International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 21 (2015) pp 42778-42782

Adaptive Subband Filtration Of Time Series

Yevgeniy G. Zhilyakov, Sergei P. Belov

Belgorod State National Research University Belgorod, Pobedy Street, 85

Abstract
This article demonstrates that effective means of allocation of trends of unsteady time series' intervals may be represented by sub band analysis, which allows building sequences, Fourier transform of which in given frequency interval coincides with the interval of Fourier transform of referential row's interval. We have obtained correlations that define matrix operator of trends' allocation; we have found the conditions, at which a wide range of sequences' intervals represent their own functions (fixed points), which correspond to individual proper numbers.

Keywords: unsteady time series, trends, sub band analysis

INTRODUCTION

Let us suppose that vector $\vec{x}=(x_1,...,x_N)'$, in which dash represents conjugation, consists of real-valued components, the values of which are fixed at observation of some parameter of the process that is being investigated (time series' interval). The main aim of registration of such vectors (empirical observations) is building behavioral models of the process. For these purposes the [1– 6] performance is often used.

$$x_k = f_k + \varepsilon_k, k = 1, ..., N, \qquad (1. 1)$$

where unknown f_k is a trend, and ε_k represent random fluctuations in reference to the trend, conditioned by multiple uncontrolled factors.
Within the framework of such model, as a rule, the trend that reflects the tendency of changes in the process is of a main interest.
Let us also note that quite often with the aim of investigating (modeling) random fluctuations, the task of elimination of [2, 4] as a main source of non-stationarity is set.
Nowadays there are two main approaches to trend building. One of them uses sampling from a priori considerations of obvious functional dependences from the reference numbers (argument), parameters of which are later adjusted to available empirical observations (training sample) [1, 3, 5].
Another approach is based on smoothing [1, 6], which allows inhibiting second component in the first part (1. 1), which is naturally considered to be filtration as well.
Let us note that both approaches have their own advantages and disadvantages. The key problem of the first one lies in justification of relevance of selected functional dependence. That's why we would find it difficult to disagree with [6] the expressed opinion that application of obvious mathematical relationships makes as much sense as imposing the rules to the nature.
While building smothering procedures these or that suppositions are used regarding the character of trends' behavior and departures of them. An example is represented by so-called Spencer formulas [1], based on the assumption that a trend is a polynomial of this or that exponent (at least locally).
The least strict assumptions regarding the character of trends' behavior (majorly the requirement of local monotonicity) at building smothering procedures are used in abovementioned source [6], which shows it advantages in comparison with the others.
We would also like to note quite widely recognized technique of exponential smoothing [7], the main advantage of which is, in our opinion, the application of a principle of decreasing influence on evaluation we obtain with the reference to the trend of remote components of empirical observations' vector being handled
Within the framework of this article, we are developing the approach that is based on filtration with the usage of frequency representation, which allows reflecting consistency of behavior of trend vector's components $\vec{f}=(f_1,...,f_N)'$, which are present in all analyzed components of empirical observations' vector. At the same time, a priori considerations regarding trend's features require adaptive processing of empiric data, which allows obtaining necessary information directly on a certain interval.

RESEARCH

1. General aspect of smoothing operator

The process of smoothing is natural to be presented in the following view:

$$\hat{\vec{f}} = W(\vec{x}), \qquad (2. 1)$$

where $W()$ is a certain smoothing operator, which is continuous over all of the intervals of R^N.
It is easy to understand that in the virtue of right part's additivity (1. 1) operator of smoothing must be linear, that is why the following condition is met:

$$\hat{\vec{f}} = W(\vec{f}) + W(\vec{e}), \qquad (2. 2)$$

where $\vec{e}=(e_1,...,e_N)'$.
Another obvious requirement is satisfaction of the following inequation for Euclidean norms of vectors (symbol $\| \ \|$)

$$b = \| W(\vec{e}) \| / \| W(\vec{f}) \| < \| \vec{e} \| / \| \vec{f} \| \qquad (2.\ 3)$$

It is also known that ideally the following equation should be satisfied:

$$W(\vec{f}) = \vec{f}\,, \qquad (2.\ 4)$$

i. e., the operator that is applied must have desired trend's vector (which is unknown, in general).
Then we denote Fourier transforms with uppercase letters

$$Z(\omega) = \sum_{k=1}^{N} z_k \exp(-j\omega(k-1),\ j = (-1)^{1/2} \qquad (2.\ 5)$$

which are denoted by corresponding lowercase letters of vectors $\vec{z} = (z_1, ..., z_N)'$, that's why it the following performance is true

$$z_i = \int_{-\pi}^{\pi} Z(\omega) \exp(j\varpi(i-1)) d\omega / 2\pi\ .$$

Suppose Ω_κ is a frequency interval, symmetrically located with respect to origin of coordinates

$$\Omega_r = [-\Omega_{2r}, -\Omega_{1r}) \cup [\Omega_{1r}, \Omega_{2r})\,. \qquad (2.\ 6)$$

where condition $0 \le \Omega_{1\kappa} \le \Omega_{2\kappa} \le \pi$ is met.
Let us insert the value of sub band interval between two vectors

$$d_r(x,z) = \int_{\omega \in \Omega_r} | X(\omega) - Z(\omega) |^2\ d\omega / 2\pi\,. \qquad (2.\ 7)$$

Having plugged definitions of Fourier transforms of the form (2. 4), after obvious transformations with consideration of (2. 6), we obtain the following performance directly in the range of original values

$$d_r(x,z) = \sum_{k=1}^{N} \lambda_k^r (\alpha_k^r - \beta_k^r)^2\,, \qquad (2.\ 8)$$

where $\lambda_k^r, k = 1, ..., N$ arranged in descending order

$$\lambda_1^r \ge \lambda_2^r \ge .. \lambda_N^r \ge 0 \qquad (2.\ 9)$$

non-negative proper values of sub band matrix $A_r = \{a_{ik}^r\}$ with the elements

$$a_{ik}^r = (\sin(\Omega_{2r}(i-k)) - \sin(\Omega_{2r}(i-k))) / \pi(i-k),\ a_{ii}^r = (\Omega_{2r} - \Omega_{1r}) / \pi\,, \qquad (2.\ 10)$$

thus, the following equation is true

$$L^r G^r = A_r G^r\,, \qquad (2.\ 11)$$

$L^r = diag(\lambda_1^r, ..., \lambda_N^r)$, $G^r = \{\vec{q}_1^r ... \vec{q}_N^r\}$ is orthogonal [8] matrix, consisting of correspondingly (2. 9) arranged proper vectors:

$$\vec{\alpha}^r = (\alpha_1^r, .., \alpha_N^r)' = G^{r'} \vec{x},\ \vec{\beta}^r = (\beta_1, .., \beta_N)' = G^{r'} \vec{z}\ . \qquad (2.\ 12)$$

Thus, in (2. 12) components of vectors (see also (2. 8)) consist of projections of original vectors onto proper vectors of sub band matrix.
Calculations show that when the choice is made (square brackets represent the integral part of the content)

$$J_r = 2[N a_{ii}^r / 2] + 4 \qquad (2.\ 13)$$

equations are performed with high accuracy:

$$\lambda_{k+J_r} = 0\,, \qquad (2.\ 14)$$

Thus, at meeting conditions

$$\beta_i^r = \alpha_i^r, i = 1, .., J_r \qquad (2.\ 15)$$

There is a very high accuracy of sub band interval equaling zero, that means

$$d_r(x,z) = 0\,. \qquad (2.\ 16)$$

Thus, for exact match of intervals Fourier transform in given frequency interval it is not obligatory to perform trivial condition of equality of corresponding vectors

$$\vec{x} = \vec{z}\,. \qquad (2.\ 17)$$

It is not difficult to demonstrate that in the view of orthonormality, peculiar to the set of proper vectors of sub band matrix among all the vectors that meet the condition (2. 16), minimal Euclidean norm would belong to the following vector

$$\vec{y}_r = \sum_{k=1}^{J_r} \alpha_k^r \vec{q}_k^r\,. \qquad (2.\ 18)$$

This result is easy to be generalized in case of union of disjoint intervals of a form (2. 6)

$$\Omega_R = \bigcup_{r \in R} \Omega_r\,, \qquad (2.\ 19)$$

where R stands for some multitude of frequency intervals, besides, inequality $\min \Omega_{1r} \ge 0, \max \Omega_{2r} \le \pi$ is supposed to be performed.

At the same time, correlation for united sub band interval takes on the following form:

$$d_R(x,z) = \sum_{k=1}^{J_R} \lambda_k^R (\alpha_k^R - \beta_k^R)^2 . \qquad (2.\ 20)$$

Symbols that are present in this equation represent the same characteristics as they do in (2. 6), with the difference that here they denote summarized sub band matrix

$$A_R = \sum_{r \in R} A_r . \qquad (2.\ 21)$$

In particular, J_R stands for the number of non-zero proper values of matrix (2. 21).
Thus, the following is correct.
Statement 1. *Performance of the following equations is a necessary and sufficient condition for coincidence of Fourier transforms' intervals of two vectors* $\vec{x}$ *and* $\vec{z}$ *in frequent interval of the view of (2. 19)*

$$\beta_k^R = \alpha_k^R, k = 1,..,J_R , \qquad (2.\ 22)$$

besides, among all such vectors, $\vec{z}$ *vector*

$$\vec{y}_R = \sum_{k=1}^{J_R} \alpha_k^R \vec{q}_k^R \qquad (2.\ 23)$$

has minimal Euclidean norm
Let us assume that

$$W_R = G_J^R G_J^{R'} , \qquad (2.\ 24)$$

where $G_{J_R}^R = \{\vec{q}_1^R ... \vec{q}_{J_R}^R\}$. Then correlation (2. 23) may be transformed to the view

$$\vec{y}_R = W_R \vec{x} . \qquad (2.\ 25)$$

Directly from presentation of elements of the matrix (2. 21)

$$a_{ik}^R = \int_{\omega \in \Omega_R} \exp(-j\omega(i-k)) d\omega / 2\pi \qquad (2.\ 26)$$

It is not difficult to obtain characteristic feature of its proper numbers and vectors

$$\lambda_k^R = \int_{\omega \in \Omega_R} | Q_k^R(\omega) |^2 \, d\omega / 2\pi , \qquad (2.\ 27)$$

thus, proper values are equal to share of the energy of corresponding vector, which falls within united frequency interval (here Parseval equality has been taken into account [9], as well as the equality of norm's units of proper vectors).
Consequently, usage of only non-zero proper numbers narrows the frequency band, in which vector's energy is made it into (2. 25).
That's why it seems natural that correlation (2. 25) is an operator of sub band smoothing (SBS) or filtration, and W_R is a matrix of operator of sub band smoothing (MOSBS).
The definition (2. 23) directly allows us obtaining correlations for scalar product of original vector and result of its smoothing.

$$(\vec{x}, \vec{y}_R) = \| \vec{y}_R \|^2 = \sum_{k=1}^{J_R} (\alpha_k^R)^{\geq} \geq 0 . \qquad (2.\ 28)$$

That's why equations below are correct

$$\rho(x, y_R) = (\vec{x}, \vec{y}_R)/(\| \vec{x} \| \| \vec{y}_R \|) = (\sum_{k=1}^{J_R} (\alpha_k^R)^2 / \sum_{i=1}^{N} (\alpha_i^R)^2)^{1/2} , \qquad (2.\ 29)$$

$$\gamma_R(x, y_R) = \| \vec{x} - \vec{y}_R \|^2 = \sum_{k=J_R+1}^{N} (\alpha_k^R)^2 . \qquad (2.\ 30)$$

It seems natural to call the characteristics (2. 29) a similarity coefficient of original and smoothed factors. Its positive property of having fixed sign within the context of the problem of building trends that is being solved, is a substantial property.
The correctness follows from the orthonormality of the set of proper vectors of sub band matrix (2. 21):
Statement 2. *Fixed points of SBS* (2. 25) *are represented only by vectors of the view*

$$\vec{u} = G_J^R \vec{b}, \vec{b} = (b_1, ..., b_{J_R})' , \qquad (2.\ 31)$$

where components of vector $\vec{b}$ *are random numbers (including complex ones).*
It is obvious that Fourier transform of (2. 23) vectors may be presented in the following view

$$Y(\omega) = \sum_{ik1}^{J_R} \alpha_k^R Q_k(\omega) . \qquad (2.\ 32)$$

That's why we may consider that this is the realization of original vector's a variant of approximate analytical continuation of Fourier transform from selected frequency interval onto adjacent subareas of frequency axis.

2. Adaptive building of SBS

We should note that representation (2. 31) defines many trends, which meet the "ideal" condition (2. 3). Real trends will meet it only roughly; besides, the degree of approximation is defined by the number of vectors in representation (2. 31) and the choice of frequency interval.

Since it is supposed that desired trend is more narrow-band than time series that is being analyzed, it seems rational to require the performance of the following variation condition.

At equal total width

$$S_R = \sum_{r \in R} (\Omega_{2r} - \Omega_{1r}) \qquad (3.\ 1)$$

among several united frequency intervals of a view (2. 19) we select the one with the higher similarity coefficient (2. 29).
Let us note that it is difficult to gain an exact equality of obtained similarity coefficient to given value, that's why it is a question of some valid approximation to it.
Let us suppose that hereafter $K+1$ means total amount of frequency intervals, on which frequency axis is divided $[-\pi, \pi)$, $\Delta = \pi./(2K+1)$, $\omega_\kappa = 2(r-1)\Delta, r = 1,\ldots,K+1$, and the borders of frequency intervals are defined by the following correlations

$$\Omega_{10} = 0; \Omega_{20} = \Delta; \Omega_{1r} = \omega_r - \Delta; \Omega_{2r} = \omega_r + \Delta, r = 1,\ldots,K, \qquad (3.\ 2)$$

i. e., frequency intervals are disjoint and cover the entire particular axis.
The number of frequency intervals may be any, however it is appropriate to adhere to the following equation:

$$K = [(N-2)/4], \qquad (3.\ 3)$$

where square brackets define the integral part of the content. It corresponds to the fact that then Δ would be equal to half of the width of spectrum's main lobe of rectangular window with N duration (Dirichlet kernel). At the same time, maximal proper number of sub band matrix A_0 with elements

$$a_{ik}^0 = \sin(\Delta(i-r))/\pi(i-k); a_{ii}^0 = 1/(2K+1), i,k = 1,\ldots,N \qquad (3.\ 4)$$

would be equal 0, 98, and $J_0 = 5$.
Whereas, it is not difficult to obtain presentation for elements of sub band matrixes from (2. 10) for other frequency intervals.

$$a_{ik}^r = 2a_{ik}^0 \cos(\omega_r(i-k)). \qquad (3.\ 5)$$

Let us assume that

$$P_r(x) = \int_{\omega \in \Omega_r} |X(\omega)|^2 / 2\pi \qquad (3.\ 5)$$

Bearing the definition (2. 4) in mind, it is not difficult to obtain representation of this characteristics directly in the area of originals.

$$P_r(x) = \vec{x}' A_r \vec{x}. \qquad (3.\ 6)$$

Let us assume that

$$sredx = \| \vec{x} \|^2 / 2\pi \qquad (3.\ 7)$$

$$h_r = 1{,}378\, sredx \circ (\Omega_{2r} - \Omega_{1r}), \qquad (3.\ 8)$$

thus, we mean density of vector's energy and parts that correspond to it, which fall on the intervals of values, defined above. Coefficient 1, 378 is a median (quantile) of probability distribution of chi-square with two degrees of freedom [3] (total integrated square if real and imaginary component of Fourier transform according to corresponding frequency intervals for Gaussian noise).
Based on correlations (2. 10) and (3. 2) we obtain the correlation:

$$\sum_{r=0}^{K} A_r = I = diag(1,\ldots,1). \qquad (3.\ 9)$$

That's why the following correlation is correct:

$$\sum_{r=0}^{K} P_r(x) = \| \vec{x} \|^2. \qquad (3.\ 10)$$

Thus, in our case, the total sum of energy's parts (3. 6) may be divided into two multitudes

$$R = \{P_r(x) \geq h_r, r \in R\} \qquad (3.\ 11)$$

and $\overline{R}$, if contrary inequalities are solved.
Having plugged presentation (1. 1) into (3. 6), we obtain correlation for parts of energy of analyzed vector in frequency intervals:

$$P_r(x) = P_r(f) + P_r(\varepsilon) + 2\vec{f}' A_r \vec{\varepsilon}.$$

That's why the following correlation is correct

$$sredx = sredf + sred\varepsilon + 2(\vec{x}, \vec{\varepsilon})/2\pi.$$

Supposing that the majority of trend's energy is concentrated in a frequency area that is totally narrower than random fluctuations, it is reasonable that this is the area where inequalities of (3. 11) view will more likely be performed than the contrary ones (noon-central chi-square distribution of integrals' probabilities from squares of real and imaginary component of Fourier transform)
That's why the multitude of frequency intervals (3. 11) is naturally called the informational one, and frequency intervals it is comprised of are called informational intervals.
Further steps involve using correlation (2. 21) and calculation of proper numbers and vectors for total sub band matrix.
We should also note that the range of definition of Fourier transform of any vector of finite dimension is all frequency axis within the borders $[-\pi, \pi)$. That's why for accurate

trend's reproduction we will need all the set of proper vectors of total sub band matrix, which is a full basis for representation of vectors of this dimension [8]. However, at this, distorting influences would be to faithfully reproduced. That's why when choosing the number of proper vectors J_R for formation of SBS matrix of a view (2. 24) it is necessary to reach a compromise between the desire of having accurate trend's reproduction as a fixed point (proper vector) of operator (2. 25) and degree of its purification from distorting influences.

CONCLUSION

We have obtained correlations that allow building adaptive filter in time series (signals) of components, energies of which constitute the majority.

ACKNOWLEDGEMENTS

The research was partially financed as apart of Federal target-oriented program (FTOP) "Research and development in priority development fields of Russia's science and technology sector for 2014-2020", Agreement 14. 575. 21. 0020 dated 17. 06. 2014 (Unique identifier RFMEFI57514X0020).

REFERENCES

1 Kendall M., Stuart A. Multivariate statistical analysis and time series. M.: Nauka, 1976.
2 Anderson T. Stastistical analysis of time series. M.: Mir, 1976.
3 Draper N, Smith H. Applied Regression Analysis, vol. 2. M.: Finances and statistics, 1986.
4 Otnes R., Enochson L. Applied time series analysis. M.: Mir, 1982.
5 Bexruchko B. P., Smirnov D. A. mathematical modeling and chaotic time series. Saratov: GosUNC "College", 2005.
6 Tukey J. Exploratory Data Analysis. M.: Mir, 1981.
7 Greshilov A. A., Stakun V. A., Stakun A. A. Mathematical models of making prognoses. M.: Radio and communication, 1997.
8 Gantmacher F. R. Tjeory of matrixes. M.: Nauka, 1967.
9 Khurgin Ya. I., Yakoivlev V. P. Finite functions in physics and technics. M.: Nauka, 1971