Notes: Intro to Time Series and Forecasting – Ch6 Nonstationary and Seasonal Time Series Models

Yingbo Li

05/09/2019

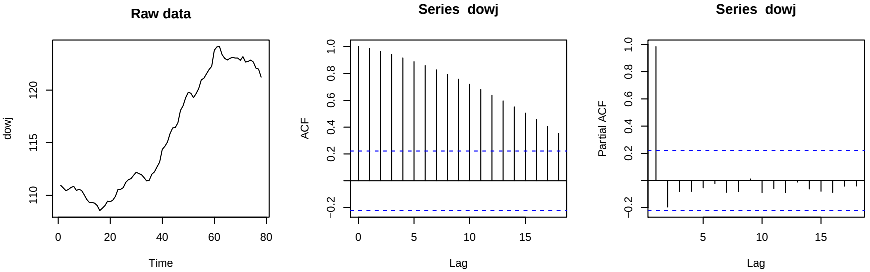
Table of Contents

When data is not stationary

- Implication of not stationary: sample ACF or sample PACF do not rapidly decrease to zero as lag increases
- What shall we do?
  - Differencing, then fit an ARMA $\rightarrow$ ARIMA
  - Transformation, then fit an ARMA
  - Seasonal model $\rightarrow$ SARIMA

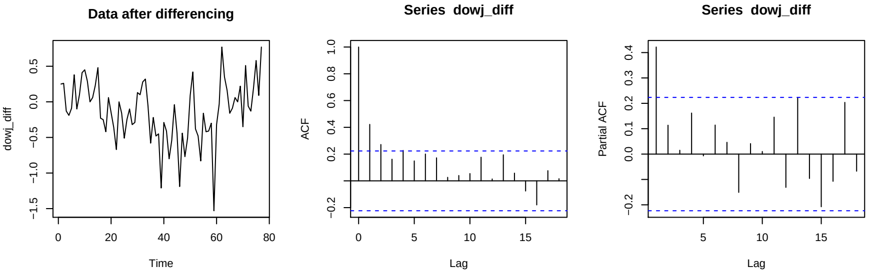
A non-stationary exmaple: Dow Jones utilities index data

```r
library(itsmr); ## Load the ITSM-R package
par(mfrow = c(1, 3));
plot.ts(dowj, main = 'Raw data');
acf(dowj); pacf(dowj);
```



After differencing

```
par(mfrow = c(1, 3));
dowj_diff = dowj[-length(dowj)] - dowj[-1];
plot.ts(dowj_diff, main = 'Data after differencing');
acf(dowj_diff); pacf(dowj_diff);
```



ARIMA model: definition

- Autoregressive integrated moving-average models (ARIMA): Let $d \in \mathbb{N}$, then series $\{X_t\}$ is an ARIMA(p, d, q) process if

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

  is a **causal** ARMA(p, q) process.
- Difference equation (DE) for an ARIMA(p, d, q) process

  $$\phi^*(B) X_t = \phi(B)(1 - B)^d X_t = \theta(B) Z_t, \quad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$$

  - $\phi(z)$: polynomial of degree p, and $\phi(z) \neq 0$ for $|z| \leq 1$
  - $\theta(z)$: polynomial of degree q
  - $\phi^*(z) = \phi(z)(1 - z)^d$: has a zero of order d at $z = 1$
- An ARIMA process with $d > 0$ is NOT stationary!

ARIMA mean is not dertermined by the DE

- $\{X_t\}$ is an ARIMA(p, d, q) process. We can add an arbitrary polynomial trend of degree $d - 1$

$$W_t = X_t + A_0 + A_1 t + \cdots + A_{d-1} t^{d-1}$$

  with $A_0, \ldots, A_{d-1}$ being any random variables, and $\{W_t\}$ still satisfies the same ARIMA(p, d, q) difference equation

- In other words, the ARIMA DE determines the second-order properties of $\{(1 - B)^d X_t\}$ but not those of $\{X_t\}$
  - For parameter estimation: $\boldsymbol{\phi}$, $\boldsymbol{\theta}$, and σ^2 are estimated based on $\{(1 - B)^d X_t\}$ rather than $\{X_t\}$
  - For forecast, we need additional assumptions

Fit data using ARIMA processes

- Whether to fit a finite time series using
  - non-stationary models (such as ARIMA), or
  - directly using stationary models (such as ARMA)?
- If the fitted stationary ARMA model's $\phi(\cdot)$ have zeros very close to unit circles, then fitting an ARIMA model is better
  - Parameter estimation is stable
  - The differenced series may only need a low-order ARMA
- Limitation of ARIMA: only permits data to be nonstationary in a very special way
  - Non-stationary: can have zeros anywhere on the unit circle $|z| = 1$
  - ARIMA model: only has a zero of multiplicity d at the point $z = 1$

Natural log transformation

- When data variance increases with mean, it's common to apply log transformation before fitting the data using ARIMA or ARMA.
- When does log transfomation work well? Suppose that

$$E(X_t) = \mu_t, \quad Var(X_t) = \sigma^2 \mu_t^2$$

Then by first-order Taylor expansion of $\log(X_t)$ at μ_t:

$$\log(X_t) \approx \log(\mu_t) + \frac{X_t - \mu_t}{\mu_t} \implies Var\left[\log(X_t)\right] \approx \frac{Var(X_t)}{\mu_t^2} = \sigma^2$$

The data after log transformation $\log(X_t)$ has a constant variance

- Note: log transformation can only be applied to positive data
- Note: If $Y_t = \log(X_t)$, then because expectation and logarithm are not interchangeable,

$$\hat{X}_t \neq \exp(\hat{Y}_t)$$

Generalize the log transformation: Box-Cox transformation

- Box-Cox transformation

$$f_\lambda(x) = \begin{cases} \frac{x^\lambda - 1}{\lambda}, & x \geq 0, \lambda > 0 \\ \log(x), & x > 0, \lambda = 0 \end{cases}$$

  - Usual range: $0 \leq \lambda \leq 1.5$
  - Common values: $\lambda = 0, 0.5$
- Note: $\lim_{\lambda \to 0} f_\lambda(x) = \log(x)$
- Box-Cox transformation can only be applied to non-negative data

Unit root test for AR(1) process

- $\{X_t\}$ is an AR(1) process: $X_t - \mu = \phi_1(X_{t-1} - \mu) + Z_t$
- Equivalent DE:

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

  – where $\phi_0^* = \mu(1 - \phi_1)$ and $\phi_1^* = \phi_1 - 1$
  – Regressing ∇X_t onto 1 and X_{t-1}, we get the OLS estimator $\hat{\phi}_1^*$ and its standard error $SE(\hat{\phi}_1^*)$
- Augmented Dickey-Fuller test for AR(1)
  – Hypotheses: $H_0 : \phi_1 = 1 \longleftrightarrow H_1 : \phi_1 < 1$
  – Equivalent hypotheses: $H_0 : \phi_1^* = 0 \longleftrightarrow H_1 : \phi_1^* < 0$
  – Test statistic: limit distribution under H_0 is not normal or t

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

  – Rejection region: reject H_0 if

$$\begin{cases} \hat{\tau} < -2.57 & (90\%) \\ \hat{\tau} < -2.86 & (95\%) \\ \hat{\tau} < -3.43 & (99\%) \end{cases}$$

Unit root test for AR(p) process

- AR(p) process: $X_t - \mu = \phi_1(X_{t-1} - \mu) + \cdots + \phi_p(X_{t-p} - \mu) + Z_t$
- Equivalent DE:

$$\nabla X_t = \phi_0^* + \phi_1^* X_{t-1} + \phi_2^* \nabla X_{t-1} + \cdots + \phi_p^* \nabla X_{t-p+1} + Z_t$$

  - where $\phi_0^* = \mu(1 - \sum_{i=1}^p \phi_i)$, $\phi_1^* = \sum_{i=1}^p \phi_i - 1$, and $\phi_j^* = -\sum_{i=j}^p \phi_i$ for $j \geq 2$
  - Regressing ∇X_t onto $1, X_{t-1}, \nabla X_{t-1}, \cdots, \nabla X_{t-p+1}$, we get the OLS estimator $\hat{\phi}_1^*$ and its standard error $SE(\hat{\phi}_1^*)$
- Augmented Dickey-Fuller test for AR(p)
  - Hypotheses: $H_0 : \phi_1^* = 0 \longleftrightarrow H_1 : \phi_1^* < 0$
  - Test statistic:

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

  - Rejection region: same as augmented Dickey-Fuller test for AR(1)

Implement augmented Dickey-Fuller test in R

```r
library(tseries);
## Note: the lag k here is the AR order p
adf.test(dowj, k = 2);
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  dowj
## Dickey-Fuller = -1.3788, Lag order = 2, p-value = 0.8295
## alternative hypothesis: stationary
```

Forecast an ARIMA$(p, 1, q)$ process

- $\{X_t\}$ is ARIMA$(p, 1, q)$, and $\{Y_t = \nabla X_t\}$ is a causal ARMA(p, q)

$$X_t = X_0 + \sum_{j=1}^{t} Y_j, \quad t = 1, 2, \dots$$

- Best linear predictor of X_{n+1}

$$P_n X_{n+1} = P_n(X_0 + Y_1 + \cdots + Y_{n+1}) = P_n(X_n + Y_{n+1}) = X_n + P_n(Y_{n+1}),$$

  - P_n means based on $\{1, X_0, X_1, \dots, X_n\}$, or equivalently, $\{1, X_0, Y_1, \dots, Y_n\}$
  - To find $P_n(Y_{n+1})$, we need to know $E(X_0^2)$ and $E(X_0 Y_j)$, for $j = 1, \dots, n+1$.

- A sufficient assumption for $P_n(Y_{n+1})$ to be the best linear predictor in term of $\{Y_1, \dots, Y_n\}$: X_0 is uncorrelated with $Y_1, Y_2, \dots$

Forecast an ARIMA(p, d, q) process

- The observed ARIMA(p, d, q) process $\{X_t\}$ satisfies

$$Y_t = (1 - B)^d X_t, \quad t = 1, 2, \ldots, \quad \{Y_t\} \sim \text{causal ARMA}(p, q)$$

- Assumption: the random vector $(X_{1-d}, \ldots, X_0)$ is uncorrelated with Y_t for all $t > 0$
- One-step predictors $\hat{Y}_{n+1} = P_n Y_{n+1}$ and $\hat{X}_{n+1} = P_n X_{n+1}$:

$$X_{n+1} - \hat{X}_{n+1} = Y_{n+1} - \hat{Y}_{n+1}$$

- Recall: the h-step predictor of ARMA(p, q) for $n > \max(p, q)$:

$$P_n Y_{n+h} = \sum_{i=1}^{p} \phi_i P_n Y_{n+h-i} + \sum_{j=h}^{q} \theta_{n+h-1,j} (Y_{n+h-j} - \hat{Y}_{n+h-j})$$

- h-step predictor of ARIMA(p, d, q) for $n > \max(p, q)$:

$$P_n X_{n+h} = \sum_{i=1}^{p+d} \phi_i^* P_n X_{n+h-i} + \sum_{j=h}^{q} \theta_{n+h-1,j} (X_{n+h-j} - \hat{X}_{n+h-j})$$

where $\phi^*(z) = (1 - z)^d \phi(z) = 1 - \phi_1^* z - \cdots - \phi_{p+d}^* z^{p+d}$

Seasonal ARIMA (SARIMA) Model: definition

- Suppose d, D are non-negative integers. $\{X_t\}$ is a seasonal ARIMA$(p, d, q) \times (P, D, Q)_s$ process with period s if the differenced series

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

is a causal ARMA process defined by

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)Z_t, \quad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$$

where

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p, \quad \Phi(z) = 1 - \Phi_1 z - \cdots - \Phi_P z^P$$
$$\theta(z) = 1 + \theta z + \cdots + \theta_q z^q, \quad \Theta(z) = 1 + \Theta z + \cdots + \Theta_Q z^Q$$

- $\{Y_t\}$ is causal if and only if neither $\phi(z)$ or $\Phi(z)$ has zeros inside the unit circle
- Usually, $s = 12$ for monthly data

Special case: seasonal ARMA (SARMA)

- Between-year model: for monthly data, each one of the 12 time series is generated by the same ARMA(P, Q) model

$$\Phi(B^{12})Y_t = \Theta(B^{12})U_t, \quad \{U_{j+12t}, t \in \mathbb{Z}\} \sim \mathrm{WN}(0, \sigma_U^2)$$

- SARMA(P, Q) with period s: in the above between-year model, the period 12 can be changed to any positive integer s
  - If $\{U_t, t \in \mathbb{Z}\} \sim \mathrm{WN}(0, \sigma_U^2)$, then the ACVF $\gamma(h) = 0$ unless h divides s evenly. But this may not be ideal for real life application! E.g., this Feb is correlated with last Feb, but not this Jan.
- General SARMA$(p, q) \times (P, Q)$ with period s: incorporate dependency between the 12 series by letting $\{U_t\}$ be ARMA:

$$\Phi(B^s)Y_t = \Theta(B^s)U_t, \quad \phi(B)U_t = \theta(B)Z_t, \quad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$$

  - Equivalent DE for the general SARMA:

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

Fit a SARIMA Model

- Period s is known

1. Find d and D to make the difference series $\{Y_t\}$ to look stationary
2. Examine the sample ACF and sample PACF of $\{Y_t\}$ at lags being multiples of s, to find orders P, Q
3. Use $\hat{\rho}(1), \ldots, \hat{\rho}(s-1)$ to find orders p, q
4. Use AICC to decide among competing order choices
5. Given orders (p, d, q, P, D, Q), estimate MLE of parameters $(\phi, \theta, \Phi, \Theta, \sigma^2)$

Regression with ARMA errors: OLS estimation

- Linear model with ARMA errors $\mathbf{W} = (W_1, \ldots, W_n)'$:

$$Y_t = \mathbf{x}_t'\boldsymbol{\beta} + W_t, \quad t = 1, \ldots, n, \quad \{W_t\} \sim \text{causal ARMA}(p, q)$$

  – Note: each row is indexed by a different time t!
  – Error covariance $\boldsymbol{\Gamma}_n = E(\mathbf{W}\mathbf{W}')$

- Ordinary least squares (OLS) estimator

$$\hat{\boldsymbol{\beta}}_{\text{OLS}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

  – Estimated by minimizing $(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)$
  – OLS is unbiased, even when errors are dependent!

Regression with ARMA errors: GLS estimation

- Generalized least squares (GLS) estimator

$$\hat{\beta}_{\mathsf{GLS}} = (\mathbf{X}'\mathbf{\Gamma}_n^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Gamma}_n^{-1}\mathbf{Y}$$

  – Estimated by minimizing the weighted sum of squares

$$(\mathbf{Y} - \mathbf{X}\beta)'\mathbf{\Gamma}_n^{-1}(\mathbf{Y} - \mathbf{X}\beta)$$

  – Covariance

$$\mathrm{Cov}\left(\hat{\beta}_{\mathsf{GLS}}\right) = (\mathbf{X}'\mathbf{\Gamma}_n^{-1}\mathbf{X})^{-1}$$

  – GLS is the best linear unbiased estimator, i.e., for any vector $\mathbf{c}$ and any unbiased estimator $\hat{\beta}$, we have

$$\mathrm{Var}(\mathbf{c}'\hat{\beta}_{\mathsf{GLS}}) \leq \mathrm{Var}(\mathbf{c}'\hat{\beta})$$

When $\{W_t\}$ is an AR(p) process

- We can apply $\phi(B)$ to each side of the regression equation and get uncorrelated, zero-mean, constant-variance errors

$$\phi(B)\mathbf{Y} = \phi(B)\mathbf{X}\boldsymbol{\beta} + \phi(B)\mathbf{W} = \phi(B)\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}$$

- Under the transformed target variable

$$Y_t^* = \phi(B)Y_t, \quad t = p+1, \ldots, n$$

  and the transformed design matrix

$$X_{t,j}^* = \phi(B)X_{t,j}, \quad t = p+1, \ldots, n, \quad j = 1, \ldots, k$$

  the OLS estimator is the best linear unbiased estimator
- Note: after the transformation, the regression sample size reduces to $n - p$

Regression with ARMA errors: MLE

- MLE of $\boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2$ can be estimated by maximizing the Gaussian likelihood with error covariance $\boldsymbol{\Gamma}_n$
- An iterative scheme
    1. Compute $\hat{\beta}_{\text{OLS}}$ and regression residuals $\mathbf{Y} - \mathbf{X}\hat{\beta}_{\text{OLS}}$
    2. Based on the estimated residuals, compute MLE of the ARMA(p, q) parameters
    3. Based on the fitted ARMA model, compute $\hat{\beta}_{\text{GLS}}$
    4. Compute regression residuals $\mathbf{Y} - \mathbf{X}\hat{\beta}_{\text{GLS}}$, and return to Step 2 until estimators stabilize
- Asymptotic properties of MLE: If $\{W_t\}$ is a causal and invertible ARMA, then
    - MLEs are asymptotically normal
    - Estimated regression coefficients are asymptotically independent of estimated ARMA parameters

References

- Brockwell, Peter J. and Davis, Richard A. (2016), *Introduction to Time Series and Forecasting, Third Edition*. New York: Springer
- Weigt, George (2018) *ITSM-R Reference Manual* http://www.eigenmath.org/itsmr-refman.pdf
- R package `tseries` https://cran.r-project.org/web/packages/tseries/index.html