Forecasting Models: Time Series Analysis
Final report

Mukhtar Urynbassarov

January 23, 2020

Abstract

Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data.

Time series forecasting is the use of a model to predict future values based on previously observed values.

While regression analysis is often employed in such a way as to test theories that the current values of one or more independent time series affect the current value of another time series, this type of analysis of time series is not called "time series analysis", which focuses on comparing values of a single time series or multiple dependent time series at different points in time. Interrupted time series analysis is the analysis of interventions on a single time series.

Abbreviations

- TSA — Time Series Analysis
- MA — Moving Average
- WMA — Weighted Moving Average
- ES — Exponential Smoothing

A1: Introduction to TSA

1 Time series

1.1 Definition

Time series — a sequence of some statistical measures taken at equally separated time intervals.

$$\{Y_1,\ Y_2,\ \dots,\ Y_T\}$$

Example:

$$Y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}\ ,$$

where $\{\varepsilon_t\}$ — independent normally distributed variables with parameters $(0,\ \sigma^2)$

1.2 Time Series operators

$$f\left(\{X_t\}, \{Y_t\}\right) = \{Z_t\}$$

- Additive: $Z_t = X_t + Y_t$
- Multiplicative: $\forall\lambda \in \mathbb{R} : Z_t = \lambda X_t$
- Lag operator: $LX_t = X_{t-1}$

$$\begin{aligned} L^2X_t &= LX_{t-1} = X_{t-2} \\ &\dots \\ L^kX_t &= X_{t-k},\ \forall k \in \mathbb{N}^* \end{aligned}$$

Those operators are commutative and distributive between each other:

$$L(X_t + Y_t) = X_{t-1} + Y_{t-1}$$

2 Numerical characteristics of time series

2.1 Autocovariance

Definition 2.1.1. Autocovariance is the covariance of the time series point with its delayed (lagged) copy.

First order autocovariance:

$$\gamma_{1t} = \mathbb{E}\left[(Y_t - \mu_t)(Y_{t-1} - \mu_{t-1})\right] \tag{2.1}$$

j^{th} order autocovariance:

$$\gamma_{jt} = \mathbb{E}\left[(Y_t - \mu_t)(Y_{t-j} - \mu_{t-j})\right] \tag{2.2}$$

In case we have joint density distribution function f:

$$\gamma_{jt} = \underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{j\ \text{times}} (Y_t - \mu_t)(Y_{t-j} - \mu_{t-j})f(Y_t,\ \dots,\ Y_{t-j})\ \mathrm{d}Y_t\ \dots\ \mathrm{d}Y_{t-j} \tag{2.3}$$

2.2 Autocorrelation

Definition 2.2.1. Autocorrelation is the correlation of the time series point with its delayed (lagged) copy.

$$\rho_j = \mathrm{Corr}(Y_t,\, Y_{t-j}) = \frac{\mathrm{Cov}(Y_t,\, Y_{t-j})}{\sqrt{\mathrm{Var}(Y_t)\mathrm{Var}(Y_{t-j})}} = \frac{\gamma_j}{\sqrt{\gamma_0^2}} \tag{2.4}$$

2.3 Weak and strong stationarity

Definition 2.3.1. Weak or Covariant stationarity

Process Y_t is called weakly (covariant) stationary, if neither expectation nor autocovariances γ_{jt} depends on the time t.

Definition 2.3.2. Strong or Strict stationarity

Process Y_t is called strongly (strictly) stationary, if any subsequence of time series Y_t depends only on the intervals separating values in this subsequence.

$\forall\, j_1,\, \dots,\, j_n$ joint distribution of a random vector $(Y_t,\, Y_{t+j_1},\, \dots,\, Y_{t+j_n})$ depends only on the intervals separating those values: $j_2 - j_1,\, j_3 - j_2,\, \dots,\, j_n - j_{n-1}$.

Lemma 2.3.1. *For any weak stationary process autocovariance γ_j verifies the following symmetry property:*

$$\gamma_j = \gamma_{j-1}$$

A2: Forecasting with TSA

3 Time series classical decomposition

We consider the following decomposition of a time series Y:

$$Y = T \times S \times C \times I \tag{3.1}$$

Where:

- $[\,T\,]$ **Trend** — long-term movement in the time series over an extended period.

  Methods used for trend estimation:

  - Leas squares approximation (Regression)
  - Moving average (MA)
  - Exponential smoothing (ES)
  - Freehand method
  - Semi-averages method

- $[\,S\,]$ **Seasonal variation** — variation in a time series within one year that is repeated more or less regularly. Seasonal variation may be caused by the temperature, public holidays, cycles of seasons or holidays. So, we must estimate how the data vary from month to month or other period of time.

  A **seasonal index** is a set of numbers chowing the relative values of a variable during the months of a year. Methods used for calculating SI:

  - Average percentage method
  - Percentage or ratio to trend method
  - Percentage moving average

  We obtain the deseasonalized data by dividing every monthly entry of the initial data by the seasonal index found by one of the three methods.

  $$Y/S = T \times C \times I \tag{3.2}$$

- $[\,C\,]$ **Cyclical fluctuations** — recurring up and down movements with respect to trend that have a duration of several years. Their study is obtained after the de-trading (adjusting (3.2) by dividing by the results of one of the methods of trend analysis):

  $$\frac{Y}{S \times T} = C \times I \tag{3.3}$$

- $[\,I\,]$ **Irregular variations** — erratic variations from trend that cannot be ascribed to the cyclical or seasonal influences.

4 Steps to complete forecasting

1. Collect data and check if they are reliable. For a better forecast it is suggested to obtain related time series.

2. Plot time series.

3. Construct long term trend curve using one of the methods mentioned before.

4. If seasonal variations present, obtain a seasonal index and deseasonalize data (adjusting data to seasonal variations).

5. Adjust deseasonalized data to trend.

   Resulting data contain (in theory) only cyclic and irregular variations.

6. Plot cyclic variations obtained in step 5. No periodicities should present.

7. Create a forecast by combining results from steps 1 to 6. Identify and evaluate all possible sources of errors and their magnitude.

   However sometimes we ignore $C \times I$ for forecasts and assume:

$$Y = T \times S$$

A3: Difference equations

5 Difference equations

5.1 First order difference equations

Definition 5.1.1. A difference equation is an expression relating a variable Y_t to its previous values.

Definition 5.1.2. Linear first order difference equation (recursive)

$$Y_t = \varphi Y_{t-1} + w_t \ , \tag{D 1.1}$$

where φ — dynamic coefficient.

This is a *first order* linear difference equation because only the first lag of the variable (Y_{t-1}) appears in the equation and it expresses Y_t as a linear function of Y_{t-1} and w_t.

Recursive substitution for (D 1.1) gives us a solution:

$$Y_t = \varphi^{t+1} Y_{-1} + \sum_{0}^{t} w_i \varphi^{t-i} \tag{D 1.3}$$

Dynamic multipliers

Note that (D 1.3) expresses Y_t, as a linear function of the initial value Y_{-1}, and the historical values of w. This makes it very easy to calculate the effect of w_0 on Y_{-1}. If w_0 were to change with Y_t, and $w_1, \ \ldots, \ w_t$ taken as unaffected, the effect on Y_t, would be given by:

$$\frac{\partial Y_t}{\partial w_0} = \varphi^t \tag{D 1.4}$$

Then Y_{t+j} could be described as function of Y_{t-1} and $w_{t+1}, \ \ldots, \ w_{t+j}$:

$$Y_{t+j} = \varphi^{t+j+1} Y_{t-1} + \sum_{t}^{t+j} w_i \varphi^{t-i} \tag{D 1.5}$$

The effect of w_t on w_{t+j} is:

$$\frac{\partial Y_{t+j}}{\partial w_t} = \varphi^j \tag{D 1.6}$$

Thus the dynamic multiplier (D 1.6) depends only on (j), the length of time separating the disturbance to the input $\{w_t\}$ and the observed value of the output (Y_{t+j}). The multiplier does not depend on t; that is, it does not depend on the dates of the observations themselves. This is true for any linear difference equation.

5.2 p^{th} order difference equation

Definition 5.2.1. **p^{th} order difference equation**

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + w_t \tag{D 2.1}$$

Defining ξ_t, V_t and F vectors ($F \in \mathbb{M}_{p,p}$)

$$\xi_t = \begin{bmatrix} Y_t \\ Y_{t-1} \\ \cdots \\ Y_{t-p+1} \end{bmatrix},\ V_t = \begin{bmatrix} Y_t \\ Y_{t-1} \\ \cdots \\ Y_{t-p+1} \end{bmatrix},\ F = \begin{pmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_p \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{D 2.2, 2.3}$$

we will rewrite (D 2.1) in a vector form:

$$\xi_t = F + \xi_{t-1} + V_t \tag{D 2.4}$$

Dynamic multiplier could be found in the same way we found it for (D 1.1). (D 2.6) is a generalization for (D 1.3):

$$\xi_t = F^{t+1}\xi_{-1} + \sum_0^t V_i F^{t-i} \tag{D 2.6}$$

$$\begin{aligned} Y_t = {} & f_{11}^{(t+1)} Y_{-1} + f_{12}^{(t+1)} Y_{-2} + \cdots + f_{1p}^{(t+1)} Y_{-p} \\ & + f_{11}^{(t)} w_0 + f_{11}^{(t-1)} w_1 + \cdots + f_{11}^{(1)} w_{t-1} + w_t \end{aligned} \tag{D 2.8}$$

(D 2.9) is a generalization for (D 1.5):

$$\xi_{t+j} = F^{j+1}\xi_{t-1} + \sum_0^j V_{t+i} F^{j-i} \tag{D 2.9}$$

$$\begin{aligned} Y_{t+j} = {} & f_{11}^{(j+1)} Y_{t-1} + f_{12}^{(j+1)} Y_{t-2} + \cdots + f_{1p}^{(j+1)} Y_{t-p} \\ & + f_{11}^{(j)} w_t + f_{11}^{(j-1)} w_{t+1} + \cdots + f_{11}^{(1)} w_{t+j-1} + w_{t+j} \end{aligned} \tag{D 2.10}$$

f_{ij} in equations above represent $(i,\ j)$ element of matrix F.

Dynamic multipliers for p^{th} order difference equations

$$\frac{\partial Y_{t+j}}{\partial w_t} = f_{11}^{(j)} \tag{D 2.11}$$

Although numerical simulation may be adequate for many circumstances, it is also useful to have a simple analytical characterization of , which, we know from (D 2.11), is given by the (1, 1) element of $F^{(j)}$. This is fairly easy to obtain in terms of the eigenvalues of the matrix F. Recall that the eigenvalues of a matrix F are those numbers λ for which:

$$|F - \lambda \mathbf{I}_p| = 0 \tag{D 2.13}$$

For example, for $p = 2$ the eigenvalues are going to be:

$$\lambda_{\pm} = \frac{\varphi_1 \pm \sqrt{\varphi_1^2 + 4\varphi_2}}{2}$$

Exercises

In this section I will cover exercises from classes with following datasets:

1. Murders in the US (in thousands) for period 1985–1995.
2. Monthly new housing prices (in thousands) for the US for period January 1990 – December 1995.
3. Divorces and annulments (in thousands) in the US for period 1986–1995
4. Monthly values (in millions of dollars) of imports from Brazil for period 1994–1996.
5. Monthly values of export from the US to Canada (in billions of dollars) for period 1990–1995.

Exercise: 1

Year	Murders	Year	Murders
1985	19.5	1990	23.8
1986	21.2	1991	25.7
1987	20.8	1992	24.2
1988	21.4	1993	25.2
1989	21.9	1994	24.1
		1995	22.7

Exercise: 2

Month	1990	1991	1992	1993	1994	1995
January	99.2	52.5	71.6	70.5	76.2	84.5
February	86.9	59.1	78.8	74.6	83.5	81.6
March	108.5	73.8	111.6	95.5	134.3	103.8
April	119	99.7	107.6	117.8	137.6	116.9
May	121.1	97.7	115.2	120.9	148.8	130.5
June	117.8	103.4	117.8	128.5	136.4	123.4
July	111.2	103.5	106.2	115.3	127.8	129.1
August	102.8	94.7	109.9	121.8	139.8	135.8
September	93.1	86.6	106	118.5	130.1	122.4
October	94.2	101.8	111.8	123.2	130.6	126.2
November	81.4	75.6	84.5	102.3	113.4	107.2
December	57.4	65.6	78.6	98.7	98.5	92.8

Exercise: 3

Year	Divorces and annulments	Year	Divorces and annulments
1986	1178	1991	1189
1987	1166	1992	1215
1988	1167	1993	1187
1989	1157	1994	1191
1990	1182	1995	1169

Exercise: 4

Years	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
1994	686	569	741	645	739	762	768	783	842	801	677	671
1995	805	633	745	647	702	732	715	812	692	775	775	797
1996	741	633	686	716	723	737	729	859	732	706	747	764

Exercise: 5

Month	1990	1991	1992	1993	1994	1995
January	6.3	6.8	6.9	6.9	7.6	10.1
February	6.7	6.4	7.0	7.7	8.2	10.2
March	8.0	7.1	8.2	9.5	10.4	11.7
April	7.4	7.6	7.8	8.8	9.4	10.6
May	7.9	7.7	7.7	8.8	10.0	11.4
June	7.5	7.5	8.4	9.1	10.2	10.9
July	6.2	6.5	6.9	7.1	7.6	8.4
August	6.7	6.8	7.0	8.3	9.9	10.8
September	6.4	7.4	7.9	8.6	10.2	10.8
October	7.5	8.3	8.0	8.9	10.5	11.4
November	7.4	7.0	7.7	8.9	10.6	11.1
December	5.9	6.1	7.1	7.9	9.8	9.7