Lecture 8

Autoregressive Integrated Moving Average (ARIMA) Models and Seasonal ARIMA Models

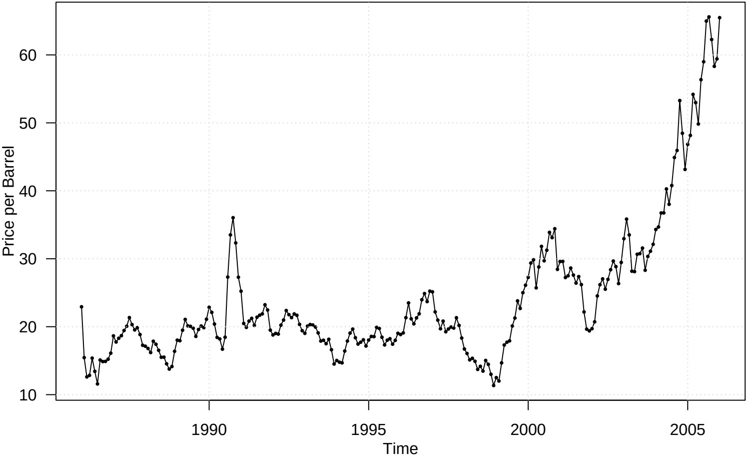
Readings: CC08 Chapter 5.1-5.3, 6.4, & 10; BD16 Chapter 6.1-6.2 & 6.4-6.5; SS17 Chapter 3.6-3.7, & 3.9

MATH 8090 Time Series Analysis
Week 8

Whitney Huang
Clemson University

Autoregressive Integrated Moving Average (ARIMA) Models

Monthly Price of Oil: January 1986–January 2006



A stationary model does not seem to be reasonable. However, it is also not clear which (deterministic) trend model is appropriate ☹

Random Walks Revisited

Recall the random walk process

$$X_t = Z_1 + Z_2 + \cdots + Z_t = \sum_{j=1}^{t} Z_j,$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$

$\{X_t\}$ is a nonstationary process

- We can obtain a stationary process by differencing

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t = Z_t$$

Random Walks Revisited

Recall the random walk process

$$X_t = Z_1 + Z_2 + \cdots + Z_t = \sum_{j=1}^{t} Z_j,$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$

$\{X_t\}$ is a nonstationary process

- We can obtain a stationary process by differencing

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t = Z_t$$

- $\{X_t\}$ is an example of an autoregressive integrated moving average (ARIMA) process– ARIMA(0, 1, 0) process

ARIMA Models

An ARIMA model is an ARMA process after differencing

- Let d be a non-negative integer. Then X_t is an ARIMA(p, d, q) process if

$$Y_t = \nabla^d X_t = (1 - B)^d X_t$$

is a causal ARMA process

ARIMA Models

An ARIMA model is an ARMA process after differencing

- Let d be a non-negative integer. Then X_t is an ARIMA(p, d, q) process if

$$Y_t = \nabla^d X_t = (1 - B)^d X_t$$

is a causal ARMA process

- Let $\phi(B)$ be the AR polynomial and $\theta(B)$ be the MA polynomial. Then for $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$

$$\phi(B) Y_t = \theta(B) Z_t,$$

and since $Y_t = (1 - B)^d X_t$, we have

$$\phi(B)(1 - B)^d X_t = \theta(B) Z_t$$

Example: ARIMA(1, 1, 0)

Let $\phi(z) = 1 - \phi_1 z$, $\theta(z) = 1$ and $d = 1$. For a causal stationary solution (after differencing) we need to assume $|\phi_1| < 1$. Then $\{X_t\}$ is an ARIMA (1, 1, 0) process,

$$(1 - \phi_1 B)(1 - B)X_t = Z_t,$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$

Example: ARIMA(1, 1, 0)

Let $\phi(z) = 1 - \phi_1 z$, $\theta(z) = 1$ and $d = 1$. For a causal stationary solution (after differencing) we need to assume $|\phi_1| < 1$. Then $\{X_t\}$ is an ARIMA (1, 1, 0) process,

$$(1 - \phi_1 B)(1 - B)X_t = Z_t,$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$

Now let $Y_t = (1 - B)X_t = X_t - X_{t-1}$, after some rearrangements we have

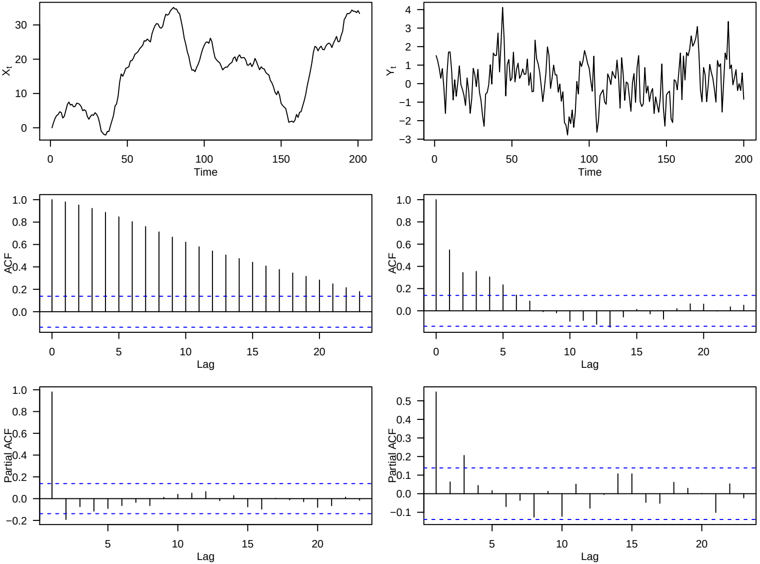
$$\begin{aligned} X_t &= X_{t-1} + Y_t \\ &= (X_{t-2} + Y_{t-1}) + Y_t \\ &\ \ \vdots \\ &= X_0 + \sum_{j=1}^{t} Y_j \end{aligned}$$

Thus $\{X_t\}$ is a "sort of random walk"–we cumulatively sum an AR(1) process, $\{Y_t\}$

Simulated ARIMA and Differenced ARMA Process

We simulate an ARIMA$(1,1,0)$:

$$(1-0.5B)(1-B)X_t = Z_t, \quad \{Z_t\} \sim \mathrm{N}(0,1)$$



Adding a Polynomial Trend

For $d \geq 1$, let $\{X_t\}$ be an ARIMA(p, d, q) process. Then $\{X_t\}$ satisfies the equation

$$\phi(B)(1-B)^d X_t = \theta(B) Z_t$$

- Let μ_t be a polynomial of degree $(d-1)$, i.e., $\mu_t = \sum_{j=0}^{d-1} a_j t^j$ for constants $\{a_j\}$

Adding a Polynomial Trend

For $d \geq 1$, let $\{X_t\}$ be an ARIMA(p, d, q) process. Then $\{X_t\}$ satisfies the equation

$$\phi(B)(1-B)^d X_t = \theta(B) Z_t$$

- Let μ_t be a polynomial of degree $(d-1)$, i.e., $\mu_t = \sum_{j=0}^{d-1} a_j t^j$ for constants $\{a_j\}$
- Now let $V_t = \mu_t + X_t$, then

$$\begin{aligned}
\phi(B)(1-B)^d V_t &= \phi(B)(1-B)^d(\mu_t + X_t) \\
&= \phi(B)(1-B)^d \mu_t + \phi(B)(1-B)^d X_t \\
&= 0 + \phi(B)(1-B)^d X_t \\
&= \theta(B) Z_t
\end{aligned}$$

Adding a Polynomial Trend

For $d \geq 1$, let $\{X_t\}$ be an ARIMA(p, d, q) process. Then $\{X_t\}$ satisfies the equation

$$\phi(B)(1-B)^d X_t = \theta(B) Z_t$$

- Let μ_t be a polynomial of degree $(d-1)$, i.e., $\mu_t = \sum_{j=0}^{d-1} a_j t^j$ for constants $\{a_j\}$
- Now let $V_t = \mu_t + X_t$, then

$$\begin{aligned}
\phi(B)(1-B)^d V_t &= \phi(B)(1-B)^d(\mu_t + X_t) \\
&= \phi(B)(1-B)^d \mu_t + \phi(B)(1-B)^d X_t \\
&= 0 + \phi(B)(1-B)^d X_t \\
&= \theta(B) Z_t
\end{aligned}$$

Adding a Polynomial Trend

For $d \geq 1$, let $\{X_t\}$ be an ARIMA(p, d, q) process. Then $\{X_t\}$ satisfies the equation

$$\phi(B)(1-B)^d X_t = \theta(B) Z_t$$

- Let μ_t be a polynomial of degree $(d-1)$, i.e., $\mu_t = \sum_{j=0}^{d-1} a_j t^j$ for constants $\{a_j\}$
- Now let $V_t = \mu_t + X_t$, then

$$\begin{aligned}
\phi(B)(1-B)^d V_t &= \phi(B)(1-B)^d(\mu_t + X_t) \\
&= \phi(B)(1-B)^d \mu_t + \phi(B)(1-B)^d X_t \\
&= 0 + \phi(B)(1-B)^d X_t \\
&= \theta(B) Z_t
\end{aligned}$$

Takeaway: ARIMA(p, d, q) are useful for modeling data with polynomial trends, due to the inherent differencing that can be used to remove trends

Steps for Modeling ARIMA Processes: Exploratory Analysis

- Plot the data, ACF, PACF and Q-Q plots
  - Check for unusual features of the data
  - Check for stationarity
  - Do we need to transform the data?
- Eliminate trend
  - Estimating the trend and removing it from the series
  - Or, differencing the series (i.e., select d in the ARIMA model)
- Plot the sample ACF/PACF for the stationary component
  - Identify candidate values of p and q

Steps for Modeling ARIMA Processes: Estimation and Model Checking

- Estimate the ARMA process parameters for the candidate models
- Check the goodness of fit: Are the time series residuals, $\{r_t\}$ a sample of *i.i.d.* noise?
- Model selection:
  - Using information criteria such as AIC and AICc
  - Test model parameters to compare between the "full" model and the "subset" model

Forecasting ARIMA Processes

We need more assumptions to forecast ARIMA(p, d, q) processes. Let us start with the case of $d = 1$, i.e.,

$$\phi(B)(1 - B)X_t = \theta(B)Z_t,$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$

- **Note:** $Y_t = (1 - B)X_t = X_t - X_{t-1}$ is an ARMA(p, q) process
- We want to find the best linear predictor (BLP) of X_{n+1} based on $X_0, X_1, \cdots, X_n$
  - We konw that $X_{n+1} = X_n + Y_{n+1} \Rightarrow$ only need to figure out the BLP of Y_{n+1} based on $\{X_0, Y_1, \cdots, Y_n\}$
  - We need to know $\mathbb{E}(X_0^2)$ and $\mathbb{E}(X_0 Y_j)$ for $j = 1, \cdots, n + 1$

Forecasting ARIMA(p, 1, q) Processes (Cont'd)

Problem: What is $\mathbb{E}(X_0 Y_j)$?

- We assume that X_0 is uncorrelated with $Y_1, Y_2, \cdots$

Forecasting ARIMA(p, 1, q) Processes (Cont'd)

Problem: What is $\mathbb{E}(X_0 Y_j)$?

- We assume that X_0 is uncorrelated with $Y_1, Y_2, \cdots$
- Then the BLP of X_{n+1} based on $\{X_0, X_1, \cdots, X_n\}$ is the same as the BLP of X_{n+1} based on $\{Y_1, Y_2, \cdots, Y_n\}$

Forecasting ARIMA(p, 1, q) Processes (Cont'd)

Problem: What is $\mathbb{E}(X_0 Y_j)$?

- We assume that X_0 is uncorrelated with $Y_1, Y_2, \cdots$
- Then the BLP of X_{n+1} based on $\{X_0, X_1, \cdots, X_n\}$ is the same as the BLP of X_{n+1} based on $\{Y_1, Y_2, \cdots, Y_n\}$
- This extends to ARIMA(p, d, q) processes:

  If we assume that $\{X_{1-d}, \cdots, X_0\}$ is uncorrelated with $Y_1, Y_2, \cdots$, then the BLP of Y_{n+1} based on $\{X_{1-d}, \cdots, X_0, \cdots, X_n\}$ is the same as the BLP based on $\{Y_1, Y_2, \cdots, Y_n\}$

Percentage Changes and Logarithms

Suppose X_t tends to have relatively stable percentage changes from one time period to the next. Specifically, assume that

$$X_t = (1 + Y_t)X_{t-1},$$

where $100Y_t$ is the percentage change from X_{t-1} to X_t. Then

$$\log(X_t) - \log(X_{t-1}) = \log\left(\frac{X_t}{X_{t-1}}\right) = \log(1 + Y_t).$$

If Y_t is restricted to, say, $|Y_t| < 0.2$ (ie., the percentage changes are at most $\pm 20\%$), then, to a good approximation, $\log(1 + Y_t) \approx Y_t$. Consequently

$$\Delta[\log(X_t)] \approx Y_t$$

will be relatively stable and perhaps well-modeled by a stationary process.

Percentage Changes and Logarithms

Suppose X_t tends to have relatively stable percentage changes from one time period to the next. Specifically, assume that

$$X_t = (1 + Y_t)X_{t-1},$$

where $100Y_t$ is the percentage change from X_{t-1} to X_t. Then

$$\log(X_t) - \log(X_{t-1}) = \log\left(\frac{X_t}{X_{t-1}}\right) = \log(1 + Y_t).$$

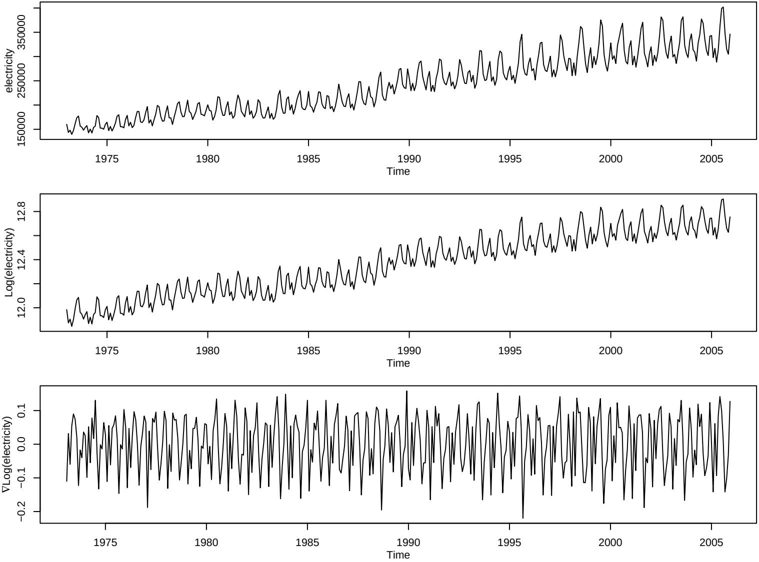
If Y_t is restricted to, say, $|Y_t| < 0.2$ (ie., the percentage changes are at most ± 20%), then, to a good approximation, $\log(1 + Y_t) \approx Y_t$. Consequently

$$\Delta[\log(X_t)] \approx Y_t$$

will be relatively stable and perhaps well-modeled by a stationary process.

> In the financial literature, the differences of the (natural) logarithms are usually called returns

Time Series Plots of Monthly US Electricity Production



Seasonal ARIMA (SARIMA) Model

Modeling Trend, Seasonality, and Noise

Recall the trend, seasonality, noise decomposition:

$$Y_t = \mu_t + s_t + \eta_t,$$

where

- μ_t: (deterministic) trend component;
- s_t: (deterministic) seasonal component with mean 0;
- η_t: random noise with $\mathbb{E}(\eta_t) = 0$

We have already described ways to estimate each component both separately and jointly (via likelihood-based method). But what about if $\{s_t\}$ is a "random" function of t?

Modeling Trend, Seasonality, and Noise

Recall the trend, seasonality, noise decomposition:

$$Y_t = \mu_t + s_t + \eta_t,$$

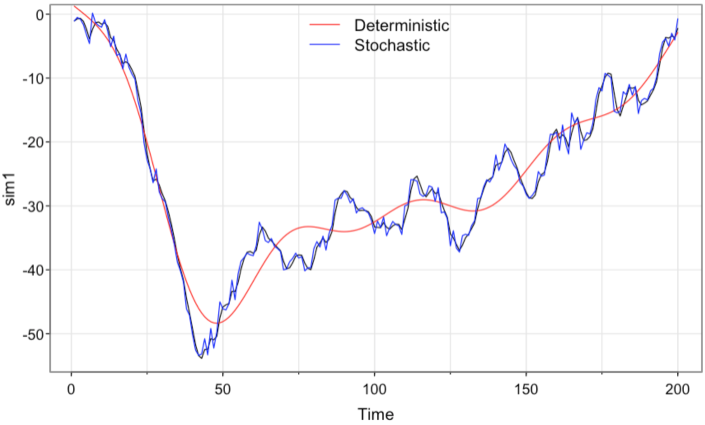
where

- μ_t: (deterministic) trend component;
- s_t: (deterministic) seasonal component with mean 0;
- η_t: random noise with $\mathbb{E}(\eta_t) = 0$

We have already described ways to estimate each component both separately and jointly (via likelihood-based method). But what about if $\{s_t\}$ is a "random" function of t?

$\Rightarrow$ The seasonal ARIMA model allows us to model the case when s_t itself varies randomly from one cycle to the next

Digression: Using ARIMA for Stochastic Trend Modeling



For a given time series, it may be challenging to identify the exact form of a deterministic trend μ_t. However, ARIMA models can effectively capture and account for a "stochastic" trend

The Seasonal ARIMA (SARIMA) Model

Let d and D be non-negative integers. Then $\{X_t\}$ is a seasonal ARIMA$(p, d, q) \times (P, D, Q)_s$ process with period s if

$$Y_t = \nabla^d \nabla_s^D X_t = (1 - B)^d (1 - B^s)^D X_t,$$

is a casual ARMA process define by

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)Z_t,$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$.

$\{Y_t\}$ is causal if $\phi(z) \neq 0$ and $\Phi(z) \neq 0$, for $|z| \leq 1$, where

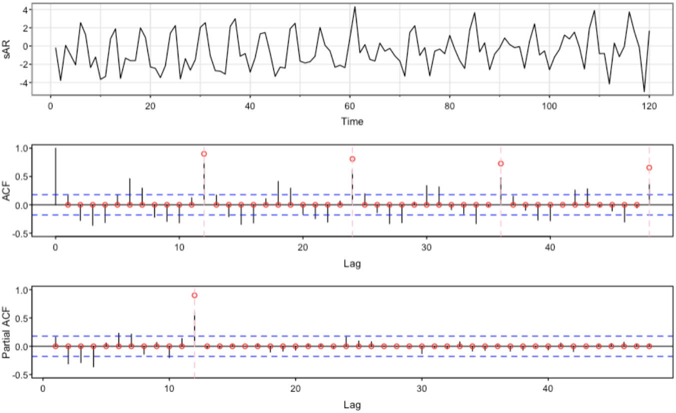
$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p;$$
$$\Phi(z) = 1 - \Phi_1 z - \cdots - \Phi_P z^P.$$

All roots of the AR and SAR characteristic equations must be greater than 1 in modulus

An Example of a Seasonal AR Model

$$Y_t = 0.9Y_{t-12} + Z_t,$$

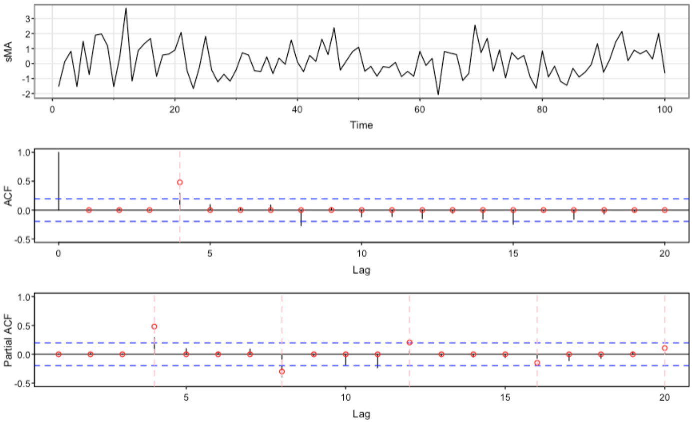
$$\Rightarrow p = q = d = D = Q = 0,\ P = 1,\ \Phi_1 = 0.9, s = 12.$$



An Example of a Seasonal MA Model

$$Y_t = Z_t + 0.75Z_{t-4},$$

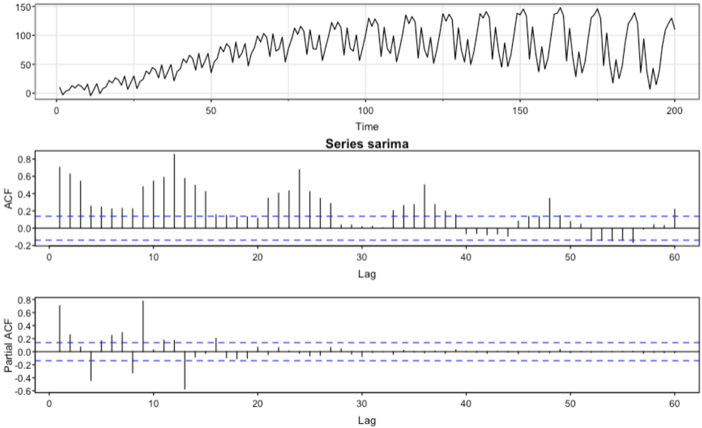
$$\Rightarrow p = q = d = D = P = 0,\ Q = 1,\ \Theta_1 = 0.75, s = 4.$$



Example of a SARIMA Model

$$(1-B)(1-B^{12})X_t = Y_t$$
$$(1+0.25B)(1-0.9B^{12})Y_t = (1+0.75B^{12})Z_t$$

$$\Rightarrow p = P = Q = d = D = 1,\ \phi = -0.25, \Phi = 0.9, \Theta_1 = 0.75, s = 12.$$



An Illustration of Seasonal Model

Consider a monthly time series $\{X_t\}$ with both a trend, and a seasonal component of period $s = 12$.

- Suppose we know the values of d and D such that $Y_t = (1 - B)^d (1 - B^{12})^D X_t$ is stationary

An Illustration of Seasonal Model

Consider a monthly time series $\{X_t\}$ with both a trend, and a seasonal component of period $s = 12$.

- Suppose we know the values of d and D such that $Y_t = (1 - B)^d (1 - B^{12})^D X_t$ is stationary
- We can arrange the data this way:

	Month 1	Month 2	$\cdots$	Month 12
Year 1	Y_1	Y_2	$\cdots$	Y_{12}
Year 2	Y_{13}	Y_{14}	$\cdots$	Y_{24}
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
Year r	$Y_{1+12(r-1)}$	$Y_{2+12(r-1)}$	$\cdots$	$Y_{12+12(r-1)}$

The Inter-annual Model

Here we view each column (month) of the data table from the previous slide as a separate time series

- For each month m, we assume the same ARMA(P, Q) model. We have

$$Y_{m+12y} - \sum_{i=1}^{P} \Phi_i Y_{m+12(y-i)} = U_{m+12y} + \sum_{j=1}^{Q} \Theta_j U_{m+12(y-j)},$$

  for each $y = 0, \cdots, r-1$, where $\{U_{m+12y:y=0,\cdots,r-1}\} \sim \mathrm{WN}(0, \sigma_U^2)$ for each m

The Inter-annual Model

Here we view each column (month) of the data table from the previous slide as a separate time series

- For each month m, we assume the same ARMA(P, Q) model. We have

$$
\begin{aligned}
&Y_{m+12y} - \sum_{i=1}^{P} \Phi_i Y_{m+12(y-i)} \\
&= U_{m+12y} + \sum_{j=1}^{Q} \Theta_j U_{m+12(y-j)},
\end{aligned}
$$

  for each $y = 0, \cdots, r-1$, where $\{U_{m+12y:y=0,\cdots,r-1}\} \sim \mathrm{WN}(0, \sigma_U^2)$ for each m
- We can write this as

$$
\Phi(B^{12}) Y_t = \Theta(B^{12}) U_t,
$$

  and this defines the inter-annual model

The Intra-Annual Model

We induce correlation between the months by letting the process $\{U_t\}$ follow an ARMA(p, q) model,

$$\phi(B)U_t = \theta(B)Z_t,$$

where $Z_t \sim \mathrm{WN}(0, \sigma^2)$

- This is the intra-annual model

The Intra-Annual Model

We induce correlation between the months by letting the process $\{U_t\}$ follow an ARMA(p, q) model,

$$\phi(B)U_t = \theta(B)Z_t,$$

where $Z_t \sim \mathrm{WN}(0, \sigma^2)$

- This is the intra-annual model
- The combination of the inter-annual and intra-annual models for the differenced stationary series,

  $$Y_t = (1 - B)^d(1 - B^{12})^D X_t,$$

  yields a SARIMA model for $\{X_t\}$

Steps for Modeling SARIMA Processes

1. Transform data is necessary

Steps for Modeling SARIMA Processes

1. Transform data is necessary

2. Find d and D so that

$$Y_t = (1 - B)^d (1 - B^s)^D X_t$$

is stationary

Steps for Modeling SARIMA Processes

1. Transform data is necessary

2. Find d and D so that

$$Y_t = (1 - B)^d (1 - B^s)^D X_t$$

is stationary

3. Examine the sample ACF/PACF of $\{Y_t\}$ at lags that are multiples of s for plausible values of P and Q

Steps for Modeling SARIMA Processes

1. Transform data is necessary

2. Find d and D so that

$$Y_t = (1 - B)^d (1 - B^s)^D X_t$$

is stationary

3. Examine the sample ACF/PACF of $\{Y_t\}$ at lags that are multiples of s for plausible values of P and Q

4. Examine the sample ACF/PACF at lags $\{1, 2, \cdots, s - 1\}$, to identify possible values of p and q

Modeling SARIMA Processes (Cont'd)

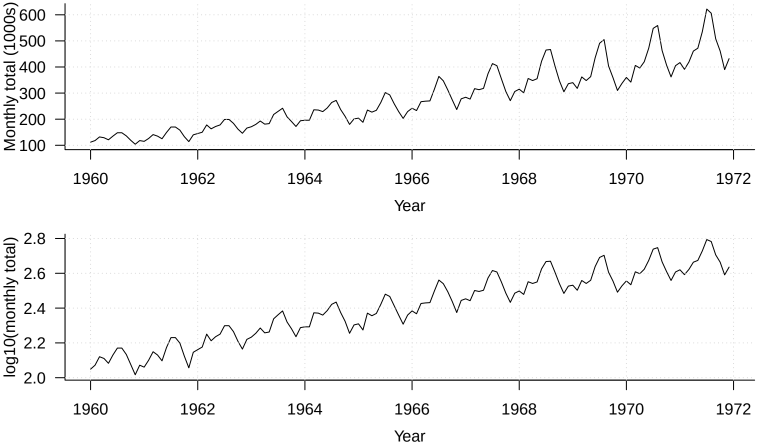
5. Use maximum likelihood method to fit the models

6. Use model summaries, diagnostics, AIC (AICc) to determine the best SARIMA model

7. Conduct forecast

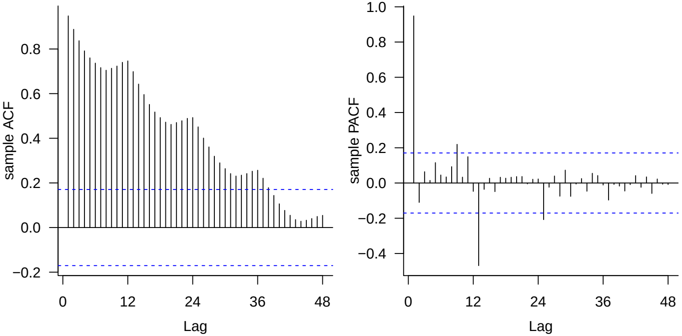
Airline Passengers Example

We consider the data set `airpassengers`, which are the monthly totals of international airline passengers from 1960 to 1971.



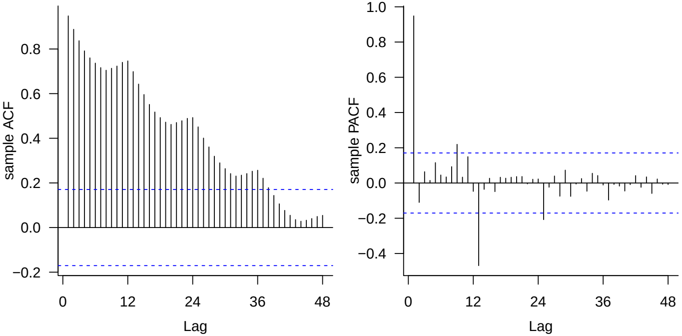
Here we stabilize the variance with a $\log_{10}$ transformation

Sample ACF/PACF Plots



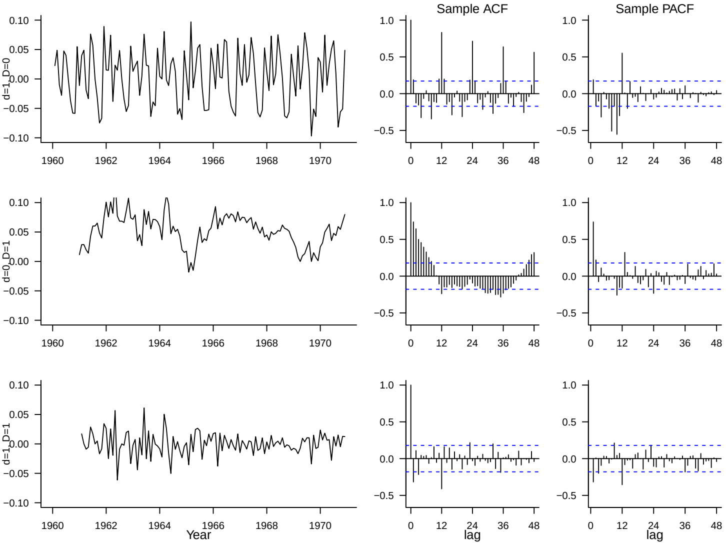
- The sample ACF decays slowly with a wave structure ⇒ seasonality

Sample ACF/PACF Plots



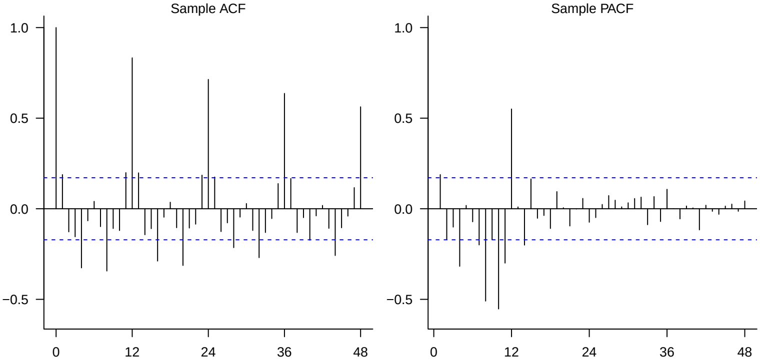
- The sample ACF decays slowly with a wave structure ⇒ seasonality
- The lag one PACF is close to one, indicating that differencing the data would be reasonable

Trying Different Orders of Differencing



Choosing Candidate SARIMA Models

We choose a SARIMA$(p, 1, q) \times (P, 0, Q)_{12}$ model. Next we examine the sample ACF/PACF of the process $Y_t = (1 - B)X_t$

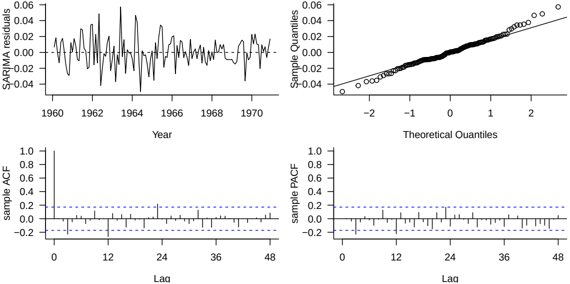


Now we need to choose P, Q, p, and q

Fitting a SARIMA$(1,1,0)\times(1,0,0)$ Model

```
> fit1 <- arima(diff.1.0, order = c(1, 0, 0), seasonal = list(order = c(1, 0, 0), period = 12))
> fit1

Call:
arima(x = diff.1.0, order = c(1, 0, 0), seasonal = list(order = c(1, 0, 0),
    period = 12))

Coefficients:
          ar1    sar1  intercept
      -0.2667  0.9291     0.0039
s.e.   0.0865  0.0235     0.0096

sigma^2 estimated as 0.0003298:  log likelihood = 327.27,  aic = -646.54
```

```
> Box.test(fit1$residuals, lag = 48, type = "Ljung-Box")

        Box-Ljung test

data:  fit1$residuals
X-squared = 55.372, df = 48, p-value = 0.2164
```



A Discussion of the Model Fit

- Residuals show greater spread in 1949-1955 and have heavier-than-normal tails
- The Ljung-Box test result indicates the fitted SARIMA $(1,1,0)\times(1,0,0)_{12}$ has sufficiently account for the temporal dependence
- 95% CI for ϕ_1 and Φ_1 do not contain zero $\Rightarrow$ no need to go with simpler model

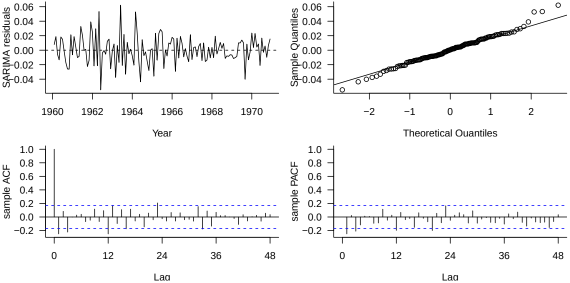
Our estimated model is:

$$X_t = \log_{10}(\text{\#Passengers})$$
$$Y_t = (1-B)X_t = X_t - X_{t-1}$$
$$(1+0.2667B)(1-0.9291B^{12})(Y_t - 0.0039) = Z_t,$$

where $\{Z_t\} \overset{i.i.d.}{\sim} \mathrm{N}(0,\sigma^2)$ with $\hat{\sigma}^2 = 0.00033$

Comparing with a SARIMA$(0,1,0)\times(1,0,0)$ Model

```
> (fit2 <- arima(diff.1.0, seasonal = list(order = c(1, 0, 0), period = 12)))

Call:
arima(x = diff.1.0, seasonal = list(order = c(1, 0, 0), period = 12))

Coefficients:
        sar1  intercept
      0.9081     0.0040
s.e.  0.0278     0.0108

sigma^2 estimated as 0.0003616:  log likelihood = 322.75,  aic = -639.51
> Box.test(fit2$residuals, lag = 48, type = "Ljung-Box")

        Box-Ljung test

data:  fit2$residuals
X-squared = 80.641, df = 48, p-value = 0.002209
```



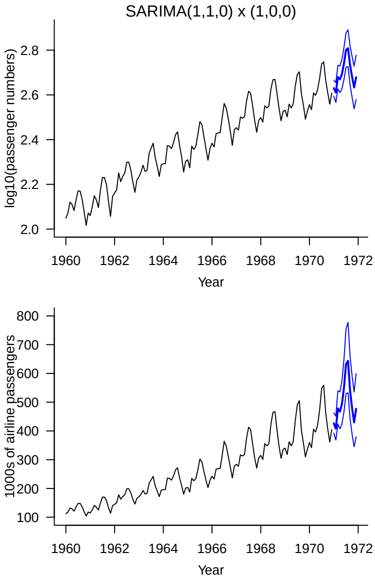
A Discussion of SARIMA$(0, 1, 0) \times (1, 0, 0)$ Model Fit

Here we drop the AR(1) term

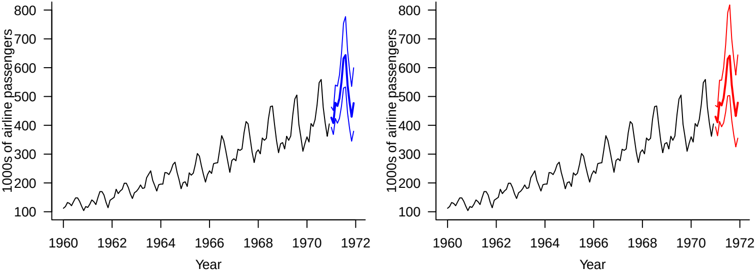
- Residual plots are similar to before, with greater spread in 1949-1955 and heavy tails
- Both $\hat{\sigma}^2$ and AIC increase (compared with model fit1)
- The lag 1 of ACF and PACF now lies outside the IID noise bounds. The Ljung-Box p-value of 0.0022, leads us to reject the IID residual assumption

In conclusion, the SARIMA$(1, 1, 0) \times (1, 0, 0)_{12}$ model fits better than SARIMA$(0, 1, 0) \times (1, 0, 0)_{12}$

Forecasting the 1971 Data



Evaluating Forecast Performance



Metrics	Model Fit1	Model Fit2
Root Mean Square Error	30.36	31.32
Mean Relative Error	0.057	0.060
Empirical Coverage	0.917	1.000

The SARIMA$(1,1,0) \times (1,0,0)$ Model is Equivalent To?

Our model for the log passenger series $\{X_t\}$ is

$$\phi(B)\Phi(B^{12})(1-B)X_t = Z_t,$$

where $\phi(B) = 1 - \phi_1 B$ and $\Phi(B) = 1 - \Phi_1(B)$

Note that

$$\begin{aligned}\phi(B)\Phi(B^{12}) &= (1-\phi_1 B)(1-\Phi_1 B^{12}) \\ &= 1 - \phi_1 B - \Phi_1 B^{12} + \phi_1\Phi_1 B^{13}\end{aligned}$$

The SARIMA$(1,1,0)\times(1,0,0)$ Model is Equivalent To?

Our model for the log passenger series $\{X_t\}$ is

$$\phi(B)\Phi(B^{12})(1-B)X_t = Z_t,$$

where $\phi(B) = 1 - \phi_1 B$ and $\Phi(B) = 1 - \Phi_1(B)$

Note that

$$\begin{aligned}\phi(B)\Phi(B^{12}) &= (1-\phi_1 B)(1-\Phi_1 B^{12}) \\ &= 1 - \phi_1 B - \Phi_1 B^{12} + \phi_1\Phi_1 B^{13}\end{aligned}$$

Question: What is this SARIMA model equivalent to?