

Forecasting the Trend Component of Time Series Using a Trigonometric Basis

V.M. Balyk¹, E.V. Sudilina², K. H. D. Nguyen³, M. T. Nguyen⁴, T. C. Vu⁵,
H. T. T Nguyen⁶

¹Moscow Aviation Institute (National Research University), Moscow, Russia, balikv@gmail.ru

²Moscow Aviation Institute (National Research University), Moscow, Russia, hitechsun@yandex.ru

³Irkutsk National research technical university, Irkutsk, Russia, hd04011996@gmail.com

⁴MIREA – Russian Technological University, Moscow, Russia, Nguen_m@mirea.ru

⁵MIREA – Russian Technological University, Moscow, Russia, Vutrichien00@gmail.com

⁶Vietnam Maritime University, Hai Phong, Viet Nam, trangnht@vimaru.edu.vn

ABSTRACT

This paper examines the problem of constructing a nonlinear forecast for time series using a trigonometric polynomial framework. Apart from the linear case, time series forecasting remains a complex and insufficiently resolved task. In the development of a technical system, three closely interrelated types of forecasts are typically distinguished:

1. *forecasting the operational conditions under which the prediction will be applied;*
2. *forecasting optimal system parameters;*
3. *forecasting the functional characteristics of the technical system.*

Although this conventional forecasting scheme is conceptually simple, it lacks reliability for medium- and especially long-term predictions. Furthermore, the need for prototypes when forming statistical datasets leads to the rapid obsolescence of information concerning new technical solutions by the time the data sample is constructed.

In this study, we propose a new variational approach to forecasting, in which the regularity criterion serves as the variational functional and trigonometric polynomials are used as the variational basis. The optimal structure of the basis functions is determined through statistical synthesis, followed by the construction of the most effective basis components. We demonstrate that statistical synthesis enables the development of simple and efficient algorithms for generating basis functions. Additionally, we introduce specialized operations on statistical samples that give rise to fundamentally new algorithms for solving multidimensional forecasting problems.

A model experiment confirms the practical convergence and effectiveness of the proposed statistical forecasting method.

Keywords: statistical synthesis; dataset; variational polynomial; forecasting; basis function; regularity criterion.

Mathematics Subject Classification: 41xx, 37M10, 62M10

INTRODUCTION

It is well known that variational methods are an effective tool for solving complex problems in numerical analysis, particularly in multidimensional settings where grid-based approaches and the finite element method face the well-known “curse of dimensionality” [1]. However, even within variational frameworks, a number of methodological difficulties arise. One of the common challenges is the slow convergence

of the variational series to the exact solution as the size of the variational basis increases, which typically results in the inversion or diagonalization of extremely large matrices. Such slow convergence is generally due to the fact that the exact solution is not sufficiently similar in form to the preselected set of basis functions [2,3]. Kuklin (1975) proposed a fundamentally new approach was proposed to overcome these difficulties—namely, a step-by-step stochastic optimization of a nonminimal variational basis, referred to as the *stochastic variational method*. In this approach, the variational basis is not chosen universally but is instead constructed through random trials specifically tailored to the boundary-value problem under consideration [4]. This strategy allows for a substantial reduction in the basis size and enables fully analytical computation of the required matrix elements. Nonetheless, for many complex problems, a significant issue remains: a large number of random trials are required to optimize the basis. This leads to considerable increases in computational time, especially when the analytical expressions for the matrix elements involve multidimensional sums or special functions [5-7].

This work further develops this new line of research. The proposed method significantly reduces the number of random trials required to evaluate the variational functional by employing an algorithm that analyzes the distribution of functional values obtained during these trials and based on this analysis, predicts the most optimal form of the basis components. This approach, referred to as the *statistical forecasting method*, allows for a substantial improvement in the efficiency of constructing the variational basis [8].

Statistical Synthesis of Forecasting in a Gaussian Basis [9-12]

Let the forecasting problem be reduced to minimizing the functional $E[\psi]$ in the space L_2 . We expand the desired solution $\psi(x)$ — where x denotes a symbolic coordinate, either one- or multidimensional — into a series with respect to a non-minimal basis $\{\varphi_i = \varphi(x, \alpha_i), i = 1, 2, \dots\}$, where α_i are the nonlinear parameters of the basis:

$$\psi(x) = \sum_{i=1} C_i \varphi_i(x), \quad (1)$$

The coefficients C_i represent the linear parameters of the trial function ψ . For forecasting problems, it is convenient to use the Ritz variational principle, which reduces the problem to minimizing the variational functional $E[\psi]$ in the space spanned by the vectors of the chosen variational basis $\{\varphi_i(x)\}$. As is well known, this leads to a system of linear algebraic equations for the linear superposition coefficients C_i in (1) ($i = 1, 2, \dots$), whose nontrivial solutions yield the desired approximate solution of the variational problem associated with the basis functions $\varphi_i(x)$.

We then proceed analogously to optimize α_2 , that is, the second component in (1), followed by α_3 , and so forth. Upon reaching α_N , this procedure corresponds to the so-called coordinate descent (or coordinate relaxation) in the established classification of multivariable function minimization methods. Alternatively, one may retain the N optimized basis components $\{\varphi(x, \alpha_i^{(m)}, i = 1, \dots, N)\}$, add N (or $M \neq N$) new components — thus expanding the basis to dimension $N + M$ (or $2N$) — and then optimize these new components using the same method described above.

In any case, the algorithm described above for determining the optimal basis components ultimately reduces to selecting a single best element (in a certain sense) from a given random sample and discarding all other “nonoptimal” components. It is clear, however, that significant potential for

improvement lies in the effective use of the nonoptimal components that were previously discarded under this approach.

The essence of the new method proposed in this work for optimizing the basis is as follows. The set of random values of the functional E_p ($p = 1, 2, \dots, m$), generated during random trials, is interpreted as the outcome of a set of experiments performed on the system. Based on these experimental results, a statistical prediction is then made regarding the optimal values of the nonlinear parameters of the variational basis. This is achieved using well-developed methods and algorithms from the theory of statistical synthesis of technical systems in automatic control [13].

The main elements of statistical synthesis are statistical samples, from which the reconstructed dependencies are formed, as well as external criteria that determine the optimal order of the reconstructed regression polynomial [14-16]. A statistical sample serves as the primary carrier of information in statistical synthesis problems. Such samples are typically divided into two unequal parts, referred to as the training sequence and the validation sequence [15]. Using the training sequence, the linear parameters of the regression polynomial are computed via the least-squares (LS) method, while the external criteria are evaluated at the validation points to determine the optimal degree of the reconstructed polynomial [17, 18].

As established in [15], the optimal degree of the reconstructed polynomial corresponds to the minimum of the external criterion — in our case, the variational functional $E[\psi]$.

Suppose based on statistical data, it is required to reconstruct an n -dimensional function $J = J(x_1, x_2, \dots, x_n)$. For such a problem, the corresponding statistical sample takes the following form:

Table 1: Multidimensional statistical sample

x_1	x_2	$\dots$	x_n	J	
$x_{1,1}$	$x_{2,1}$	$\dots$	$x_{n,1}$	J_1	} A
$x_{1,2}$	$x_{2,2}$	$\dots$	$x_{n,2}$	J_2	
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	} B
$x_{1,N}$	$x_{2,N}$	$\dots$	$x_{n,N}$	J_N	

Here, the statistical sample of size N contains n input variables x_i , $i = 1, \dots, n$, and a single output characteristic J . In our case, this output corresponds to the values of the variational functional $J[\psi]$ evaluated at random values of the parameters $x_1, x_2, \dots, x_n$. The sample is conventionally divided into two parts: A — the training sequence, B — the validation sequence.

The main idea behind constructing a regular method for determining the optimal values of the nonlinear parameters α_j , $j = 1, 2, \dots$ of the variational basis is to establish a correlation relationship between these nonlinear parameters and the functional of the variational problem $J[\psi] = J(\alpha_1, \dots, \alpha_N)$. This relationship is subsequently used to predict the values of α_j , $j = 1, 2, \dots$, that minimize the functional $J[\psi]$ at a given step of the optimization process.

Such a relationship can be constructed by means of a specially formed statistical sample.

In the considered algorithm, statistical samples are constructed at the stage of determining each term of the variational series. At the k -th step of the algorithm, the statistical sample for the problem takes the following form:

Table 2: Sample for Constructing the Wave Function

α	$\psi(x, \alpha_k)$	$E[\psi]$
$(\alpha_k)_1$	$\psi_i(x, \alpha_k)_1$	E_1
$(\alpha_k)_2$	$\psi_i(x, \alpha_k)_2$	E_2
...	...	...
...	...	...
$(\alpha_k)_N$	$\psi_i(x, \alpha_k)_N$	E_N

Here, $k = 1, \dots, m$ denotes the step number, and $\psi(x, \alpha_k)_r$ represents a trial function of the form (1), consisting of N terms, in which all nonlinear parameters except α_k are fixed, while α_k is being optimized.

Table 3: Reduced sample of the form:

α	E_T
$(\alpha_k)_1$	$E_{T,1}$
$(\alpha_k)_2$	$E_{T,2}$
...	...
...	...
$(\alpha_k)_N$	$E_{T,N}$

Based on this reduced sample, a power polynomial of optimal complexity is constructed in the form:

$$E_M(\alpha_i) = b_0 + b_1\alpha_i + b_2\alpha_i^2 + \dots + b_p\alpha_i^p, i = 1, 2, \dots, N \quad (2)$$

where the order p is determined by minimizing an external criterion.

Unlike internal statistical criteria, external criteria are evaluated at the points of the statistical sample that were not used in determining the linear parameters b_i , $i = 1, 2, \dots, p$, of the polynomial obtained by the LS method.

The main types of external criteria used in models for constructing the variational basis are as follows:

1. Regularity criterion:

$$\Delta^2(B) = \frac{\sum_{i=1}^{N_b} (J_T - J_M)_i^2}{\sum_{i=1}^{N_b} J_{T,i}^2} \Rightarrow \min \quad (3)$$

where $J_{T,i}$ are the tabulated (true) values of the functional, $J_{M,i}$ are the model values of the functional at point i , computed from polynomial (2), and N_b is the number of sample points not used in constructing model (2).

2. Bias-Minimization Criterion:

$$n_{\text{CM}}^2 = \frac{\sum_{i=1}^N (J_A - J_B)_i^2}{a \sum_{i=1}^N J_{T,i}^2} \Rightarrow \min \quad (4)$$

where

- J_A denotes the values of the functional obtained from the points with larger variance of the output variable,
- J_B denotes the corresponding values for the points with smaller variance,
- a is the extrapolation coefficient.

The bias-minimization criterion requires the maximum possible agreement between the output values produced by the two regression polynomials constructed from two different parts (A and B) of the statistical sample. According to the Widrow theorem [3], the optimal partition of the statistical sample consists of splitting it into two equal subsets, $N_A = N/2$ and $N_B = N/2$, each of which may contain different data points [15, 19].

Criterion (4) makes it possible to choose the polynomial that is least sensitive to random errors in the input data used to construct it. According to this criterion, the polynomial should provide nearly identical predictions on the data sequences of both groups A and B. Thus, criterion (4) is suitable for solving the problem of reconstructing an underlying law from noisy experimental data. It is therefore reasonable to apply this criterion when constructing improved regression polynomials by artificially assigning weights to experimental points in order to refine the shape of the regression polynomial.

Let the regularity criterion be chosen as the functional of the problem:

$$E(\psi) = \frac{\sum_{i=1}^{N_B} (E_T - E_M)_i^2}{\sum_{i=1}^{N_B} E_{T,i}^2} \quad (5)$$

where N_B is the size of the validation subset, $E_T(i)$ denotes the tabulated (true) values of the functional, and $E_M(i)$ denotes the corresponding values produced by the variational polynomial.

The variational polynomials are represented in the form

$$E_M \cong \sum_{i=1}^N C_i \varphi(x, \alpha_i) \quad (6)$$

where $\varphi(x, \alpha_i) = \exp(-\alpha_i x^2)$, $i = 1, 2, \dots$. Here, $\varphi(x, \alpha_i)$ is a one-dimensional Gaussian basis function. The quantities x_i , $i = 1, \dots, N_p$, form the sample used to reconstruct the variational polynomial, where N_p is the size of this sample, and x_i is the value of the i -th forecasted characteristic. The coefficients C_i are the linear parameters of the function E_M .

The minimization of the variational functional $E(\psi)$ in the space spanned by the vectors of the chosen variational basis leads to a system of linear algebraic equations with respect to the linear coefficients C_i , $i = 1, 2, \dots$. The solution of this system yields the desired approximate solution of the variational problem that minimizes the functional $E(\psi)$. The non-minimality of the complete basis $\{\varphi_i(x, \alpha_i)\}$ implies, in particular, that there exists an infinite number of possible choices of the nonlinear parameters α_i , $i = 1, 2, \dots$, that preserve the completeness of the basis (i.e., satisfy the conditions of the Müntz

theorem [20-22]). Therefore, the nonlinear parameters of the basis should be chosen such that the variational expansion (5) converges with the highest possible rate.

To compute the series (5), a statistical sample is constructed for each of its terms. Such a sample, corresponding to the coefficient α_1 , is shown in table 4.

Table 4: Statistical Data for Determining α_1

α_1	E_T
$(\alpha_1)_1$	$E_{M,1}$
$(\alpha_1)_2$	$E_{M,2}$
$\dots$	$\dots$
$\dots$	$\dots$
$(\alpha_1)_N$	$E_{M,N}$

With

$$E_{M,1} = C_0 + C_1 \exp(-\alpha_1 x^2)$$

Below is the process of constructing the variational polynomial (6), starting from the second term.

α_2	$E_{M,2}$
α_3	$E_{M,3}$

$$\Leftrightarrow E_{M,2} = C_0 + C_1 \exp(-\alpha_1 x^2) + C_2 \exp(-\alpha_2 x^2)$$

$$\Leftrightarrow E_{M,3} = C_0 + C_1 \exp(-\alpha_1 x^2) + C_2 \exp(-\alpha_2 x^2) + C_3 \exp(-\alpha_3 x^2)$$

...

...

α_N	$E_{M,N}$
------------	-----------

$$\Leftrightarrow E_{M,N} = C_0 + C_1 \exp(-\alpha_1 x^2) + C_2 \exp(-\alpha_2 x^2) + C_3 \exp(-\alpha_3 x^2) + \dots + C_N \exp(-\alpha_N x^2)$$

To illustrate the numerical construction of the variational polynomial, we present a fragment of the solution to a problem discussed in [23].

We consider the problem of constructing the seventh term of the variational polynomial, i.e., determining the nonlinear parameter α_7 based on the statistical data of the following form.

Table 5: Statistical data for computing the seventh term of the variational polynomial

$E_T = \alpha_7$	E_M
0.16	-0.467
0.48	-0.18
57	-0.462
0.06	-0.462
0.002	-0.462
1.65	-0.498
5.469	-0.484
18	-0.467
34	-0.463
68	-0.462
72	-0.462

The polynomial of optimal smoothness is reconstructed in the form

$$J(x) = \frac{C_0}{2} + \sum_{i=1}^m [a_i \cos(w_i x) + b_i \sin(w_i x)] \quad (7)$$

where m is the number of selected harmonics, C_0 is the mean value of the time series,

$$C_0 = \frac{1}{N} \sum_{j=0}^{2N-1} J(x_j)$$

where N is the size of the statistical sample, a_i and b_i are the Fourier coefficients,

$$a_i = \frac{1}{N} \sum_{j=0}^{2N} J(x_j) \cos(w_i x_j)$$

$$b_i = \frac{1}{N} \sum_{j=0}^{2N-1} J(x_j) \sin(w_i x_j)$$

where w_i are the frequencies of the process, and $j = 1, \dots, m$.

From the solution of the minimization problem (7), the forecasted parameter x_{N+1} is determined as

$$\min J(x_{N+1}) = J_T(x_{N+1})$$

Model Problem

We consider a model problem of forecasting the wholesale price of AI-92 gasoline. Statistical data on the price dynamics of AI-92 for the period from 1997 to 2002 were taken from [25]. Since the prices are provided on a monthly basis, the total size of the initial dataset is 72 observations ($N = 72$).

Table 5 presents the original statistical data, and figure 1 shows the corresponding plot. Here, t denotes time, and C represents the cost of one liter of AI-92 gasoline in conventional units.

Table 5: Dynamics of Wholesale Prices for AI-92 Gasoline

t	C	t	C	t	C
1	1.31	28	2	55	5.48
2	1.39	29	2.62	56	4.68
3	1.46	30	3.23	57	5.98
4	1.46	31	3.31	58	4.12
5	1.46	32	4.37	59	4.32
6	1.46	33	5.76	60	5.65
7	1.46	34	5.66	61	5.24
8	1.54	35	5.61	62	4.84
9	1.54	36	5.37	63	4.58
10	1.62	37	5.37	64	5.47
11	1.62	38	5.18	65	6.56
12	1.62	39	5.11	66	8.09
13	1.69	40	5.05	67	8.29
14	1.69	41	5.01	68	8.2
15	1.58	42	4.93	69	8.35
16	1.62	43	5.43	70	8.35
17	1.54	44	5.74	71	7.89
18	1.39	45	6.85	72	6.72
19	1.46	46	7.12		
20	1.39	47	7.45		
21	1.42	48	7.05		
22	1.46	49	6.7		
23	1.5	50	6.16		
24	1.54	51	6.14		
25	1.58	52	6.04		
26	1.69	53	5.78		
27	1.88	54	5.78		

At this stage, the input data (nonlinear parameters) are generated using a random number generator over the interval $[0,100]$ with a normal distribution.

Based on the given statistical data, the following polynomial of optimal smoothness was obtained:

$$E(\alpha_7) = -0.473 - 0.01\alpha_7 + 0.000067\alpha_7^2$$

from which the optimal value is determined as $\alpha_7^{opt} = 13.972$.

Considering the calculated value α_7^{opt} , the variational polynomial takes the form:

$$\begin{aligned} E_{M_7}(x) = & -8.431 \exp(-0.282x^2) + 0.00017 \exp(-0.001x^2) - 0.0014 \exp(-0.01x^2) \\ & + 0.0635 \exp(-0.123x^2) \\ & - 0.012 \exp(-23.775x^2) + 9.565 \exp(-0.25x^2) + 0.00414 \exp(-13.972x^2) \end{aligned}$$

Statistical Synthesis of Forecasting in a Trigonometric Basis

The construction of forecasts in the class of trigonometric polynomials is of considerable importance, since any continuous function $f(x)$ defined on a circle Q can be represented as the limit of a uniformly convergent sequence of trigonometric polynomials

$$J(x) = \sum_{k=0}^h (a_k \cos k\varphi + b_k \sin k\varphi)$$

with real coefficients [24].

At the first stage of the algorithm, a forecast is constructed based on a statistical sample of the form presented in table 6.

Table 6: Statistical Sample

x	$J(x)$
x_1	$J(x_1)$
x_2	$J(x_2)$
x_3	$J(x_3)$
...	...
...	...
x_N	$J(x_N)$
x_{N+1}	$J(x_{N+1})$

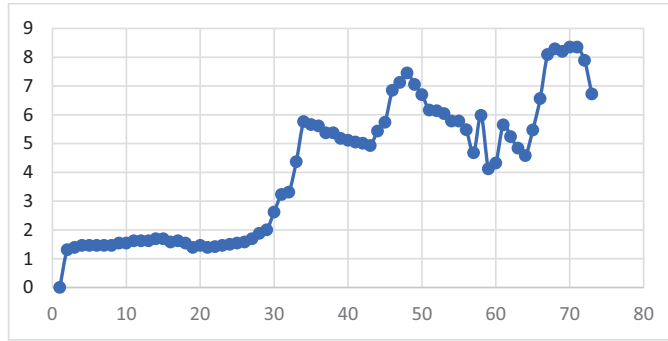


Figure 1. Graph of the original statistical sample

From the plot, it is clear that the function $C = C(t)$ exhibits a complex, highly multiextremal dependence, which makes it practically impossible to construct a reliable forecast directly from the raw data. As a variational basis we employ the harmonic basis

$$\varphi(x, w_i) = a_i \cos\left(\frac{2\pi x_i}{w_i}\right) + b_i \sin\left(\frac{2\pi x_i}{w_i}\right), i = 1, 2, \dots$$

where x_i is the argument of the process, w_i is the process frequency, and a_i, b_i are the Fourier coefficients.

The size of the given sample is $N_A = 64$. It is required to provide a forecast for the 65-th observation. It is assumed that the value at the 65-th point lies in the interval $[4, 10]$.

A supplementary statistical sample of size $N = 11$ is constructed (Table 7). Each row of this sample represents a candidate value of $S_t(65)$. Here, S_t is the regularity criterion defined as

$$St = \frac{\sum_{i=1}^N [C_i \varphi(x, w_i) - C_{Ti}]^2}{\sum_{i=1}^N C_{Ti}^2} \Leftrightarrow \min$$

where C_{Ti} denotes the observed (tabulated) cost of one liter of AI-92 gasoline. The elements of the supplementary statistical sample x_i are randomly distributed over the interval $[4, 10]$.

Table 7: Supplementary Statistical Sample

N_i	$ar(i)$	$St(i)$
1	4	$3.175 \cdot 10^{-5}$
2	4.5	$3.092 \cdot 10^{-5}$
3	5	$3.304 \cdot 10^{-5}$
4	5.5	$3.389 \cdot 10^{-5}$
5	6.5	$3.502 \cdot 10^{-5}$
6	7	$2.851 \cdot 10^{-5}$
7	7.5	$4.501 \cdot 10^{-5}$
8	8	$3.563 \cdot 10^{-5}$
9	8.5	$7.231 \cdot 10^{-5}$
10	9	$5.809 \cdot 10^{-5}$
11	9.5	$7.198 \cdot 10^{-5}$

Here, $ar(i), i = \overline{1,11}$, denotes the candidate values of the cost of one liter at the 65-th time point, while $S_t(i)$ represents the cost (or total value) of the entire system under the assumption that the candidate value $ar(i)$ is taken.

The variational polynomial obtained from the statistical data in Table 7 has the form:

$$S_t = 4.328 \cdot 10^{-5} + 1.451 \cdot 10^{-5} \cos\left(\frac{6.28 \cdot x}{8.675}\right) + 1.251 \cdot 10^{-5} * \sin\left(\frac{6.28 \cdot x}{8.675}\right) + 2.191 \cdot 10^{-6} * \cos\left(\frac{6.28 \cdot x}{4.396}\right) + 8.232 \cdot 10^{-6} * \sin\left(\frac{6.28 \cdot x}{4.396}\right) \quad (8)$$

From the minimization condition of polynomial (8), we obtain

$$x^{opt} = ar^{opt} = 6.656$$

This value yields an error of approximately 1.5% when compared with the exact experimental value.

Conclusion

1. A new approach to constructing a variational basis has been examined. This approach is grounded in the statistical synthesis of a polynomial whose extremum depends on the parameter being forecast. The framework of dynamic statistical samples has a significantly broader range of applications than merely constructing optimal basis functions.
2. The capability of statistical synthesis to construct inverse dependencies using remarkably simple algorithms is demonstrated. This feature is particularly important, as it leads to fundamentally new models of optimization and forecasting.

References

1. Petromichelakis, I., Kougioumtzoglou, I.A. (2020). Addressing the curse of dimensionality in stochastic dynamics: A Wiener path integral variational formulation with free boundaries. *Proceedings of the Royal Society A*, 476, 20200385.
2. Sinha, V.K., Maroju, P. (2023). New development of variational iteration method using quasilinearization for solving nonlinear problems. *Mathematics*, 11(4), 935.
3. Ahmad, H., Khan, T.A., Stanimirović, P.S., Chu, Y.M., Ahmad, I. (2020). Modified variational iteration algorithm-II: Convergence and applications to diffusion models. *Complexity*, 2020, 1–14.
4. Kuklin V.I. *Izvestiya AN SSSR, Series Physics*, 1975, Vol. 39, No. 3, pp. 535–542.
5. Beutel, M., Ahrens, A., Huang, C., Suzuki, Y., Varga, K. (2021). Deformed explicitly correlated Gaussians. *Journal of Chemical Physics*, 155, 214103.
6. Muolo, A., Mátyus, E., Reiher, M. (2018). Explicitly correlated Gaussian functions with shifted center and projection techniques in pre-Born–Oppenheimer calculations. *Journal of Chemical Physics*, 149, 184105.
7. Hermann, J., Schätzle, Z., Noé, F. (2020). Deep-neural-network solution of the electronic Schrödinger equation. *Nature Chemistry*, 12, 891–897.
8. Khoo, Y., Lu, J., Ying, L. (2021). Solving parametric PDE problems with artificial neural networks. *European Journal of Applied Mathematics*, 32(3), 421–435.
9. Basson, M., Louw, T. M., Smith, T. R. (2023). Variational Tobit Gaussian process regression. *Statistics & Probability Letters*, Article 10225-3. <https://doi.org/10.1007/s11222-023-10225-3>
10. Mukherjee, S., Sen, S. (2022). Variational inference in high-dimensional linear regression. *Journal of Machine Learning Research*, 23, 1-56.

11. Wynne, G., et al. (2022). Variational Gaussian Processes: A Functional Analysis View. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics (AISTATS 2022)*, PMLR 151, 231–239.
12. Cavoretto, R., De Rossi, A., Lancellotti, S. (2023). Bayesian Approach for Radial Kernel Parameter Tuning. arXiv preprint arXiv:2311.04311
13. Bulygin V.S., Krasovskaya M.A., Khakhulin G.F. *Theoretical Foundations of Automated Control*. Moscow: MAI Publishing, 2005.
14. Lyu, B., Li, J. (2024). Multiple-model polynomial regression and efficient algorithms for data analysis. *Theoretical Computer Science*, 1021, 114878. <https://doi.org/10.1016/j.tcs.2024.114878>
15. Ivakhnenko A.G., Müller I.A. *Self-Organization of Predictive Models*. Kiev: Tekhnika, 1983.
16. Antil, Zh., Vainberg-Allen, A.L. (2007). D-optimal planning for polynomial regression: Degree selection and robustness (*D-optimal'noe planirovanie dlya polinomial'noi regressii: vybor stepeni i roblastnost'*). *Prikladnaya Ekonometrika (Applied Econometrics)*, 4, 65–66.
17. Lin, Scott Shu-Cheng, Analytic Hierarchy Process by Least Square Method Revisit, *Mathematical Problems in Engineering*, 2019, 2797515, 5 pages, 2019. <https://doi.org/10.1155/2019/2797515>
18. Wolberg J. *Data Analysis Using the Method of Least Squares: Extracting the Most Information from Experiments*. Berlin; Heidelberg: Springer, 2006. XIII, 250 p. DOI: 10.1007/3-540-31720-1.
19. Rinaldo, A., Wasserman, L., G'Sell, M. (2019). Bootstrapping and sample splitting for high-dimensional, assumption-lean inference. *Annals of Statistics*, 47(6), 3438–3469. DOI:10.1214/18-AOS1784.
20. Trefethen, L.N. (2023). Spectacularly large expansion coefficients in Müntz's theorem. *Analyses and Applications* 21, 357–378. DOI: 10.1007/s44007-022-00039-6.
21. Baklouti, A., Ben Ayed, S. (2021). Müntz–Szász analogues for compact extensions of Heisenberg groups. *Applicable Analysis* 100, 1151–1167. DOI: 10.1007/s12044-021-00617-8
22. Kalman S., Shtengauz G. *Theory of Orthogonal Series*. Moscow: Fizmatlit, 1958.
23. Balyk V.M., Kostomarov D.P., Kukulin V.I., et al. *Self-Organizational Approach to Constructing a Variational Basis*. Mathematical Modelling, 2002, Vol. 14, No. 10, pp. 43–58.
24. Shilov G.E. *Mathematical Analysis (Functions of a Single Variable), Part 3*. 1970, p. 352.
25. *Dynamics of Wholesale and Retail Prices for AI-92 Gasoline and Oil Prices (1997–2003)*. Available at: www.gk3.ru/dbscript3/cbsd/dbinet.