Journal of Mathematical Sciences, Vol. 228, No. 5, February, 2018

NONPARAMETRIC METHOD OF LEAST SQUARES: ACCOUNTING FOR SEASONALITY

A. I. Orlov[1]

We consider the problem of restoring a nonparametric dependence described by the sum of linear trend and seasonal component, i.e., a periodic function with the known period. We obtain the asymptotic distribution of the parameter estimates and the trend component. We find the mathematical expectation of the residual sum of squares. We also develop the methods of estimation of the seasonal component and construction of the interval forecast.

1. Introduction. The statement of the problem

1.1. The problem of restoring linear dependence

Let t be an independent variable (for example, time), and x be a dependent variable (for example, inflation index, dollar exchange rate, monthly production volume, or size of daily revenue of a retail outlet). Consider the problem of restoring the dependence $x = x(t)$ from n pairs of numbers (t_k, x_k), $k = 1, 2, \ldots, n$, where t_k are the values of independent variable and x_k are corresponding values of dependent variable.

The restoration of dependence can be based on different models. In the simplest model it is assumed that the variable x depends linearly on the variable t up to the measurement errors, i.e.,

$$x_k = a(t_k - t_{\text{avg}}) + b + e_k,\ k = 1, 2, \ldots, n, \tag{1}$$

where a and b are the parameters unknown to the statistician and are subjected to estimation, while e_k are the errors, distorting the dependence. The average of time moments

$$t_{\text{avg}} = (t_1 + t_2 + \ldots + t_n)/n$$

is introduced in the model to simplify further calculations.

Usually the parameters a and b of linear dependence are estimated by the least squares method. Then the restored dependence is used for point and interval forecast estimation.

The least squares method was developed by Gauss in 1795 [1, p.37]. (As stated in [2, p.181], "Gauss points out two dates: 1794 and 1795. Modern researchers tend to assume that the correct date is 1794".) According to this method, to calculate the best function, in a linear manner approximating the dependency x on t in model (1), one should consider the function of two variables

$$f(a, b) = \sum_{i=1}^{n} (x_i - a(t_i - t_{\text{avg}}) - b)^2.$$

The least squares estimates are the values a^* and b^* for which the function $f(a, b)$ attains its minimum over all values of arguments. In order to find these estimates, one should calculate the partial derivatives of $f(a, b)$ in a and b, equate them to 0, and then obtain the estimates from the resulting equations. Making the corresponding calculations (see, e.g., [3, Section 5.1]), we see that the least squares estimates

[1] Bauman Moscow State Technical University and Moscow Physics-Technical Institute, Moscow, Russia, e-mail: prof-orlov@mail.ru

Translated from *Statisticheskie Metody Otsenivaniya i Proverki Gipotez*, Vol. 21, pp. 88–97, 2008

(LSE) of the linear dependence parameters have the form

$$a^* = \frac{\sum\limits_{i=1}^{n} x_i(t_i - t_{\text{avg}})}{\sum\limits_{i=1}^{n}(t_i - t_{\text{avg}})^2} = \frac{\sum\limits_{i=1}^{n}(x_i - x_{\text{avg}})(t_i - t_{\text{avg}})}{\sum\limits_{i=1}^{n}(t_i - t_{\text{avg}})^2}, \qquad b^* = x_{\text{avg}} = \frac{x_1 + x_2 + \ldots + x_n}{n} \tag{2}.$$

A number of other equivalent expressions [3, Section 5.1] for the considered estimates are known.

Note that LSE (2) are obtained in the deterministic statement.

1.2. The nonparametric linear model

It is natural to study the properties of LSE on the basis of some probabilistic and statistical data model. In [3, Section 5.1]) the following model is studied.

Let the values of the independent variable t be determined and the errors e_k, $k = 1, 2, \ldots, n$, be i.i.d. random variables with zero mean and variance σ^2 unknown to the statistician.

In the literature, there is still a conventional assumption concerning the normal distribution of errors. However, it is long known that the actual distribution of data tends to differ from normal [3, Section 4.1]. That is why we consider a nonparametric model in which there is no assumption that the distribution of errors belongs to a certain parametric family.

Further we will repeatedly use the Central limit theorem (CLT) for the values e_k, $k = 1, 2, \ldots, n$ (with weights), therefore it is necessary to assume, for example, that the errors e_k, $k = 1, 2, \ldots, n$, are finite of have a finite third absolute moment. However, it is not necessary here to draw attention to these internal mathematical "regularity conditions".

The above model was considered in detail in [3, Section 5.1]. In particular, the asymptotic normality and independence of LSE (2) were proved, which allowed the development of methods for confidence estimation of the function $x(t)$ and testing statistical hypotheses about the parameters of a linear dependence.

Various generalizations of the considered model also have practical value. For example, the variances of the errors may be different (which corresponds to weighted sum of squares); errors may be dependent, for example, they may be described by stochastic processes, as is common in the time series models; the dependence itself can be arbitrary, and then for its estimation it is natural to use nonparametric kernel density estimates [3, Section 5.2].

Note the fundamental difference of all of these models from those in which t is considered to be a random variable. For example, when the initial data (t_k, x_k), $k = 1, 2, \ldots, n$, are a sample from a two-dimensional distribution, i.e., they are independent identically distributed random vectors. Or when each coordinate is measured with a random error (the model of confluence analysis). In this row we must also mention the models of restoring dependencies within the frames of interval data statistics [4].

From the variety of models of restoring dependencies we consider the models with periodic components.

1.3. Accounting for seasonality in the model

In the analysis of economic data it is necessary to use time series models, which include three components: trend (T), periodic or cyclic (S), and random (E). There are the additive model $T + S + E$ and the multiplicative model $T \times S \times E$ [5].

The simplest additive model has the form

$$x_k = a(t_k - t_{\text{avg}}) + b + e_k = a(t_k - t_{\text{avg}}) + b + f(t_k) + E_k, \ k = 1, 2, \ldots, n. \tag{3}$$

Here the trend component is a linear function $a(t_k - t_{\text{avg}}) + b$; the periodic component $f(t)$ usually describes the seasonality, i.e., the period is known and equals one year; the random component is described by the terms E_k, which can be considered to be independent identically distributed random

variables with zero mean and variance σ^2 unknown to the statistician. Model (3) reduces to model (1), if we set

$$e_k = f(t_k) + E_k,\ k = 1, 2, \ldots, n.$$

Unlike the model studied in [3, Section 5.1], the errors e_k in model (3) are not identically distributed. However, their distributions differ only by the shifts (the values of the deterministic seasonal component).

The corresponding multiplicative model has the form

$$y_k = [Bt_k^a] \times f_1(t_k) \times [1 + \varepsilon_k],\ k = 1, 2, \ldots, n. \tag{4}$$

In (4) the factors have the meaning described above. After taking the logarithm, model (4) reduces to an analog of model (3), hence, it suffices to consider model (3).

The practical value of this model is obvious. However, the computational methods described in [5] are heuristic. This determines the purpose of the paper: to build a nonparametric theory of time series forecasting based on a linear trend with the additive or multiplicative model of seasonality.

Following the heuristic approach of [5], let us study the asymptotic behavior of the LSE a^* and b^* defined by (2). We will prove their asymptotic normality, and then consistently estimate the periodic component $f(t)$ and construct the interval forecast for $x(t)$. In particular, we will identify the feasibility of the data analysis for the total number of years (periods). Unlike [6] (see also [3, Section 6.3; 4, Section 10.2]), the period length should not be estimated, since it is given from the context (usually, it is one year).

2. The asymptotic theory

2.1. The asymptotic distribution of the parameter estimates

From (2) it follows that under the above assumptions and notations we have

$$b^* = \frac{a}{n}\sum_{i=1}^{n}(t_i - t_{\text{avg}}) + b + \frac{1}{n}\sum_{i=1}^{n} e_i = b + \frac{1}{n}\sum_{i=1}^{n} e_i = b + \frac{1}{n}\sum_{i=1}^{n} f(t_i) + \frac{1}{n}\sum_{i=1}^{n} E_i. \tag{5}$$

According to the CLT, the estimate b^* is asymptotically normal with the expectation $b + \frac{1}{n}\sum\limits_{i=1}^{n}(t_i)$ and the variance σ^2/n, whose estimate is given below.

From (2) and (5) it follows that

$$x_i - x_{\text{avg}} = a(t_i - t_{\text{avg}}) + b + e_i - b - \frac{1}{n}\sum_{i=1}^{n} e_i,$$

$$(x_i - x_{\text{avg}})(t_i - t_{\text{avg}}) = a(t_i - t_{\text{avg}})^2 + e_i(t_i - t_{\text{avg}}) - \frac{(t_i - t_{\text{avg}})}{n}\sum_{i=1}^{n} e_i.$$

The last term in the second relation vanishes after the summation with respect to i, therefore

$$a^* = a + \sum_{i=1}^{n} c_i e_i = a + \sum_{i=1}^{n} c_i f(t_i) + \sum_{i=1}^{n} c_i E_i,\ c_i = \frac{(t_i - t_{\text{avg}})}{\sum\limits_{i=1}^{n}(t_i - t_{\text{avg}})^2}. \tag{6}$$

Formulas (6) show that the estimate a^* is asymptotically normal with the expectation $a + \sum\limits_{i=1}^{n} c_i f(t_i)$ and the variance

$$D(a^*) = \sum_{i=1}^{n} c_i^2 D(E_i) = \frac{\sigma^2}{\sum\limits_{i=1}^{n}(t_i - t_{\text{avg}})^2}.$$

Note that the multidimensional normality holds when each term in (6) is small as compared to the whole sum, i.e.

$$\lim_{n\to\infty} \frac{\max\limits_{i=1,\dots,n} |t_i - t_{\text{avg}}|}{\sqrt{\sum\limits_{i=1}^{n} (t_i - t_{\text{avg}})^2}} = 0. \tag{7}$$

Condition (7) is satisfied if t_i form an arithmetic progression with infinitely growing number of members.

So, the variances of the LSE a^* and b^* of the linear trend are the same as in the case where there are no seasonal distortions. But their mathematical expectations depend on the periodic component. However, if

$$\sum_{i=1}^{n} f(t_i) = 0, \ \sum_{i=1}^{n} (t_i - t_{\text{avg}}) f(t_i) = 0, \tag{8}$$

then the estimates a^* and b^* are unbiased.

Conditions (8) are crucial. They are necessary and sufficient for the unbiasedness and consistency of the estimates considered in this article.

The first condition in (8) can be assumed satisfied if t_i form an arithmetic progression and the integer number of steps forms a single period (for example, if the measurements are made monthly or quarterly, and the period is one year), and, in addition, the data are measured for an integer number of periods. Indeed, then it is natural to assume that the sum of the values of the periodic component over the period is equal to zero, since otherwise the constant term could be adjusted (i.e., by the same reasons that the mathematical expectations of the random components E_i are assumed to be zeros).

For the validity of the second condition in (8) it suffices to add the assumptions of the symmetry of the set $\{t_k, \ k = 1, 2, \dots, n\}$ with respect to t_{avg} (for example, the beginning of the year) and the evenness of the periodic component $f(t)$ with respect to the same point. The latter holds if, for example, the graph of $f(t)$ is symmetric with respect to the middle of the year.

The unbiasedness (under assumptions (8)) and asymptotic normality of the LSE allow us to easily construct the confidence intervals for them and to test statistical hypotheses, for example, about their equality to specific values, particularly zero.

2.2. The asymptotic distribution of the trend component

From (5) and (6) it follows that under the validity of (8) we have

$$M\{a^*(t - t_{\text{avg}}) + b^*\} = M(a^*)(t - t_{\text{avg}}) + M(b^*) = a(t - t_{\text{avg}}) + b,$$

i.e., the estimate $y^*(t) = a^*(t_k - t_{\text{avg}}) + b^*$ of the trend component $y(t) = a(t - t_{\text{avg}}) + b$ of the considered dependence is unbiased. Therefore,

$$D(y^*(t)) = D(a^*)(t - t_{\text{avg}})^2 + 2M\{(a^* - a)(b^* - b)(t - t_{\text{avg}})\} + D(b^*).$$

Since the errors E_i are independent and $M(E_i) = 0$, then

$$M\{(a^* - a)(b^* - b)(t - t_{\text{avg}})\} = \frac{1}{n} \sum_{i=1}^{n} c_i (t - t_{\text{avg}}) M(E_i^2) = \frac{1}{n} (t - t_{\text{avg}}) \sigma^2 \sum_{i=1}^{n} c_i = 0.$$

Hence,

$$D(y^*(t)) = \sigma^2 \left\{ \frac{1}{n} + \frac{(t - t_{\text{avg}})^2}{\sum\limits_{i=1}^{n} (t_i - t_{\text{avg}})^2} \right\}. \tag{9}$$

Thus, the estimate $y^*(t)$ is unbiased and asymptotically normal. For its practical application (for interval estimation or testing statistical hypotheses), one should be able to estimate the residual variance $M(E_i^2) = \sigma^2$.

In particular, it is not hard to obtain the lower and the upper bounds for the trend component of the forecasting function:

$$y_{\text{low}}(t) = a^*(t - t_{\text{avg}}) + b^* - \delta(t), \quad y_{\text{up}}(t) = a^*(t - t_{\text{avg}}) + b^* + \delta(t),$$

where the half-width of the confidence interval $\delta(t)$ has the form

$$\delta(t) = U(\gamma)\sqrt{D^*(y^*(t))} = U(\gamma)\sigma^* \sqrt{\frac{1}{n} + \frac{(t - t_{\text{avg}})^2}{\sum\limits_{i=1}^{n}(t_i - t_{\text{avg}})^2}}. \tag{10}$$

Here γ is the confidence probability, and $U(\gamma)$ is the $\dfrac{1+\gamma}{2}$-order quantile of the normal distribution, i.e.,

$$U(\gamma) = \Phi^{-1}\left(\frac{1+\gamma}{2}\right),$$

where $\Phi(x)$ is the distribution function of the standard normal law with zero mean and variance 1. When $\gamma = 0.95$ (the most frequently used value), we have $U(\gamma) = 1.96$. In formula (10) $D^*(y^*(t))$ is a consistent estimate of the variance $y^*(t)$. In accordance with (9) it is the product of the consistent estimate σ^* of the mean square deviation σ of the random errors E_i and the deterministic function of t that is known to the statistician.

3. The periodic component and the forecast

3.1. The mathematical expectation of the residual sum of squares

In the points t_k, $k = 1, 2, \ldots, n$, there are the initial values of the dependent component x_k and the recovered values $y^*(t_k)$. Consider the residual sum of squares

$$SS = \sum_{i=1}^{n} (y^*(t_i) - x_i)^2 = \sum_{i=1}^{n} \{(a^* - a)(t_i - t_{\text{avg}}) + (b^* - b) - f(t_i) - E_i\}^2.$$

Recall that if there is no periodic component, then we use [3, Sections 5.1, 5.2] the consistent estimates σ^* of the mean square deviation σ of random errors constructed on the basis of the residual sum of squares

$$\sigma^* = \sqrt{\frac{SS}{n}} \ \text{ or } \ \sigma^* = \sqrt{\frac{SS}{n-2}}.$$

In accordance with (5) and (6), under the validity of conditions (8), we obtain

$$SS = \sum_{i=1}^{n} \left\{ (t_i - t_{\text{avg}}) \sum_{j=1}^{n} c_j E_j + \frac{1}{n} \sum_{j=1}^{n} E_j - f(t_i) - E_i \right\}^2 =$$

$$= \sum_{i=1}^{n} \left\{ \sum_{j=1}^{n} \left[c_j (t_i - t_{\text{avg}}) + \frac{1}{n} \right] E_j - f(t_i) - E_i \right\}^2 = \sum_{i=1}^{n} SS_i.$$

Let us find the mathematical expectation of each term. We have

$$M(SS_i) = M \left\{ \sum_{j=1}^{n} \left[c_j (t_i - t_{\text{avg}}) + \frac{1}{n} \right] E_j - f(t_i) - E_i \right\}^2 =$$

$$= M\left\{\sum_{j=1}^{n}\left[c_j(t_i - t_{\text{avg}}) + \frac{1}{n}\right]E_j\right\}^2 -$$

$$-2M\left\{\sum_{j=1}^{n}\left[c_j(t_i - t_{\text{avg}}) + \frac{1}{n}\right]E_j\right\}(f(t_i) + E_i) + M(f(t_i) - E_i)^2.$$

Since E_i are independent identically distributed and have zero mean, then

$$M\left\{\sum_{j=1}^{n}\left[c_j(t_i - t_{\text{avg}}) + \frac{1}{n}\right]E_j\right\}^2 = \sum_{j=1}^{n}\left\{c_j(t_i - t_{\text{avg}}) + \frac{1}{n}\right\}^2\sigma^2.$$

Next,

$$-2M\left\{\sum_{j=1}^{n}\left[c_j(t_i - t_{\text{avg}}) + \frac{1}{n}\right]E_j\right\}(f(t_i) + E_i) = -2\left\{c_i(t_i - t_{\text{avg}}) + \frac{1}{n}\right\}\sigma^2.$$

Finally,

$$M(f(t_i) - E_i)^2 = f^2(t_i) + \sigma^2.$$

On the basis of the last three equalities we can show that when the condition of the asymptotic normality (7) holds,

$$\lim_{x\to\infty} M(SS_i) = f^2(t_i) + \sigma^2.$$

Hence,

$$M\left(\frac{SS}{n}\right) = \sigma^2 + \frac{1}{n}\sum_{i=1}^{n} f^2(t_i). \tag{11}$$

On the right-hand side of (11) the first term corresponds to the contribution of the random component, and the second one corresponds to the contribution of the periodic component.

In some cases, the second term on the right-hand side of (9) may be known from previous experience or it may be estimated by experts; however, in most situations, it is advisable to start from the estimation of the periodic component.

3.2. The estimation of the seasonal component

Both parametric and nonparametric approaches are considered. A popular method is based on the fact that a sufficiently smooth function can be expanded in a Fourier series and a good approximation can be obtained using a small number of harmonics. In the simplest case it is a single harmonic. Thus, the dynamics of the inflation index can be studied using the model

$$x_k = a(t_k - t_{\text{avg}}) + b + f(t_k) + E_k = a(t_k - t_{\text{avg}}) + b + d\cos(2\pi t_k) + E_k,$$

$k = 1, 2, \ldots, n$ (the time t is measured in years). Then the unknown parameters a, b, d are estimated using the method of least squares.

However, there is usually no reason to believe that the periodic component belongs to a particular parametric family of functions, so one has to build non-parametric estimates. We describe one of the possible settings.

Suppose that in accordance with assumptions (8) we consider an integer number of periods, i.e., $n = mq$, where n is the number of observations, m is the number of periods, and q is the number of

observations in a single period. Then, in accordance with the definition of the periodic component, the following equalities hold:

$$f(t_k) = f(t_{q+k}) = f(t_{2q+k}) = \ldots = f(t_{(m-1)q+k}),\ k = 1, 2, \ldots, q. \tag{12}$$

Let g_k be the common value in (12). We have to to estimate $g_1, g_2, \ldots, g_q$.

A natural approach is to average m values of $x_i - y^*(t_i)$ (i.e., the initial data "cleared" from the trend component), corresponding to the moments of time, spaced apart by an integer number of periods. Here we deal with the estimates

$$g_k^* = \frac{1}{m}\sum_{i=1}^{m}(x_{k+(j-1)q} - y^*(t_{k+(j-1)})q),\ k = 1, 2, \ldots, q. \tag{13}$$

The estimate of the periodic component extends over the whole observation period in an obvious way:

$$f^*(t_k) = f^*(t_{q+k}) = f^*(t_{2q+k}) = \ldots = f^*(t_{(m-1)q+k}) = g_k^*,\ k = 1, 2, \ldots, q. \tag{14}$$

Summing the recovered values of the trend and periodic components, we estimate the dependence, "cleared" from the random component

$$x^*(t) = y^*(t) + f^*(t) = a^*(t - t_{\text{avg}}) + b^* + f^*(t). \tag{15}$$

Here the estimates a^* and b^* are obtained by formulas (2), whereas the estimates $f^*(t)$ are obtained by (13),(14).

With the use of (15) one can construct a point forecast, using it outside the range of observations. It suffices to extend the seasonal component $f^*(t)$ up to the considered point of time according to (14) and sum it with the forecast of the trend component $y^*(t)$. The interpolation and extrapolation for the time moments t that are not included in the original set $\{t_i,\ i = 1, 2, \ldots, n\}$ and the sets obtained by shifting it by an integer number of periods can be performed by the linear interpolation between the nearest values or by another smoothing method.

Let us discuss the properties of estimates (13)–(15).

When the number of observations grows infinitely and conditions (7) and (8) hold, the estimates a^* and b^* of the parameters of the trend component are consistent and unbiased; thus, it can be shown that under the conditions of this paper, sums (13) estimate the periodic component in a consistent and unbiased manner. Hence,

$$\frac{1}{n}\sum_{i=1}^{n}[f^*(t_i)]^2 - \frac{1}{n}\sum_{i=1}^{n} f^2(t_i) \to 0 \tag{16}$$

in probability as $n \to \infty$. In accordance with (11) the last relation makes it possible to estimate σ^2 and then to construct an interval forecast for the trend component according to (10).

Note that in this situation, as a rule, n grows, with its increments being multiples of q, the number of observations in a single period. As a consequence, the minuend in (16) is a constant, and there is no dependence on n. These features are related to the fact that conditions (8) involve the consideration of an integer number of periods.

Let us consider estimates (13) in detail. From (3), (12), and (13) it follows that

$$g_k^* = f(t_k) - (a^* - a)\frac{1}{m}\sum_{j=1}^{m}(t_{k+(j-1)q} - t_{\text{avg}}) - (b^* - b) + \frac{1}{m}\sum_{j=1}^{m} E_{k+(j-1)g},\ k = 1, 2, \ldots, q.$$

In view of (5), (6), and (8) we obtain

$$g_k^* = f(t_k) - \left(\sum_{i=1}^{n} c_i E_i\right)\left(\frac{1}{m}\sum_{j=1}^{m}(t_{k+(j-1)q} - t_{\text{avg}})\right) - \frac{1}{n}\sum_{i=1}^{n} E_i + \frac{1}{m}\sum_{j=1}^{m} E_{k+(j-1)q},\ k = 1, 2, \ldots, q.$$

Thus,

$$g_k^* = f(t_k) + \sum_{i=1}^{n} h_i E_i,\ k = 1, 2, \ldots, q, \tag{17}$$

where $h_i = -c_i r_k - \frac{1}{n} + \frac{1}{m}$, if $i \in \{k + (j-1)q,\ j = 1, 2, \ldots, m\}$, and $h_i = -c_i r_k - \frac{1}{n}$ for all other values of the summation index i, where $r_k = \frac{1}{m}\sum_{j=1}^{m}(t_{k+(j-1)q} - t_{\text{avg}})$.

Relation (17) means that the considered estimates are sums of independent random variables, and therefore using the central limit theorem, we can build confidence intervals for the considered values of the periodic component (under the assumption that conditions (7) hold).

3.3. An interval forecast

The point forecast is constructed by formula (12) on the basis of $x^*(t)$, the dependence estimate "cleared" from the random component but including trend and periodic components. If conditions (8) hold, then

$$Mx^*(t) = x(t) = a(t - t_{\text{avg}}) + b + f(t),$$

i.e., the estimate $x^*(t)$ is unbiased.

Under conditions (8) in view of (5), (6), and (17) we conclude that for the moment of time t included in the initial set $\{t_i,\ i = 1, 2, \ldots, n\}$ or the sets obtained by shifting it by an integer number of periods, we have

$$x^*(t) - x(t) = (t - t_{\text{avg}})\sum_{i=1}^{n} c_i E_i + \frac{1}{n}\sum_{i=1}^{n} E_i + \sum_{i=1}^{n} h_i E_i. \tag{18}$$

In (18), when calculating the values of the coefficients h_i, for k one should take the number of minimum initial moment of time $\{t_i, i = 1, 2, \ldots, n\}$, spaced apart from the considered moment t by an integer number of periods. Using (17) we conclude that

$$x^*(t) - x(t) = \sum_{i=1}^{n} w_i E_i,$$

where $w_i = c_i(t - t_{\text{avg}} - r_k) + \frac{1}{m}$, if $i \in \{k + (j-1)q,\ j = 1, 2, \ldots, m\}$, and $w_i = c_i(t - t_{\text{avg}} - r_k)$ for all other values of the summation index i, where r_k is the same as in (17). On the right-hand side of (18) there is a sum of independent random variables, so the estimate $x^*(t)$ is asymptotically normal (under conditions (7)) with the mathematical expectation $x(t)$ and the variance

$$D(x(t)) = \sum_{i=1}^{n} w_i^2 D(E_i) = \sigma^2 \sum_{i=1}^{n} w_i^2. \tag{19}$$

Hence, the lower and the upper confidence bounds for the forecasting function (considering both the trend and the periodic components) have the form

$$y_{\text{low}}(t) = a^*(t - t_{\text{avg}}) + b^* - \Delta(t),\ \ y_{\text{up}}(t) = a^*(t - t_{\text{avg}}) + b^* + \Delta(t),$$

where

$$\Delta(t) = U(\gamma)\sqrt{D^*(x^*(t))} = U(\gamma)\sigma^*\sqrt{\sum_{i=1}^{n} w_i^2}. \tag{20}$$

Here γ is the confidence probability, and $U(\gamma)$ is the $\frac{1+\gamma}{2}$-quantile of the normal distribution. In (20) $D^*(x^*(t))$ is a consistent estimate of the variance of the point forecast $x^*(t)$. In accordance with (19) it is the product of the consistent estimate σ^* of the mean-square deviation σ of random errors E_i and the deterministic function of t, which is known to the statistician. The value of σ^* is calculated in accordance with (11) and (16).

3.4. Comparison of parametric and non-parametric approaches

In many sources a parametric probabilistic model of the least squares method is considered. It is assumed that the errors have the normal distribution. This assumption allows a rigorous mathematical derivation of a number of conclusions. Thus, the distributions of the statistics are calculated exactly. Accordingly, instead of the quantiles of the normal distribution in the calculation of confidence intervals, the quantiles of the Student distribution are used. It is clear that the differences disappear as the amount of data increases.

The above non-parametric approach does not use the unrealistic assumption of the normality of errors. The price for this is the asymptotic nature of the results. In the case of the simplest model of the least squares method, both approaches give almost identical recommendations. This is not always so; the two approaches sometimes yield similar results. For example, in the problem of outlier detection, methods based on the normal distribution cannot be justified, and this was discovered with the use of a non-parametric approach (see [3, Section 4.2; 4, Section 7.2]).

Let us briefly state some general principles of the construction, description, and use of statistical and data analysis techniques. First, presuppositions should be clearly formulated, i.e., the adopted probabilistic-statistical model should be fully described. Second, calculation algorithms must be correct from the point of view of mathematical and statistical theory. Third, the algorithms must give conclusions that are useful in practice. With regard to the problem of recovering the dependence, this means that it is advisable to use a non-parametric approach, as was done above. However, the assumption of normality, although greatly reducing the possibility of sound practical application, from a purely mathematical point of view allows moving further. Therefore, for the initial study of the situation, so to say, "in vitro," the normal model may prove useful.

4. Concluding remarks

As compared to heuristic algorithms [5], the theory developed in this paper makes it possible:

1) to give a general justification for these algorithms within the framework of asymptotic methods of mathematical statistics and point out their applicability conditions (formula (7));

2) to identify fundamentally important conditions (8) that are necessary and sufficient for the unbiasedness and consistentcy of the considered estimates;

3) to construct confidence intervals for the dependence (forecasting function) and its trend component.

Within the framework of mathematical statistics one cannot analyze all the common heuristic algorithms. Thus, it is quite often recommended, first, to perform a smoothing ("alignment") of the time series, for example, using the moving average method [5, p.137]. In this case the periodic (seasonal) component changes, and the errors (deviations from the sum of the trend and periodic components) become dependent, which makes it impossible to apply the methods described in this paper.

REFERENCES

1. F. Klein, *Development of Mathematics in the Nineteenth Century. Part I.*, Joint Scientific and Technical Publishing NKTP USSR, Moscow (1937).
2. L. E Maystrov, *Theory of Probability: Historical Essay*, Nauka, Moscow (1967).
3. A. I. Orlov, *Econometrics*, Examen, Moscow (2004).
4. A. I. Orlov, *Applied Statistics*, Examen, Moscow (2006).
5. I. I. Eliseeva (ed.), *Workshop on Econometrics: Proc. Manual*, Finance and Statistics, Moscow (2001).
6. A. I. Orlov, "Method for estimating the length of the period and the periodic component of the signal ," *J. Math. Sci.*, **126**, No. 1, 955–960 (2005).