Uneven time series forecasting using a modified exponential smoothing method

Abstract. The article is devoted to the problem of forecasting time series with an uneven distribution of observations over time. The exponential smoothing model is used as the basic forecasting model, in which the variable weights of observations decrease exponentially. The exponential smoothing model allows us to take into account the attenuation of the correlation of cross sections of a random process of time series change over time. However, this does not take into account the factors of temporal unevenness of the results of observations and the finiteness of the sample of observations. The article describes a method for predicting an uneven time series based on a modified exponential smoothing model, in which the transition from exponential smoothing to decreasing non-exponential smoothing is carried out. The modified sequence of the weights of the observations is determined by adjusting the classically calculated exponential weights, taking into account the actual irregularity of the observations.

1 Introduction

The results of periodic measurements of parameters of various processes in economics, medicine, engineering and other applied fields are usually presented in the form of time series (TS) [1-5]. A time series is understood as time-ordered measurement results of values of a certain controlled parameter at discrete time intervals $h$. Let's imagine TS as:

$$y(t) = f(t) + \epsilon(t),$$

where $y(t)$ is the trend or the deterministic (non-random) component of the random TS; $\epsilon(t)$ is the random component of TS with a mathematical expectation equal to zero; $y(t)$, $\epsilon(t)$, etc., are the moments of measurement of TS values; $n$ is the number of measurements (observations) of the parameter.

In the traditional formulation [6-9], the task of predicting TS consists in extrapolating TS formed by parameter values $c$, obtained at regular intervals using a segment of the Taylor series, i.e. a polynomial of the degree of decomposition of the function $f$ in the neighborhood $t = t_i$, where $i$ is the moment of the last observation. In this case, the problem reduces to the problem of time series extrapolation [10].

Corresponding author: mironov-anik@yandex.ru

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).
In the vicinity of a point \( t_0 \), if this rule were true over the entire observation interval, estimates of the derivatives could be obtained using the least squares method (LSM) [10, 11].

For TS with a significant random component \( \varepsilon(t) \), the use of LSM to determine the parameters of the forecast model can lead to significant errors. To avoid these errors, in a number of papers [12-14] it is proposed to consider observations as unequal. Weights are attributed to them, decreasing depending on the prehistory of observations of TS values according to the law of geometric progression. Such methods are called TS forecasting methods with discounting of observational results [6].

Among the methods with discounted observations, the exponential smoothing (ES) method or the Brown method is the most well-known [7-9, 10-14]. Formulas for estimating derivatives determined by the ES method are derived on the assumption that there is not a finite, but an infinite series of observations

\[
\begin{align*}
\text{E} & \approx \frac{1}{\alpha} \sum_{n=0}^{\infty} \alpha^n c \left( n + 1 \right) \\
\text{E} & \approx \frac{1}{\alpha} \sum_{n=0}^{\infty} \alpha^n c \\
\end{align*}
\]

where \( \alpha \to \infty \).

In the infinite series (2), an observation made at a time \( t_k \) is assigned a weight equal to \( \alpha^{k-1} \), where \( \alpha \) is a constant number, called the smoothing parameter and equal to the weight of the last observation \( \alpha \), \( \alpha < \alpha < \alpha \).

It can be shown that the estimates of the derivatives of the function at the last observation point, obtained using the ES method, coincide with the estimates of the coefficients when solving the following:

\[
\begin{align*}
\hat{c}_a & = \alpha \sum_{n=0}^{\infty} \alpha^n c \\
\end{align*}
\]
The process is fundamentally uneven in time. Secondly, part of the measurements may be lost when transmitting the results of parameter monitoring over lines and communication channels due to interference, which transforms uniform TS to an uneven appearance [15-17].

In order to avoid forecasting errors due to the finiteness of the observations, as well as in order to generalize the ES method to the case of unequal observations, we give a brief conclusion of the formulas used in the ES method.

2 Materials and methods

The ES method introduces quantities called exponential averages of the k order for TS).}

Next, a system of linear equations with unknown values of derivative estimates is compiled. The right-hand sides of the equations contain the values of exponential averages. By solving these equations relatively, an expression for the estimate is obtained in the form of a linear combination of exponential averages. The exponential averages themselves are found recursively, based on the values of TS).}

Knowing the estimates of the derivatives of the function at the last observation point, the following expression is taken as the parameter value predicted at the time \( \tau + n \tau \).
The recurrent formula (7) gives an algorithm for generating values, which is conveniently presented in the form of Table 1.

Table 1. An algorithm for generating values based on a recurrent formula (7)

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A specially selected value \( k \) is taken instead of the entire sum of terms corresponding to observations \( c(\tau m) \) with \( m \leq 0 \).

The choice of these initial values is, in our opinion, the most difficult and at the same time the most vulnerable place in the ES method. It seems advisable to use an algorithm to find the optimal value \( \hat{k} \) according to the criterion of the minimum error of the forecast. This procedure can be implemented by dividing some training data set into two parts: basic and verification. In the first part of the scale, the parameters of the forecast model are estimated, in the second the forecast error. By varying the value \( k \), its value \( \hat{k} \) is determined such that the forecast error is minimal.

In accordance with the main theorem of the ES method proved by Brown, any \( k \)-th derivative \((k = 0, 1, 2, \ldots, p)\) in equation (1) can be expressed in terms of linear combinations of exponential averages up to \((p+1)\) th order.

The linear system (8) can be written in matrix form. To do this, we denote \( f(j) (t) \), \( j = 0, 1, \ldots, p \), has the form: \( 9 \):

\[
\sum_{j=0}^{p} \sum_{k=n-j}^{n} \alpha_{jk} f(j) (t) = \sum_{j=0}^{p} \alpha_{jk} f(j) (t) + \alpha_{jk} f(j) (t) = \sum_{j=0}^{p} \alpha_{jk} f(j) (t) + \sum_{j=0}^{p} \alpha_{jk} f(j) (t)
\]

\( p \)
Let's find a vector $\mathbf{m} \equiv \sum_{j=1}^{\infty} \mathbf{m}_j$. The elements of the matrix $\mathbf{M}$ are functions of $t$, $k$, and $\mathbf{m}_j$.

$$\mathbf{M} = \sum_{j=1}^{\infty} \mathbf{m}_j = \sum_{j=1}^{\infty} \mathbf{m}_j.$$

$$\mathbf{\alpha} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \frac{1}{\alpha_1} \sum_{j=1}^{\infty} \mathbf{m}_j.$$
\[ \partial \Phi = -\alpha \sum_{k=1}^{\infty} \alpha \sum_{j=1}^{\infty} c_{ij} - \alpha \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \alpha - \alpha \sum_{j=1}^{\infty} \alpha - \alpha \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \alpha - \alpha. \]

\[ \sum_{k=1}^{\infty} \alpha - \alpha = \frac{1}{\alpha} \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \alpha - \alpha. \]
Observations that must be taken into account as much as possible in the forecast. Then expression (16) will take the form

\[
\left(\sum_{i}^{\infty} \alpha_i \right) - \alpha \sum_{i=0}^{\infty} \left( \sum_{m=0}^{\infty} \alpha^m \right) c_{m, n} = \sum_{i=0}^{\infty} \alpha_i \left( \sum_{m=0}^{\infty} \alpha^m \right) c_{m, n} - \sum_{i=0}^{\infty} \alpha_i \sum_{m=0}^{\infty} \alpha^m c_{m, n} = \sum_{i=0}^{\infty} \alpha_i \left( \sum_{m=0}^{\infty} \alpha^m \right) c_{m, n} + \sum_{i=0}^{\infty} \alpha_i \alpha^m c_{m, n} = \sum_{i=0}^{\infty} \left( \alpha_i - \sum_{m=0}^{\infty} \alpha^m c_{m, n} \right) = \left( \frac{1}{\alpha} - \sum_{m=0}^{\infty} \alpha^m c_{m, n} \right) / \alpha
\]

Let's choose \( \alpha = 0 \)

\[
\sum_{i=0}^{\infty} \alpha_i \left( \sum_{m=0}^{\infty} \alpha^m \right) c_{m, n} + \sum_{i=0}^{\infty} \alpha_i \sum_{m=0}^{\infty} \alpha^m c_{m, n} = \sum_{i=0}^{\infty} \alpha_i \left( \sum_{m=0}^{\infty} \alpha^m \right) c_{m, n} + \sum_{i=0}^{\infty} \alpha_i \sum_{m=0}^{\infty} \alpha^m c_{m, n} = \sum_{i=0}^{\infty} \left( \alpha_i - \sum_{m=0}^{\infty} \alpha^m c_{m, n} \right) = \frac{1}{\alpha} - \sum_{m=0}^{\infty} \alpha^m c_{m, n} = \frac{1}{\alpha} - \sum_{m=0}^{\infty} \alpha^m c_{m, n}
\]

To represent the exponential average, we can use the following formulas:

\[
\begin{align*}
\alpha_i &= \frac{- \alpha_{i-1}}{\alpha} \\
\alpha_{i-1} &= \frac{- \alpha_i}{\alpha}
\end{align*}
\]

\[
\alpha = \int \left( \frac{n_{i, n}}{n_{i+1, n}} \right) = n_{i, n} - n_{i+1, n} = n_{i, n} - n_{i+1, n}.
\]
\[ \alpha = \int_{t_j}^{t_n} \alpha = \alpha \]

\[ \alpha_j = \frac{a(t_j - t_n)}{t_n - t_1} - 1, \]

where \( a \) is some positive constant.

If \( t = t_j \), then \( x = x_j = a(t_j - t_n)(t_n - t_1) \).

Note that \( x_1 = -a; x_p = 0 \).

Thus, formula (17) defines the linear transformation of the observation interval \((t_1 \ldots t_n)\) in the interval \((-a \ldots 0)\).

Let's determine the weight of the \( j \)-th observation using the formula

\[ \sum_{j=1}^{n} \sum_{l=0}^{p} \sum_{k=0}^{n} \sum_{j=1}^{l} \sum_{n}^{j} \sum_{k=1}^{p} \sum_{c}^{a} \sum_{0}^{0} \sum_{1}^{1} \sum_{*}^{*} \sum_{+}^{+} = \]

The expressions (17) and (18) obtained make it possible to calculate the weights of unequally spaced observations. It is easy to verify that all the properties of the weights listed above are fulfilled, i.e. the proposed method is indeed a generalization (modification) of the ES method. After the weights of the observations are found, the optimization problem is solved:

\[ \min_{\alpha} \text{arg} \left( \sum_{t} \right) \]

It can be shown that the coefficient estimates obtained from condition (19) satisfy the following system of linear algebraic equations

\[ \sum_{t} \left( -\sum_{a} c - \sum_{\tau} \right) = \sum_{a} \tau + \sum_{c} \tau, \]

\[ c_{\tau + \tau} = \tau + \tau + \tau. \]

3 Results
Step 1

Step 2

Step 3

Step 4

Step 5

Step 6

Step 7

Fig. 1. Calculation of the predicted parameter values. The structure of the method for predicting changes in uneven TS of limited duration.
The matrices $V^{-1}$, $K$, $K^{-1}$ are determined by the formulas (the notation is the same):

\begin{align*}
V^{-1} & = \left[ \alpha, \delta \right] \left[ \alpha, \delta \right] = \left[ \alpha \right] \left[ \alpha \right] = \begin{cases} 0 & \text{for } \alpha = \alpha \\ 1 & \text{for } \alpha \neq \alpha \end{cases} \\
K & = \left[ \delta, \delta \right] \left[ \sqrt{\alpha}, \sqrt{\alpha} \right] = \left[ \sqrt{\alpha} \right] \left[ \sqrt{\alpha} \right] \\
K^{-1} & = \left[ \sqrt{\alpha}, \sqrt{\alpha} \right] \left[ \alpha, \alpha \right] = \left[ \alpha \right] \left[ \alpha \right] = \begin{cases} 0 & \text{for } \alpha = \alpha \\ 1 & \text{for } \alpha \neq \alpha \end{cases}
\end{align*}

The matrix $X$ has the form:

\[ X = \begin{bmatrix}
\alpha & \alpha & \cdots & \alpha \\
\delta & \delta & \cdots & \delta \\
\vdots & \vdots & \ddots & \vdots \\
\alpha & \alpha & \cdots & \alpha 
\end{bmatrix} \]

Step 4. For the given (known) values of the parameters measured at time $t_1$, $t_2$, ..., $t_n$ and forming a vector $(c_1, c_2, ..., c_p)$, we determine the estimates of the coefficients $a_0, ..., a_p$ of the segment of the Taylor series according to the formula:

\[ \hat{c} = B^T B \hat{c} = B^T U = X^T V^{-1} X \hat{c} \]

The remaining sum of squares is determined by the formula:

\[ \overline{c} \approx \overline{c} + \frac{X^T V^{-1} X \hat{c} - X \hat{c}}{t} \]

The remaining sum of squares is determined by the formula:

\[ \overline{c} \approx \overline{c} + \frac{X^T V^{-1} X \hat{c} - X \hat{c}}{t} \]
The residual sum of squares has \((n - p - 1)\) degrees of freedom and can be used to estimate the \(S^2\) value of \(\sigma^2\).

\[
\begin{align*}
\hat{\sigma}^2 & = \frac{1}{n-p} \sum_{i=1}^{n-p} (y_i - \hat{y}_i)^2 \\
& = \frac{1}{n-p} \sum_{i=1}^{n-p} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 \\
& = \frac{1}{n-p} \sum_{i=1}^{n-p} (y_i - \hat{\beta}_0)^2
\end{align*}
\]

Step 6. Calculate the covariance matrix of estimate \(\hat{a}\) of the vector \(a\) using the formula (29).

\[
D[\hat{a}] = M[D[a]] = \sigma^2 (B^T B)^{-1} = \sigma^2 (X^T V X)^{-1}
\]

Step 7. Using the formula (4), we calculate the predicted values of the parameters.

4 Discussions

Figure 2 and 3 show experimental curves of the dependence of the relative methodological error of the TS forecast results by one step on the degree of unevenness and the sample size of the initial time series for three forecasting models: MLS with a linear trend without differentiation; exponential smoothing and modified exponential smoothing based on the developed method for predicting parameters from uneven time series of observations of limited duration.

Fig. 2. Experimental dependences of the relative methodological error of the TS forecast results by one step on the degree of unevenness of the initial time series...
5 Conclusion
References

2. A. A. Akhmetshina, A young scientist, 50 (288), 161-163 (2019)
17. O. V. Russkov, et al., Eurasian Union of Scientists, 9, 28-33 (2014)

https://doi.org/10.1051/e3sconf/202453103020