ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2023, Vol. 84, No. 4, pp. 412–433.
Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, 2023.
Russian Text © The Author(s), 2023, published in Avtomatika i Telemekhanika, 2023, No. 2, pp. 122–149.

STOCHASTIC SYSTEMS

Optimal Recovery of a Square Integrable Function from Its Observations with Gaussian Errors

S. A. Bulgakov[*,a] **and V. M. Khametov**[*,b]

[*] *National Research University Higher School of Economics, Moscow, Russia*
e-mail: [a]*s.a.bulgakov@gmail.com,* [b]*khametovvm@mail.ru*

Received November 9, 2020
Revised August 9, 2022
Accepted September 29, 2022

Abstract—This paper is devoted to the mean-square optimal stochastic recovery of a square integrable function with respect to the Lebesgue measure defined on a finite-dimensional compact set. We justify an optimal recovery procedure for such a function observed at each point of its compact domain with Gaussian errors. The existence of the optimal stochastic recovery procedure as well as its unbiasedness and consistency are established. In addition, we propose and justify a near-optimal stochastic recovery procedure in order to: i) estimate the dependence of the standard deviation on the number of orthogonal functions and the number of observations and ii) find the number of orthogonal functions that minimizes the standard deviation.

Keywords: orthogonal functions, Fourier coefficients, observation error, projection estimator, unbiasedness, consistency

DOI: 10.25728/arcRAS.2023.21.74.002

1. INTRODUCTION

This paper considers the theory of optimal recovery of square integrable functions with respect to the Lebesgue measure defined on a finite-dimensional compact set that are observed with Gaussian errors. We establish the existence of an optimal recovery procedure in terms of the minimum standard deviation criterion using the minimum number of orthonormal functions.

The stochastic recovery of an unknown function from some class is usually understood as follows: the value of this function can be observed with errors at any point of its domain and the problem lies in estimating (recovering) it from observations in terms of a given optimality criterion. Note that this problem belongs to the theory of nonparametric (infinite-dimensional) estimation. Numerous researches were devoted to nonparametric estimation; for example, see [1–22].

Let us overview the results on the stochastic recovery of functions in chronological order. Many books by Russian researchers were first published in Russian and then translated into English. For the ease of foreign readers, we present both the Russian originals and their translated versions in the References section.

The papers [3, 4] considered the problem of estimating a one-dimensional unknown square-integrable probability density function from its independent observations. As was demonstrated, the problem can be solved using kernel estimators [3], which are asymptotically unbiased and consistent. Subsequently, this class of estimators was called the Parzen–Rosenblatt estimators. For brevity, probability density functions will be called densities below.

In [5], the Parzen–Rosenblatt estimators were generalized to the multivariate case.

The publication [6] was focused on recovering a scalar unknown function from its observations with uncorrelated Gaussian errors in a finite number of its domain points. Stratonovich described optimal recurrence-based recovery algorithms and established their rate of convergence to the unknown function.

In [7], Watson studied the problem of recovering an unknown density defined on the real line. The standard deviation of his projection estimator of the density, proposed therein, is equivalent to that of the projection estimator introduced by Chentsov [9].

The research work [8] solved the nonparametric estimation problem for an unknown density of an absolutely continuous random variable by observing m independent random variables with an unknown density. The author established conditions ensuring the existence of a kernel estimator of this unknown density. Moreover, these estimators were proved to be asymptotically unbiased and consistent. The cited study generalizes the well-known results of Parzen and Rosenblatt and Murphy [3–5, 7].

The monograph [9] outlined a new nonparametric estimation method. Chentsov was the first to introduce the concept of projection estimator. His approach to recovering an unknown density consists in estimating the Fourier coefficients of this density by a suitable system of orthonormal functions. As it turned out, such estimators are linear functionals of the observations. They are used to construct optimal estimators in terms of the minimum standard deviation criterion. Following this approach, the author proved the existence of a finite number of Fourier components ensuring convergence to the unknown density with an optimal rate by the order of magnitude.

In his book [10], Vapnik proposed a method for recovering an unknown function based on Kolmogorov's theory of n-widths and the Glivenko–Cantelli theorem.

The monograph [11] was devoted to asymptotic methods in the theory of point and nonparametric estimation. It also presented a theory for estimating an unknown smooth square-integrable signal observed against additive white Gaussian noise. This problem was shown to be the one of nonparametric estimation. The minimum standard deviation was adopted as an optimality criterion. Ibragimov and Has'minskii proved the existence of a signal estimator whose standard deviation is (by the order of magnitude) equivalent to that of the optimal estimator. In addition, this estimator was demonstrated to be unimprovable.

The book [12] was concerned with the nonparametric estimation of an unknown density. Asymptotic unbiasedness and consistency were established for the Parzen–Rosenblatt kernel estimators of densities. In addition, the limiting properties of the deviations of these estimators from the true density were investigated. The author also presented a method for constructing a nonparametric estimator of the regression curve.

In [13], the rates of convergence were found for the maximum likelihood estimation of an unknown L_2-function from its observations in finitely many points. The conditions obtained in the paper ensure the unimprovable rate of convergence. In particular, the authors proved that for a monotonic unknown function from the class L_2, the nonlinear maximum likelihood estimator has a better rate of convergence by the order of magnitude than any linear nonparametric estimator.

Darkhovsky [14, 15] surveyed in detail the results on nonparametric estimation using the minimax approach.

The paper [16] considered a two-component random vector with a random element taking values in some measurable space with a probability measure (the first component) and a random variable (the second component). The problem was to estimate the regression function for the first component by n independent observations of the second one. The regression function was assumed to belong to some class of smooth square-integrable functions with known metric characteristics such as the Kolmogorov ε-entropy or Kolmogorov n-widths. The asymptotic properties of its standard deviation were studied.

The three-chapter monograph [17] was devoted to the theory of nonparametric estimation. The following branches of the theory were described:

(i) design methods for nonparametric estimators;

(ii) statistical properties of nonparametric estimators (convergence and the rate of convergence);

(iii) adaptive nonparametric estimation procedures.

Branches (i) and (ii) were discussed in detail in Chapter 1. Branch (iii) constituted the main content of the monograph (Chapters 2 and 3).

The publication [18] dealt with recovering an unknown scalar square-integrable function from its observations at each point of a finite-dimensional compact domain with independent Gaussian errors. Using the spectral representation, the authors derived conditions for the existence of an optimal recovery procedure in terms of the minimum standard deviation criterion. Moreover, the recovery procedure was shown to be unbiased and consistent.

The paper [19] considered the application of optimal interpolation methods based on the properties of Abel–Jacobi elliptic functions for estimating nonparametric regressions. It described optimal interpolation methods for statistical data in terms of some optimality criteria, particularly their use in stochastic recovery problems.

In [20], Juditsky and Nemirovski developed a stochastic recovery method for general linear models. As was shown, this problem can be reduced to monotonic variational inequalities, numerically solvable by well-known and efficient computational procedures. The authors proved that strongly monotonic variational inequalities have an upper bound.

The research work [21] considered the problem of estimating a linear functional from its observations, i.e., an additive mixture of this functional and white Gaussian noise. Projection estimators were used for this functional. Golubev described a methodology for selecting the best estimator in this class and detailed the idea of constructing such estimators.

The paper [22] was concerned with the stochastic recovery of scalar, smooth, deterministic, square-integrable functions with respect to the Lebesgue measure on the real line from their observations with independent Gaussian errors at each point of the domain. Existence conditions were established for optimal estimators in terms of the minimum standard deviation criterion. The problem statement considered therein has the following peculiarity: both the observed sequence and the performance criterion were described not in the coordinate form but in terms of the Fourier coefficients of the observations, recovered function, and observation errors. Such a representation with the trigonometric basis yields results in a simple and convenient form. Easily verifiable conditions were derived for the existence of the optimal recovery procedure and its important properties (unbiasedness and consistency) were proved. In addition, for smooth functions from the Sobolev space, the authors constructed a recovery procedure equivalent to the optimal one. Note that in this case, the constructed procedure has the standard deviation with the following properties:

(i) Due to the presence of bias, the deviation is smaller than the optimal one.

(ii) It does not depend on the recovered function.

(iii) It is unimprovable.

Topicality. According to the survey above, stochastic recovery has been studied in numerous research works. In [3–17, 19, 20], this problem was stated in the coordinate representation; see formulas (1)–(4) and (15) in Section 2 for the corresponding mathematical description. It follows from [26] that the solution of the optimal problem (15) exists under the conditions of the Yankov–von Neumann lemma; furthermore, it is an analytical function. Hence, the solution of problem (15) is not a Borel function and, therefore, there is no nonparametric statistical estimator of an unknown square-integrable function observed against additive white noise.

Let us also mention the publication [21], which showed the following fact: in the coordinate representation, there is no maximum likelihood estimator for an unknown square-integrable function observed against noises described by a Gaussian random function in some Hilbert space.

Another approach was introduced in [18] under the assumption of observing the Fourier coefficients of the additive mixture (7) of an unknown function and the Gaussian error function. Such a representation is called spectral. In particular, formula (7) yields existence conditions and explicit-form estimates of each Fourier coefficient and its variance. Using this fact below, we derive an explicit form of the optimal nonparametric estimator and the standard deviation of an unknown function, for the first time in the literature; see Theorem 1 and Corollary 1. Based on these results, the unbiasedness and consistency of the estimators are established (Theorems 2 and 3). Meanwhile, how close are the nonparametric estimators proposed in [3–17, 19–21] to the optimal ones? This issue is not clear.

Also, we obtain the following results:

(i) an explicit dependence for the standard deviation of an unknown function estimated by a random function with only the first $N \in \mathbb{Z}^+$ terms in (23) and $m \in \mathbb{Z}^+ \backslash 0$ observations (Theorem 4);

(ii) for each m, the existence of $N^0(m) \in \mathbb{Z}^+$ that minimizes the standard deviation (Theorem 5). In addition, a constructive method is proposed to find $N^0(m)$ (Theorem 6). This result gives conditions for the equivalence of the standard deviation and m^{-1} as well as the equivalence of $N^0(m)$ and m; see Corollary 3. With these assertions, we define the Chentsov projection estimator (formula (40)) and determine its standard deviation. Moreover, we estimate the rate of convergence of this procedure to the unknown function (Theorem 8) and its independence from the unknown function (Theorem 10).

At the end of this paper, we provide an example with the number $N^0(m) \in \mathbb{Z}^+$ calculated explicitly.

The remainder of this paper is organized as follows.

Section 2 states the stochastic recovery problem in the spectral representation.

In Section 3, we establish an existence condition for the solution of the stochastic recovery problem (Theorem 1), which is in turn a nonparametric estimation problem. In addition, we demonstrate important statistical properties of optimal recovery, namely, unbiasedness (Theorem 2) and consistency (Theorem 3).

Section 4 is devoted to finding the dependence of the standard deviation $V_m(N)$ on the number of orthonormal functions and the number of observations (Theorem 4). Here, (i) for each $m \in \mathbb{Z}^+ \setminus 0$, we prove the existence of $N^0(m)$ minimizing $V_m(N)$ (Theorem 5) and (ii) we describe a constructive method for finding $N^0(m)$. In addition, the conditions for the equivalence of $V_m(N^0)$ and $\frac{N^0(m)}{m}$ as well as the equivalence of $N^0(m)$ and m are derived.

Section 5 describes the properties of optimal projection estimators of the unknown function with the standard deviation $V_m(N^0(m))$ (see formula (40)). They are called the Chentsov projection estimators (CPEs). In addition, the following results are proved here:

1) CPE converges to the unknown function with a rate of $m^{-\frac{1}{2}}$.

2) CPE is asymptotically unbiased and consistent.

The proofs of all assertions are postponed to Appendices A and B.

2. STOCHASTIC RECOVERY: PROBLEM STATEMENT, DEFINITIONS, NOTATIONS, AND JUSTIFICATION OF THE SPECTRAL REPRESENTATION

2.1. Let K be a finite-dimensional compact set, $\mathcal{B}(\mathrm{K})$ be a Borel σ-algebra in K, and $L_2(\mathrm{K}, \Lambda)$ be the set of square integrable functions $f\colon \mathrm{K} \to \mathbb{R}^1$ with respect to the Lebesgue measure Λ on K, i.e., $\int_{\mathrm{K}} f^2(x)dx < \infty$. Since $L_2(\mathrm{K},\Lambda)$ is a separable Hilbert space, it has a complete orthonormal system of functions (generally speaking, nonunique), which will be denoted by $\{\varphi_j(x)\}_{j\geqslant 0}$. In other words, for any $j, j' \in \overline{\mathbb{Z}}^+$ ($\overline{\mathbb{Z}}^+ \triangleq \mathbb{Z}^+ \cup \{\infty\}$), there exist functions $\varphi_j(x), \varphi_{j'}(x) \in L_2(\mathrm{K},\Lambda)$ such that $\int_{\mathrm{K}} \varphi_j(x)\varphi_{j'}(x)dx = \delta_{j,j'}$, where $\delta_{j,j'}$ is the Kronecker delta, and $\sum\limits_{j=0}^{\infty}\int\limits_{\mathrm{K}} \varphi_j^2(x)dx < \infty$ [24]. Hence, for almost all $x \in \mathrm{K}$,

$$f(x) = \sum_{j=0}^{\infty} c_j \varphi_j(x), \tag{1}$$

where $\{c_j\}_{j\in\overline{\mathbb{Z}}^+}$ are the Fourier coefficients of a function $f(x)$, i.e., $c_j \triangleq \int\limits_{\mathrm{K}} f(x)\varphi_j(x)dx$.

In a complete probability space $(\Omega, \mathcal{F}, \mathsf{P})$ we define a measurable function $n\colon \mathbb{Z}^+ \times \Omega \times \mathrm{K} \to \mathbb{R}^1$, further denoted by $n_m(x)$, that satisfies the following conditions for any $x \in \mathrm{K}$ and $m \in \mathbb{Z}^+ \setminus 0$.

Conditions (n_1):

$$\mathsf{E}n_m(x) = 0, \qquad \sigma^2 \triangleq \mathsf{E}\int\limits_{\mathrm{K}} n_m^2(x)dx < \infty, \tag{2}$$

and for any $y, x \in \mathrm{K}$, $y \neq x$, $m \neq q$,

$$\mathsf{E}n_m(x)n_q(x) = 0, \qquad \mathsf{E}n_m(y)n_m(x) = 0\,. \tag{3}$$

Here, $\mathsf{E}(\cdot)$ stands for the Lebesgue integral with respect to the probability measure P.

Obviously, the measure $\Lambda \times \mathrm{P}$, denoted by $\widetilde{\mathrm{P}}$, is defined on the σ-algebra $\mathcal{B}(\mathrm{K}) \otimes \mathcal{F}$.

At any point $x \in \mathrm{K}$ we observe a function $y_m(x)$ as the sum of functions $f(x) \in L_2(\mathrm{K},\Lambda)$ and $n_m(x)$. In other words, the observations of $f(x)$ contain the additive errors $n_m(x)$:

$$y_m(x) = f(x) + n_m(x) = \sum_{j=0}^{\infty} c_j\varphi_j(x) + n_m(x) \quad \widetilde{\mathrm{P}}\text{-a.e.}, \tag{4}$$

where $m \in \mathbb{Z}^+ \setminus 0$ is the observation number.

2.2. For the further presentation, we also need the Fourier coefficients of the random functions $y_m(x)$ and $n_m(x)$:

$$y_m^j \triangleq \int\limits_{\mathrm{K}} y_m(x)\varphi_j(x)dx, \tag{5}$$

$$n_m^j \triangleq \int\limits_{\mathrm{K}} n_m(x)\varphi_j(x)dx. \tag{6}$$

Due to (5)–(6), the Fourier coefficients are random variables for each $m \in \mathbb{Z}^+ \setminus 0$ and $j \in \overline{\mathbb{Z}}^+$. It also follows from (1), (4), (5), (6) that

$$y_m^j = c_j + n_m^j \quad \mathsf{P}\text{-a.s.} \tag{7}$$

Throughout the paper, we adopt the conventional abbreviations “a.s.” and “a.e.” for “almost surely” and “almost everywhere,” respectively, involving an appropriate probability measure.

Since $f(x) \in L_2(\mathrm{K},\Lambda)$ and conditions (n_1) hold, equality (7) implies

$$\sum_{j=0}^{\infty}\left[c_j^2 + \mathsf{E}(n_m^j)^2\right] = \sum_{j=0}^{\infty}\mathsf{E}(y_m^j)^2 = \mathsf{E}\sum_{j=0}^{\infty}(y_m^j)^2 < \infty \tag{8}$$

for any $m \in \mathbb{Z}^+ \setminus 0$.

Conditions (n_2):

(i) $\sigma_j^2 \triangleq \mathsf{E}(n_m^j)^2$, $j \in \overline{\mathbb{Z}}^+$, *i.e., the variance of the Fourier coefficients of the errors is independent of the observation number.*

(ii) $\sigma^2 \triangleq \sum\limits_{j=0}^{\infty} \sigma_j^2 < \infty$.

According to [23], given the Fourier coefficients of the set $\{y_k^j\}_{\substack{j\in\overline{\mathbb{Z}}^+\\ k=\overline{1,m}}}$, $m \in \mathbb{Z}^+ \setminus 0$, any observed function from the set $\{y_m(x)\}_{\substack{x\in\mathrm{K}\\ m\in\mathbb{Z}^+\setminus 0}}$ can be represented as

$$y_m(x) \triangleq \sum_{j\in\overline{\mathbb{Z}}^+} y_m^j \varphi_j(x)\ \widetilde{\mathsf{P}}\text{-a.e.} \tag{9}$$

Let $\mathcal{F}_m^{y^j}$ and $\mathcal{F}_m^{y}$ be the σ-algebras induced by the families of random variables $\{y_k^j\}_{k=\overline{1,m}}$ and $\{y_k^j\}_{\substack{j\in\overline{\mathbb{Z}}^+\\ k=\overline{1,m}}}$, respectively, i.e.,

$$\mathcal{F}_m^{y^j} \triangleq \sigma\left\{y_1^j, \ldots, y_m^j\right\}, \tag{10}$$

$$\mathcal{F}_m^{y} \triangleq \sigma\left\{y_1^j, \ldots, y_m^j, j \in \overline{\mathbb{Z}}^+\right\}. \tag{11}$$

Obviously, $y_m(x)$ is an $\mathcal{F}_m^y \otimes \mathcal{B}(\mathrm{K})$-measurable random function.

Definition 1. Any $\mathcal{F}_m^y \otimes \mathcal{B}(\mathrm{K})$-measurable function $\overline{f}_m(x)$ taking values in $\mathbb{R}^1$ is called an <u>*estimator*</u> of an unknown function $f(x) \in L_2(\mathrm{K}, \Lambda)$ from $m \in \mathbb{Z}^+ \setminus 0$ observations.

Definition 2. An estimator $\overline{f}_m(x)$ is said to be <u>*admissible*</u> if

$$\mathsf{E}\int_{\mathrm{K}} |\overline{f}_m(x)|^2 dx < \infty. \tag{12}$$

We denote by $\mathcal{M}_{2,m}(\widetilde{\mathsf{P}})$ the set of admissible estimators $\overline{f}_m(x)$. Obviously, $\mathcal{M}_{2,m}(\widetilde{\mathsf{P}}) \neq \varnothing$ and is a Hilbert space.

Definition 3. An admissible estimator $\widehat{f}_m(x) \in \mathcal{M}_{2,m}(\widetilde{\mathsf{P}})$ is said to be a <u>*projection estimator*</u> if it can be represented as

$$\widehat{f}_m(x) = \sum_{j=0}^{\infty} c_{j,m}\varphi_j(x)\ \widetilde{\mathsf{P}}\text{-a.e.}, \tag{13}$$

where for each $j \in \overline{\mathbb{Z}}^+$ and $m \in \mathbb{Z}^+ \setminus 0$, the Fourier coefficient $c_{j,m}$ of the estimator $\widehat{f}_m(x)$ is an $\mathcal{F}_m^{y^j}$-measurable random variable such that $\mathsf{E}\sum\limits_{j=0}^{\infty} c_{j,m}^2 < \infty$.

We denote by $\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$ the set of projection estimators. Obviously, $\mathbb{M}_{2,m}(\widetilde{\mathsf{P}}) \subseteq \mathcal{M}_{2,m}(\widetilde{\mathsf{P}})$.

This paper considers the problem of constructing projection estimators $\widehat{f}_m(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$ of an unknown function $f(x) \in L_2(\mathrm{K}, \Lambda)$ from its observations (7) such that

$$\mathsf{E}\int_{\mathrm{K}} \left[f(x) - \widehat{f}_m(x)\right]^2 dx \to \inf_{\widehat{f}_m(x)\in\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})}. \tag{14}$$

The optimality criterion (14) is the minimum deviation with respect to the measure $\widetilde{\mathsf{P}}$.

Now, we define an optimal projection estimator.

Definition 4. A projection estimator $\widehat{f}_m^{\,0}(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$ is said to be <u>*optimal*</u> if

$$\inf_{\widehat{f}_m(x)\in\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})} \mathsf{E}\int\limits_{\mathrm{K}} \left[f(x)-\widehat{f}_m(x)\right]^2 dx = \mathsf{E}\int\limits_{\mathrm{K}} \left[f(x)-\widehat{f}_m^{\,0}(x)\right]^2 dx. \tag{15}$$

The representation (5)–(15) of the stochastic recovery problem data is called the <u>spectral</u> representation.

2.3. This paper aims at:

1) proving the existence of a recovery procedure of an unknown function $f(x) \in L_2(\mathrm{K}, \Lambda)$ with finitely many orthonormal functions and finding its standard deviation;

2) investigating the statistical properties of the recovery procedure.

3. EXISTENCE CONDITIONS FOR OPTIMAL ESTIMATORS OF AN UNKNOWN FUNCTION

We denote by $\widetilde{\mathbb{M}}_{2,m}(\mathsf{P})$ the set of infinite-dimensional random vectors $\widehat{c}_m \triangleq (\widehat{c}_{0,m}, \widehat{c}_{1,m}, \dots)$ such that:

i) For any $j \in \overline{\mathbb{Z}}^+$ and each $m \in \mathbb{Z}^+ \setminus 0$, the random variable $\widehat{c}_{j,m}$ is $\mathcal{F}_m^{y^j}$-measurable.

ii) $\mathsf{E}\sum\limits_{j=0}^{\infty} \left|\widehat{c}_{j,m}\right|^2 < \infty$.

3.1. If $\widehat{f}_m(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$, then

$$\widehat{f}_m(x) = \sum_{j=0}^{\infty} \widehat{c}_{j,m}\varphi_j(x) \ \widetilde{\mathsf{P}}\text{-a.e.}, \tag{16}$$

where $\widehat{c}_{j,m} \in \widetilde{\mathbb{M}}_{2,m}(\mathsf{P})$, $j \in \overline{\mathbb{Z}}^+$, $m \in \mathbb{Z}^+ \setminus 0$, is an $\mathcal{F}_m^{y^j}$-measurable random variable specifying the Fourier coefficient of the estimator $\widehat{f}_m(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$, i.e.,

$$\widehat{c}_{j,m} \triangleq \int\limits_{\mathrm{K}} \widehat{f}_m(x)\varphi_j(x)dx \ \mathsf{P}\text{-a.s.} \tag{17}$$

From (16) we have the relation

$$\mathsf{E}\int\limits_{\mathrm{K}} \left|\widehat{f}_m(x)\right|^2 dx = \mathsf{E}\sum_{i=0}^{\infty} \widehat{c}_{j,m}^{\,2}, \tag{18}$$

which generalizes the well-known Parseval identity [24, 25].

Due to (18), $\widetilde{\mathbb{M}}_{2,m}(\mathsf{P})$ is a Hilbert space.

In addition, equality (18) leads to the following result.

Proposition 1. $\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$ *and* $\widetilde{\mathbb{M}}_{2,m}(\mathsf{P})$ *are isomorphic.*

Proposition 2. *For any* $m \in \mathbb{Z}^+ \setminus 0$, *the optimal projection estimator* $\widehat{f}_m^{\,0}(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$ *exists if and only if there are* $\{\widehat{c}_m^{\,0}\} \in \widetilde{\mathbb{M}}_{2,m}(\mathsf{P})$ *such that*

$$\inf_{\{\widehat{c}_{j,m}\}_{j\geqslant 0}\in\widetilde{\mathbb{M}}_{2,m}(\mathsf{P})} \mathsf{E}\sum_{j=0}^{\infty} \left[c_j-\widehat{c}_{j,m}\right]^2 = \mathsf{E}\sum_{j=0}^{\infty} \left[c_j-\widehat{c}_{j,m}^{\,0}\right]^2. \tag{19}$$

The proof of Proposition 2 is given in Appendix A; see item A.1.

Remark 1. According to Proposition 2, the existence of an optimal projection estimator of an unknown function from the class $\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$ is equivalent to the existence of the optimal estimators of its Fourier coefficients.

3.2. This subsection presents existence conditions for a solution of the optimal stochastic recovery problem in the class $L_2(\mathrm{K}, \Lambda)$ of unknown functions.

Conditions (n_3).

For any $j \in \overline{\mathbb{Z}}^+$ and $m \in \mathbb{Z}^+ \setminus 0$, the family $\{n_m^j\}_{\substack{j\in\overline{\mathbb{Z}}^+\\ m\in\mathbb{Z}^+\setminus 0}}$ is a Gaussian system of uncorrelated random variables with $\mathrm{Law}(n_m^j) = \mathcal{N}(0, \sigma_j^2)$.

We denote by $V_m(\infty)$ the standard deviation of the optimal estimator.

Existence conditions can be formulated as follows.

Theorem 1. *Let $f(x) \in L_2(\mathrm{K}, \Lambda)$ and conditions (n_i), $i = \overline{1,3}$, be satisfied. Then for almost all $x \in \mathrm{K}$ and $m \in \mathbb{Z}^+ \setminus 0$, there exists an optimal projection estimator $\widehat{f}_m^{\,0}(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$ that can be represented as*

$$\widehat{f}_m^{\,0}(x) = \sum_{j=0}^{\infty} \widehat{c}_{j,m}^{\,0} \varphi_j(x) \ \widetilde{\mathsf{P}}\text{-}a.e., \tag{20}$$

where

$$\widehat{c}_{j,m}^{\,0} = \frac{1}{m} \sum_{k=1}^{m} y_k^j, \tag{21}$$

and its standard deviation has the form

$$V_m(\infty) = \inf_{\widehat{f}_m(x)\in\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})} \mathsf{E} \int_{\mathrm{K}} \left[f(x) - \widehat{f}_m(x)\right]^2 dx = \frac{1}{m} \sum_{j=0}^{\infty} \sigma_j^2. \tag{22}$$

The proof of Theorem 1 is given in Appendix A; see item A.2.

Remark 2. In contrast to [3–17, 19–21], Theorem 1 provides sufficient conditions for the existence of an optimal projection estimator of an unknown function from the class $L_2(\mathrm{K}, \Lambda)$ from its observations with independent Gaussian errors.

3.3. Theorem 1 leads to a simple form of the estimator $\widehat{f}_m^{\,0}(x)$.

Corollary 1. *Under the hypotheses of Theorem 1, for each $m \in \mathbb{Z}^+ \setminus 0$, the optimal estimator $\widehat{f}_m^{\,0}(x)$ can be represented as*

$$\widehat{f}_m^{\,0}(x) = \frac{1}{m} \sum_{k=1}^{m} y_k(x) \ \widetilde{\mathsf{P}}\text{-}a.e. \tag{23}$$

The proof of Corollary 1 is given in Appendix A; see item A.3.

3.4. The Properties of Optimal Estimators

Theorem 2. *Under the hypotheses of Theorem 1, the estimator* (20) *is unbiased.*

The proof of Theorem 2 is given in Appendix A; see item A.4.

Theorem 3. *Under the hypotheses of Theorem 1, let the series $\sum\limits_{j=0}^{\infty} \sigma_j^2 \varphi_j^2(x)$ be convergent for almost all $x \in \mathrm{K}$ with respect to the Lebesgue measure. Then the estimator $\widehat{f}_m^{\,0}(x)$ is consistent.*

The proof of Theorem 3 is given in Appendix A; see item A.5.

4. THE STANDARD DEVIATION OF THE ESTIMATOR OF AN UNKNOWN FUNCTION $f(x) \in L_2(\mathrm{K}, \Lambda)$: DEPENDENCE OF THE NUMBER OF ORTHOGONAL FUNCTIONS AND THE NUMBER OF OBSERVATIONS AND OTHER PROPERTIES

4.1. Let $f_N(x) \triangleq \sum\limits_{j=0}^{N} c_j\varphi_j(x)$, where $c_j \triangleq \int\limits_{\mathrm{K}} f(x)\varphi_j(x)dx$ and $N \in \overline{\mathbb{Z}}^+$ is the number of orthogonal functions used. Obviously, $f_N(x) \in L_2(\mathrm{K}, \Lambda)$.

We consider $\widehat{f}^{\,0}_{m,N}(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$, i.e., the optimal projection estimator of an unknown function $f_N(x) \in L_2(\mathrm{K}, \Lambda)$. By Theorem 1, it has the form

$$\widehat{f}^{\,0}_{m,N}(x) \triangleq \sum_{j=0}^{N} \widehat{c}^{\,0}_{j,m}(x)\varphi_j(x). \tag{24}$$

Due to Fubini's theorem, from (24) it follows that

$$\mathsf{E}\int\limits_{\mathrm{K}} |\widehat{f}^{\,0}_{m,N}(x)|^2 dx = \mathsf{E}\sum_{j=0}^{N} |\widehat{c}^{\,0}_{j,m}|^2, \tag{25}$$

where $\widehat{c}^{\,0}_{j,m}$ is given by (21).

Let $V_m(N)$ be the standard deviation of the estimator $\widehat{f}^{\,0}_{m,N}(x)$ of an unknown function $f(x) \in L_2(\mathrm{K}, \Lambda)$ constructed using a "segment" of N orthonormal functions with m observations. Obviously,

$$V_m(N) \triangleq \mathsf{E}\int\limits_{\mathrm{K}} \left[f(x) - \widehat{f}^{\,0}_{m,N}(x)\right]^2 dx. \tag{26}$$

Now, we formulate the main result of this subsection.

Theorem 4. *Under the hypotheses of Theorem 1, the following assertions are true for any* $m \in \mathbb{Z}^+ \setminus 0$ *and* $N \in \mathbb{Z}^+$*:*

1) $V_m(N)$ *can be represented as*

$$V_m(N) = \sum_{j=0}^{N} \left[\frac{1}{m}\sigma_j^2 - c_j^2\right] + \|f\|^2_{L_2(\mathrm{K},\Lambda)}. \tag{27}$$

2) *Assume that* $\|f\|_{L_2(\mathrm{K},\Lambda)} \geqslant \sigma_0^2$ *and there exists a constant* $C_{10} > 0$ *such that* $\sigma_0^2 \geqslant C_{10}$. *Then*

$$0 < C_{10} \leqslant V_m(N) \leqslant \max\left(\|f\|^2_{L_2(\mathrm{K},\Lambda)}, \frac{\sigma^2}{m}\right). \tag{28}$$

The proof of Theorem 4 is given in Appendix B; see item B.1.

4.2. Theorem 4 implies a simple result.

Corollary 2. *Under the hypotheses of Theorem 4, the following assertions are true.*

1) *For each* $m \in \mathbb{Z}^+ \setminus 0$, *the sequence* $\{V_m(N)\}_{N\in\mathbb{Z}^+}$ *satisfies the recurrent relation*

$$\begin{cases} V_m(N+1) = V_m(N) - c^2_{N+1} + \dfrac{1}{m}\sigma^2_{N+1} \\ V_m(N)\Big|_{N=0} = \sum\limits_{j=0}^{\infty} c_j^2, \end{cases} \tag{29}$$

and its solution has the form (22).

2) *For each $N \in \mathbb{Z}^+$, the partial sequence $\{V_m(N)\}_{m\in\mathbb{Z}^+\setminus 0}$ satisfies the recurrent relation*

$$\begin{cases} V_{m+1}(N) = V_m(N) - \dfrac{1}{(m+1)m}\displaystyle\sum_{j=0}^{N}\sigma_j^2 \\ V_m(N)\Big|_{m=1} = -\displaystyle\sum_{j=N+1}^{\infty} c_j^2 + \sum_{j=0}^{N}\sigma_j^2, \end{cases} \tag{30}$$

and its solution has the form (22).

The proof of Corollary 2 is given in Appendix B; see item B.2.

Remark 3. Due to assertion 2 of Corollary 2 and (27), we have

$$\lim_{m\to\infty} V_{m+1}(N) = \lim_{m\to\infty} V_m(N) = \sum_{j=N+1}^{\infty} c_j^2$$

for any $N \in \mathbb{Z}^+$. Therefore,

$$\lim_{N\to\infty}\lim_{m\to\infty} V_m(N) = 0\ .$$

4.3. In this section, we establish that for each $m \in \mathbb{Z}^+ \setminus 0$, the sequence $\{V_m(N)\}_{N\in\mathbb{Z}^+}$ has a unique minimum.

Theorem 5. *Under the hypotheses of Theorem 4, for each $m \in \mathbb{Z}^+ \setminus 0$, there exists a unique number $N^0(m) \in \mathbb{Z}^+$ such that*

$$\inf_{N\in\mathbb{Z}^+} V_m(N) = V_m(N^0(m)). \tag{31}$$

The proof of Theorem 5 is given in Appendix B; see item B.3.

Remark 4. Theorem 5 implies the following results:

(a) For any $N \in \overline{\mathbb{Z}}^+$,

$$V_m(N) \geqslant V_m(N^0(m)). \tag{32}$$

(b) For all $s \in \mathbb{Z}^+$ such that $N^0(m) - s \geqslant 0$ and $N^0(m) + s \in \mathbb{Z}^+$,

$$V_m(N^0(m) - s) - 2V_m(N^0(m)) + V_m(N^0(m) + s) \geqslant 0. \tag{33}$$

4.4. This subsection describes a constructive method for finding $N^0(m)$.

Theorem 6. *Under the hypotheses of Theorem 4, the number $N^0(m) \in \mathbb{Z}^+$ can be represented as*

$$N^0(m) = \inf\left\{N \in \mathbb{Z}^+ \colon \sum_{j=0}^{N}\frac{\sigma_j^2}{m} \geqslant \sum_{j=N+1}^{\infty} c_j^2\right\}. \tag{34}$$

The proof of Theorem 6 is given in Appendix B; see item B.4.

4.5. Based on Theorem 6, we establish equivalence relations between the standard deviation of the CPE and $\sum\limits_{j=0}^{N^0(m)}\frac{\sigma_j^2}{m}$.

Theorem 7. *Under the hypotheses of Theorem 6,*

$$\sum_{j=0}^{N^0(m)} \frac{\sigma_j^2}{m} \leqslant V_m(N^0(m)) \leqslant 2 \sum_{j=0}^{N^0(m)} \frac{\sigma_j^2}{m} \tag{35}$$

for any $m \in \mathbb{Z}^+ \setminus 0$. *In other words,*

$$V_m(N^0(m)) \asymp \sum_{j=0}^{N^0(m)} \frac{\sigma_j^2}{m}. \tag{36}$$

The proof of Theorem 7 is given in Appendix B; see item B.5.

4.6. Theorem 7 leads to important results.

Corollary 3. *Under the hypotheses of Theorem 7, assume that:*

(i) $\sup_{j \in \overline{\mathbb{Z}}^+} \sigma_j^2 \leqslant C_{11}$.

(ii) *There exists a number* $j_0 \in \{0, \dots, N^0(m)\}$ *such that* $\sigma_{j_0}^2 \geqslant C_{12} > 0$.

Then the following conditions hold:

1)

$$C_{12} \frac{N^0(m)}{m} \leqslant V_m(N^0(m)) \leqslant C_{11} \frac{N^0(m)}{m}, \tag{37}$$

i.e., $V_m(N^0(m)) \asymp \frac{N^0(m)}{m}$.

2)

$$N^0(m) \asymp m. \tag{38}$$

The proof of Corollary 3 is given in Appendix B; see item B.6.

5. CHENTSOV PROJECTION ESTIMATORS AND THEIR PROPERTIES

According to Section 4, for any $m \in \mathbb{Z}^+ \backslash 0$, the standard deviation $V_m(N)$ of the projection estimator $\widehat{f}_{m,N}^{\,0}(x)$ achieves a unique minimum (Theorem 5) at the point $N^0(m) \in \mathbb{Z}^+$. Theorem 6 therein provides a constructive method to find $N^0(m)$. In addition, Section 4 has established the equivalence relations $V_m(N^0(m)) \asymp \sum_{j=0}^{N^0(m)} \frac{\sigma_j^2}{m} \asymp \frac{N^0(m)}{m} \asymp \text{const}$; see Theorem 7 and Corollary 3.

5.1. In this subsection, we define the Chentsov projection estimator of an unknown function $f(x) \in L_2(\mathrm{K}, \Lambda)$.

Let us denote

$$\widetilde{f}_m^{\,0}(x) \triangleq \widehat{f}_{m,N}^{\,0}(x)\Big|_{N=N^0(m)}. \tag{39}$$

Obviously, $\widetilde{f}_m^{\,0}(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$ and, due to (20), it can be represented as

$$\widetilde{f}_m^{\,0}(x) = \sum_{j=0}^{N^0(m)} \widehat{c}_{j,m}^{\,0} \varphi_j(x) \ \widetilde{\mathsf{P}}\text{-a.e.}, \tag{40}$$

where $\widehat{c}_{j,m}^{\,0} = \frac{1}{m} \sum_{k=1}^{m} y_k^j$ stands for the jth component of the infinite-dimensional vector $\widehat{c}_m^{\,0} \in \widetilde{\mathbb{M}}_{2,m}(\mathsf{P})$ (the Fourier coefficient of the optimal estimator $\widehat{f}_m^{\,0}(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$).

Definition 5. An $\mathcal{F}_m^y \otimes \mathcal{B}(\mathrm{K})$-measurable function $\widetilde{f}_m^0(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$ is called *the Chentsov projection estimator (CPE)* of an unknown function $f(x) \in L_2(\mathrm{K}, \Lambda)$ if it admits the representation (40).

By this definition and (25), the standard deviation $V_m(N^0(m))$ of the CPE from the corresponding function $f(x) \in L_2(\mathrm{K}, \Lambda)$ is given by

$$V_m(N^0(m)) = \mathsf{E}\int\limits_{\mathrm{K}} \left[f(x) - \widetilde{f}_m^0(x)\right]^2 dx = \sum_{j=N^0(m)+1}^{\infty} c_j^2 + \sum_{j=0}^{N^0(m)} \frac{\sigma_j^2}{m}. \tag{41}$$

5.2. Theorem 7 allows estimating the rate of convergence of the CPE to the corresponding function $f(x) \in L_2(\mathrm{K}, \Lambda)$ as $m \to \infty$. Since $N^0(m) \in \mathbb{Z}^+$, from Corollary 3 we have

$$V_m(N^0(m)) \asymp m^{-1}. \tag{42}$$

Thus, the following result is true.

Theorem 8. *Under the hypotheses of Theorem 7 and Corollary 3,*

$$\|f - \widetilde{f}_m^0\|_{\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})} = O(m^{-\frac{1}{2}}). \tag{43}$$

5.3. This subsection focuses on statistical properties of the CPE (40).

Theorem 9. *The CPE* (40) *possesses the following properties:*

1) *For any* $m \in \mathbb{Z}^+ \setminus 0$, *the estimator* (40) *is biased.*
2) *Under the hypotheses of Theorem 5, the standard deviation of the CPE satisfies the inequality*

$$V_m(N^0(m)) \leqslant V_m(\infty). \tag{44}$$

3) *The estimator* (40) *is asymptotically unbiased, i.e.,*

$$\lim_{m\to\infty} \mathsf{E}\widetilde{f}_m^0(x) = f(x) \tag{45}$$

for almost all $x \in \mathrm{K}$.

4) *For any* $x \in \mathrm{K}$, *let* $\sum\limits_{j=0}^{\infty} \sigma_j^2\varphi_j^2(x) < \infty$ *and* $\left|\sum\limits_{j=0}^{\infty} c_j\varphi_j(x)\right| < \infty$. *Then*

$$\widetilde{f}_m^0(x) \xrightarrow[m\to\infty]{\mathsf{P}} f(x).$$

The proof of Theorem 9 is given in Appendix B; see item B.7.

5.4. Here, we establish conditions under which the standard deviation of the CPE is independent of the function $f(x) \in L_2(\mathrm{K}, \Lambda)$ under estimation.

Theorem 10. *Under the hypotheses of Theorems 5 and 6, we have the inequality*

$$\sup_{f(x)\in L_2(\mathrm{K},\Lambda)} \inf_{N\in\mathbb{Z}^+} \sup_{f_{m,N}(x)\in\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})} \mathsf{E}\int\limits_{\mathrm{K}} [f(x) - f_{m,N}(x)]^2 dx \leqslant 2\sum_{j=0}^{\infty} \frac{\sigma_j^2}{m} \xrightarrow[m\to\infty]{} 0. \tag{46}$$

The proof of Theorem 10 is given in Appendix B; see item B.8.

5.5. This subsection gives one example with an explicit solution of the stochastic recovery problem. Due to Theorems 8 and 9 and the CPE formula, optimal stochastic recovery is reduced to finding $N^0(m)$; see below.

Example. Suppose that the elements of the sequences $\{c_j^2\}_{j\geqslant 0}$ and $\{\sigma_j^2\}_{j\geqslant 0}$ can be represented as

1) $\sigma_j^2 = \sigma_0^2 q_1^j$,

2) $c_j^2 = c_0^2 q_2^j$,

where $\sigma_0^2 > 0$, $c_0^2 > 0$, $q_i \in (0,1)$ for $i = 1, 2$, and $j \in \overline{\mathbb{Z}}^+$. According to Theorems 5 and 6, there exists a number $N^0(m) \in \mathbb{Z}^+$ such that $\inf\limits_{N\in\overline{\mathbb{Z}}^+} V_m(N) = V_m(N^0)$; furthermore, for any $N \geqslant N^0(m)$,

$$\sum_{j=0}^{N} \frac{\sigma_j^2}{m} \geqslant \sum_{j=N+1}^{\infty} c_j^2.$$

Under assumptions 1) and 2), this inequality can be written as

$$\frac{\sigma_0^2}{m} \sum_{j=0}^{N} q_1^j \geqslant c_0^2 \sum_{j=N+1}^{\infty} q_2^j.$$

By Theorem 5, for $N = N^0(m)$ we have the equality

$$\frac{\sigma_0^2}{m}\frac{1}{1-q_1} = \frac{c_0^2 q_2^{N^0(m)+1}}{1-q_2} + \frac{\sigma_0^2}{m}\frac{q_1^{N^0(m)+1}}{1-q_1}. \tag{47}$$

Consider two special cases in which the number $N^0(m) \in \mathbb{Z}^+$ can be found explicitly.

We denote by $\lfloor a \rfloor$ the integer part of $a \in \mathbb{Z}^+ \setminus 0$.

Case 1. Let $q_1 = q_2$. If $\frac{\sigma_0^2}{c_0^2 m + \sigma_0^2} > q_1$, from (47) it follows that

$$N^0(m) = \left\lfloor \frac{\ln \frac{\sigma_0^2}{c_0^2 m + \sigma_0^2}}{\ln q_1} \right\rfloor.$$

Case 2. Let $q_2 = (q_1)^2$. In this case, from (47) it follows that

$$\frac{\sigma_0^2}{m} = \frac{c_0^2 q_1^{2(N^0(m)+1)}}{1+q_1} + \frac{\sigma_0^2}{m} q_1^{N^0(m)+1}. \tag{48}$$

The expression (48) is a quadratic equation for $q_1^{N^0(m)+1}$. Under the condition

$$\frac{\sigma_0^2(1+q_1)}{2mc_0^2 q_1}\left[\sqrt{1 + \frac{4mc_0^2}{\sigma_0^2(1+q_1)}} - 1\right] > 1,$$

we obtain

$$N^0(m) = \left\lfloor \frac{\ln \frac{\sigma_0^2(1+q_1)}{2mc_0^2}\left[\left(1 + \frac{4mc_0^2}{\sigma_0^2(1+q_1)}\right)^{\frac{1}{2}} - 1\right]}{\ln q_1} \right\rfloor.$$

APPENDIX A

A.1. *The proof of Proposition 2.* Let $f(x) \in L_2(\mathrm{K}, \Lambda)$ and $\widehat{f}_m(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$ be some projection estimator. Since the system $\{\varphi_j(x)\}_{j\geqslant 0}$ is complete and orthonormal, from (1), (13), and Fubini's theorem we have the equalities

$$\mathsf{E}\int\limits_{\mathrm{K}} \left[f(x) - \widehat{f}_m(x)\right]^2 dx = \mathsf{E}\sum_{j=0}^{\infty} \left[c_j - \widehat{c}_{j,m}\right]^2 = \sum_{j=0}^{\infty} \mathsf{E}\left[c_j - \widehat{c}_{j,m}\right]^2$$

for any $m \in \mathbb{Z}^+ \setminus 0$.

Due to Proposition 1,

$$\inf_{\widehat{f}_m(x)\in\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})} \mathsf{E}\int\limits_{\mathrm{K}} \left[f(x) - \widehat{f}_m(x)\right]^2 dx = \inf_{\widehat{c}_{j,m}\in\widetilde{\mathbb{M}}_{2,m}(\mathsf{P})} \sum_{j=0}^{\infty} \mathsf{E}\left[c_j - \widehat{c}_{j,m}\right]^2. \tag{A.1}$$

Recall that $\widetilde{\mathbb{M}}_{2,m}(\mathsf{P})$ is the set of $\mathcal{F}_m^{y_j}$-measurable square-integrable random variables. Therefore, from (A.1) it follows that

$$\inf_{\widehat{f}_m(x)\in\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})} \mathsf{E}\int\limits_{\mathrm{K}} \left[f(x) - \widehat{f}_m(x)\right]^2 dx \geqslant \sum_{j=0}^{\infty} \inf_{\widehat{c}_{j,m}\in\widetilde{\mathbb{M}}_{2,m}(\mathsf{P})} \mathsf{E}\left[c_j - \widehat{c}_{j,m}\right]^2.$$

Hence, the estimator $\widehat{c}_{j,m}^{\,0}$ is optimal if and only if

$$\inf_{\widehat{c}_{j,m}\in\widetilde{\mathbb{M}}_{2,m}(\mathsf{P})} \mathsf{E}\left[c_j - \widehat{c}_{j,m}\right]^2 = \mathsf{E}\left[c_j - \widehat{c}_{j,m}^{\,0}\right]^2.$$

Thus, given the existence of $\widehat{c}_{j,m}^{\,0}$,

$$\inf_{\widehat{f}_m(x)\in\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})} \mathsf{E}\int\limits_{\mathrm{K}} \left[f(x) - \widehat{f}_m(x)\right]^2 dx \geqslant \sum_{j=0}^{\infty} \mathsf{E}\left[c_j - \widehat{c}_{j,m}^{\,0}\right]^2.$$

In view of (A.1) and this inequality, we have

$$\inf_{\widehat{f}_m(x)\in\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})} \mathsf{E}\int\limits_{\mathrm{K}} \left[f(x) - \widehat{f}_m(x)\right]^2 dx \geqslant \sum_{j=0}^{\infty} \mathsf{E}\left[c_j - \widehat{c}_{j,m}^{\,0}\right]^2$$

$$= \inf_{\widehat{c}_{j,m}\in\widetilde{\mathbb{M}}_{2,m}(\mathsf{P})} \mathsf{E}\sum_{j=0}^{\infty} \left[c_j - \widehat{c}_{j,m}\right]^2 = \mathsf{E}\int\limits_{\mathrm{K}} \left[f(x) - \widehat{f}_m^{\,0}(x)\right]^2 dx,$$

where $\widehat{f}_m^{\,0}(x) \triangleq \sum\limits_{j=0}^{\infty} \widehat{c}_{j,m}^{\,0}\varphi_j(x)$. The proof of this proposition is complete.

A.2. *The proof of Theorem 1.* By Proposition 2, there exists an optimal projection estimator $\widehat{f}_m^{\,0}(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$ if and only if (19) holds. Therefore,

$$\inf_{\widehat{f}_m(x)\in\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})} \mathsf{E}\int\limits_{\mathrm{K}} \left[f(x) - \widehat{f}_m(x)\right]^2 dx = \mathsf{E}\int\limits_{\mathrm{K}} \left[f(x) - \widehat{f}_m^{\,0}(x)\right]^2 dx.$$

The main content of Theorem 1 is equalities (20)–(22).

To prove them, we consider $\mathsf{E}\int\limits_{\mathrm{K}}\left[f(x)-\widehat{f}_m^{\,0}(x)\right]^2dx$. From Proposition 2 (see formulas (A.1) and (25)) it follows that

$$\mathsf{E}\int\limits_{\mathrm{K}}\left[f(x)-\widehat{f}_m^{\,0}(x)\right]^2dx=\sum_{j=0}^{\infty}\mathsf{E}\left|c_j-\widehat{c}_{j,m}^{\,0}\right|^2. \tag{A.2}$$

Hence, for each $j\in\overline{\mathbb{Z}}^+$, it is required to construct a mean-square optimal estimate of the Fourier coefficient c_j from the observations $(y_1^j,\ldots,y_m^j)$. Note that due to (7), the random variable y_m^j has the Gaussian distribution: $\mathrm{Law}(y_m^j)=\mathcal{N}(c_j,\sigma_j^2)$. As is well known [1, 2], in this case, the optimal estimate $\widehat{c}_{j,m}^{\,0}$ of the Fourier coefficient c_j from the error-containing observations $(y_1^j,\ldots,y_m^j)$ coincides with the maximum likelihood estimate. Thus, $\widehat{c}_{j,m}^{\,0}$ has the form (19). We multiply both sides of (19) by $\varphi_j(x)$ and perform summation over all j to obtain (18).

Now, we find the value $\mathsf{E}\int\limits_{\mathrm{K}}\left[f(x)-\widehat{f}_m^{\,0}(x)\right]^2dx$. Due to (A.2), (7), (19) and Proposition 1, we have

$$\mathsf{E}\int\limits_{\mathrm{K}}\left[f(x)-\widehat{f}_m^{\,0}(x)\right]^2dx=\sum_{j=0}^{\infty}\mathsf{E}\left|\widehat{c}_{j,m}^{\,0}-c_j\right|^2$$

$$=\sum_{j=0}^{\infty}\mathsf{E}\left|\frac{1}{m}\sum_{k=1}^{m}y_k^j-c_j\right|^2=\sum_{j=0}^{\infty}\mathsf{E}\left(\frac{1}{m}\sum_{k=1}^{m}n_k^j\right)^2=\sum_{j=0}^{\infty}\frac{\sigma_j^2}{m}.$$

The proof of this theorem is complete.

A.3. *The proof of Corollary 1.* From (20)–(22) and Fubini's theorem we obtain (23) since

$$\widehat{f}_m^{\,0}(x)=\sum_{j=0}^{\infty}\widehat{c}_{j,m}^{\,0}\varphi_j(x)=\sum_{j=0}^{\infty}\frac{1}{m}\sum_{k=1}^{m}y_k^j\varphi_j(x)=\frac{1}{m}\sum_{k=1}^{m}\sum_{j=0}^{\infty}y_k^j\varphi_j(x)=\frac{1}{m}\sum_{k=1}^{m}y_k(x).$$

The proof of this corollary is complete.

A.4. *The proof of Theorem 2.* From (7), (20), and (21), by Fubini's theorem, we have

$$\mathsf{E}\widehat{f}_m^{\,0}(x)=\mathsf{E}\sum_{j=0}^{\infty}\widehat{c}_{j,m}^{\,0}\varphi_j(x)=\sum_{j=0}^{\infty}\varphi_j(x)\mathsf{E}\widehat{c}_{j,m}^{\,0}=\sum_{j=0}^{\infty}\varphi_j(x)\frac{1}{m}\mathsf{E}\sum_{k=1}^{m}y_k^j$$

$$=\sum_{j=0}^{\infty}\varphi_j(x)\frac{1}{m}\sum_{k=1}^{m}\left(c_j+\mathsf{E}n_k^j\right)=\sum_{j=0}^{\infty}\varphi_j(x)c_j=f(x)$$

for any $x\in\mathrm{K}$ and $m\in\mathbb{Z}^+\setminus 0$. The proof of this theorem is complete.

A.5. *The proof of Theorem 3.* It is required to establish that

$$\widehat{f}_m^{\,0}(x)\xrightarrow[m\to\infty]{\mathsf{P}}f(x)$$

for almost all $x\in\mathrm{K}$. It suffices to demonstrate that the variance of the estimator $\widehat{f}_m^{\,0}(x)$ vanishes as $m\to\infty$.

For each $x \in \mathrm{K}$, we calculate the variance $\mathsf{D}\widehat{f}^{\,0}_m(x)$ of the estimator $\widehat{f}^{\,0}_m(x)$. From (20), (7), and (21), by Fubini's theorem, we have

$$\mathsf{D}\widehat{f}^{\,0}_m(x) = \mathsf{E}\big[\widehat{f}^{\,0}_m(x) - f(x)\big]^2 = \mathsf{E}\left[\sum_{j=0}^{\infty}(\widehat{c}_{j,m} - c_j)\varphi_j(x)\right]^2$$

$$= \mathsf{E}\left[\sum_{j=0}^{\infty}\left(\frac{1}{m}\sum_{k=1}^{m} y_k^j - c_j\right)\varphi_j(x)\right]^2 = \mathsf{E}\left[\sum_{j=0}^{\infty}\frac{1}{m}\sum_{k=1}^{m} n_k^j\varphi_j(x)\right]^2 = \frac{1}{m}\sum_{j=0}^{\infty}\varphi_j^2(x)\sigma_j^2.$$

Since the series $\sum\limits_{j=0}^{\infty}\varphi_j^2(x)\sigma_j^2$ converges for almost all $x \in \mathrm{K}$, the latter equality yields the desired result. The proof of this theorem is complete.

APPENDIX B

B.1. *The proof of Theorem 4.* We begin with the first assertion. According to the definition of $V_m(N)$,

$$V_m(N) = \mathsf{E}\int\limits_{\mathrm{K}}\big[f(x) - f_N(x) + f_N(x) - \widehat{f}^{\,0}_{m,N}(x)\big]^2 dx$$

$$= \mathsf{E}\int\limits_{\mathrm{K}}\left[\sum_{j=0}^{\infty}c_j\varphi_j(x) - \sum_{j=0}^{N}c_j\varphi_j(x) + \sum_{j=0}^{N}c_j\varphi_j(x) - \sum_{j=0}^{N}\widehat{c}^{\,0}_{j,m}\varphi_j(x)\right]^2 dx$$

$$= \mathsf{E}\int\limits_{\mathrm{K}}\left[\sum_{j=N+1}^{\infty}c_j\varphi_j(x) + \sum_{j=0}^{N}(c_j - \widehat{c}^{\,0}_{j,m})\varphi_j(x)\right]^2 dx.$$

Hence, for any $m \in \mathbb{Z}^+ \setminus 0$ and $N \in \mathbb{Z}^+$, the standard deviation $V_m(N)$ is given by

$$V_m(N) = \sum_{j=N+1}^{\infty}c_j^2 + \mathsf{E}\sum_{j=0}^{N}(c_j - \widehat{c}^{\,0}_{j,m})^2. \tag{B.1}$$

Since $\widehat{c}^{\,0}_{j,m} = \frac{1}{m}\sum\limits_{k=1}^{m}y_k^j$, by (7), we have

$$\widehat{c}^{\,0}_{j,m} = \frac{1}{m}\sum_{k=1}^{m}(c_j + n_k^j) = c_j + \frac{1}{m}\sum_{k=1}^{m}n_k^j. \tag{B.2}$$

In view of (B.2) and conditions (n_i), $i = \overline{1,3}$, by Fubini's theorem, formula (B.1) reduces to (23). Indeed,

$$V_m(N) = \sum_{j=N+1}^{\infty}c_j^2 + \mathsf{E}\sum_{j=0}^{N}\left(\frac{1}{m}\sum_{k=1}^{m}n_k^j\right)^2$$

$$= \sum_{j=N+1}^{\infty}c_j^2 + \sum_{j=0}^{N}\frac{1}{m^2}\mathsf{E}\sum_{k=1}^{m}n_k^j\sum_{k'=1}^{m}n_{k'}^j \tag{B.3}$$

$$= \sum_{j=N+1}^{\infty}c_j^2 + \sum_{j=0}^{N}\frac{\sigma_j^2}{m} = \sum_{j=N+1}^{\infty}c_j^2 + \frac{1}{m}\sum_{j=0}^{N}\sigma_j^2.$$

Thus, the first assertion is true.

Now, we prove the second assertion of Theorem 4. According to the first assertion, for any $N \in \mathbb{Z}^+$, the standard deviation $V_m(N)$ of the estimator $\widehat{f}^{\,0}_{m,N}(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$ has the form (B.3).

For each m, it is required to derive a lower bound for $V_m(N)$. Obviously, $V_m(N)$ consists of two terms, namely, a monotonically decreasing sequence (the first term) and a monotonically increasing sequence (the second term). Therefore,

$$\inf_{N\in\overline{\mathbb{Z}}^+} V_m(N) = \inf_{N\in\overline{\mathbb{Z}}^+}\left(\sum_{j=N+1}^{\infty} c_j^2 + \sum_{j=0}^{N}\frac{\sigma_i^2}{m}\right)$$
$$= \max\left(\lim_{N\to\infty}\sum_{j=N+1}^{\infty} c_j^2, \lim_{N\to 0}\sum_{j=0}^{N}\frac{\sigma_i^2}{m}\right) = \max\left(0, \frac{\sigma_0^2}{m}\right) \geqslant C_{10} > 0.$$

The proof of this theorem is complete.

B.2. *The proof of Corollary 2.* The desired result obviously follows from Theorem 4 (see formula (27)).

B.3. *The proof of Theorem 5.* Due to Theorem 4, $V_m(N)$ can be represented as (23). Hence, it consists of two terms:

—The first term is the series $\sum\limits_{j=N+1}^{\infty} c_j^2$, which converges by the convergence of the series $\sum\limits_{j=0}^{\infty} c_j^2 = \|f\|^2_{L_2(\mathrm{K},\Lambda)} < \infty$. Obviously, the sequence $\left\{\sum\limits_{j=N}^{\infty} c_j^2\right\}_{N\geqslant 0}$ is nonincreasing with increasing N, i.e., $\sum\limits_{j=N+1}^{\infty} c_j^2 \leqslant \sum\limits_{j=N}^{\infty} c_j^2$; as a result,

$$\lim_{N\to\infty}\sum_{j=N}^{\infty} c_j^2 = 0\,.$$

—The second term is the convergent nondecreasing sequence $\left\{\sum\limits_{j=0}^{N}\sigma_j^2\right\}_{N\geqslant 0}$ $\left(\sum\limits_{j=1}^{\infty}\sigma_j^2 = \sigma^2 < \infty\right)$.

Therefore, for each $m \in \mathbb{Z}^+\setminus 0$, we have the sets

$$A_m^1 \triangleq \left\{j\in\overline{\mathbb{Z}}^+ \colon \frac{\sigma_j^2}{m} - c_j^2 \geqslant 0\right\} \neq \varnothing,$$
$$A_m^2 \triangleq \left\{j\in\overline{\mathbb{Z}}^+ \colon \frac{\sigma_j^2}{m} - c_j^2 \leqslant 0\right\} \neq \varnothing.$$

If $N \in A_m^1$ $(N \in A_m^2)$, Corollary 2 implies the inequality

$$V_m(N+1) \geqslant V_m(N)$$
$$(V_m(N-1) \geqslant V_m(N),\ \text{respectively}).$$

Obviously, $A_m^1 \cap A_m^2 \neq \varnothing$ and there exists a number $N^0(m) \in \mathbb{Z}^+$ such that $A_m^1 \cap A_m^2 = \{N^0(m)\}$. This result immediately leads to (31). The proof of this theorem is complete.

B.4. *The proof of Theorem 6.* According to the proof of Theorem 3, for any $m \in \mathbb{Z}^+\setminus 0$ and $N \in \overline{\mathbb{Z}}^+$, the standard deviation of the estimator $\widehat{f}^{\,0}_{m,N}(x) \in \mathbb{M}_{2,m}(\widetilde{\mathsf{P}})$ is given by (B.3). Due to Theorem 5, there exists a function $N^0(m) = N^0 \colon (\mathbb{Z}^+\setminus 0) \to \mathbb{Z}^+$ such that

$$V_m(N) \geqslant V_m(N^0(m))$$

for each $m \in \mathbb{Z}^+\setminus 0$ and any $N \in \mathbb{Z}^+$.

Let us denote

$$\mathbf{1}_{\{N\geqslant N^0(m)\}} \triangleq \begin{cases} 1, & N \geqslant N^0(m) \\ 0, & N < N^0(m). \end{cases} \tag{B.4}$$

According to the proof of Theorem 5, $N \geqslant N^0(m)$ if and only if $\sum\limits_{j=0}^{N} \frac{\sigma_j^2}{m} \geqslant \sum\limits_{j=N+1}^{\infty} c_j^2$. Therefore,

$$\mathbf{1}_{\{N\geqslant N^0(m)\}} = \mathbf{1}_{\left\{N\in\mathbb{Z}^+:\ \sum\limits_{j=0}^{N} \frac{\sigma_j^2}{m} \geqslant \sum\limits_{j=N+1}^{\infty} c_j^2\right\}}.$$

Let us denote

$$\ell_m(N) \triangleq \left(\sum_{j=0}^{N} \frac{\sigma_j^2}{m} - \sum_{j=N+1}^{\infty} c_j^2\right) \mathbf{1}_{\left\{N\in\mathbb{Z}^+:\ \sum\limits_{j=0}^{N} \frac{\sigma_j^2}{m} \geqslant \sum\limits_{j=N+1}^{\infty} c_j^2\right\}}. \tag{B.5}$$

Obviously, for each $m \in \mathbb{Z}^+ \setminus 0$ and any $N \in \mathbb{Z}^+$,

$$\ell_m(N) \geqslant 0. \tag{B.6}$$

From (B.5) and (B.6) it follows that $\ell_m(N)$ can be represented as

$$\ell_m(N) = \max\left(\sum_{j=0}^{N} \frac{\sigma_j^2}{m} - \sum_{j=N+1}^{\infty} c_j^2, 0\right). \tag{B.7}$$

The graphs of the functions $V_m(N)$ and $\ell_m(N)$ for each $m \in \mathbb{Z}^+ \setminus 0$ and any $N \in \mathbb{Z}^+$ demonstrate the properties:

(i)

$$V_m(N) \geqslant \ell_m(N), \tag{B.8}$$

(ii)

$$N^0(m) = \operatorname*{argmin}_{N\in\overline{\mathbb{Z}}^+} V_m(N) = \operatorname*{argmin}_{N\in\overline{\mathbb{Z}}^+} \ell_m(N). \tag{B.9}$$

From (B.7) and (B.9) we finally arrive at the assertion of Theorem 6. The proof of this theorem is complete.

B.5. *The proof of Theorem 7.* First, Theorem 4, Corollary 2, and (41) imply the representation

$$V_m(N^0(m)) = V_m(N)\Big|_{N=N^0(m)} = \sum_{j=N^0(m)+1}^{\infty} c_j^2 + \sum_{j=0}^{N^0(m)} \frac{\sigma_j^2}{m}. \tag{B.10}$$

From (B.10) we obtain the inequality

$$V_m(N^0(m)) \geqslant \sum_{j=0}^{N^0(m)} \frac{\sigma_j^2}{m}, \tag{B.11}$$

expressing a lower bound for $V_m(N^0(m))$.

Due to Theorem 6,

$$\sum_{j=0}^{N^0(m)} \frac{\sigma_j^2}{m} \geqslant \sum_{j=N^0(m)+1}^{\infty} c_j^2 \tag{B.12}$$

for any $m \in \mathbb{Z}^+ \setminus 0$.

Therefore, (B.10) and (B.12) lead to

$$V_m(N^0(m)) \leqslant 2 \sum_{j=0}^{N^0(m)} \frac{\sigma_j^2}{m}. \tag{B.13}$$

The desired result finally follows from inequalities (B.11) and (B.13):

$$V_m(N^0(m)) \asymp \sum_{j=0}^{N^0(m)} \frac{\sigma_j^2}{m}.$$

The proof of this theorem is complete.

B.6. *The proof of Corollary 3.* It is immediate from conditions (i) and (ii) of the corollary and the proof of Theorem 7.

B.7. *The proof of Theorem 9.*

1) From (40) it follows that

$$\widetilde{f}_m^0(x) = \sum_{j=0}^{N^0(m)} \left[c_j\varphi_j(x) + \sum_{k=1}^{m} n_k^j \varphi_j(x) \right]^2 = f_{N^0(m)}(x) + \sum_{j=0}^{N^0(m)} \frac{1}{m} \sum_{k=1}^{m} n_k^j \varphi_j(x).$$

Taking the expectation of both sides of this equality yields

$$\mathsf{E}\widetilde{f}_m^0(x) = f_{N^0(m)}(x) + \mathsf{E} \sum_{j=0}^{N^0(m)} \frac{1}{m} \sum_{k=1}^{m} n_k^j \varphi_j(x) = f_{N^0(m)}(x) \neq f(x). \tag{B.14}$$

Thus, the estimator (40) is biased.

2) For proving the second assertion of this theorem, we have to show inequality (44). According to Theorem 4,

$$V_m(N^0(m)) \leqslant V_m(N)$$

for any $N \geqslant N^0(m)$. Passing to the limit as $N \to \infty$ gives

$$V_m(N^0(m)) \leqslant \lim_{N\to\infty} V_m(N) = V_m(\infty).$$

3) Next, we establish the third assertion of Theorem 9. Let us consider (40) and pass to the limit as $m \to \infty$. For almost all $x \in \mathrm{K}$, we obtain

$$\lim_{m\to\infty} \mathsf{E}\,\widetilde{f}_m^0(x) = \lim_{m\to\infty} f_{N^0(m)}(x).$$

By item (ii) of Corollary 3, $N^0(m) \xrightarrow[m\to\infty]{} \infty$. Hence,

$$\lim_{m\to\infty} f_{N^0(m)}(x) = f(x).$$

This means that the estimator (40) is asymptotically unbiased.

4) Finally, we demonstrate the consistency of the estimator (40). Due to the Chebyshev inequality, for any $m \in \mathbb{Z}^+ \setminus 0$ and $\varepsilon > 0$,

$$\mathsf{P}\Big(\big|\widetilde{f}_m^{\,0}(x) - f(x)\big|^2 \geqslant \varepsilon\Big) \leqslant \frac{1}{\varepsilon}\mathsf{E}\big|\widetilde{f}_m^{\,0}(x) - f(x)\big|^2. \tag{B.15}$$

We analyze the right-hand side of inequality (43). From (40) it follows that

$$\widetilde{f}_m^{\,0}(x) - f(x) = -\sum_{j=N^0(m)+1}^{\infty} c_j\varphi_j(x) + \frac{1}{m}\sum_{j=0}^{N^0(m)}\sum_{k=1}^{m} n_k^j\varphi_j(x).$$

Therefore, $\mathsf{E}\big|\widetilde{f}_m^{\,0}(x) - f(x)\big|^2$ can be represented as

$$\begin{aligned}\mathsf{E}\big|\widetilde{f}_m^{\,0}(x) - f(x)\big|^2 &= \mathsf{E}\left|-\sum_{j=N^0(m)+1}^{\infty} c_j\varphi_j(x) + \frac{1}{m}\sum_{j=0}^{N^0(m)}\sum_{k=1}^{m} n_k^j\varphi_j(x)\right|^2 \\ &= \left|\sum_{j=N^0(m)+1}^{\infty} c_j\varphi_j(x)\right|^2 + \frac{1}{m}\sum_{j=0}^{N^0(m)}\sigma_j^2\varphi_j^2(x).\end{aligned} \tag{B.16}$$

According to Corollary 3 and the conditions of this theorem, the series $\left|\sum\limits_{j=N^0(m)+1}^{\infty} c_j\varphi_j(x)\right|$ and $\sum\limits_{j=0}^{N^0(m)}\sigma_j^2\varphi_j^2(x)$ are convergent for almost all $x \in \mathrm{K}$. As a result, we have

$$\left|\sum_{j=N^0(m)+1}^{\infty} c_j\varphi_j(x)\right|^2 \xrightarrow[m\to\infty]{} 0,$$

$$\frac{1}{m}\sum_{j=0}^{N^0(m)}\sigma_j^2\varphi_j^2(x) \xrightarrow[m\to\infty]{} 0.$$

Consequently, for any $\varepsilon > 0$,

$$\lim_{m\to\infty}\mathsf{P}\Big(\big|\widetilde{f}_m^{\,0}(x) - f(x)\big|^2 \geqslant \varepsilon\Big) = 0.$$

The proof of this theorem is complete.

B.8. *The proof of Theorem 10.* According to the proof of Theorem 7, the standard deviation of the CPE satisfies the inequalities

$$2\sum_{j=0}^{N^0(m)}\frac{\sigma_j^2}{m} \geqslant V_m(N^0(m)) \geqslant \sum_{j=0}^{N^0(m)}\frac{\sigma_j^2}{m}.$$

From Theorem 4, Corollary 2, and Theorems 5 and 6 we have the inequality

$$\begin{aligned}V_m(N^0(m)) &= \inf_{N\in\mathbb{Z}^+}\ \inf_{f_{m,N}(x)\in\mathbb{M}_{2,m}(\widetilde{\mathsf{P}})}\mathsf{E}\int_{\mathrm{K}}\big[f(x) - f_{N,m}(x)\big]^2dx \\ &= \sum_{j=0}^{N^0(m)}\frac{\sigma_j^2}{m} + \sum_{j=N^0(m)+1}^{\infty}c_j^2 \leqslant 2\sum_{j=0}^{N^0(m)}\frac{\sigma_j^2}{m} \leqslant 2\sum_{j=0}^{\infty}\frac{\sigma_j^2}{m} = \frac{2\sigma^2}{m}.\end{aligned} \tag{B.17}$$

Since the right-hand side of (B.17) is independent of $f(x) \in L_2(\mathrm{K}, \Lambda)$, the desired conclusion follows. The proof of this theorem is complete.

REFERENCES

1. Borovkov, A.A., *Matematicheskaya statistika*, Novosibirsk: Nauka, 1997. Translated under the title *Mathematical Statistics*, 1st ed., Gordon and Breach, 1999.
2. Ivchenko, G.I. and Medvedev, Yu.I., *Vvedenie v matematicheskuyu statistiku*, Moscow: LKI, 2010. Translated under the title *Mathematical Statistics*, URSS, 1990.
3. Parzen, E., On Estimation of a Probability Density Function and Mode, *Ann. Math. Statist.*, 1962, vol. 33, no. 3, pp. 1065–1076. https://doi.org/10.1214/aoms/1177704472.
4. Rosenblatt, M., Curve Estimates, *Ann. Math. Statist.*, 1971, vol. 42, no. 6, pp. 1815–1842. https://doi.org/10.1214/aoms/1177693050.
5. Murthy, V.K., Nonparametric Estimation of Multivariate Densities with Applications, in *Multivariate Analysis*, 1966, pp. 43–56.
6. Stratonovich, R.L., The Efficiency of Mathematical Statistics Methods in the Design of Algorithms to Recover an Unknown Function, *Izv. Akad. Nauk SSSR. Tekh. Kibern.*, 1969, no. 1, pp. 32–46.
7. Watson, G.S., Density Estimation by Orthogonal Series, *Ann. Math. Statist.*, 1969, vol. 40, no. 4, pp. 1496–1498. https://doi.org/10.1214/aoms/1177697523
8. Konakov, V.D., Non-Parametric Estimation of Density Functions, *Theory of Probability & Its Applications*, 1973, vol. 17, iss. 2, pp. 361–362. https://doi.org/10.1137/1117042.
9. Chentsov, N.N., *Statisticheskie reshayushchie pravila i optimal'nye vyvody*, Moscow: Fizmatlit, 1972. Translated under the title *Statistical Decision Rules and Optimal Inference*, American Mathematical Society, 1982.
10. Vapnik, V.N., *Vosstanovlenie zavisimostei po empiricheskim dannym*, Moscow: Nauka, 1979. Translated under the title *Estimation of Dependences Based on Empirical Data*, New York: Springer, 2010.
11. Ibragimov, I.A. and Has'minskii, R.Z., *Assimptoticheskaya teoriya otsenivaniya*, Moscow: Nauka, 1979. Translated under the title *Statistical Estimation: Asymptotic Theory*, New York: Springer, 1981.
12. Nadaraya, E.A., *Neparametricheskoe otsenivanie plotnostei veroyatnostei i krivoi regressii*, Tbilisi: Tbilissk. Gos. Univ., 1983. Translated under the title *Nonparametric Estimation of Probability Densities and Regression Curves*, Dordrecht: Springer, 1988. https://doi.org/10.1007/978-94-009-2583-0.
13. Nemirovskij, A.S., Polyak, B.T., and Tsybakov, A.B., Signal Processing by the Nonparametric Maximum-Likelihood Method, *Problems of Information Transmission*, 1984, vol. 20, no. 3, pp. 177–192.
14. Darkhovskii, B.S., On a Stochastic Renewal Problem, *Theory of Probability & Its Applications*, 1999, vol. 43, no. 2, pp. 282–288. https://doi.org/10.1137/S0040585X9797688X.
15. Darkhovsky, B.S., Stochastic Recovery Problem, *Problems of Information Transmission*, 2008, vol. 44, no. 4, pp. 303–314. https://doi.org/10.1134/S0032946008040030.
16. Ibragimov, I.A., Estimation of Multivariate Regression, *Theory of Probability & Its Applications*, 2004, vol. 48, no. 2, pp. 256–272. https://doi.org/10.1137/S0040585X9780385.
17. Tsybakov, A.B., *Introduction to Nonparametric Estimation*, New York: Springer, 2009.
18. Bulgakov, S.A. and Khametov, V.M., Recovery of a Square Integrable Function from Observations with Gaussian Errors, *Upravl. Bol'sh. Sist.*, 2015, vol. 54, pp. 45–65.
19. Levit, B., On Optimal Cardinal Interpolation, *Mathematical Methods of Statistics*, 2018, vol. 27, no. 4, pp. 245–267. https://doi.org/10.3103/S1066530718040014.
20. Juditsky, A.B. and Nemirovski, A.S., Signal Recovery by Stochastic Optimization, *Autom. Remote Control*, 2019, vol. 80, no. 10, pp. 1878–1893. https://doi.org/10.1134/S0005231019100088.
21. Golubev, G.K., On Adaptive Estimation of Linear Functionals from Observations against White Noise, *Problems of Information Transmission*, 2020, vol. 56, no. 2, pp. 185–200. https://doi.org/10.31857/S0555292320020047.

22. Bulgakov, S.A., Gorshkova, V.M., and Khametov, V.M., Stochastic Recovery of Square-Integrable Functions, *Herald of the Bauman Moscow State Technical University*, 2020, vol. 93, no. 6, pp. 4–22. https://doi.org/10.18698/1812-3368-2020-6-4-22.

23. Shiryaev, A.N., *Veroyatnost'*, Moscow: Nauka, 1980. Translated under the title *Probability*, Graduate Texts in Mathematics, vol. 95, New York: Springer, 1996. https://doi.org/10.1007/978-1-4757-2539-1.

24. Kolmogorov, A.N. and Fomin, S.V., *Elementy teorii funktsii i funktsional'nogo analiza*, 4th. ed., Moscow: Nauka, 1976. Translated under the title *Elements of the Theory of Functions and Functional Analysis*, Dover, 1957.

25. Kashin, B.S. and Saakyan, A.A., *Ortogonal'nye ryady* (Orthogonal Series), Moscow: Nauka, 1984.

26. Bertsekas, D. and Shreve, S.E., *Stochastic Optimal Control: The Discrete-Time Case*, Athena Scientific, 1996.

This paper was recommended for publication by O.N. Granichin, a member of the Editorial Board