Adaptive prediction methods

Methods of smoothing of a time series

- The methods of the first type ***(analytical)***
- The methods of the second type ***(algorithmic)***

Moving average

Consider a time series $y_1, y_2, \dots, y_n$.

We select the smoothing interval $m = 2p + 1$. Usually, $p \le n/3$.

n is the length of the time series.

The smoothed value $\widehat{f}(t)$ is calculated by the formula:

$$\widehat{f}(t) = \sum_{i=-p}^{p} w_i y_{t+i},$$

$$t = p+1, p+2, \dots, n-p,$$

where w_i is the weight of the i-th value of the time series, and

$$w_i \ge 0, \quad \sum_i w_i = 1.$$

Moving average

Any smooth function $f(t)$ can be locally (in a limited range) represented by a polynomial of degree q:

$$f(t)=a_0+\sum_{i=1}^{q}a_i t^i .$$

1) Let $m=2p+1$ and the local behavior of the function $f(t)$ be described by a polynomial of the 1st degree:

$$f(t)=a_0+a_1 t.$$

The task is:

$$\sum_{i=-p}^{p}\left(y_i-a_0-a_1 i\right)^2\rightarrow\min$$

We find the partial derivatives over parameters a_0 and a_1.

We get the system:

$$\begin{cases}(2m+1)a_0+a_1\sum_{i=-p}^{p}i=\sum_{i=-p}^{p}y_i\\ a_0\sum_{i=-p}^{p}i+a_1\sum_{i=-p}^{p}(i)^2=\sum_{i=-p}^{p}iy_i\end{cases}$$

$\boxed{w_i}$

$$\hat{f}(t)=\frac{1}{2p+1}\sum_{i=-p}^{p}y_{t+i}=\frac{1}{m}\sum_{i=t-p}^{t+p}y_i,\quad t=p+1,p+2,\ldots,n-p,$$

Forecast for (t + 1)-th period:

$$y^*_{t+1}=\hat{f}(t)$$

the last design value of the moving average

HW 2:

1) Let $f(t)=a_0.$

Find a moving average $\hat{f}(t)$ of a time series.

HW 2:

1) Let $f(t)=a_0+a_1 t+a_2 t^2.$

Find weights w_i of the smoothed time series for smoothing intervals with 5, 7 and 9 points.

Exponential smoothing
Brown's method

Consider a time series $y_1, y_2, \dots, y_t$.

1) Let $y_\tau = a_0 + \varepsilon_\tau, \quad \tau = 1,2,\dots,t,$ $\quad a_0$ is an unknown constant.

The following recurrence relation is true:

$$S_t = \alpha y_t + (1-\alpha) S_{t-1}, \qquad 0 < \alpha < 1.$$

smoothing parameter (adaptation parameter)

Sequentially substituting the values $S_{t-1}, \ S_{t-2}, \ \dots$ expressed in terms of previous values and till $S_0 = y_0$, we obtain:

$$S_t = \alpha y_t + (1-\alpha) S_{t-1} = \alpha y_t + (1-\alpha)\left[\alpha y_{t-1} + (1-\alpha) S_{t-2}\right] = \alpha y_t + a(1-\alpha) y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} +$$
$$\dots + \alpha(1-\alpha)^k y_{t-k} + \dots + (1-\alpha)^t y_0$$

or

$$S_t = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i y_{t-i} + (1-\alpha)^t y_0.$$

With $t \to \infty$ we have $$S_t = \alpha \sum_{i=0}^{\infty} (1-\alpha)^i y_{t-i}.$$

for long time series with "infinite past"

Brown R.G. Smoothing, Forecasting and prediction. Prentice-Hall, Englewood Cliffs, N.Y., 1963

Forecast for 1 periods ahead : $\hat{y}_{t+1} = S_t = \hat{a}_0(t;\alpha).$

Simple exponential smoothing

Forecast for τ periods ahead: $\hat{y}_{t+\tau} = \sum_{k=0}^{n} \frac{\hat{a}_{k+1}(t,\alpha)}{k!}\tau^k.$

Smoothing with the polynomial

2) A generalization of the method of exponential smoothing for a polynomial non-random component.

Let the considered time series is described by a polynomial of order n, i.e.

$$y_\tau = a_0 + a_1\tau + \ldots + a_n\tau^n + \varepsilon_\tau,$$

Multiple exponential smoothing procedure has the form

$$S_t^{[p]} = \alpha S_t^{[p-1]} + (1-\alpha)S_{t-1}^{[p]}, \qquad p = 1,2,\ldots,l \text{ is an order of smoothing.}$$

Here $S_t^{[0]} = y_t, \quad S_0, S_0^{[2]}, \ldots, S_0^{[l]}$ are initial values of the exponential average of the corresponding order.

$p = 0$ – simple exponential smoothing,

$p = 1$ – linear exponential smoothing,

$p = 2$ – quadratic exponential smoothing.

Basic formulas for predicting adaptive polynomial models

1	2	3	4	5
n=0	$S_0^{(1)} = \hat{a}_{1,0}$	$S_t^{(1)} = \alpha \cdot y_t + \beta \cdot S_{t-1}^{(1)}$	$\hat{a}_{1,t} = S_t^{(1)}$	$\hat{y}_\tau(t) = \hat{a}_{1,t}$
n=1	$S_0^{(1)} = \hat{a}_{1,0} - \frac{\beta}{\alpha} \cdot \hat{a}_{2,0}$ $S_0^{(2)} = \hat{a}_{1,0} - \frac{2\beta}{\alpha} \cdot \hat{a}_{2,0}$	$S_t^{(1)} = \alpha \cdot y_t + \beta \cdot S_{t-1}^{(1)}$ $S_t^{(2)} = \alpha \cdot S_t^{(1)} + \beta \cdot S_{t-1}^{(2)}$	$\hat{a}_{1,t} = 2 \cdot S_t^{(1)} - S_t^{(2)}$ $\hat{a}_{2,t} = \frac{\alpha}{\beta} \cdot \left[S_t^{(1)} - S_t^{(2)}\right]$	$\hat{y}_\tau(t) =$ $= \hat{a}_{1,t} + \tau \cdot \hat{a}_{2,t}$
n=2	$S_0^{(1)} = \hat{a}_{1,0} - \frac{\beta}{\alpha} \cdot \hat{a}_{2,0} + \frac{\beta \cdot (2-\alpha)}{2\alpha^2} \cdot \hat{a}_{3,0}$ $S_0^{(2)} = \hat{a}_{1,0} - \frac{2\beta}{\alpha} \cdot \hat{a}_{2,0} + \frac{\beta(3-2\alpha)}{\alpha^2} \cdot \hat{a}_{3,0}$ $S_0^{(3)} = \hat{a}_{1,0} - \frac{3\beta}{\alpha} \cdot \hat{a}_{2,0} + \frac{3\beta(4-3\alpha)}{2\alpha^2} \hat{a}_{3,0}$	$S_t^{(1)} = \alpha \cdot y_t + \beta \cdot S_{t-1}^{(1)}$ $S_t^{(2)} = \alpha \cdot S_t^{(1)} + \beta \cdot S_{t-1}^{(2)}$ $S_t^{(3)} = \alpha \cdot S_t^{(2)} + \beta \cdot S_{t-1}^{(3)}$	$\hat{a}_{1,t} = 3 \cdot (S_t^{(1)} - S_t^{(2)}) + S_t^{(3)}$ $\hat{a}_{2,t} = \frac{\alpha}{2\beta^2} \cdot \left[(6-5\alpha) \cdot S_t^{(1)} - 2(5-4\alpha) \cdot S_t^{(2)} + (4-3\alpha) \cdot S_t^{(3)}\right]$ $\hat{a}_{3,t} = \frac{\alpha^2}{\beta^2} \cdot (S_t^{(1)} - 2 \cdot S_t^{(2)} + S_t^{(3)})$	$\hat{y}_\tau(t) =$ $= \hat{a}_{1,t} + \tau \cdot \hat{a}_{2,t} +$ $+ \frac{1}{2} \cdot \tau^2 \cdot \hat{a}_{3,t}$

Holt model

This method is used when the data shows a trend.

$\hat{y}_{t+\tau} = \hat{a}_t + \hat{b}_t \tau$ - the model for forecasting, where

$$\hat{a}_t = \alpha y_t + (1-\alpha)(\hat{a}_{t-1} + \hat{b}_{t-1}),$$

$$\hat{b}_t = \beta(\hat{a}_t - \hat{a}_{t-1}) + (1-\beta)\hat{b}_{t-1}, \qquad 0 < \alpha < 1, \quad 0 < \beta < 1.$$

Here

$\hat{a}_t$ is the smoothed estimate of the level of the time series at the end of each period.

$\hat{b}_t$ is the estimate of the trend (which is an average growth at the end of each period).

Initial values

There are several methods to choose the initial values for $\hat{a}_t$ and $\hat{b}_t$.
$\hat{a}_0$ is in general set to y_1. Three suggestions for $\hat{b}_t$:

Holt-Winters model (HW-model)

This method is used when the data shows a trend and seasonality. There are two main HW-models, depending on the type of seasonality.

Multiplicative Seasonal Model

$$\hat{y}_{t+\tau} = (\hat{a}_t + \hat{b}_t \tau) c_{t-s+\tau},$$

where

$$\hat{a}_t = \alpha \frac{y_t}{c_{t-s}} + (1-\alpha)(\hat{a}_{t-1} + \hat{b}_{t-1}), \qquad 0 < \alpha < 1,$$

$$\hat{b}_t = \beta(\hat{a}_t - \hat{a}_{t-1}) + (1-\beta)\hat{b}_{t-1}, \qquad 0 < \beta < 1,$$

$$\hat{c}_t = \gamma \frac{y_t}{\hat{a}_t} + (1-\gamma)\hat{c}_{t-s}. \qquad 0 < \gamma < 1,$$

Smoothing parameters

Let the length of season be s periods.

$\hat{a}_t$ is the deseasonalized smoothed estimate of the level of the time series (*called constant component*).

$\hat{b}_t$ is a linear trend component. That is simply the difference between two successive estimates of the deseasonalized level.

$\hat{c}_t$ is a multiplicative seasonal factor.

The seasonal factors are defined so that sum to the length of the season, i.e.

$$\sum_{t=1}^{s} c_t = s.$$

Holt-Winters model (HW-model)

This model is used when the data exhibits additive seasonality.

Additive Seasonal Model

$$\hat{y}_{t+\tau} = \hat{a}_t + \hat{b}_t \tau + c_{t-s+\tau},$$

where

$$\hat{a}_t = \alpha(y_t - \hat{c}_{t-s}) + (1-\alpha)(\hat{a}_{t-1} + \hat{b}_{t-1}),$$

$$\hat{b}_t = \beta(\hat{a}_t - \hat{a}_{t-1}) + (1-\beta)\hat{b}_{t-1},$$

$$\hat{c}_t = \gamma(y_t - \hat{a}_t) + (1-\gamma)\hat{c}_{t-s}.$$

Let the length of season be *s* periods.

$\hat{a}_t$ is the deseasonized smoothed estimate of the level of the time series (*called constant component*).

$\hat{b}_t$ is a linear trend component.

The trend component, if deemed unnecessary, maybe deleted from the model

$\hat{c}_t$ is an additive seasonal factor.

The seasonal factors are defined so that the sum to the length of the season, i.e.

$$\sum_{t=1}^{s} c_t = 0.$$

Genesis of the formation of a time series:

In general, a time series model may be represented as

$$y_t = f(t) + \varepsilon_t$$

systematic component

random component with $E(\varepsilon_t) = 0, \quad \text{Var}(\varepsilon_t) = \sigma^2$

T		long-term component or ***trend***
S		***seasonal*** component
C		***cyclic*** component
E		random component or ***error***

Systematic (non-random) component of a time series and the methods of smoothing

The main tasks of time series analysis are:

- Determining the composition of non-random component $f(t)$ of a time series;
- Estimates construction for the non-random components that are present in the decomposition;
- Selection a model that adequately describes the behavior of random component ε_t and evaluation of the parameters of this model.

Detection of non-random components of a time series

1. You should reveal the fact about the presence/absence of a non-random component by testing the hypotheses

- Testing the hypothesis of the constancy of the average values of the series based on the Student's t-test;
- Verifying the uniformity on the basis of two samples F-test;
- Checking the homogeneity of samples based on Cochran's criterion;
- The criterion of the squares of consecutive differences (Abbe criterion);
- Series test, based on the median of the sample;
- The criterion of "upstream" and "downstream" series.

2. You should construct the estimate (approximation) for the unknown integral non-random component

$$f(t) = T + S \quad \text{or} \quad f(t) = T \times S,$$

i.e. the problem of smoothing of the time series (elimination of random residuals ε_t) is solved.

Seasonal decomposition method (Census II) Trend-seasonal models

Additive model is as follows:

$$Y = T + S + E\ .$$

This model assumes that each level of the time series can be represented as the sum of trend T, seasonal component S and random component E.

Multiplicative model is:

$$Y = T \times S \times E\ .$$

This model assumes that each level of the time series can be represented as the product of trend T, seasonal component S and random component E.

Selecting the type of decomposition is based on analysis of the structure of seasonal fluctuations.

- ***If the oscillation amplitude is approximately constant we should build an additive decomposition model,*** in which the values of the seasonal component are assumed to be constant for different cycles.
- ***If the amplitude of the seasonal variation increases or decreases, we build a multiplicative model.***

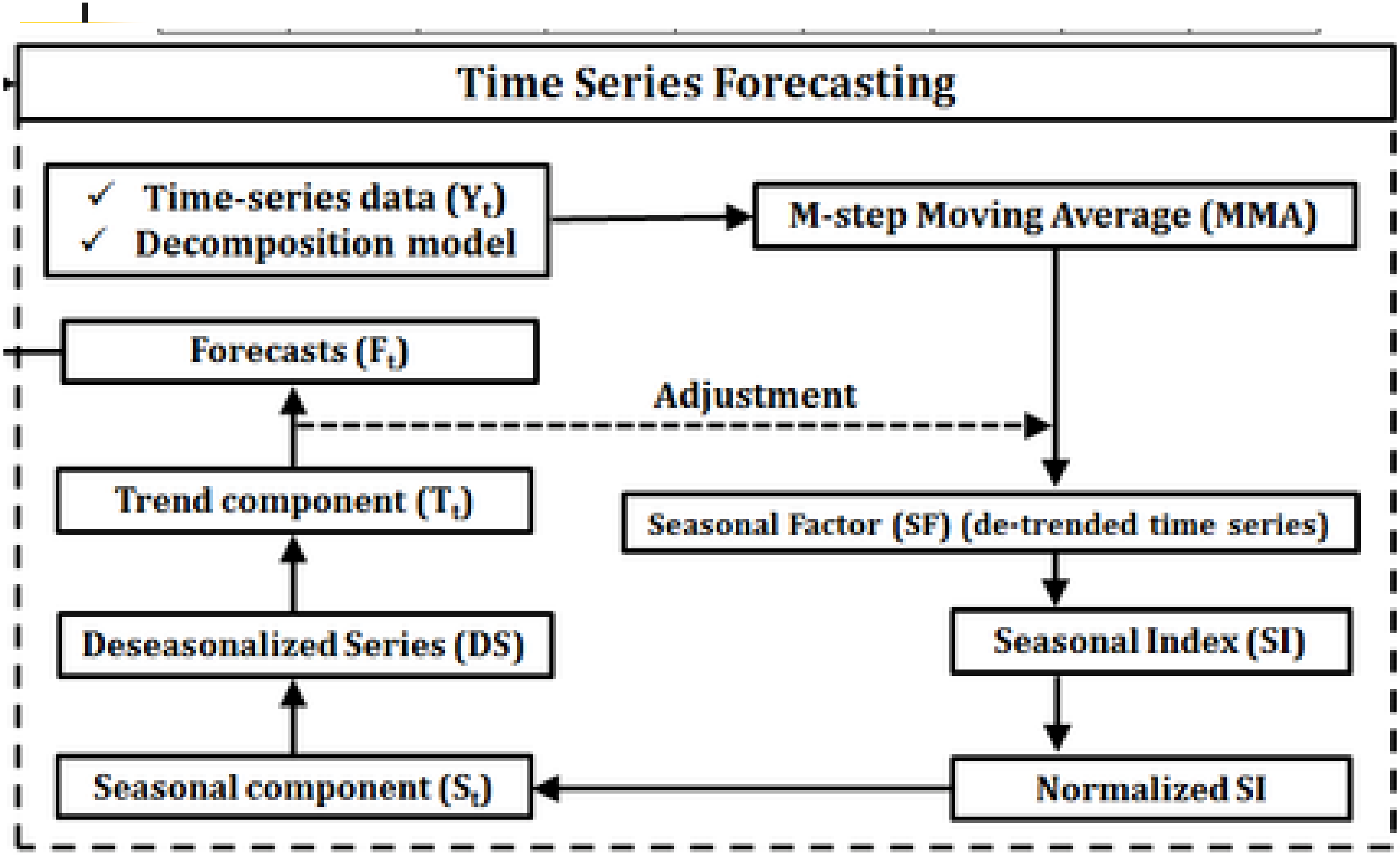
Seasonal decomposition method (Census II) Trend-seasonal models

Construction of additive and multiplicative models involves calculation of values T, S and E for each level of the time series.

Stages of model building

1. Smoothing an original time series by the moving average.
2. Calculation of the seasonal component S.
3. Eliminating the seasonal component S and getting series ($T+E$) in case of the additive model and ($T \times E$) in case of the multiplicative model.
4. Analytical leveling ($T+E$) or ($T \times E$) and calculation of the T series using a trend model.
5. Calculation of ($T+S$) or ($T \times S$) series.
6. Calculation of absolute and relative errors (MAD and MAPE).

Seasonal decomposition method (Census II)
Trend-seasonal models

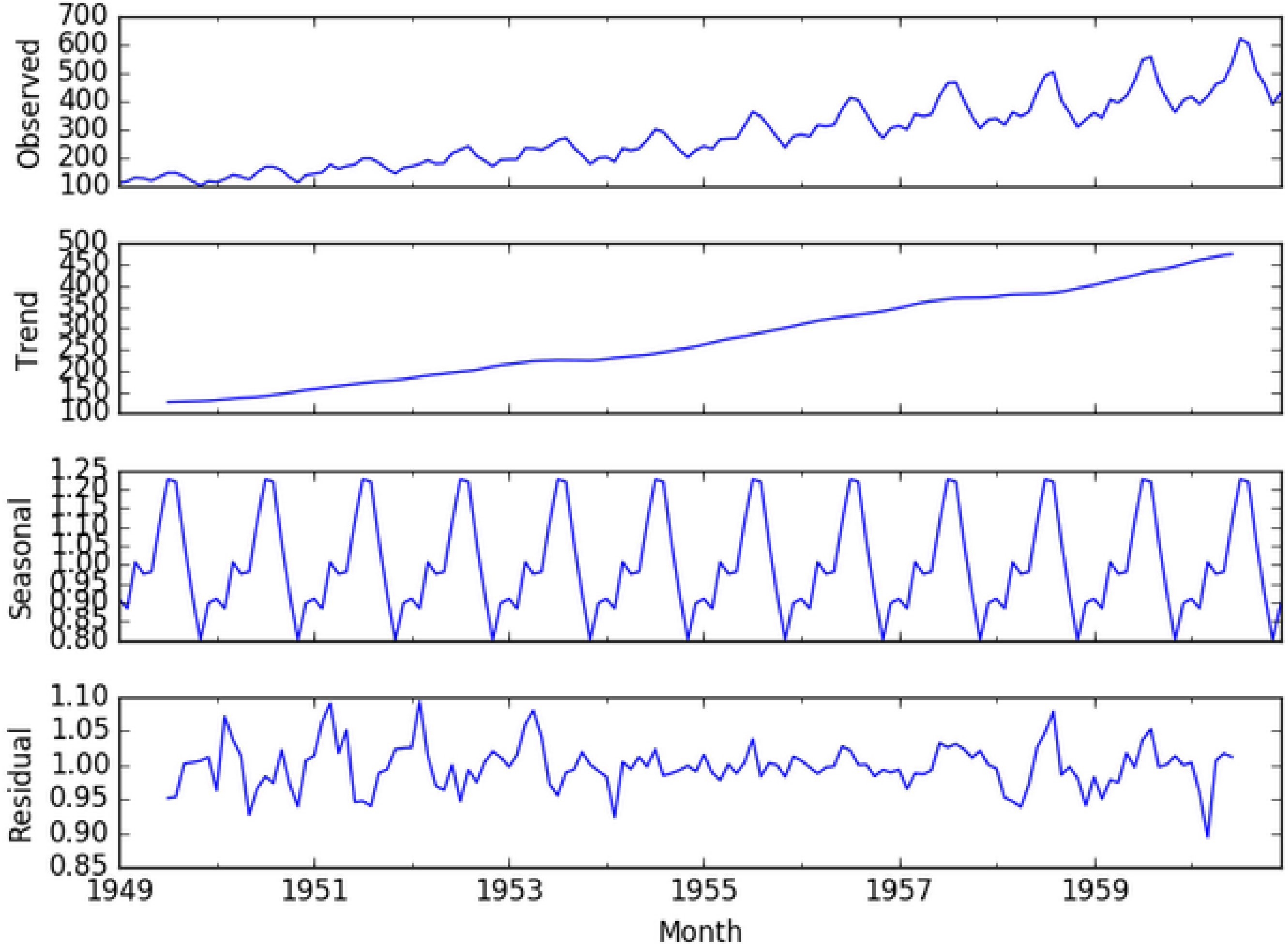


1. Calculation of absolute and relative errors (MAD and MAPE).

Seasonal decomposition method (Census II)
Trend-seasonal models

The Airline Passengers dataset describes the total number of airline passengers over a period of time.

The units are a count of the number of airline passengers in thousands. There are 144 monthly observations from 1949 to 1960.



Multiplicative Decomposition of Airline Passenger Dataset