Notes: Intro to Time Series and Forecasting – Ch1 Introduction

Yingbo Li

01/29/2019

Table of Contents

Objective of time series models

- Seasonal adjustment: recognize seasonal components and remove them to study long-term trends
- Separate (or filter) noise from signals
- Prediction
- Test hypotheses
- Predicting one series from observations of another

A general approach to time series modeling

1. Plot the series and check main features:
   - Trend
   - Seasonality
   - Any sharp changes
   - Outliers
2. Remove trend and seasonal components to get **stationary** residuals
   - May need data transformation first
3. Choose a model to fit the residuals

Definitions: stationary

- Series $\{X_t\}$ has
  - Mean function $\mu_X(t) = E(X_t)$ and
  - Covariance function $\gamma_X(r, s) = \mathrm{Cov}(X_r, X_s)$
- $\{X_t\}$ is **(weakly) stationary** if
  - $\mu_X(t)$ does not depend on t
  - $\gamma_X(t+h, t)$ does not depend on t, for each h
  - (Weakly) stationary is defined based on the first and second order properties of a series
- $\{X_t\}$ is strictly stationary if $(X_1, \ldots, X_n)$ and $(X_{1+h}, \ldots, X_{n+h})$ have the same joint distributions for all integers h and $n > 0$
  - If $\{X_t\}$ is strictly stationary, and $E(X_t^2) < \infty$ for all t, then $\{X_t\}$ is weakly stationary
  - Weakly stationary does not imply strictly stationary

Definitions: autocovariance and autorrelation

- $\{X_t\}$ is a stationary time series
- Autocovariance function (ACVF) of at lag h

$$\gamma_X(h) = \mathrm{Cov}(X_{t+h}, X_t)$$

- Autocorrelation function (ACF) of at lag h

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \mathrm{Cor}(X_{t+h}, X_t)$$

- Note that $\gamma(h) = \gamma(-h)$ and $\rho(h) = \rho(-h)$

Definitions: sample ACVF and sample ACF

$x_1, \ldots, x_n$ are observations of a time series with sample mean $\bar{x}$

- Sample autocovariance function: for $-n < h < n$,

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} \left(x_{t+|h|} - \bar{x}\right)(x_t - \bar{x})$$

  - Use n in the denominator ensures the sample covariance matrix $\hat{\Gamma}_n = [\hat{\gamma}(i-j)]_{i,j=1}^{n}$ is nonnegative definite

- Sample autocorrelation function: for $-n < h < n$,

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

  - Sample correlation matrix $\hat{R}_n = [\hat{\rho}(i-j)]_{i,j=1}^{n}$ is also nonnegative definite

iid noise and white noise

- White noise: uncorrelated, with zero mean and variance σ^2

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

- IID$(0, \sigma^2)$ sequences is WN$(0, \sigma^2)$, but not conversely

Binary process and random walk

- Binary process: an example of iid noise $\{X_t, t = 1, 2, \ldots\}$

$$P(X_t = 1) = p, \quad P(X_t = -1) = 1 - p$$

- Random walk: $\{S_t, t = 0, 1, 2, \ldots\}$, with $S_0 = 0$ and iid noise $\{X_t\}$

$$S_t = X_1 + X_2 + \cdots + X_t, \text{ for } t = 1, 2, \ldots$$

  - $\{S_t\}$ is a simple symmetric random walk if $\{X_t\}$ is a binary process with $p = 0.5$
  - Random walk is not stationary: if $\mathrm{Var}(X_t) = \sigma^2$, then $\gamma_S(t + h, t) = t\sigma^2$ depends on t

First-order moving average, MA(1) process

Let $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$, and $\theta \in \mathbb{R}$, then $\{X_t\}$ is a MA(1) process:

$$X_t = Z_t + \theta Z_{t-1}, \quad t = 0, \pm 1, \ldots$$

- ACVF: does not depend on t, stationary

$$\gamma_X(t+h, t) = \begin{cases} (1+\theta^2)\sigma^2, & \text{if } h = 0, \\ \theta\sigma^2, & \text{if } h = \pm 1, \\ 0, & \text{if } |h| > 1. \end{cases}$$

- ACF:

$$\rho_X(h) = \begin{cases} 1, & \text{if } h = 0, \\ \theta/(1+\theta^2), & \text{if } h = \pm 1, \\ 0, & \text{if } |h| > 1. \end{cases}$$

First-order autoregression, AR(1) process

Let $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$, and $|\phi| < 1$, then $\{X_t\}$ is a AR(1) process:

$$X_t = \phi X_{t-1} + Z_t, \quad t = 0, \pm 1, \ldots$$

- ACVF:

$$\gamma_X(h) = \frac{\sigma^2}{1 - \phi^2} \cdot \phi^{|h|}$$

- ACF:

$$\rho_X(h) = \phi^{|h|}$$

Classcial decomposition

Observation $\{X_t\}$ can be decomposed into

- a (slowly changing) trend component m_t,
- a seasonal component s_t with period d and $\sum_{j=1}^{d} s_j = 0$,
- a zero-mean series Y_t

$$X_t = m_t + s_t + Y_t$$

- Method 1: estimate s_t first, then m_t, and hope the noise component Y_t is stationary (to model)
- Method 2: differencing
- Method 3: trend and seasonality can be estimated together in a regression, whose design matrix contains both polynomial and harmonic terms

Estimate trend: polynomial regression fitting

Observation $\{X_t\}$ can be decomposed into a trend component m_t and a zero-mean series Y_t:

$$X_t = m_t + Y_t$$

- Least squares polynomial regression

$$m_t = a_0 + a_1 t + \cdots + a_p t^p$$

Estimate trend: smoothing with a finite MA filter

- Linear filter

$$\hat{m}_t = \sum_{j=-\infty}^{\infty} a_j X_{t-j}$$

- Two-sided moving average filter, with $q \in \mathbb{N}$

$$W_t = \frac{\sum_{j=-q}^{q} X_{t-j}}{2q+1}$$

  - $W_t \approx m_t$ for $q + 1 \leq t \leq n - q$, if X_t only has the trend component m_t but not seasonality s_t, and m_t is approximately linear in t
  - W_t is a low-pass filter: remove the rapidly fluctuating (high frequency) component Y_t, and let the slowly varying component m_t pass

Estimate trend: exponential smoothing

For any fixed $\alpha \in [0, 1]$, the one-sided MA $\hat{m}_t : t = 1, \ldots, n$ defined by recursions

$$\hat{m}_t = \begin{cases} X_1, & \text{if } t = 1 \\ \alpha X_t + (1-\alpha)\hat{m}_{t-1}, & \text{if } t = 2, \ldots, n \end{cases}$$

- Equivalently,

$$\hat{m}_t = \sum_{j=0}^{t-2} \alpha(1-\alpha)^j X_{t-j} + (1-\alpha)^{t-1} X_1$$

Eliminate trend by differencing

- Backward shift operator

$$BX_t = X_{t-1}$$

- Lag-1 difference operator

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

  – If ∇ is applied to a linear trend function $m_t = c_0 + c_1 t$, then $\nabla m_t = c_1$

- Powers of operators B and ∇:

$$B^j(X_t) = X_{t-j}, \quad \nabla^j(X_t) = \nabla\left[\nabla^{j-1}(X_t)\right] \text{ with } \nabla^0(X_t) = X_t$$

  – ∇^k reduces a polynomial trend of degree k to a constant

$$\nabla^k\left(\sum_{j=0}^{k} c_j t^j\right) = k! c_k$$

Estimate seasonal component: harmonic regression

Observation $\{X_t\}$ can be decomposed into a seasonal component s_t and a zero-mean series Y_t:

$$X_t = s_t + Y_t$$

- s_t: a periodic function of t with period d, i.e., $s_{t-d} = s_t$
- Harmonic regression: a sum of harmonics (or sine waves)

$$s_t = a_0 + \sum_{j=1}^{k} \left[a_j \cos\left(\lambda_j t\right) + b_j \sin\left(\lambda_j t\right)\right]$$

- Unknown (regression) parameters: a_j, b_j
- Specified parameters:
  - Number of harmonics: k
  - Frequencies λ_j, each being some integer multiple of $\frac{2\pi}{d}$
  - Sometimes λ_j are instead specified through Fourier indices $f_j = \frac{n \cdot j}{d}$

Estimate trend and seasonal components

1. Estimate $\hat{m}_t$: use a MA filter chosen to elimate the seasonality
   - If d is odd, let $d = 2q$
   - If d is even, let $d = 2q$ and

$$\hat{m}_t = (0.5x_{t-q} + x_{t-q+1} + \cdots + x_{t+q-1} + 0.5x_{t+q})/d$$

2. Estimate $\hat{s}_t$: for each $k = 1, \ldots, d$
   - Compute the average $w_k = \operatorname{avg}_j(x_{k+jd} - \hat{m}_{k+jd})$
   - To ensure $\sum_{k=1}^{d} s_k = 0$, let $\hat{s}_k = w_k - \bar{w}$, where $\bar{w} = \sum_{k=1}^{d} w_k/d$
3. Re-estimate $\hat{m}_t$: based on the deseasonalized data

$$d_t = x_t - \hat{s}_t$$

Eliminate trend and seasonal components: differencing

- Lag-d differencing

$$\nabla_d X_t = X_t - X_{t-d} = (1 - B^d) X_t$$

  - Note: the operators ∇_d and $\nabla^d = (1 - B)^d$ are different

- Apply ∇_d to $X_t = m_t + s_t + Y_t$

$$\nabla_d X_t = m_t - m_{t-d} + Y_t - Y_{t-d}$$

  - Then the trend $m_t - m_{t-d}$ can be eliminated using methods discussed before, e.g., applying a power of the operator ∇

Test series $\{Y_1, \ldots, Y_n\}$ for iid: sample ACF based

Test name	Test statistic	Distribution under H_0
Sample ACF	$\hat\rho(h)$, for all $h \in \mathbb{N}$	$\mathrm{N}(0, 1/n)$
Portmanteau	$Q = n \sum_{j=1}^{h} \hat\rho^2(j)$	$\chi^2(h)$

- Under H_0, about 95% of the sample ACFs should fall between $\pm 1.96\sqrt{n}$
- The Portmanteau test has some refinements
  - Ljung and Box $Q_{LB} = n(n+2) \sum_j \hat\rho^2(j)/(n-j)$
  - McLeod and Li $Q_{ML} = n(n+2) \sum_j \hat\rho^2_{WW}(j)/(n-j)$, where $\hat\rho^2_{WW}(h)$ is the sample ACF of squared data

Test series $\{Y_1, \ldots, Y_n\}$ for iid: also detect trends

Test name	Test statistic	Distribution under H_0
Turning point	T: number of turning points	$\mathrm{N}(\mu_T, \sigma_T^2)$
Difference-sign	S: number of i that $y_i > y_{i-1}$	$\mathrm{N}(\mu_S, \sigma_S^2)$

- Time i is a turning point, if $y_i - y_{i-1}$ and $y_{i+1} - y_i$ have flipped signs
  - $\mu_T = 2(n-2)/3 \approx 2/3$
- A large positive (or negative) value of $S - \mu_S$ indicates increasing (or decreasing) trend
  - $\mu_S = (n-1)/2 \approx 1/2$

Test series $\{Y_1, \ldots, Y_n\}$ for iid: other methods

- Fitting an AR model
  - Using Yule-Walker algorithm and choose order using AICC statistic
  - If the selected order is zero, then the series is white noise
- Normal qq plot: check of normality
- A general strategy is to check all above mentioned tests, and proceed with caution if any of them suggests not iid

References

- Brockwell, Peter J. and Davis, Richard A. (2016), *Introduction to Time Series and Forecasting, Third Edition*. New York: Springer