Notes for Time Series Data Analysis

Oliver Y. Chén

Last update: August 9, 2016

Introduction

The notes are summarized and organized from the tremendously wonderful textbook (Brockwell and Davis, 2006) in time series analysis . We recommend our readers to use the notes in accordance with the book. Advanced theoretical properties and mathematical derivations can be found in (Brockwell and Davis, 2013).

The study of time series data is essential in analysing brain functional data over time; it also provides insight in studying structual brain data, as the temporal space could be easily adapated to spatial space. Readers could refer to *Time Series Modeling of Neuroscience Data* (Ozaki, 2012) for further reference.

Finally, we have to confess that the notes include only the most important building blocks that we think are helpful for elementary readers, and are, inevatably, subjective.

1 Basic Definitions

- **Time series**: a time series of a set of observations x_t, each one being recorded at a specific time t.

- A **time series model** for the observed data $\{x_t\}$ (a realization of a sequence of random variable $\{X_t\}$) is a specification of the joint distributions (means, covariance, etc.) of

$\{X_t\}$.

- Define the **mean function** of $\{X_t\}$ as $\mu_X(t) = \mathbb{E}(X_t)$; then the **covariance function** of $\{X_t\}$ is $\gamma_X(r,s) = Cov(X_r, X_s) = \mathbb{E}\big(X_r - \mu_X(r)\big)\big(X_s - \mu_X(s)\big)$, for all r and s in $\mathbb{Z}$.

- **Autocovariance function** (ACVF) of X_t at **lag** h is $\gamma_X(h) = Cov(X_{t+h}, X_t\}$.

- **Autocorrelation function** (ACF) of X_t at **lag** h is $\rho(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = Cor(X_{t+h}, X_t)$.

- **(Weakly) stationary**: $\{X_t\}$ is weakly stationary if (a) $\mu_X(t) \perp t$, and (b) $\gamma_X(t+h,t) \perp t$, $\forall h$.

- **(Strictly) stationary**: $\{X_t\}$ is strictly stationary if $F(X_1, \ldots, X_n) = F(X_{1+h}, \ldots, X_{n+h})$, for all $h \in \mathbb{Z}$ and $n \geq 1$. Here $F(\cdot)$ indicates the joint distribution function.

2 Linear Processes

Definition 1. *$\{X_t\}$ is a **linear process** if*

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$$
$$= \psi(B) Z_t,$$

for all t, where $Z_{t-j} \sim WN(0, \sigma^2)$, $\{\psi_j\}$ is a sequence of constants with $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$, $\psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$ is a linear filter (that is applied to the white noise "input" $\{Z_t\}$ and produces an "output" $\{X_t\}$), and B^j is the backward shift operator such that $B^j X_t = X_{t-j}$, shifting X_t back to X_{t-j}.

2.1 Moving Average Models

Definition 2. *A **linear process** $\{X_t\}$ is called a **moving average** if*

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}.$$

Definition 3. *A **linear process** $\{X_t\}$ is called an **MA(q)** if*

$$X_t = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$$
$$= \theta(B) Z_t,$$

where $\theta(B) = 1 + \theta_1 B + \ldots + \theta_q B^q$ with B as the backshifting operator.

2.2 Autoregressive Models

Definition 4. *A **linear process** $\{X_t\}$ is called an **AR(p)** if*

$$X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = Z_t$$
$$= \phi(B) X_t,$$

where $\phi(B) = 1 - \phi_1 B - \ldots - \phi_q B^q$ with B as the backshifting operator.

2.3 Autoregressive Moving Average Models

Definition 5. *A **linear process** $\{X_t\}$ is called an **ARMA(p,q)** if*

$$X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q},$$

which, for simplicity, can be written as

$$\phi(B) X_t = \theta(B) Z_t, \tag{1}$$

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

2.4 Properties

Definition 6. *An **ARMA(p,q)** process is **causal** (function) of $\{z_t\}$ if there exist constants $\{\psi_j\}$ such that $\sum_{j=0}^{\infty} |\psi_j| < \infty$ and*

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \forall t.$$

Causality is equivalent to the condition

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_z z^p \neq 0, \forall\, |z| \leq 1.$$

3 Nonlinear Processes

3.1 Bilinear Models

Definition 7. *A* ***bilinear model*** *of order* (p, q, r, s) *is defined by*

$$X_t = Z - t + \sum_{i=1}^{p} a_i X_{t-i} + \sum_{j=1}^{q} b_j Z_{t-j} + \sum_{i=1}^{r} \sum_{j=1}^{s} c_{ij} X_{t-i} Z_{t-j}$$

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

3.2 Autoregressive Models with Random Coefficients

Definition 8. *A* ***random coefficients autoregressive process*** $\{X_t\}$ *of order* p *is defined by*

$$X_t = \sum_{i=1}^{p} (\phi_i + U_t^{(i)}) X_{t-i} + Z_t,$$

where $\{Z_t\} \sim iid(0, \sigma^2)$, $U_t^{(i)} \sim iid(0, \nu^2)$, $\{Z_t\}$ *is independent of* U_t, *and* $\phi_1, \ldots, \phi_p \in \mathbb{R}$.

3.3 Threshold Models

Definition 9. ***Threshold models*** *are piecewise linear models where the linear relationship varies with the values (which are realizations of different field partitions) of the process. For example, if* $R^{(i)}, i = 1, \ldots, k$, *is a partition of* $\mathbb{R}^p$, *and* $\{Z_t\} \sim iid(0, 1)$, *then the* k *difference equations*

$$X_t = \sigma^{(i)} Z_t + \sum_{j=1}^{p} \phi_j^{(i)} X_{t-j},$$

where $X_{t-1}, \ldots, X_{t-p} \in \mathbb{R}^{(i)}$, *for* $i = 1, \ldots, k$, *define a threshold* $AR(p)$ *model.*

3.4 ARCH(p) and GARCH(p,q) Models

When time series are "less predictable" (or more "volatile"), such as financial time series, depending on the past history of the series, the predictability (i.e. the size of the presdiction mean squared error) is dependent on the past of the series. Nonlinear processes such as ARCH and GARCH models account for circumstances where past histories may permit more accurate forcasting (whereas linear models do not) by considering the dependence of the conditional variance of the process on its past history.

Definition 10. *A* ***autoregressive conditional heteroscedasticity model****, or* ***ARCH(p) model*** $\{Z_t\}$ *is defined as*

$$Z_t = \sqrt{h_t} e_t, \{e_t\} \sim iid \;\; N(0,1),$$

where

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Z_{t-i}^2$$

with $\alpha_0 > 0$ *and* $\alpha_j \geq 0$ *for* $j = 1, \ldots, p$*. Note that* h_t *is the conditional variance of* Z_t *given* $\{Z_s, s < t\}$.

Definition 11. *A* ***generalized autoregressive conditional heteroscedasticity model****, or* ***GARCH(p,q) model*** $\{Z_t\}$ *is defined as*

$$Z_t = \sqrt{h_t} e_t, \{e_t\} \sim iid \;\; N(0,1),$$

where

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Z_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}^2$$

with $\alpha_0 > 0$ *and* $\alpha_j, \beta_j \geq 0$ *for* $j = 1, \ldots, p$*.*

4 Modeling and Forecasting with ARMA Process

For $AR(p)$ models, we choose *Yule-Walker* and *Burg* estimation; for $ARMA(p,q)$, we use *Innovating* and *Hanna-Rissanen* algorithms.

Definition 12. *Yule-Walker Estimtion*

For a pure autoregressive model, we may write X_t *in the form*

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \tag{2}$$

where $\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{1}{\phi(z)}$.

Multiply each side of (1) by X_{t-j} $\forall j = 0, 1, 2, \ldots, p$, taking expectation, where the right-hand side (RHS) of (1) is evaluated using RHS of (2), we have the *Yule-Walker* equations

$$\Gamma_p \boldsymbol{\phi} = \boldsymbol{\gamma}_p$$

and

$$\sigma^2 = \gamma(0) - \boldsymbol{\phi}' \boldsymbol{\gamma}_p,$$

where $\Gamma_p = [\gamma(i-j)]_{i,j=1}^{p}$ is the covariance matrix, $\boldsymbol{\gamma}_p = (\gamma(1), \ldots, \gamma(p))'$, and $\gamma(h) = \mathbb{E}(X_{t+h}X_t) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+|h|}$ (see 3.2.3, Brockwell and Davis (2006)).

Using their sample estimates, and rearrange, we have:

$$\hat{\phi} = (\hat{\phi}_1, \ldots, \hat{\phi}_p)' = \hat{R}_p^{-1} \hat{\boldsymbol{\rho}}_p$$

and

$$\hat{\sigma}^2 = \hat{\gamma}(0) \left[1 - \hat{\boldsymbol{\rho}}_p' \hat{R}_p^{-1} \hat{\boldsymbol{\rho}}_p\right]$$

where $\hat{\boldsymbol{\rho}}_p = (\hat{\rho}(1), \ldots, \hat{\rho}(p))' = \frac{\hat{\boldsymbol{\gamma}}_p}{\hat{\gamma}(0)}$.

Finally, when the sample size is sufficiently large, the asymptotic distribution of *Yule-Walker* estimators is

$$\hat{\boldsymbol{\phi}} \sim N(\boldsymbol{\phi}, n^{-1}\sigma^2\Gamma_p^{-1}).$$

References

Brockwell, P. J. and R. A. Davis (2006). *Introduction to time series and forecasting*. Springer Science & Business Media.

Brockwell, P. J. and R. A. Davis (2013). *Time series: theory and methods*. Springer Science & Business Media.

Ozaki, T. (2012). *Time series modeling of neuroscience data*. CRC Press.