

Lecture 8

Seasonal Time Series Models

Readings: CC08 Chapter 10 & Chapter 6.4; BD16 Chapter 6.5
& Chapter 6.3; SS17 Chapter 3.9 & Chapter 5.2

MATH 8090 Time Series Analysis
Week 8

Whitney Huang
Clemson University

Modeling Trend, Seasonality, and Noise

Recall the trend, seasonality, noise decomposition mentioned in Week 2:

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

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 ?

⇒ The **seasonal ARIMA** model allows us to model the case when s_t itself varies **randomly** from one cycle to the next

The Seasonal ARIMA (SARIMA) Model

Let d and D be non-negative integers. Then $\{X_t\}$ is a **seasonal ARIMA** $(p, d, q) \times (P, D, Q)$ 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 $\{Z_t\} \sim \text{WN}(0, \sigma^2)$.

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

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p;$$

$$\Phi(z) = 1 - \Phi_1 z - \dots - \Phi_P z^P.$$

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	...	Month 12
Year 1	Y_1	Y_2	...	Y_{12}
Year 2	Y_{13}	Y_{14}	...	Y_{24}
⋮	⋮	⋮	...	⋮
Year r	$Y_{1+12(r-1)}$	$Y_{2+12(r-1)}$...	$Y_{12+12(r-1)}$

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 $y = 0, \dots, r-1$, where

$\{U_{m+12y:y=0,\dots,r-1}\} \sim \text{WN}(0, \sigma_U^2)$ 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**

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

- 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 $\{X_t\}$

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

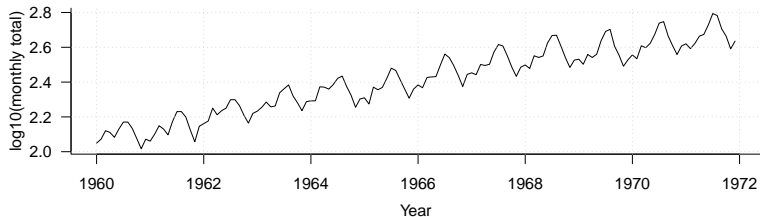
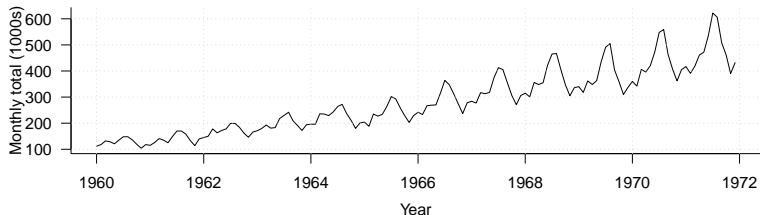
3. Examine the sample ACF/PACF of $\{Y_t\}$ at lags that are multiples of s for plausible values of P and Q

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

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

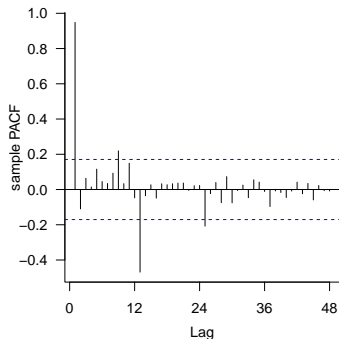
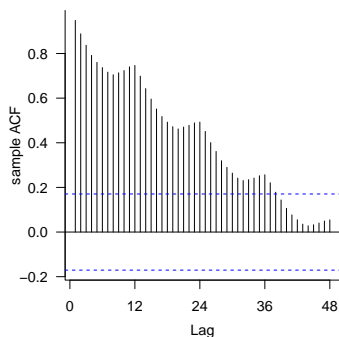
Airline Passengers Example

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



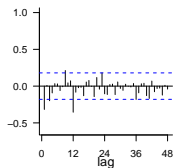
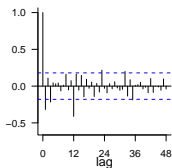
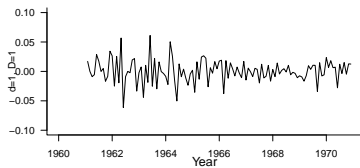
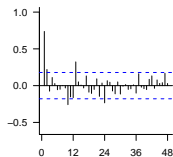
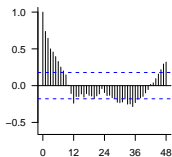
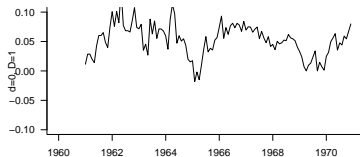
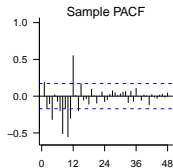
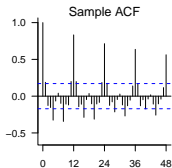
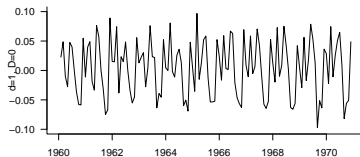
Here we stabilize the variance with a \log_{10} transformation

Sample ACF/PACF Plots



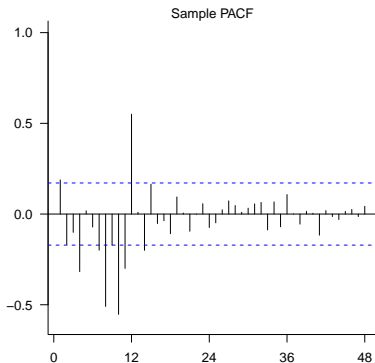
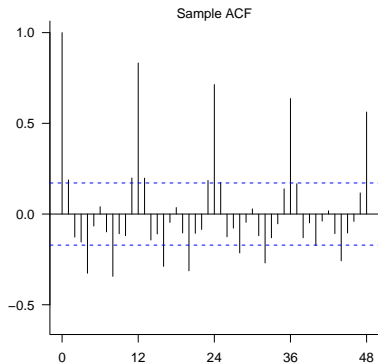
- The sample ACF decays slowly with a wave structure \Rightarrow seasonality
- The lag one PACF is close to one, indicating that differencing the data would be reasonable

Trying Different Orders of Differencing



Choosing Candidate SARIMA Models

We choose a SARIMA($p, 1, q$) \times ($P, 0, Q$) 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) × (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

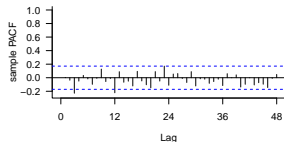
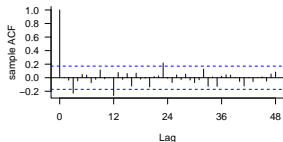
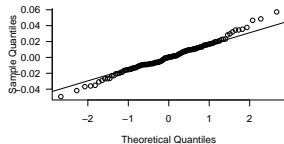
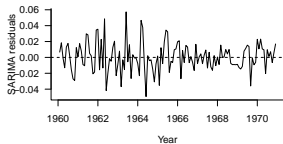
sigma² estimated as 0.0003298: log likelihood = 327.27, aic = -646.54

```
> 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

- The spread of the residuals is larger in 1949-1955 compared to the later years and the residual distribution has heavy tails
- The Ljung-Box test result indicates the fitted SARIMA $(1, 1, 0) \times (1, 0, 0)$ 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

$$(1 + 0.2667B)(1 - 0.9291B^{12})(X_t - 0.0039) = Z_t,$$

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

Comparing with a SARIMA(0,1,0) × (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

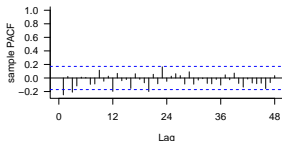
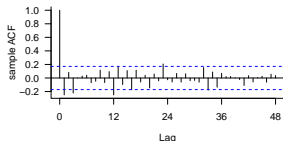
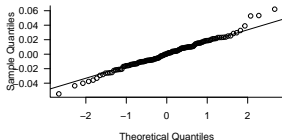
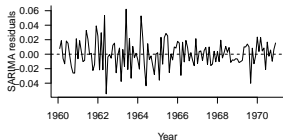
sigma² estimated as 0.0003616: log likelihood = 322.75, aic = -639.51

```
> 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



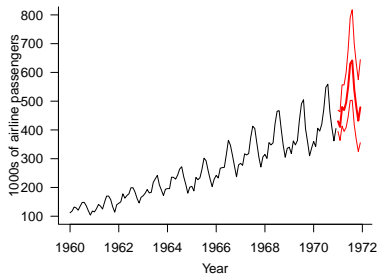
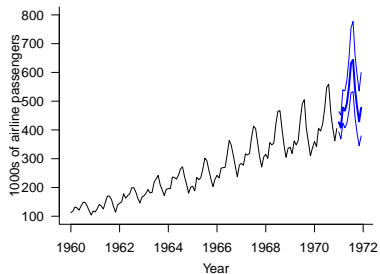
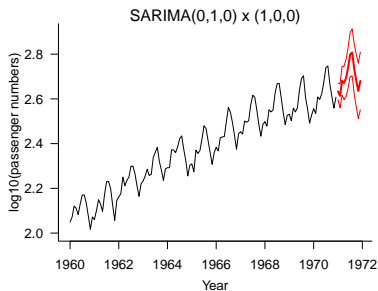
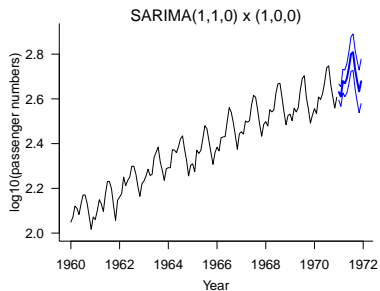
A Discussion of SARIMA(0, 1, 0) \times (1, 0, 0) Model Fit

Here we drop the AR(1) term

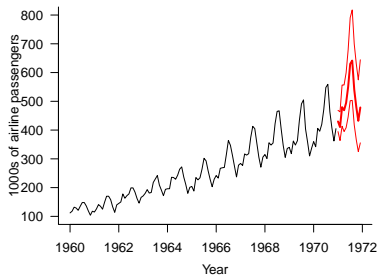
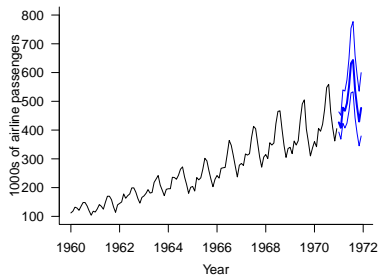
- The residual plots looks quite similar to before: The spread of the residuals is larger in 1949-1955 compared to the later years and the residual distribution has heavy tails
- Both $\hat{\sigma}^2$ and AIC increase (compared with model fit1)
- The lag 1 of ACF and PACF now lies outside the IID noise 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) model fits better than SARIMA(0, 1, 0) \times (1, 0, 0)

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

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

Question: What is this SARIMA model equivalent to?

Unit Root Tests

Suppose we have X_1, \dots, X_n that follow the model

$$(X_t - \mu) = \phi(X_{t-1} - \mu) + Z_t,$$

where $\{Z_t\}$ is a $WN(0, \sigma^2)$ process

- A **unit root test** considers the following hypotheses:

$$H_0 : \phi = 1 \text{ versus } H_a : |\phi| < 1$$

- Note that where $|\phi| < 1$ the process is **stationary** (and causal) while $\phi = 1$ leads to a nonstationary process
- **Exercise:** Letting $Y_t = \nabla X_t = X_t - X_{t-1}$, show that

$$\begin{aligned} Y_t &= (1 - \phi)\mu + (\phi - 1)X_{t-1} + Z_t \\ &= \phi_0^* + \phi_1^* X_{t-1} + Z_t, \end{aligned}$$

where $\phi_0^* = (1 - \phi)\mu$ and $\phi_1^* = (\phi - 1)$

Unit Root Tests (Cont'd)

- We can estimate ϕ_0^* and ϕ_1^* using ordinary least squares
- Using the estimate of ϕ_1^* , $\hat{\phi}_1^*$, and its standard error, $\widehat{SE}(\hat{\phi}_1^*)$, the **Dickey-Fuller statistics** is

$$T = \frac{\hat{\phi}_1^*}{\widehat{SE}(\hat{\phi}_1^*)}$$

- Under H_0 this statistic follows a **Dickey-Fuller distribution**. For a level α test we reject if the observed test statistic is smaller than a critical value C_α

α	0.01	0.05	0.10
C_α	-3.43	-2.86	-2.57

- We can extend to other processes (AR(p), ARMA(p, q), and MA(q))—see Brockwell and Davis [2016, Section 6.3] for further details

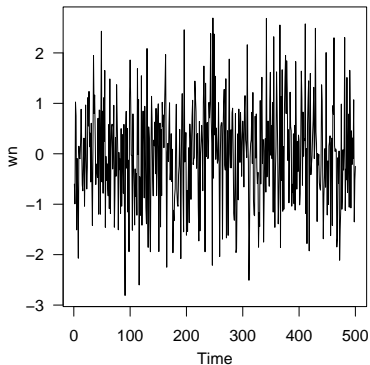
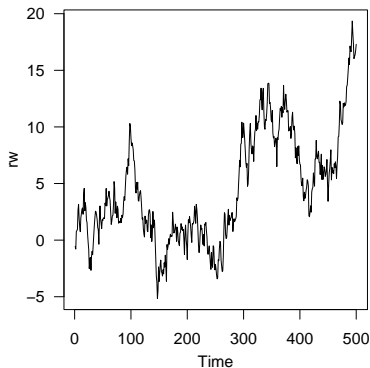
Unit Root Test: Simulated Examples

Recall

$$\nabla X_t = \phi_0^* + \phi_1^* X_{t-1} + Z_t,$$

where $\phi_0^* = (1 - \phi)\mu$ and $\phi_1^* = (\phi - 1)$

Let's demonstrate the test with a simulated **random walk** ($\phi = 1$)
and a simulated **white noise** ($\phi = 0$)



```
> diff.rw <- diff(rw); n <- length(rw)
> ys <- diff.rw; xs <- rw[1:(n-1)]
> ols.rw <- lm(ys ~ xs); summary(ols.rw)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.10125	0.05973	1.695	0.0906 .
xs	-0.01438	0.00899	-1.600	0.1102

```
> diff.wn <- diff(wn)
> ys <- diff.wn; xs <- wn[1:(n-1)]
> ols.wn <- lm(ys ~ xs); summary(ols.wn)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.001138	0.045329	-0.025	0.98
xs	-1.002420	0.044843	-22.354	<2e-16