Time-series Analysis Notes

Nong Minh Hieu[1]

[1] School of Physical and Mathematical Sciences, Nanyang Technological University (NTU - Singapore)

Contents

1 Introduction

1.1 Time-series definitions

Definition 1.1 (Time series).

A time-series is a sequence of observations of time-indexed random variables $(X_t)_{t\geq 0}$. Typically, A time-series consists of three components.

- *Trend* (T_t).
- *Cycle* (C_t).
- *Seasonal* (S_t).

Finally, there is a random error (residual) term (R_t). Note that for the sake of this course, we will hardly consider the cycle (C_t) term.

Definition 1.2 (Multiplicative & Additive models).

*The components of a time-series can be combined either **additively** or multiplicatively:*

- ***Additive model*** : $X_t = T_t + C_t + S_t + R_t$.
- ***Multiplicative model*** : $X_t = T_t \times C_t \times S_t \times R_t$.

1.2 Trend model

Definition 1.3 (Trend model).

A trend model is defined as:

$$X_t = T_t + R_t$$

Remark : Consider the linear trend:

$$T_t = \beta_0 + \beta_1 t, \quad \beta_1 \neq 0$$

We can then estimate the time-series $(X_t)_{t\geq 0}$ using the following methods:

- **Method 1** : Least square estimate the trend component T_t by minimizing the sum squared error

$$SSE = \sum_{t\geq 0}(X_t - T_t)^2$$

- **Method 2** : Eliminate the trend component via differencing operator.

Definition 1.4 (Differencing operator).

We define the first difference operator ∇ by:

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

Where I is the identity operator and B is the backward shift operator:

$$BX_t = X_{t-1},\ B^i X_t = X_{t-i}$$

The i^{th} order differencing operator is:

$$\nabla^i X_t = \nabla(\nabla^{i-1} X_t)$$

Note that $\nabla^0 X_t = I$.

Proposition 1.1: Binomial formula for $\nabla^n X_t$

Given a time-series $(X_t)_{t\geq 0}$, we can write the differenced data to the n^{th} order as followed:

$$\nabla^n X_t = \sum_{k=0}^{n} \binom{n}{k} (-1)^k B^k X_t = \sum_{k=0}^{n} \binom{n}{k} (-1)^k X_{t-k}$$

As a result, we can write:

$$X_t = \nabla^n X_t - \sum_{k=1}^{n} \binom{n}{k} (-1)^k X_{t-k}$$

Proof (Proposition 1.1).

Using the Binomial theorem, we have:

$$\begin{aligned}
\nabla^n X_t &= (I - B)^n X_t \\
&= \sum_{k=0}^{n} \binom{n}{k} (I)^{n-k} (-B)^k X_t \quad (\textit{Binomial theorem}) \\
&= \sum_{k=0}^{n} \binom{n}{k} (-1)^k B^k X_t \\
&= \sum_{k=0}^{n} \binom{n}{k} (-1)^k X_{t-k} \\
&= X_t + \sum_{k=1}^{n} \binom{n}{k} (-1)^k X_{t-k} \\
\implies X_t &= \nabla^n X_t - \sum_{k=1}^{n} \binom{n}{k} (-1)^k X_{t-k}
\end{aligned}$$

$\square$.

1.3 Trend + Seasonality model

Definition 1.5 (Trend + Seasonality model).

The model that displays both trend and seasonality is defined as:

$$X_t = T_t + S_t + R_t$$

There are two types of seasonality:

- ***Constant seasonality*** *: degree of seasonality does not depend on time.*
- ***Varying (increasing) seasonality*** *: degree of seasonality depends on time.*

Remark : When the time-series displays increasing seasonality, it is a common practice to transform the data so that the series display constant seasonality.

Definition 1.6 (Box-Cox transformation).
For some $\lambda > 0$, the Box-Cox transformation is defined as:

$$z_t = \frac{x_t^\lambda - 1}{\lambda}$$

Proposition 1.2: Box-Cox as $\lambda \to 0$

As $\lambda \to 0$, $z_t \to \log(x_t)$.

Proof (Proposition 1.2).
We can write:

$$\begin{aligned}
\lim_{\lambda \to 0} e^{z_t} &= \lim_{\lambda \to 0} \lim_{n \to \infty} \left(1 + \frac{z_t}{n}\right)^n \\
&= \lim_{\lambda \to 0} \left(1 + \lambda z_t\right)^{1/\lambda} \quad \left(\textit{Letting } n = \frac{1}{\lambda}\right) \\
&= \lim_{\lambda \to 0} \left(1 + (x^\lambda - 1)\right)^{1/\lambda} \quad = x_t \\
\implies \lim_{\lambda \to 0} z_t &= \log(x_t)
\end{aligned}$$

□.

1.4 Modelling seasonality using dummy variables

Definition 1.7 (Dummy variables for seasonality).
The seasonal factor (supposing that there are L seasons) expressed using dummy variables can be written as:

$$S_t = \beta_1^{(s)} D_1(t) + \cdots + \beta_{L-1}^{(s)} D_{L-1}(t)$$

Where we have:

$$D_k(t) = \begin{cases} 1, & \textit{if } t \textit{ belongs to season } k \\ 0, & \textit{otherwise} \end{cases}$$

We can write the time-series whose seasonality is modelled using dummy variables as:

$$X_t = T_t + \sum_{k=1}^{L-1} \beta_k^{(s)} D_k(t) + R_t$$

Remark : Take a time-series with linear trend and seasonality modelled using dummy variables as an example. Suppose that we have 4 seasons, the time-series will take the form:

$$X_t = \underbrace{\beta_0 + \beta_1 t}_{T_t} + \underbrace{\beta_1^{(s)} D_1(t) + \beta_2^{(s)} D_2(t) + \beta_3^{(s)} D_3(t)}_{S_t} + R_t$$

We can estimate the parameters $\beta = \left(\beta_0, \beta_1, \beta_1^{(s)}, \beta_2^{(s)}, \beta_3^{(s)}\right)$ the same way we do for linear regression models:

$$\hat{\beta} = \left(X^T X\right)^{-1} \left(X^T y\right), \quad X \in \mathbb{R}^{T \times 5}$$

Where T is the length of the time-series.

1.5 Autocorrelation & autocorrelation function (ACF)

Definition 1.8 (Autocovariance & Autocorrelation functions).
Given a time-series $(X_t)_{t\geq 0}$, the autocovariance $\gamma(h)$ and autocorrelation $\rho(h)$ are defined as followed:

$$\gamma(h) = Cov\Big(X_{t+h}, X_t\Big)$$
$$\rho(h) = \frac{\gamma(h)}{Var(X_t)} = \frac{\gamma(h)}{\gamma(0)}$$

When we are given a set of observations $(x_t)_{t=1}^n$, we can estimate γ, ρ by using their sample counterparts:

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (x_{t+h} - \overline{x})(x_t - \overline{x})$$
$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{n-h} (x_{t+h} - \overline{x})(x_t - \overline{x})}{\sum_{t=1}^{n} (x_t - \overline{x})^2}$$

We have the following useful identities:

- $\gamma_0 = Var(X_t)$.
- $\gamma_k = \gamma_{-k}$ *(symmetric).*
- $\rho_0 = 1$.
- $\rho_k = \rho_{-k}$ *(symmetric).*

Definition 1.9 (Autocorrelation function (ACF) plot).
The autocorrelation function plot, or correlogram is the plot of sample autocorrelation coefficients $\hat{\rho}(h)$ against $h = 0, 1, 2, \ldots$.

2 Stationary time-series

2.1 Strict vs. Weak stationarity

Definition 2.1 (Mean, covariance and variance functions).
Given a time-series $(X_t)_{t\in T}$ (where T is a time domain), we have the following functions:

- *Mean :* $\mu(t) = \mathbb{E}[X_t],\ t \in T$.
- *Covariance :* $\gamma(t,s) = Cov(X_t, X_s),\ t,s \in T$.
- *Variance :* $\sigma^2(t) = \gamma(t,t) = Var(X_t),\ t \in T$.

Definition 2.2 (Strict stationarity).
A time-series is said to be ***strictly stationary*** *if,* $\forall n \in \mathbb{Z}_+, h \in \mathbb{N}$, *we have:*

$$(X_1, \dots, X_n) \overset{d}{\simeq} (X_{1+h}, \dots, X_{n+h})$$

Meaning the two tuples are equal in distribution.

Remark : Note that in practice, strict stationarity is often hard to achieve. Hence, we have some relaxed conditions for stationarity.

Definition 2.3 (Weak stationarity).
A time-series $(X_t)_{t\geq 0}$ is said to be ***weakly stationary*** *if the following two conditions are met:*

- $\mu(t)$ *does not depend on* t.
- $\gamma(t,h)$ *does not depend on t for all* $h \in \mathbb{N}$.

2.2 Moving Average - MA(q) model

Definition 2.4 (White noise).
A process $(Z_t)_{t\in T}$ is called ***white noise*** *if it satisfies:*

- $Cov(Z_t, Z_s)$ *for all* $s \neq t$.
- $Var(Z_t) = \sigma^2,\ \forall t \in T$.

We denote that $Z_t \sim WN(0, \sigma^2)$.

Definition 2.5 (**MA(q)** model).
The moving average of order **q** *is defined as followed:*

$$X_t = Z_t + \sum_{j=1}^{q} \theta_j Z_{t-j},\ Z_t \sim WN(0, \sigma^2)$$

Remark : The following propositions are proven to provide the insight into the properties of the $MA(a)$ model.

Proposition 2.1: Stationarity of MA(q)

The moving average model **MA(q)** is stationary. Specifically, for a time-series X_t that follows the **MA(q)** model, we have:

- $\mathbb{E}[X_t] = 0$.
- $\gamma(k) = \left(\theta_k + \sum_{j=1}^{q-k} \sum_{i=k+j}^{q} \theta_i\theta_j\right)\sigma^2$.

Proof (Proposition 2.1).

To calculate the mean of X_t, we have:

$$\mathbb{E}[X_t] = \mathbb{E}\left[Z_t + \sum_{j=1}^{q} \theta_j Z_{t-j}\right] = \mathbb{E}[Z_t] + \sum_{j=1}^{q} \theta_j \mathbb{E}[Z_{t-j}] = 0$$

To calculate the covariance, we have:

$$\begin{aligned}
\gamma(k) &= Cov\left(Z_t + \sum_{i=1}^{q} \theta_i Z_{t-i}, Z_{t-k} + \sum_{j=1}^{q} \theta_j Z_{t-k-j}\right) \\
&= \underbrace{Cov(Z_t, Z_{t-k})}_{=0} + \underbrace{\sum_{j=1}^{q} \theta_j Cov(Z_t, Z_{t-k-j})}_{=0} + \sum_{i=1}^{q} \theta_i Cov(Z_{t-i}, Z_{t-k}) \\
&+ \sum_{i=1}^{q} \sum_{j=1}^{q} \theta_i\theta_j Cov(Z_{t-i}, Z_{t-k-j})
\end{aligned}$$

When $k > q$, we know that $Cov(Z_{t-i}, Z_{t-k-j}) = 0$ and $Cov(Z_{t-i}, Z_{t-k}) = 0$ for all $1 \le i, j \le q$. Therefore, suppose that $1 \le q \le k$, we have:

$$\begin{aligned}
\gamma(k) &= \sum_{i=1}^{q} \theta_i Cov(Z_{t-i}, Z_{t-k}) + \sum_{i=1}^{q} \sum_{j=1}^{q} \theta_i\theta_j Cov(Z_{t-i}, Z_{t-k-j}) \\
&= \theta_k Cov(Z_{t-k}, Z_{t-k}) + \sum_{j=1}^{q-k} \sum_{i=k+j}^{q} \theta_i\theta_j Cov(Z_{t-k-j}, Z_{t-k-j}) \\
&= \sigma^2 \left(\theta_k + \sum_{j=1}^{q-k} \sum_{i=k+j}^{q} \theta_i\theta_j\right)
\end{aligned}$$

Remark *: When we set $k = 0$, we obtain the variance and the correlation functions for X_t:*

$$\begin{aligned}
Var(X_t) &= \sigma^2 \left(1 + \sum_{j=1}^{q} \theta_j^2\right) \\
\implies \rho(k) &= \frac{\theta_k + \sum_{j=1}^{q-k} \sum_{i=k+j}^{q} \theta_i\theta_j}{1 + \sum_{j=1}^{q} \theta_j^2}, \quad 1 < k \le q
\end{aligned}$$

$\square$.

Definition 2.6 (**MA**(∞) model)**.**
*Assume that $Z_t \sim WN(0, \sigma^2)$ and $\psi_0 = 1$. A time-series X_t following the **MA**(∞) model is defined as:*

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad \sum_{j=0} |\psi_j| < \infty$$

*X_t is called a **Generalized linear process** or an **MA**(∞) process.*

2.3 Autoregressive - AR(p) model

Definition 2.7 (**AR**(**p**) model)**.**
*The autoregressive model of order **p** is defined as followed:*

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

Defining the p^{th} degree polynomial:

$$\phi(z) = 1 - \sum_{i=1}^{p} \phi_i z^i$$

*We can re-write the **AR**(**p**) model as followed:*

$$Z_t = \phi(B) X_t, \quad B^i X_t = X_{t-i}$$

Theorem 2.1: Yule-Walker Equation

Let $(X_t)_{t \geq 0}$ follows the **AR**(**p**) model. Let $k \geq 1$, the Yule-Walker equations are defined as:

$$\rho_k = \phi_1 \rho_{k-1} + \cdots + \phi_p \rho_{k-p} = \sum_{j=1}^{p} \phi_j \rho_{k-j}$$

Proof (Theorem 2.1)**.**
We have:

$$X_t = Z_t + \sum_{j=1}^{p} \phi_j X_{t-j}$$

Recall the fact that:

$$Cov(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Hence, we have:

$$
\begin{aligned}
\gamma_k &= Cov(X_t, X_{t-k}) \\
&= \mathbb{E}[X_t X_{t-k}] - \mathbb{E}[X_t]\mathbb{E}[X_{t-k}] \quad (\mathbb{E}[X_t] = 0, \quad \forall t \geq 0) \\
&= \mathbb{E}[X_t X_{t-k}] \\
&= \mathbb{E}\left[Z_t X_{t-k} + \sum_{j=1}^{p} \phi_j X_{t-j} X_{t-k}\right] \\
&= \underbrace{\mathbb{E}[Z_t X_{t-k}]}_{=0} + \sum_{j=1}^{p} \phi_j \mathbb{E}[X_{t-j} X_{t-k}] \\
&= \sum_{j=1}^{p} \phi_j \gamma_{k-j}
\end{aligned}
$$

Dividing both sides with γ_0, we have:

$$
\rho_k = \sum_{j=1}^{p} \phi_j \rho_{k-j}
$$

$\square$.

Theorem 2.2: Stationarity of AR(1)

AR(1) is stationarity if all the roots of the polynomial $\phi(z)$ are outside the unit circle. Furthermore, the ACF decays exponentially.

Proof (Theorem 2.2).

Consider the **AR(1)** *model, we have:*

$$
X_t = \phi_1 X_{t-1} + Z_t, \;\; Z_t \sim WN(0, \sigma^2)
$$

Then, we have:

$$
\phi(z) = 1 - \phi_1 z
$$

Hence, we have the root $z = \frac{1}{\phi_1}$. We will show that X_t is stationary if and only if $|\phi_1| < 1$. Hence, $|z| > 1$ and the root lies outside the unit circle. We have:

$$
\begin{aligned}
X_1 &= \phi_1 X_0 + Z_1 \\
X_2 &= \phi_1(\phi_1 X_0 + Z_1) + Z_2 = \phi_1^2 X_0 + \phi_1 Z_1 + Z_2 \\
&\;\;\ldots \\
X_t &= \phi_1^t X_0 + \sum_{j=1}^{t} \phi_1^j Z_{t-j}
\end{aligned}
$$

Hence, if $|\phi_1| > 1 \implies \mathbb{E}[X_t] = \phi_1^t \mathbb{E}[X_0]$, which is dependent on t. On the other hand, $|\phi_1| < 1 \implies \mathbb{E}[X_t] \to 0$ as $t \to \infty$. Therefore, X_t is stationary.
To prove that the ACF decays exponentially, by the Yule-Walker equations, we have:

$$
\begin{aligned}
\rho_k &= \phi_1^k \rho_0 = \phi_1^k \\
\implies |\rho_k| &= |\phi_1|^k = e^{k \log|\phi_1|}
\end{aligned}
$$

Hence, when $|\phi_1| < 1$, $|\rho_k|$ decays exponentially. $\square$.

Theorem 2.3: Stationarity of $AR(p)$

$\mathbf{AR(p)}$ is stationary if all the roots of the polynomial $\phi(z)$ are outside the unit circle.

Proof (Theorem 2.3).

We will revise the proof for this theorem later. $\square$.

A List of Definitions

B Important Theorems

C Important Corollaries

D Important Propositions

E References

References

[1] Rick Durrett. *Probability: Theory and Examples*. 4th. USA: Cambridge University Press, 2010. ISBN: 0521765390.

[2] Erhan undefinedinlar. *Probability and Stochastics*. Springer New York, 2011. ISBN: 9780387878591. DOI: 10.1007/978-0-387-87859-1. URL: http://dx.doi.org/10.1007/978-0-387-87859-1.

[3] Wikipedia. *Vitali set — Wikipedia, The Free Encyclopedia*. http://en.wikipedia.org/w/index.php?title=Vitali%20set&oldid=1187241923. [Online; accessed 24-December-2023]. 2023.