Autoregressive & Moving Average Modeling & Forecasting

Steps of Building A Time Series Model

- Use ACF and PACF to identify which model should be built (AR? MA? ARIMA?);

- Use software packages to identify the coefficients of a model and model performance factors such as error terms of the model;

- If multiple types of models could work, compare error terms of each model, and choose the model with the smallest error terms;

- Check the performance of error terms to ensure the quality of the chosen model.

Example

An analyst plans to use sales revenues of previous days to estimate today's sales revenues.

Identify the best time series model, and use it to make the estimation.

An Easier Way to Write AR Model

y_t= β_0 + $\beta_1 y_{t-1}$+ ε_t (Eq.1)

β_0= $\mu(1-\beta_1)$

An easier way to build an AR model:

y_t- μ = $\beta_1(y_{t-1}$- μ) + ε_t (Eq.2)

y_t- 14.59 = 0.68*$(y_{t-1}$- 14.59) + ε_t

Therefore,

$$y_t = 4.74 + 0.68 y_{t-1} + \varepsilon_t$$

What About MA Model?

- General form of an MA model

$y_t = \mu + \varepsilon_t + \theta_1\,\varepsilon_{t-1} + \theta_2\,\varepsilon_{t-2} + \cdots$

Therefore,

$y_t = 14.56 + \varepsilon_t + 0.56 * \varepsilon_{t-1}$

What About ARMA Model?

- Since ARMA(1,1) is a combination of AR & MA, the model is

y_t- μ = $\beta_1 (y_{t-1}$- μ) + $\varepsilon_t + \theta_1 \, \varepsilon_{t-1}$

Thus,

y_t- 14.59 = 0.72*$(y_{t-1}$- 14.59) + $\varepsilon_t - 0.09 * \varepsilon_{t-1}$

Stationarity of Time Series

Stationarity

- (1) The mean of a time series y_t is the same for all *t*. Sometimes, the mean may be written as *E*(y_t) *or μ* for simplicity purpose.

- (2) The variance of y_t is the same for all *t*. Sometimes, the variance is written as *Var*(y_t) for simplicity purpose. In other words, *Var*(y_t) is constant.

- (3) The covariance (and also correlation) between y_t and y_{t-1} is the same for all *t*.

What About ARIMA?

- A constant value is reported instead of the μ value, because the variable z_t is a result of differencing. For example,

z_t= (y_t- μ) - (y_{t-1}- μ)

z_{t-1}= (y_{t-1}- μ) - (y_{t-2}- μ)

…

So, μ is gone. I use z_t to represent the variable after one lag differencing. Remember "I" in ARIMA is the differencing we want to do? If you forgot how to do differencing, see video here: https://www.youtube.com/embed/Zu1iimmsKD0?start=350

ARIMA Model Example

$$z_t = 0.06 + 0.62 * z_{t-1} + \varepsilon_t - 1.00 * \varepsilon_{t-1}$$

```
    Q), period = S), xreg = constant,
    optim.control = list(trace = trc,

Coefficients:
         ar1      ma1  constant
      0.6228  -1.0000    0.0643
s.e.  0.1325   0.0736    0.0392

sigma^2 estimated as 1.469:  log likel

$degrees_of_freedom
[1] 36
```

Model Diagnostics

- Choose the model with smaller error terms

AR(1) Model

```
Coefficients:
         ar1    xmean
      0.6754  14.5947
s.e.  0.1121   0.5716

sigma^2 estimated as 1.519:  log likelihood = -65.42,  aic = 136.84

$degrees_of_freedom
[1] 38

$ttable
      Estimate     SE t.value p.value
ar1     0.6754 0.1121  6.0232       0
xmean  14.5947 0.5716 25.5338       0

$AIC
[1] 3.421001

$AICc
[1] 3.429109

$BIC
[1] 3.547667
```

MA(1) Model

```
Coefficients:
         ma1    xmean
      0.5578  14.5585
s.e.  0.1311   0.3371

sigma^2 estimated as 1.907:  log likelihood = -69.85,  aic = 145.71

$degrees_of_freedom
[1] 38

$ttable
      Estimate     SE t.value p.value
ma1     0.5578 0.1311  4.2538   1e-04
xmean  14.5585 0.3371 43.1832   0e+00

$AIC
[1] 3.642655

$AICc
[1] 3.650763

$BIC
[1] 3.769321
```

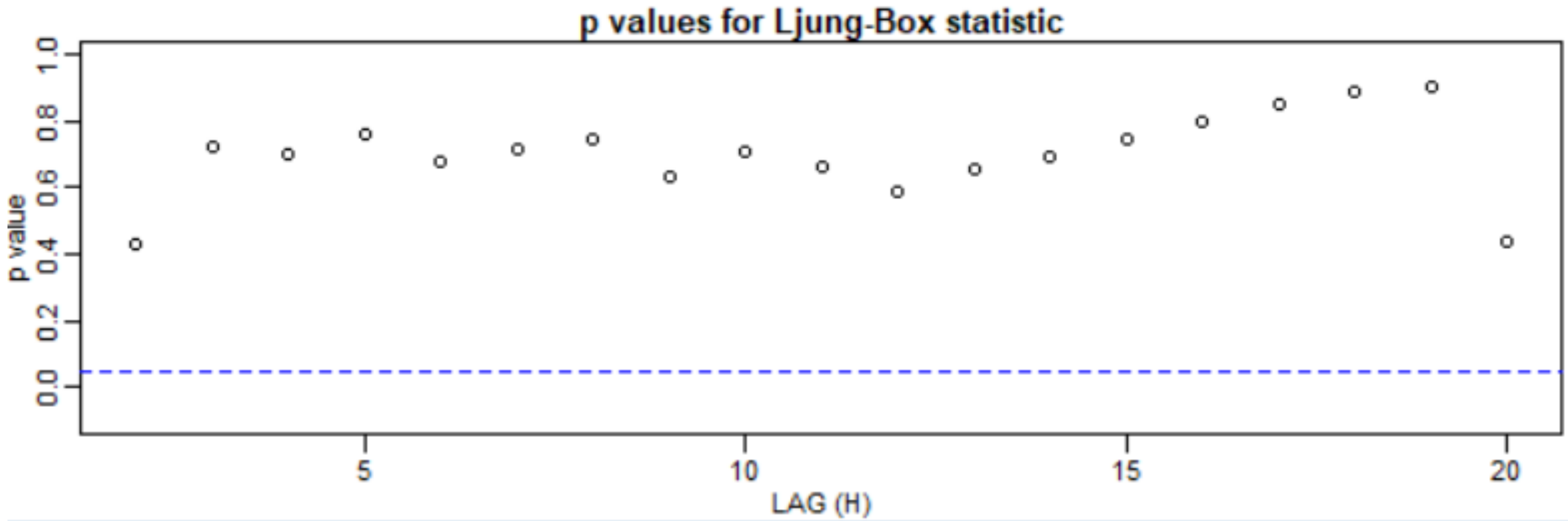
Model Diagnostics

- Ljung-Box Statistic

The p-value of the Ljung-Box test is used to validate the following hypotheses:

$$H_0: The\ population\ error\ term\ is\ 0;$$

$$H_a: The\ population's\ error\ term\ is\ not\ 0.$$



ACF vs. PACF

ACF vs. PACF

- ACF (**A**uto-**C**orrelation **F**unction)

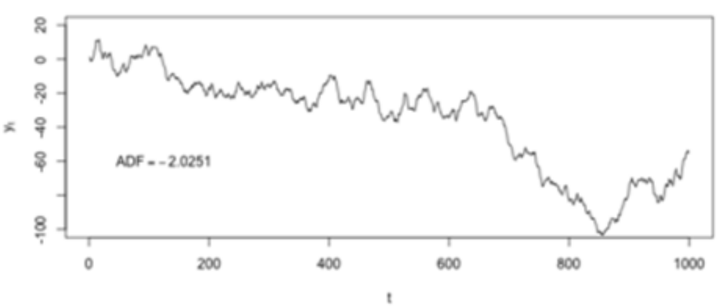
  The correlation between the observation at the current time spot and the observations at previous time spots.

- PACF (**P**artial **A**uto-**C**orrelation **F**unction)

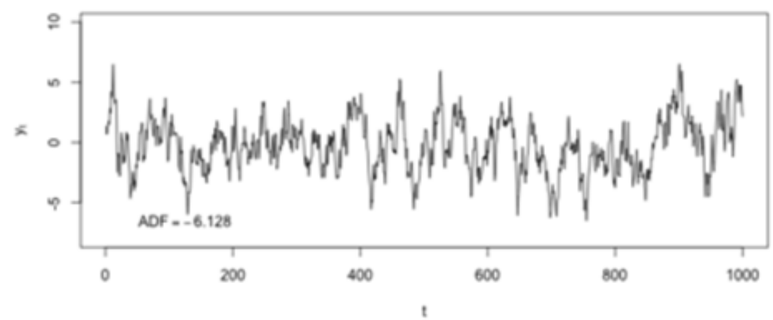
  The correlation between observations at two time spots given that we consider both observations are correlated to observations at other time spots. For example, today's stock price can be correlated to the day before yesterday, and yesterday can also be correlated to the day before yesterday. Then, PACF of yesterday is the "real" correlation between today and yesterday after taking out the influence of the day before yesterday.

Rules of Using ACF and PACF(1)

- Determine if there is an obvious trend in the dataset. If there is, use differencing (i.e., diffing) to “detrend” data. Usually, one-lag differencing is used.



Data with trend



Data after one-lag differencing

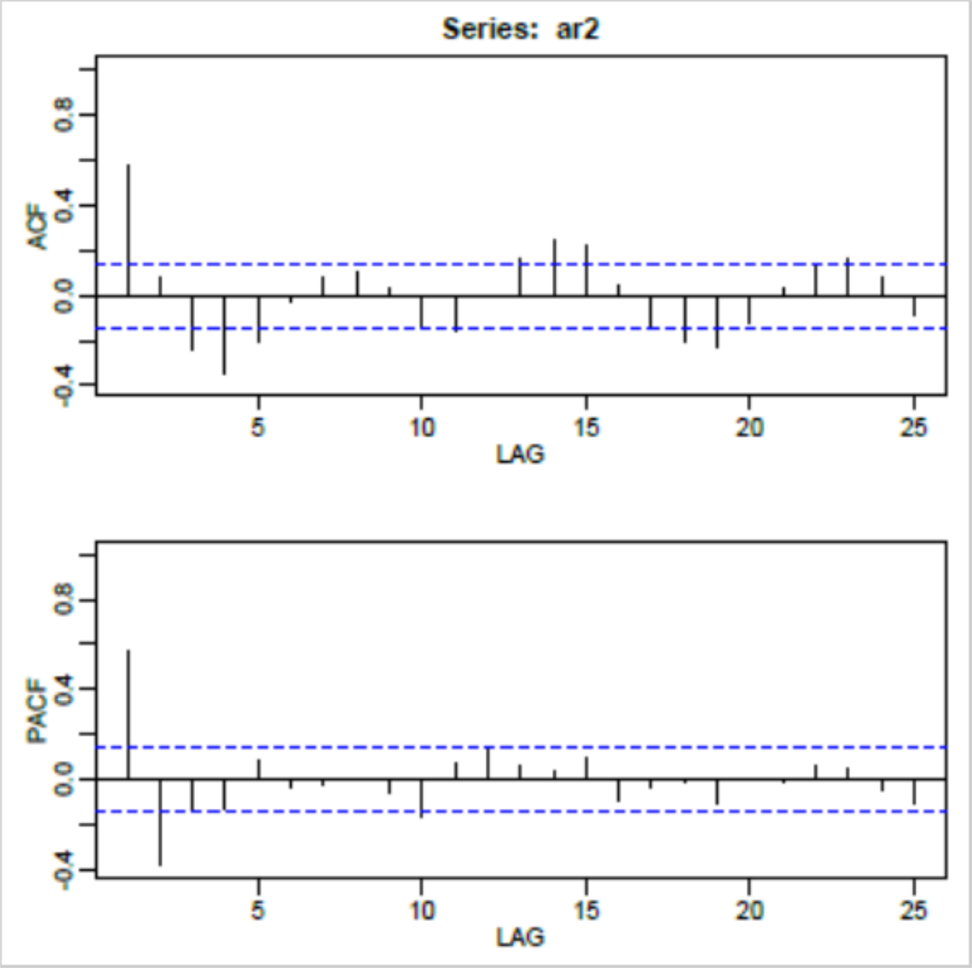
Rules of Using ACF and PACF(2)

- Use PACF to determine the terms used in the AR model. Only significant terms will be chosen. The number of terms determines the order of the model.

  For example, if the PACF of yesterday's stock price is significant. All other days' PACFs are not significant. Then, yesterday's stock price will be used to predict today's stock price. The AR model is called the first-order autoregression model.

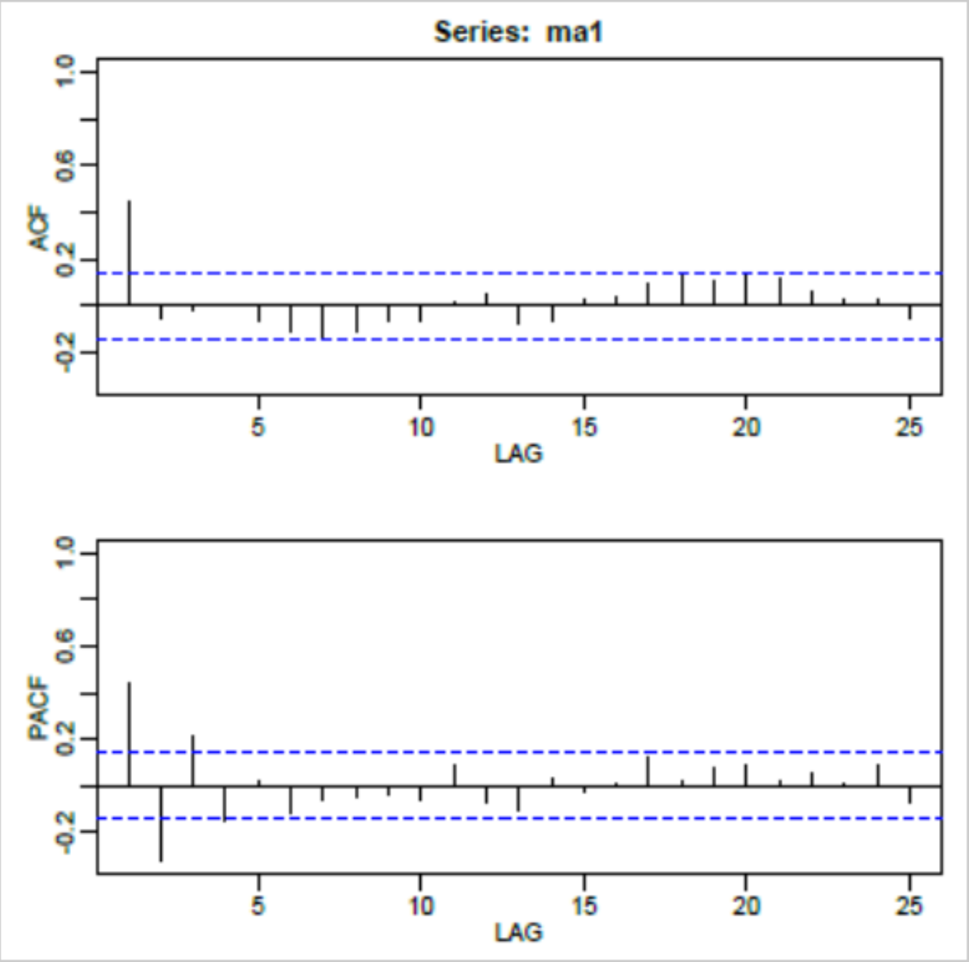
- Use ACF to determine the terms used in the MA model.
- Choose a model by using PACF and ACF charts together.
- Use the "simpler" model if several models could work.

Examples



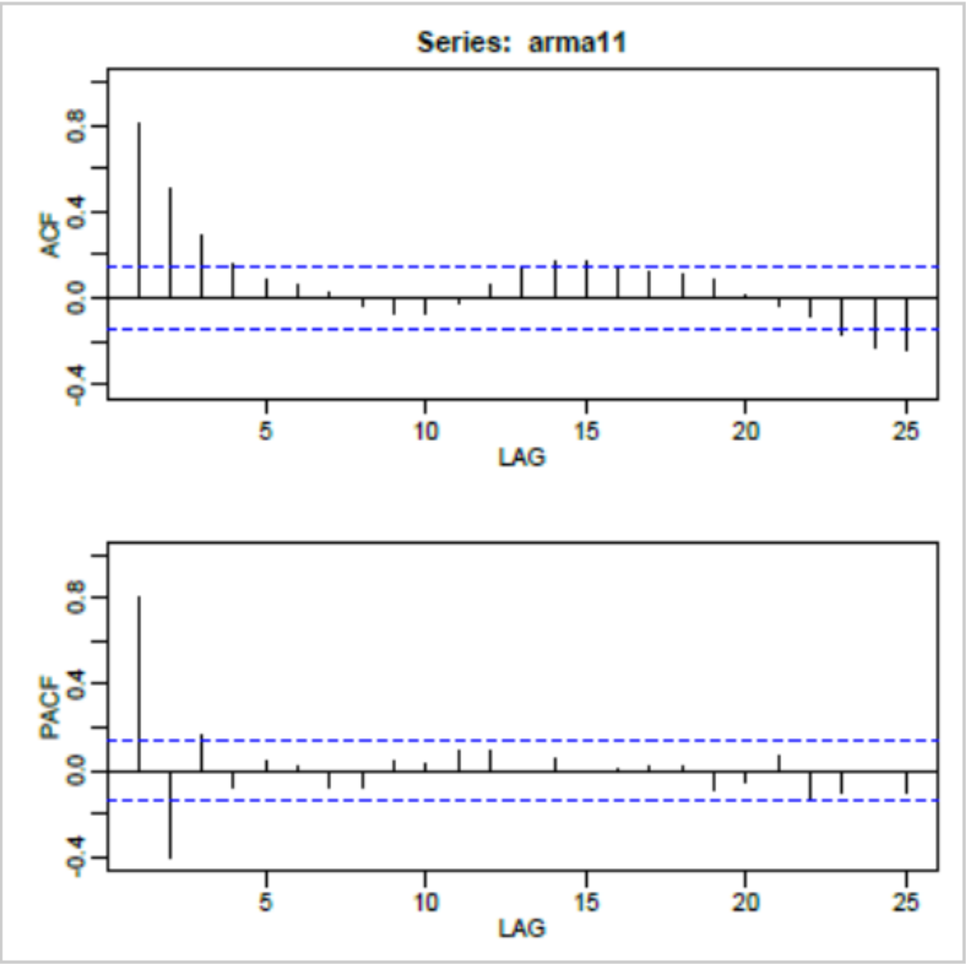
AR(2) Example

Examples



MA(1) Example

Examples



ARMA(1,1) Example

An alternative way to write ARMA is ARIMA. "I" stands for "Integrated". It represents the differencing used to handle non-stationary data (i.e., time series with trends). If the charts on the left are received after one-lag differencing, ARMA(1,1) can be written as ARIMA(1,1,1). If the charts on are received after one-lag differencing twice, ARMA(1,1) can be written as ARIMA(1,2,1)

What If …

- If all autocorrelations are non-significant, then the series is random (white noise; the ordering matters, but the data are independent and identically distributed.)

- If you have taken first differences and all autocorrelations are non-significant, then the series is called a random walk and you are done. (Both the mean and variance are increasing over time.)