : t \in \mathcal{I}\}$ is an $ARMA(p,q)$ process, let

$$W_t^* = W_t + c$$

$\Rightarrow E(W_t^*) = c$

$\Rightarrow Cov(W_t^*, W_{t-k}^*) = Cov(W_t, W_{t-k})$.

$$W_t = \sum_{j=1}^{p} \phi_j W_{t-j} - \sum_{j=0}^{q} \theta_j e_{t-j}$$

$$W_t^* = \theta_0 + \sum_{j=1}^{p} \phi_j W_{t-j}^* - \sum_{j=0}^{q} \theta_j e_{t-j}$$

We see that

$$\theta_0 = c - \sum_{j=1}^{p} c\phi_j$$

$$c = \frac{\theta_0}{1 - \sum_{j=1}^{p} \phi_j}$$

$\Rightarrow$ if we include an intercept term in an $ARMA(p,q)$ model, we can model stationary processes with non-zero means.

- **Approach 1**: Define $W_t = \nabla^d Y_t = \nabla(\nabla^{d-1} Y_t)$, where $d = 1$ or $d = 2$.
- **Approach 2**: Define $W_t = f(Y_t)$, for some function $f(.)$.

If we want to transform our data, how to choose $f(.)$?

Box-Cox Power Transformations: For a given value of λ and for $Y_t > 0$ for all $t \in \mathcal{I}$, a *power transformation* with parameter λ is defined by

$$g(x) = \begin{cases} \frac{x^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \log x & \text{if } \lambda = 0 \end{cases}.$$

We see that if

- $\lambda = 0 \Rightarrow$ logarithm transformation
- $\lambda = \frac{1}{2} \Rightarrow$ square-root transformation
- $\lambda = -1 \Rightarrow$ inverse transformation
- $\lambda = 1 \Rightarrow$ no transformation

$\Rightarrow$ use an estimate $\hat{\lambda}$ to help us specify $f(x)$

In R: `BoxCox.ar`

- Computes a log-likelihood function for a grid of λ-values based on a normal likelihood function.
- Generates a 95% confidence interval for λ, where the centre is $\hat{\lambda}$.
- Use the 95% confidence interval to guide us in selecting a proper λ.

Using the techniques we practiced in Video 20, write out the differenced equation form of Y_t for each of the following models. Show all your work.

(a) $IMA(1,2)$

(b) $ARI(1,2)$

(c) $ARIMA(0,1,2)$

Using the techniques we practiced in Video 20, write out the differenced equation form of Y_t for each of the following models. Show all your work.

(a) $IMA(1, 2)$

(b) $ARI(1, 2)$

(c) $ARIMA(0, 1, 2)$

Let $\{Y_t : t \in \mathcal{I}\}$ be an $ARIMA(p, d, q)$ process.

(a) We have $W_t = Y_t - Y_{t-1}$, where $\{W_t : t \in \mathcal{I}\}$ is an $MA(2)$ process. Then we can write W_t as

$$\begin{aligned} W_t &= e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \\ (Y_t - Y_{t-1}) &= e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \\ Y_t &= Y_{t-1} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}. \end{aligned}$$

Note: $\{Y_t : t \in \mathcal{I}\}$ looks like an $ARMA(1, 2)$ process, but we see that it is not stationary since $\phi = 1$.

Using the techniques we practiced in Video 20, write out the differenced equation form of Y_t for each of the following models. Show all your work.

(a) $IMA(1,2)$

(b) $ARI(1,2)$

(c) $ARIMA(0,1,2)$

Let $\{Y_t : t \in \mathcal{I}\}$ be an $ARIMA(p,d,q)$ process.

(b) We have $W_t = Y_t - 2Y_{t-1} + Y_{t-2}$, where $\{W_t : t \in \mathcal{I}\}$ is an $AR(1)$ process. Then we can write W_t as

$$\begin{aligned} W_t &= \phi W_{t-1} + e_t \\ (Y_t - 2Y_{t-1} + Y_{t-2}) &= \phi(Y_{t-1} - 2Y_{t-2} + Y_{t-3}) + e_t \\ Y_t &= (2+\phi)Y_{t-1} + (-1-2\phi)Y_{t-2} + \phi Y_{t-3} + e_t. \end{aligned}$$

Note: $\{Y_t : t \in \mathcal{I}\}$ looks like an $AR(3)$ process, but we see that it is not stationary since $(2+\phi) + (-1-2\phi) + \phi = 1$.

(c) Since an $ARIMA(0,1,2)$ process is an $IMA(1,2)$ process, we have from part (a) that

$$Y_t = Y_{t-1} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}.$$

Properties of the Sample Autocorrelation Function

For a *stationary* time series $\{Y_t : t \in \mathcal{I}\}$, where $\mathcal{I} = \{1, 2, \cdots, n\}$, recall the autocorrelation function (ACF) for lag $k \geq 0$:

$$\rho_k = Corr(Y_t, Y_{t-k}) = \frac{Cov(Y_t, Y_{t-k})}{\sqrt{Var(Y_t)Var(Y_{t-k})}} = \frac{\gamma_k}{\gamma_0}.$$

$\Rightarrow$ Estimate ρ_k with

$$r_k = \hat{\rho}_k = \widehat{Corr}(Y_t, Y_{t-k}) = \frac{\sum_{t=k+1}^{n} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n} (Y_t - \bar{Y})^2}$$

where $\bar{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t$ is an estimate for $E(Y_t) = E(Y_{t-k}) = \mu$.

Turns out that if $\{Y_t : t \in \mathcal{I}\}$ is a stationary process:

$$r_k \xrightarrow{d} \mathcal{N}\left(\rho_k, \frac{c_{kk}}{n}\right), \quad \text{as } n \to \infty,$$

where

$$c_{kk} = \sum_{i=-\infty}^{\infty} (\rho_{i+k}^2 + \rho_{i-k}\rho_{i+k} - 4\rho_k\rho_i\rho_{i+k} + 2\rho_k^2\rho_i^2).$$

A useful hypothesis test for each $k > 0$:

$$H_0 : \rho_k = 0$$
$$H_a : \rho_k \neq 0.$$

$\Rightarrow$ This can be done by computing for each $k > 0$ an approximate 95% confidence interval for ρ_k:

$$r_k \pm \underbrace{2}_{\approx 1.96} \widehat{SE}(r_k),$$

where $\widehat{SE}(r_k)$ is an estimate of $SE(r_k) = \sqrt{Var(r_k)}$.

- **Example 1**: If $\{Y_t : t \in \mathcal{I}\}$ is white noise, a 95% confidence interval for ρ_k is

$$r_k \pm \frac{2}{\sqrt{n}}.$$

  Note: The `acf()` function in R plots the 95% confidence interval error bounds $\pm \frac{2}{\sqrt{n}}$ by default!

- **Example 2**: If $\{Y_t : t \in \mathcal{I}\}$ is an $MA(q)$ process, a 95% confidence interval for ρ_k is

$$r_k \pm 2\sqrt{\frac{1}{n}\left[1 + 2\sum_{j=1}^{k-1} r_j^2\right]}.$$

  Note: Specify `ci.type = "ma"` with `acf()` in R to plot the corresponding 95% confidence interval error bounds.

Partial Correlation: Measure of association between random variables X and Y upon removing the effect of controlling variables $\boldsymbol{Z} = (Z_1, \cdots, Z_m)$', for some m:

$$\rho_{XY.\boldsymbol{Z}} = Corr(\hat{\varepsilon}_X, \hat{\varepsilon}_Y),$$

where

$$\hat{\varepsilon}_X = X - \boldsymbol{Z}'\hat{\boldsymbol{\beta}}$$
$$\hat{\varepsilon}_Y = Y - \boldsymbol{Z}'\hat{\boldsymbol{\alpha}},$$

where $\hat{\boldsymbol{\beta}}$, and $\hat{\boldsymbol{\alpha}}$ are estimated regression vectors.

For a time series: $\{Y_t : t \in \mathcal{I}\}$...

Partial Autocorrelation Function (PACF): The partial correlation between Y_t and Y_{t-k} upon removing the effect of $(Y_{t-1}, Y_{t-2}, \cdots, Y_{t-(k-1)})'$.

$\Rightarrow$ For a stationary time series $\{Y_t : t \in \mathcal{I}\}$, the PACF at lag k, denoted by ϕ_{kk}, is

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_j}$$

$$\phi_{k,j} = \phi_{k-1,j} - \phi_{kk}\phi_{k-1,k-j}, \quad \text{for } j = 1, 2, \cdots, k-1,$$

and $\phi_{11} = \rho_1$.

Sample (PACF): An estimator for ϕ_{kk} is

$$\hat{\phi}_{kk} = \frac{r_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_j}$$

$$\hat{\phi}_{k,j} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk}\hat{\phi}_{k-1,k-j}, \quad \text{for } j = 1, 2, \cdots, k-1,$$

and $\hat{\phi}_{11} = r_1$.

$\Rightarrow$ Turns out that if $\{Y_t : t \in \mathcal{I}\}$ is an $AR(p)$ process:

$$\hat{\phi}_{kk} \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{n}\right) \quad \text{as } n \to \infty.$$

Remark: $\phi_{kk} = 0$ for $k > p$. That is, the PACF *cuts off* after lag p if $\{Y_t : t \in \mathcal{I}\}$ is an $AR(p)$ process.

A useful hypothesis test for each $k > 0$:

$$H_0 : \phi_{kk} = 0$$
$$H_a : \phi_{kk} \neq 0.$$

$\Rightarrow$ an approximate 95% confidence interval for ϕ_{kk} is

$$\hat{\phi}_{kk} \pm \frac{2}{\sqrt{n}}.$$

Note 1: R will produce estimates $\hat{\phi}_{kk}$ for each $k > 0$, as well as plot $\hat{\phi}_{kk}$ and the corresponding 95% confidence error bounds with `pacf()`; ie $\pm \frac{2}{\sqrt{n}}$.

Note 2: The following table is useful for model identification purposes:

	$AR(p)$	$MA(q)$	$ARMA(p, q)$
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

Question: Since the ACF and PACF "tails off" for an $ARMA(p, q)$ process, how can we identify p and q?

$\Rightarrow$ A variety of graphical tools have been provided to identify p and q

- The Extended Autocorrelation Function (EACF) is what we will consider using, as it has been shown to have good sampling properties for moderately large sample sizes.

Basic idea: Suppose that we know p, and define

$$W_{t,p,j} = Y_t - \tilde{\phi}_1 Y_{t-1} - \tilde{\phi}_2 Y_{t-2} - \cdots - \tilde{\phi}_p Y_{t-p},$$

where $\{\tilde{\phi}_\ell\}_{\ell=1}^p$ are estimates of the $AR(p)$ coefficients.

$\Rightarrow$ $\{W_{t,p,j} : t \in \mathcal{I}\}$ should "behave" like an $MA(q)$ process

$\Rightarrow$ Specify $j = q$ by looking at the ACF of $\{W_{t,p,j} : t \in \mathcal{I}\}$.

Problem: We don't know p!

Solution: For each $k \in \{0, 1, 2, \cdots\}$, set $p = k$ and then determine q.

$\Rightarrow$ Have a variety of time series $\{W_{t,k,j} : t \in \mathcal{I}\}$.

We can summarize the information into a table:

- Values of k down the rows $\downarrow$
- Values of j across the columns $\rightarrow$

$\Rightarrow$ The (k, j)th element of the table corresponds to the sample ACF value with the time series $\{W_{t,k,j} : t \in \mathcal{I}\}$.

$\Rightarrow$ Use an "X" if the sample ACF value is statistically significant.

- **Recall**: the distribution of the sample ACF is approximately $\mathcal{N}\left(0, \frac{1}{\sqrt{n-k-j}}\right)$ if the process $\{W_{t,k,j} : t \in \mathcal{I}\}$ is (approximately) an $MA(j)$ process. Here, $\frac{1}{\sqrt{n-k-j}}$ is the (asymptotic) standard error of the sample ACF.

  $\Rightarrow$ Construct a 95% confidence interval for the *true* ACF value at lag k.

  $\Rightarrow$ An $ARMA(p, q)$ process should *theoretically* give a triangle of zeros, with the upper left-hand corner corresponding to the orders of the process:

Exhibit 6.4 Theoretical Extended ACF (EACF) for an ARMA(1,1) Model

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1	x	0*	0	0	0	0	0	0	0	0	0	0	0	0
2	x	x	0	0	0	0	0	0	0	0	0	0	0	0
3	x	x	x	0	0	0	0	0	0	0	0	0	0	0
4	x	x	x	x	0	0	0	0	0	0	0	0	0	0
5	x	x	x	x	x	0	0	0	0	0	0	0	0	0
6	x	x	x	x	x	x	0	0	0	0	0	0	0	0
7	x	x	x	x	x	x	x	0	0	0	0	0	0	0

Suppose that a certain time series dataset has the following sample ACF values: $r_1 = 0.3$, $r_2 = 0.4$, $r_3 = -0.04$, and $r_4 = 0$. Obtain the following sample PACF values. Show all your work.

(a) $\hat{\phi}_{11}$

(b) $\hat{\phi}_{22}$

(c) $\hat{\phi}_{33}$

(d) $\hat{\phi}_{44}$

Suppose that a certain time series dataset has the following sample ACF values: $r_1 = 0.3$, $r_2 = 0.4$, $r_3 = -0.04$, and $r_4 = 0$. Obtain the following sample PACF values. Show all your work.

(a) $\hat{\phi}_{11}$
(b) $\hat{\phi}_{22}$
(c) $\hat{\phi}_{33}$
(d) $\hat{\phi}_{44}$

Recall the partial ACF:
$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_j}$$

$$\phi_{k,j} = \phi_{k-1,j} - \phi_{kk}\phi_{k-1,k-j} \quad \text{for } j = 1, 2, \cdots, k-1.$$

(a) We have that

$$\phi_{11} = \rho_1$$
$$\Rightarrow \hat{\phi}_{11} = \hat{\rho}_1 = r_1 = 0.3$$

(b) We have that

$$\phi_{22} = \frac{\rho_2 - \phi_{11}\rho_1}{1 - \phi_{11}\rho_1}$$
$$\Rightarrow \hat{\phi}_{22} = \frac{r_2 - \hat{\phi}_{11} r_1}{1 - \hat{\phi}_{11} r_1} = \frac{0.4 - 0.3^2}{1 - 0.3^2} = 0.3407$$

SFU

Suppose that a certain time series dataset has the following sample ACF values: $r_1 = 0.3$, $r_2 = 0.4$, $r_3 = -0.04$, and $r_4 = 0$. Obtain the following sample PACF values. Show all your work.

(a) $\hat{\phi}_{11}$

(b) $\hat{\phi}_{22}$

(c) $\hat{\phi}_{33}$

(d) $\hat{\phi}_{44}$

Recall the partial ACF:

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_j}$$

$$\phi_{k,j} = \phi_{k-1,j} - \phi_{kk}\phi_{k-1,k-j} \quad \text{for } j = 1, 2, \cdots, k-1.$$

(c) We have that

$$\phi_{33} = \frac{\rho_3 - \phi_{21}\rho_2 - \phi_{22}\rho_1}{1 - \phi_{21}\rho_1 - \phi_{22}\rho_2}$$

$$\phi_{21} = \phi_{11} - \phi_{22}\phi_{11}.$$

$$\Rightarrow \hat{\phi}_{21} = \hat{\phi}_{11} - \hat{\phi}_{22}\hat{\phi}_{11} = 0.3 - (0.3407)(0.3) = 0.1978$$

$$\hat{\phi}_{33} = \frac{r_3 - \hat{\phi}_{21}r_2 - \hat{\phi}_{22}r_1}{1 - \hat{\phi}_{21}r_1 - \hat{\phi}_{22}r_2} = \frac{(-0.04) - (0.1978)(0.4) - (0.3407)(0.3)}{1 - (0.1978)(0.3) - (0.3407)(0.4)} = -0.2752$$

Suppose that a certain time series dataset has the following sample ACF values: $r_1 = 0.3$, $r_2 = 0.4$, $r_3 = -0.04$, and $r_4 = 0$. Obtain the following sample PACF values. Show all your work.

(a) $\hat{\phi}_{11}$

(b) $\hat{\phi}_{22}$

(c) $\hat{\phi}_{33}$

(d) $\hat{\phi}_{44}$

Recall the partial ACF:
$$\phi_{kk} = \frac{\rho_k - \sum\limits_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum\limits_{j=1}^{k-1} \phi_{k-1,j}\rho_j}$$

$$\phi_{k,j} = \phi_{k-1,j} - \phi_{kk}\phi_{k-1,k-j} \quad \text{for } j = 1, 2, \cdots, k-1.$$

(d) We have that
$$\phi_{44} = \frac{\rho_4 - \phi_{31}\rho_3 - \phi_{32}\rho_2 - \phi_{33}\rho_1}{1 - \phi_{31}\rho_1 - \phi_{32}\rho_2 - \phi_{33}\rho_3}$$
$$\phi_{31} = \phi_{21} - \phi_{33}\phi_{22}$$
$$\phi_{32} = \phi_{22} - \phi_{33}\phi_{21}.$$
$$\Rightarrow \hat{\phi}_{31} = \hat{\phi}_{21} - \hat{\phi}_{33}\hat{\phi}_{22} = (0.1978) - (-0.2572)(0.3407) = 0.2854$$
$$\hat{\phi}_{32} = \hat{\phi}_{22} - \hat{\phi}_{33}\hat{\phi}_{21} = (0.3407) - (-0.2572)(0.1978) = 0.3916$$
$$\hat{\phi}_{44} = \frac{r_4 - \hat{\phi}_{31}r_3 - \hat{\phi}_{32}r_2 - \hat{\phi}_{33}r_1}{1 - \hat{\phi}_{31}r_1 - \hat{\phi}_{32}r_2 - \hat{\phi}_{33}r_3}$$
$$= \frac{(0) - (0.2854)(-0.04) - (0.3916)(0.4) - (-0.2752)(0.3)}{1 - (0.2854)(0.3) - (0.3916)(0.4) - (-0.2752)(-0.04)} = -0.0839$$

Office hour Tomorrow:

Tuesday, November 10, 7:00-8:00 PM (PT)

See Canvas for the Zoom link

Good luck on your midterm!