B.VIII. Exponential Smoothing Methods

Olivier Scaillet

University of Geneva and Swiss Finance Institute

Outline

Introduction

Exponential smoothing methods belong to the category of fast computational forecasting methods.

Advantages:

1. Simple to implement: forecasting equations are easy to understand and compute.
2. No underlying models.
3. Methods often as performant as more sophisticated ones.
4. Robust methods: may work on short series with possible structural changes.

Disadvantages:

1. Not adapted to certain type of data.
2. Arbitrary choice of smoothing constant

Setup

Data: $Y_1, ..., Y_t, ..., Y_T$

We are at date T and wish to forecast the value Y_{T+h} at date $T+h$.

The *forecast value* is denoted $\hat{Y}_T(h)$, and h is called the *forecast horizon.*

Simple Exponential Smoothing

Parameter: α = *smoothing constant* satisfying $0 < \alpha < 1$.

Definition:

$$\hat{Y}_T(h) = (1-\alpha)\sum_{j=0}^{T-1}\alpha^j Y_{T-j}$$

Remarks:

1. Most recent observations have the greatest influence (exponential decay of influence of past data)
2. α close to 1: large influence of past data (rigidity of forecast).
3. α close to 0: small influence of past data (adaptivity to most recent data)
4. Definition does not depend on h, thus we use notation $\hat{Y}_T$.

Simple Exponential Smoothing

Updating formulas:

$$\hat{Y}_T = \alpha \hat{Y}_{T-1} + (1-\alpha) Y_T$$

This is a weighted average of forecast at date $T-1$ and of the latest observation

If α is small we get more weight to the latest observation

$$\hat{Y}_T = \hat{Y}_{T-1} + (1-\alpha)(Y_T - \hat{Y}_{T-1})$$

We have an error correction mechanism.

We need to initialize the algorithm. To do so, we may choose $\hat{Y}_1 = Y_1$, for example

Simple Exponential Smoothing

Interpretation:
Assume that the series is approximately constant on the observation period

$$Y_t = a + u_t, \qquad t = 1, ..., T$$

We may estimate the constant a by weighted least squares

$$\min_a \sum_{j=0}^{T-1} \alpha^j (Y_{T-j} - a)^2$$

The weights decline exponentially when move back in time.

Simple Exponential Smoothing

Solution of minimization problem:

$$\hat{a} = \frac{1-\alpha}{1-\alpha^T} \sum_{j=0}^{T-1} \alpha^j Y_{T-j}$$

For T large enough, we get $\hat{a} \approx \hat{Y}_T$.

Hence $\hat{Y}_T$ takes the interpretation of the constant which approximates best the series around T since the weight declines when moves away from T.

The method is only relevant when the series is approximately constant next to T (locally constant) and should be avoided otherwise, for example, in the presence of a trend.

Simple Exponential Smoothing

Choice of the smoothing constant:

1. Subjective: depending of willingness to have fast adaptivity or more rigidity.
2. Choice advocated by Brown (inventor of the method): $\alpha = 0.7$
3. Objective: constant chosen to minimize the sum of squared forecast errors

$$\sum_{t=1}^{T-1} \left(Y_{t+1} - \hat{Y}_t \right)^2$$

Double Exponential Smoothing

Simple exponential smoothing adapted to locally constant series (constant = horizontal line)

Generalization:
Take a line with a slope (trend)

$$Y_t = a_1 + (t - T)a_2 + u_t$$

We may use to forecast

$$\hat{Y}_T(h) = \hat{a}_1(T) + h\hat{a}_2(T)$$

The coefficients $\hat{a}_1(T), \hat{a}_2(T)$ are obtained by solving :

$$\min_{a_1, a_2} \sum_{j=0}^{T-1} \alpha^j (Y_{T-j} - a_1 + a_2 j)^2.$$

Double Exponential Smoothing

For large T, we get

$$\hat{a}_1(T) = 2S_1(T) - S_2(T),$$
$$\hat{a}_2(T) = \frac{1-\alpha}{\alpha}\left(S_1(T) - S_2(T)\right)$$

where $S_1(T)$ is the smoothed series

$$S_1(T) = (1-\alpha)\sum_{j=0}^{T-1} \alpha^j Y_{T-j}$$

and $S_2(T)$ is the doubly smoothed series

$$S_2(T) = (1-\alpha)\sum_{j=0}^{T-1} \alpha^j S_1(T-j)$$

Double Exponential Smoothing

Updating formulas used in practice:

$$\hat{a}_1(T) = \hat{a}_1(T-1) + \hat{a}_2(T-1) + (1-\alpha^2)\left(Y_T - \hat{Y}_{T-1}(1)\right),$$
$$\hat{a}_2(T) = \hat{a}_2(T-1) + (1-\alpha)^2\left(Y_T - \hat{Y}_{T-1}(1)\right)$$

Last terms are proportional to *error forecast* $Y_T - \hat{Y}_{T-1}(1)$
If perfect forecast $Y_T = \hat{Y}_{T-1}(1)$, i.e., forecast error is zero, then there is no need to update, and we get:

$$\hat{a}_1(T) = \hat{a}_1(T-1) + \hat{a}_2(T-1),$$
$$\hat{a}_2(T) = \hat{a}_2(T-1)$$

Initialization values:

$$\hat{a}_1(2) = Y_2, \qquad \hat{a}_2(2) = Y_2 - Y_1$$

Double Exponential Smoothing

Remark:

There exist other exponential smoothing methods which use two smoothing parameters:

- More adaptive such as the Holt-Winters method
- May take into account the presence of seasonality, etc ...