Известия НАН Армении, Математика, том 54, н. 6, 2019, стр. 28 – 41

ASYMPTOTIC BEHAVIOR OF THE VARIANCE OF THE BEST LINEAR UNBIASED ESTIMATOR FOR THE MEAN OF A DISCRETE-TIME SINGULAR STATIONARY PROCESS

N. M. BABAYAN, M. S. GINOVYAN

*Russian-Armenian University, Yerevan, Armenia**
Institute of Mathematics Yerevan, Armenia, Boston University, Boston, USA
E-mails: *nmbabayan@gmail.com; ginovyan@math.bu.edu*

Abstract. It is known that for a wide class of discrete-time stationary processes possessing spectral densities f, the variance $\sigma_n^2(f)$ of the best linear unbiased estimator for the mean depends asymptotically only on the behavior of the spectral density f near the origin, and behaves hyperbolically as $n \to \infty$. In this paper, we obtain necessary as well as sufficient conditions for exponential rate of decrease of $\sigma_n^2(f)$ as $n \to \infty$. In particular, we show that a necessary condition for $\sigma_n^2(f)$ to decrease to zero exponentially is that the spectral density f vanishes on a set of positive measure in any vicinity of zero, and if f vanishes only at the origin, then it is impossible to obtain exponential decay of $\sigma_n^2(f)$, no mater how high the order of the zero of f at the origin.

MSC2010 numbers: 60G10, 62M15, 62F12.

Keywords: linear estimation; discrete-time stationary process; mean; variance; spectral density; exponential rate.

1. Introduction

Consider the following, possibly complex-valued, stochastic model:

$$Y(t) = m + X(t), \quad t \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\},$$

where m is the constant unknown mean of $Y(t)$, and the noise $X(t)$ is assumed to be a zero-mean, wide-sense stationary process with a spectral density function $f(\lambda)$, $\lambda \in \Lambda := [-\pi, \pi]$, and a covariance function $r(t)$, $t \in \mathbb{Z}$, so that

$$r(t) = E[X(t+s)\overline{X}(s)] = \int_{-\pi}^{\pi} e^{it\lambda} f(\lambda) d\lambda. \tag{1.1}$$

In this paper we consider the problem of estimation of the unknown mean m for this model by unbiased linear estimators $\widehat{m}_n$, based on a random sample $\{Y(t), t =$

*The research of M. S. Ginovyan was partially supported by National Science Foundation Grant #DMS-1309009 at Boston University.

$0, 1, \ldots, n\}$:

$$\widehat{m}_n = \sum_{k=0}^{n} c_k Y(k), \qquad \sum_{k=0}^{n} c_k = 1, \tag{1.2}$$

where the condition $\sum_{k=0}^{n} c_k = 1$ is needed for unbiasedness of $\widehat{m}_n$. Of particular interest is the best linear unbiased estimator (BLUE) $\widehat{m}_{n,BLU}$, that is, the estimator of the form (1.2), where the weights c_k, $k = 0, 1, \ldots, n$, are chosen so that the variance

$$\mathrm{Var}\,(\widehat{m}_n) = E|\widehat{m}_n - m|^2 = \sum_{j,k=0}^{n} c_j \bar{c}_k r(j-k) \tag{1.3}$$

is minimal under the condition $\sum_{k=0}^{n} c_k = 1$.

We assume that the noise process $X(t)$ is non-degenerate, that is, $E|X(t)|^2 = r(0) = \int_{-\pi}^{\pi} f(\lambda)d\lambda > 0$, which implies that the BLUE exists and is unique (see, e.g., Adenstedt [1], or Grenander and Szegö [14], Sections 11.1). Also, under our assumptions the BLUE $\widehat{m}_{n,BLU}$ is a mean square consistent estimator for m, that is, $\lim_{n\to\infty} E|\widehat{m}_{n,BLU} - m|^2 = 0$ (see Grenander and Szegö [14], Sections 11.2).

Typically, calculation of $\widehat{m}_{n,BLU}$ and its variance is difficult, because they involve the inverse of covariance matrix, and hence adequate approximations are needed in terms of more easily calculated estimators.

There is a substantial literature comparing the BLUE with other estimators, especially the least squares estimator (LSE) $\widehat{m}_{n,LS} = (n+1)^{-1}\sum_{k=0}^{n} Y_k$ (see, e.g., Adenstedt [1], Adenstedt and Eisenberg [2], Beran [5], Beran and Künsch [6], Beran et al. [7], Grenander [10]-[12], Grenander and Rosenblatt [13], Grenander and Szegö [14], Samarov and Taqqu [18], Vitale [19], Yajima [20], and references therein).

We are concerned here with asymptotic behavior of the variance:

$$\sigma_n^2(f) := \mathrm{Var}\,(\widehat{m}_{n,BLUE}, f) \quad \text{as} \quad n \to \infty. \tag{1.4}$$

To recall some known results in this direction, we first recall the definitions of short memory, anti-persistent and long memory processes (see, e.g., Beran et al. [7], Sections 1.3.1). We say that the process $X(t)$ displays *short memory* if the covariance function $r(t)$ satisfies the condition: $0 < \sum_{t\in\mathbb{Z}} r(t) < \infty$. In this case the spectral density $f(\lambda)$ is bounded away from zero and infinity at frequency $\lambda = 0$, that is, $0 < f(0) < \infty$. The process $X(t)$ is said to be *anti-persistent* if $\sum_{t\in\mathbb{Z}} r(t) = 0$. In this case the spectral density $f(\lambda)$ vanishes at frequency zero: $f(0) = 0$. We say that the process $X(t)$ displays *long memory* or *long-range dependence* if $\sum_{t\in\mathbb{Z}} r(t) = \infty$. In this case the spectral density $f(\lambda)$ has a pole at frequency zero, that is, it is unbounded at the origin.

The first result on asymptotic behavior of the variance $\sigma_n^2(f)$, defined by (1.4), as $n \to \infty$, goes back to the classical works by Grenander (see [10]-[12], and [14], Sections 11.1-11.3), who described the asymptotic behavior of $\sigma_n^2(f)$ for short-memory models. More precisely, he showed that, as $n \to \infty$, both $\widehat{m}_{n,BLU}$ and $\widehat{m}_{n,LS}$ have asymptotic variance $2\pi f(0)/n$ as long as $f(\lambda)$ is positive and continuous. Then Vitale [19] considered the case of anti-persistent models, when $f(\lambda)$ is continuous and positive except at the origin, where $f(\lambda) \cong L\lambda^2$ $(L > 0)$ as $\lambda \to 0$. He showed that in this case $\sigma_n^2(f) \cong 24\pi L/n^3$ as $n \to \infty$, and the estimator (1.2) with coefficients $c_k = 6k(n-k)/[n(n^2-1)]$ is asymptotically efficient, while $\widehat{m}_{n,LS}$ is not. The asymptotic behavior of $\sigma_n^2(f)$ for long-memory models was studied by Adenstedt [1]. Let

$$f_\alpha(\lambda) = \frac{1}{2\pi}\left|1 - e^{i\lambda}\right|^{2\alpha} = \frac{2^{2\alpha-1}}{\pi}\left(\sin^2\frac{\lambda}{2}\right)^\alpha, \quad \alpha > -\frac{1}{2}, \tag{1.5}$$

$$g(\lambda) = h(\lambda)|\lambda - \lambda_1|^{\alpha_1}\cdots|\lambda - \lambda_r|^{\alpha_r}, \tag{1.6}$$

where $r \in \mathbb{N} := \{1, 2, \ldots\}$, $\lambda_1, \lambda_2, \ldots, \lambda_r$ are non-zero distinct constants in $[-\pi, \pi]$, $\alpha_1, \alpha_2, \ldots, \alpha_r$ are nonnegative constants, and $0 < C_1 \le h(\lambda) \le C_2 < \infty$, with some constants C_1 and C_2.

In [1], it was shown that the variance $\mathrm{Var}\,(\widehat{m}_{n,BLU}, f)$ depends asymptotically only on the behavior of the spectral density $f(\lambda)$ near the origin $\lambda = 0$, and among others, was proved the following result: if the noise $X(t)$ has a spectral density of the form $f(\lambda) = f_\alpha(\lambda)g(\lambda)$ with $f_\alpha(\lambda)$ and $g(\lambda)$ as in (1.5) and (1.6), respectively, then

$$\sigma_n^2(f) \simeq n^{-2\alpha-1}\frac{\Gamma(2\alpha+1)g(0)}{B(\alpha+1, \alpha+1)},$$

where $B(p, q) = \Gamma(p)\cdot\Gamma(q)/\Gamma(p+q)$, and $\Gamma(p)$ is the gamma function.

Thus, the variance $\sigma_n^2(f)$ of the best linear unbiased estimator for the mean depends asymptotically only on the behavior of the spectral density f near the origin, and for some classes of spectral densities satisfying the condition $f(\lambda) \sim \lambda^\nu$ $(\nu > -1)$ as $\lambda \to 0$, the variance $\sigma_n^2(f)$ decreases hyperbolically, that is,

$$\sigma_n^2(f) \sim n^{-\nu-1} \quad \text{as} \quad n \to \infty. \tag{1.7}$$

In this paper we obtain necessary as well as sufficient conditions for exponential rate of decrease of $\sigma_n(f)$ as $n \to \infty$, that is, for fulfillment of the equality:

$$\sigma_n(f) = \gamma_n\tau^n, \tag{1.8}$$

where $0 < \tau < 1$ and $\{\gamma_n,\, n \in \mathbb{N}\}$ is a sequence of positive numbers satisfying the condition: $\lim_{n\to\infty}\sqrt[n]{\gamma_n} = 1$.

Throughout the paper we will use the following notation.

By $\mathbb{T}$ we denote the unit circle $|z| = 1$ in the complex plane $\mathbb{C}$, that is, $\mathbb{T} = \{e^{i\lambda},\ \lambda \in [-\pi, \pi]\}$. By Γ_α $(0 < \alpha < \pi)$ we denote the arc of the unit circle $\mathbb{T}$ of length 2α with center at $z = 1$, that is, $\Gamma_\alpha = \{e^{i\lambda},\ |\lambda| \leq \alpha, 0 < \alpha < \pi\}$. By Γ'_α we denote the arc of $\mathbb{T}$ of length $2(\pi - \alpha)$ with center at $z = -1$, that is, $\Gamma'_\alpha = \{e^{i\lambda},\ \alpha \leq |\lambda| \leq \pi, 0 < \alpha < \pi\}$. By W we denote the mapping $W : \Lambda \to \mathbb{T}$, defined by formula $W(\lambda) = e^{i\lambda}$. By E_f we denote the spectrum of process $X(t)$, that is, $E_f = \{e^{i\lambda},\ f(\lambda)) > 0\}$. By $\mathcal{Q}_n$ we denote the set of polynomials $q_n(z) = c_0 z^n + c_1 z^{n-1} + \cdots + c_{n-1} z + c_n$, $z \in \mathbb{C}$ of degree $n \in \mathbb{N}$ with leading coefficient $c_0 = 1$. For a fixed complex number $z_0 \in \mathbb{C}$, by $\mathcal{Q}_n(z_0)$ we denote the set of polynomials $q_n(z)$, $z \in \mathbb{C}$ of degree at most n, satisfying the condition $q_n(z_0) = 1$. In particular, $\mathcal{Q}_n(1) = \{q_n(z) = c_0 z^n + c_1 z^{n-1} + \cdots + c_{n-1} z + c_n :\ q_n(1) = \sum_{k=0}^{n} c_k = 1\}$.

The paper is structured as follows. In Section 2 we state the main results of the paper - Theorems 2.1 and 2.2. Section 3 contains a number of preliminary results, needed in the proofs of the main results. Section 4 is devoted to the proofs of Theorems 2.1 and 2.2.

2. Main results

The main results of the present paper are the following theorems.

Theorem 2.1. *If the spectrum $E_f = \{e^{i\lambda},\ f(\lambda) > 0\}$ of the process $X(t)$ is an arc of the unit circle or the unit circle itself, then the sequence $\{\sqrt[n]{\sigma_n(f)},\ n \in \mathbb{N}\}$ converges to some limit $\widetilde{\tau}(E_f) \leq 1$, that is,*

$$\lim_{n\to\infty} \sqrt[n]{\sigma_n(f)} = \widetilde{\tau}(E_f). \tag{2.1}$$

Theorem 2.2. *The following assertions hold:*

(a) *If the spectral density $f(\lambda)$ is positive almost everywhere in some vicinity of zero, then*

$$\lim_{n\to\infty} \sqrt[n]{\sigma_n(f)} = 1. \tag{2.2}$$

(b) *If the spectral density $f(\lambda)$ vanishes almost everywhere for $|\lambda| < \alpha$, $0 < \alpha < \pi$, then $\sigma_n(f)$ decreases at least exponentially. More precisely, we have*

$$\lim_{n\to\infty} \sqrt[n]{\sigma_n(f)} \leqslant \cos\frac{\alpha}{2}. \tag{2.3}$$

As an immediate consequence of the assertion (a) of Theorem 2.2 we have the following result.

Corollary 2.1. *A necessary condition for variance $\sigma_n^2(f)$ to decrease to zero exponentially as $n \to \infty$ is that the spectral density $f(\lambda)$ vanishes on a set of positive Lebesgue measure in any vicinity of zero.*

Remark 2.1. The relation (2.1) is equivalent to the equality (1.8), where $\tau = \widetilde{\tau}(E_f)$ and $\gamma_n = \sigma_n(f)/\tau^n$. The quantity $\widetilde{\tau}(E_f)$ is a metric characteristic of the spectrum E_f, which we call the generalized Tchebychev constant of the set E_f (see Lemma 3.3 and Remark 3.3). Thus, Theorem 2.1 shows that the question of exponential decay of $\sigma_n(f)$ in fact does not depend on the form of $f(\lambda)$ and is determined solely by the value of $\widetilde{\tau}(E_f)$, while the factor γ_n in (1.8) is determined by the behavior of spectral density $f(\lambda)$ on the spectrum E_f.

Remark 2.2. From the proof of Theorem 2.1, one conclude that the assertion of the theorem remains valid also in the case where the spectrum E_f consists of the union of a finite number of arcs of the unit circle.

Remark 2.3. From the results mentioned in Introduction (see, e.g., (1.7)), it follows that the higher the order of zero of spectral density $f(\lambda)$ at $\lambda = 0$, the higher the rate of decrease of $\sigma_n^2(f)$ to zero. Corollary 2.1 shows, in particular, that if $f(\lambda)$ vanishes only at the origin, then it is impossible to obtain exponential decay of $\sigma_n^2(f)$, no matter how high the order of the zero of $f(\lambda)$ at the origin.

Remark 2.4. For the considered estimation problem we have a complete similarity with the problem of asymptotic behavior of the best linear prediction error variance for stationary processes (see Babayan [3], [4] and Rosenblatt [17]). The only difference is that in the prediction problem the asymptotic of prediction error variance is determined by the behavior of spectral density $f(\lambda)$ on the entire interval $[-\pi, \pi]$ but not only at the origin.

3. Preliminaries

In this section we present a number of auxiliary results that will be used in the proofs of Theorems 2.1 and 2.2. We first use Kolmogorov's isometric isomorphism between the time- and frequency-domains: $X(t) \leftrightarrow e^{it\lambda}$ to reformulate the problem of finding BLUE in the frequency domain, and apply a result by Szegö to obtain a convenient formula for variance $\sigma_n^2(f) = \mathrm{Var}\,(\widehat{m}_{n,BLU}, f)$ in terms of orthogonal polynomials on the unit circle with respect to spectral density $f(\lambda)$.

Using (1.1)-(1.3) we can write

$$\begin{aligned}\operatorname{Var}(\widehat{m}_n) &= \sum_{j,k=0}^{n} c_j\bar{c}_k r(j-k) = \int_{-\pi}^{\pi}\left|\sum_{\nu=0}^{n} c_\nu e^{i\nu\lambda}\right|^2 f(\lambda)d\lambda\\ &= \int_{-\pi}^{\pi}\left|\sum_{\nu=0}^{n} c_\nu e^{i(n-\nu)\lambda}\right|^2 f(\lambda)d\lambda = \int_{-\pi}^{\pi}|q_n(e^{i\lambda}|^2 f(\lambda)d\lambda,\end{aligned}$$

where $q_n(z) = c_0z^n+c_1z^{n-1}+\cdots+c_{n-1}z+c_n$. Thus, the problem of finding $\widehat{m}_{n,BLU}$ becomes to the solution of the following minimum problem:

$$\int_{-\pi}^{\pi}|q_n(e^{i\lambda})|^2 f(\lambda)d\lambda = \min, \quad q_n(z)\in\mathcal{Q}_n(1). \tag{3.1}$$

The polynomial $p_n(z) := p_n(z,f)$ that solves the minimum problem (3.1) is called the *optimal polynomial* for $f(\lambda)$. The optimal polynomial $p_n(z,f)$ exists, is unique and can be expressed in terms of orthogonal polynomials $\varphi_n(z)$, $n\in\mathbb{Z}_+ := \{0,1,2,\ldots\}$, on the unit circle $\mathbb{T}$ with respect to $f(\lambda)$.

The system of orthogonal polynomials $\{\varphi_n(z) = \varphi_n(f;z),\ z = e^{i\lambda},\ n\in\mathbb{Z}_+\}$ is uniquely determined by the following conditions:

(i) $\varphi_n(z) = \kappa_n(f)z^n +$ lower order terms
is a polynomial of degree n, in which the coefficient $\kappa_n = \kappa_n(f)$ is real and positive;

(ii) for arbitrary nonnegative integers k and j

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\varphi_k(z)\overline{\varphi_j(z)}f(\lambda)d\lambda = \delta_{kj} = \begin{cases}1, & \text{for } k=j\\ 0, & \text{for } k\neq j,\end{cases} \qquad z = e^{i\lambda}.$$

For a fixed $z_0\in\mathbb{C}$, consider the *Szegö kernel* $G_n(z,z_0)$ defined by

$$G_n(z,z_0) = G_n(f,z,z_0) = \sum_{k=0}^{n}\varphi_k(z)\overline{\varphi_k(z_0)}.$$

The kernel $G_n(z,z_0)$ possesses the following extremal property (see Nikishin and Sorokin [16], Section 3.6, and Grenander and Szegö [14], Section 2.2).

Lemma 3.1. *Let the set of polynomials $\mathcal{Q}_n(z_0)$ and the kernel $G_n(z,z_0)$ be as above. The polynomial*

$$p_n(z) = \frac{G_n(z,z_0)}{G_n(z_0,z_0)}$$

is the unique solution of the extremal problem:

$$\min_{q_n\in\mathcal{Q}_n(z_0)}\int_{-\pi}^{\pi}\left|q_n(e^{i\lambda})\right|^2 f(\lambda)d\lambda = \int_{-\pi}^{\pi}\left|p_n(e^{i\lambda},f)\right|^2 f(\lambda)d\lambda = \frac{1}{G_n(z_0,z_0)}.$$

Applying Lemma 3.1 with $z_0 = 1$, we obtain that the optimal polynomial $p_n(z, f)$ in finding $\widehat{m}_{n,BLU}$ is given by formula:

$$p_n(z, f) = \frac{G_n(z, 1)}{G_n(1, 1)} = \frac{\sum_{v=0}^{n} \overline{\varphi_v(1)}\varphi_v(z)}{\sum_{v=0}^{n} |\varphi_v(1)|^2}.$$

Thus, for variance $\sigma_n^2(f)$ we have

$$\begin{aligned} \sigma_n^2(f) &= \min_{q_n \in \mathcal{Q}_n(1)} \int_{-\pi}^{\pi} \left|q_n(e^{i\lambda})\right|^2 f(\lambda) d\lambda = \int_{-\pi}^{\pi} \left|p_n(e^{i\lambda}, f)\right|^2 f(\lambda) d\lambda \\ &= \frac{1}{G_n(1, 1)} = \frac{1}{\sum_{v=0}^{n} |\varphi_v(1)|^2}. \end{aligned} \tag{3.2}$$

Remark 3.1. From the obvious embedding $\mathcal{Q}_n(1) \subset \mathcal{Q}_{n+1}(1)$, it follows that the sequence $\{\sigma_n^2(f),\, n \in \mathbb{N}\}$ is non-increasing in n: $\sigma_{n+1}^2(f) \leq \sigma_n^2(f)$. Also, it follows from (3.2) that $\sigma_n^2(f)$ is a non-decreasing functional of $f(\lambda)$:

$$\sigma_n^2(f) \leq \sigma_n^2(g) \quad \text{when} \quad f(\lambda) \leq g(\lambda), \quad \lambda \in \Lambda. \tag{3.3}$$

Indeed, by the definition of optimal polynomials $p_n(z, f)$ and $p_n(z, g)$, corresponding to spectral densities f and g, respectively, we have

$$\begin{aligned} \sigma_n^2(f) &= \int_{-\pi}^{\pi} \left|p_n(e^{i\lambda}, f)\right|^2 f(\lambda) d\lambda \leq \int_{-\pi}^{\pi} \left|p_n(e^{i\lambda}, g)\right|^2 f(\lambda) d\lambda \\ &\leq \int_{-\pi}^{\pi} \left|p_n(e^{i\lambda}, g)\right|^2 g(\lambda) d\lambda = \sigma_n^2(g). \end{aligned}$$

The next lemma is an analog of the assertion on existence of Tchebychev polynomial for the class $\mathcal{Q}_n(1)$ (see Goluzin [9], Section 7.1).

Lemma 3.2. *Let F be an arbitrary infinite bounded closed set in the complex plane $\mathbb{C}$, and let $\mathcal{Q}_n(1)$ be the set of polynomials $q_n(z)$, $z \in \mathbb{C}$ of degree at most n, satisfying the condition $q_n(1) = 1$. Then in the class $\mathcal{Q}_n(1)$ there exists a polynomial $\widetilde{T}_n(z) := \widetilde{T}_n(z, F)$ with least maximum modulus on F:*

$$\max_{z \in F} |\widetilde{T}_n(z)| = \min_{q_n \in \mathcal{Q}_n(1)} \max_{z \in F} |q_n(z)|. \tag{3.4}$$

Proof. Denote

$$\widetilde{m}_n(F) := \inf_{q_n \in \mathcal{Q}_n(1)} \max_{z \in F} |q_n(z)|. \tag{3.5}$$

Let for a fixed $n \in \mathbb{N}$, $\{q_{n,k}(z),\, k \in \mathbb{N}\}$ be a sequence of polynomials from the class $\mathcal{Q}_n(1)$, whose maxima of moduli on F tend to $\widetilde{m}_n(F)$ as $k \to \infty$. We fix $n+1$ points

$z_1, z_2, \ldots, z_{n+1} \in F$ and represent the polynomials $q_{n,k}(z)$ by the Lagrange formula as follows:

$$q_{n,k}(z) = \sum_{\nu=1}^{n+1} \frac{(z-z_1)(z-z_2)\cdots(z-z_{\nu-1})(z-z_{\nu+1})\cdots(z-z_{n+1})}{(z_\nu-z_1)(z_\nu-z_2)\cdots(z_\nu-z_{\nu-1})(z_\nu-z_{\nu+1})\cdots(z_\nu-z_{n+1})}\, q_{n,k}(z_\nu).$$

It follows from this representation that the polynomials $q_{n,k}(z)$ are uniformly in k bounded in modulus on an arbitrary bounded closed set of $\mathbb{C}$.

Hence, in view of known condensation principle (see Goluzin [9], Section 1.1), the sequence $\{q_{n,k}(z),\, k \in \mathbb{N}\}$ contains a subsequence of polynomials that converges uniformly on every bounded set of $\mathbb{C}$. Also, together with the polynomials from this subsequence, the sequence of their coefficients converges as well, and hence the limiting function $\widetilde{T}_n(z) := \widetilde{T}_n(z, F)$ is a polynomial from the class $\mathcal{Q}_n(1)$, and satisfies the condition (3.4), that is, we have

$$\max_{z\in F} |\widetilde{T}_n(z, F)| = \widetilde{m}_n(F) = \min_{q_n \in \mathcal{Q}_n(1)} \max_{z \in F} |q_n(z)|. \tag{3.6}$$

Lemma 3.2 is proved. □

Remark 3.2. The polynomial $\widetilde{T}_n(z, F)$ is an analog of the Tchebychev polynomial $T_n(z, F)$, which has least maximum modulus on the set F in the class $\mathcal{Q}_n$, and hence, we call $\widetilde{T}_n(z, F)$ the *Tchebychev polynomial for the set F with respect to the point $z_0 = 1$*, or simply *the generalized Tchebychev polynomial*.

The next lemma is an analog of the assertion on existence of Tchebychev constant for an arbitrary bounded closed set F of the complex plane $\mathbb{C}$ (see Goluzin [9], Section 7.1).

Lemma 3.3. *For any bounded closed set F of the complex plane $\mathbb{C}$, the sequence $\{\widetilde{\tau}_n(F) := \sqrt[n]{\widetilde{m}_n(F)},\, n \in \mathbb{N}\}$, where $\widetilde{m}_n(F)$ is as in* (3.5), *converges to some finite limit $\widetilde{\tau}(F)$, that is,*

$$\lim_{n\to\infty} \sqrt[n]{\widetilde{m}_n(F)} = \widetilde{\tau}(F) < \infty. \tag{3.7}$$

Proof. Taking into account that the set F is bounded, we have

$$R := R_F = \max_{z\in F} |z| < \infty. \tag{3.8}$$

Next, observing that the polynomial $q_n(z) = z^n$ belongs to the class $\mathcal{Q}_n(1)$, in view of (3.6) and (3.8), we conclude that the sequence $\{\widetilde{\tau}_n(F),\, n \in \mathbb{N}\}$ is bounded:

$$\widetilde{\tau}_n(F) = \sqrt[n]{\widetilde{m}_n(F)} \le \sqrt[n]{\max_{z\in F} |z^n|} = R. \tag{3.9}$$

Define

$$\liminf_{n\to\infty} \widetilde{\tau}_n(F) = a \quad \text{and} \quad \limsup_{n\to\infty} \widetilde{\tau}_n(F) = b,$$

and observe that $a \leqslant b$. So, to complete the proof, we have to show that $b \leqslant a$. To this end, for given $\varepsilon > 0$ we choose $n_0 \in \mathbb{N}$ so that $\widetilde{\tau}_{n_0}(F) < a + \varepsilon$. Then on the set F we have the following inequality

$$|\widetilde{T}_{n_0}(z)| < (a+\varepsilon)^{n_0}, \quad z \in F.$$

Next, observe that for any $q \in \mathbb{N}$ and $r \in \mathbb{Z}_+$, the polynomial $s_n(z) = z^r[\widetilde{T}_{n_0}(z)]^q$ of degree $n = n_0 q + r$ satisfies the conditions:

$$s_n(z) \in \mathcal{Q}_n(1) \quad \text{and} \quad |s_n(z)| < R^r(a+\varepsilon)^{n_0 q}, \quad z \in F. \tag{3.10}$$

Therefore, in view of (3.6) and (3.10), we have

$$\widetilde{m}_n(F) \leq \max_{z \in F} |s_n(z)| \leqslant R^r(a+\varepsilon)^{n_0 q}$$

and

$$\widetilde{\tau}_n(F) \leqslant R^{\frac{r}{n}}(a+\varepsilon)^{\frac{n_0 q}{n}}. \tag{3.11}$$

Now let the subsequence $\{\widetilde{\tau}_{n_\nu}(F),\, n_\nu \in \mathbb{N}\}$ converge to b as $\nu \to \infty$. We write the inequality (3.11) for $n_\nu := n_0 q_\nu + r_\nu$ with $0 \leq r_\nu < n_0$ to obtain

$$\widetilde{\tau}_{n_\nu}(F) \leqslant R^{\frac{r_\nu}{n_\nu}}(a+\varepsilon)^{\frac{n_0 q_\nu}{n_\nu}}. \tag{3.12}$$

Finally, letting ν tend to infinity, from (3.12) we obtain $b \leqslant a + \varepsilon$. Taking into account arbitrariness of ε, we conclude that $b \leqslant a$. □

Remark 3.3. 1. The quantity $\widetilde{\tau}(F)$ is a metric characteristic of a closed set F, similar to Tchebychev constant $\tau(F)$ (see Goluzin [9], Section 7.1). Hence, we call $\widetilde{\tau}(F)$ the *Tchebychev constant of the set F with respect to the point $z_0 = 1$*, or simply the *generalized Tchebychev constant of F*.
2. It follows from the relation (3.9) that $\widetilde{\tau}(F) \leq R_F$. In particular, for unit circle $\mathbb{T}$ we have

$$\widetilde{\tau}(\mathbb{T}) \leq 1. \tag{3.13}$$

3. The quantity $\widetilde{\tau}(F)$ is a non-decreasing set function, that is, if $F_1 \subset F_2$, then

$$\widetilde{\tau}(F_1) \leq \widetilde{\tau}(F_2). \tag{3.14}$$

Indeed, in view of (3.6) we have

$$\widetilde{m}_n(F_1) = \max_{z \in F_1} |\widetilde{T}_n(z, F_1)| \leq \max_{z \in F_1} |\widetilde{T}_n(z, F_2)| \leq \max_{z \in F_2} |\widetilde{T}_n(z, F_2)| = \widetilde{m}_n(F_2),$$

from which, after taking the root of order n and passing to the limit as $n \to \infty$, we obtain (3.14). Also, from inequalities (3.13) and (3.14) it follows that for any $F \subset \mathbb{T}$, we have

$$\widetilde{\tau}(F) \leq 1. \tag{3.15}$$

4. If the set F contains the point $z = 1$, then we obviously have

$$\widetilde{\tau}(F) \geq 1, \tag{3.16}$$

and from inequalities (3.15) and (3.16) we infer that for any set $F \subset \mathbb{T}$, containing the point $z = 1$, we have

$$\widetilde{\tau}(F) = 1.$$

In particular, for the unit circle $\mathbb{T}$ and for any its arc $\Gamma_\alpha := \{e^{i\lambda},\ |\lambda| \leq \alpha, 0 < \alpha < \pi\}$ we have

$$\widetilde{\tau}(\mathbb{T}) = \widetilde{\tau}(\Gamma_\alpha) = 1. \tag{3.17}$$

Notice that for unit circle $\mathbb{T}$, the Tchebychev constant $\tau(\mathbb{T})$ is also equal to 1 (see Goluzin [9], Section 7.1). Thus, we have $\widetilde{\tau}(\mathbb{T}) = \tau(\mathbb{T}) = 1$.

In the proof of Theorem 2.1 will be used arguments similar to those applied in Babayan [3], [4], to obtain the corresponding result concerning asymptotic behavior of the best linear prediction error variance for stationary processes, where a key role played the following result by S. Mazurkievicz (see, e.g., Geronimus [8], Mazurkievicz [15]).

Lemma 3.4. *For any $\epsilon > 0$ there is $\delta = \delta(\epsilon) > 0$ depending only on ϵ, such that for any continuum Γ of diameter d and any of its closed subset $F \subset \Gamma$ the following inequality holds:*

$$M_n := \max_{x \in \Gamma} |q_n(z)| \leq (1 + \varepsilon)^n \max_{z \in F} |q_n(z)|,$$

provided that $\mu(\Gamma \setminus F) < d\delta$, where $q_n(z)$ is an arbitrary polynomial of degree n, and $\mu(e)$ stands for the linear measure of a set e.

4. Proof of Theorems 2.1 and 2.2

Proof of Theorem 2.1. In view of the definition of optimal polynomial $p_n(z, f)$, and formulas (3.2) and (3.6), we can write

$$\begin{aligned} \sigma_n^2(f) = \int_{-\pi}^{\pi} |p_n(e^{i\lambda}, f)|^2 f(\lambda) d\lambda \quad &\leq \quad \int_{-\pi}^{\pi} |\widetilde{T}_n(e^{i\lambda}, E_f)|^2 f(\lambda) d\lambda \\ &\leq \quad \widetilde{m}_n^2(E_f) \cdot \int_{-\pi}^{\pi} f(\lambda) d\lambda. \end{aligned}$$

Hence, by Lemma 3.3 and (3.15), we get

$$\limsup_{n\to\infty} \sqrt[n]{\sigma_n(f)} \leqslant \widetilde{\tau}(E_f) \leqslant 1. \tag{4.1}$$

Now we proceed to prove the inequality:

$$\liminf_{n\to\infty} \sqrt[n]{\sigma_n(f)} \geqslant \widetilde{\tau}(E_f). \tag{4.2}$$

To this end, we first consider a sequence of subsets $\{e_n, n \in \mathbb{N}\}$ of E_f, defined by the relation:

$$e_n = \left\{ z \in E_f : |p_n(z,f)| > \sqrt{\sigma_n(f) \cdot \widetilde{m}_n(E_f)} \right\}, \tag{4.3}$$

and by μ_f denote a measure on the unit circle defined as follows:

$$\mu_f(e) = \int_{W^{-1}(e)} f(\lambda) d\lambda, \quad e \subset E_f,$$

where $W^{-1}(e) = \{\lambda \in \Lambda : e^{i\lambda} \in e\}$. It is clear that

$$\limsup_{n\to\infty} \sqrt[n]{\mu_f(e_n)} \leqslant 1.$$

Next, for a given sufficiently small number ρ, $0 < \rho < 1$, the set of natural numbers $\mathbb{N}$ we write in the form $\mathbb{N} = J_1 \cup J_2$, where

$$\begin{aligned} J_1 &= J_1(\rho) = \{n \in \mathbb{N} : \mu_f(e_n) > (1-\rho)^n\}, \\ J_2 &= J_2(\rho) = \{n \in \mathbb{N} : \mu_f(e_n) \leq (1-\rho)^n\}. \end{aligned}$$

It is clear that at least one of the sets J_1 and J_2 is infinite. Without loss of generality, we can assume that both J_1 and J_2 are infinite sets.

For $n \in J_1$, we have

$$\sigma_n^2(f) = \int_{E_f} |p_n(z,f)|^2 d\mu_f \geqslant \int_{e_n} |p_n(z,f)|^2 d\mu_f > \sigma_n(f)\widetilde{m}_n(E_f)\mu_f(e_n),$$

implying that

$$\liminf_{n\to\infty} \sqrt[n]{\sigma_n(f)} \geq \lim_{n\to\infty} \sqrt[n]{\widetilde{m}_n(E_f)} \cdot \liminf_{n\to\infty} \sqrt[n]{\mu_f(e_n)} \geq \widetilde{\tau}(E_f)(1-\rho). \tag{4.4}$$

For $n \in J_2$, we have

$$\lim_{n\to\infty} \mu_f(e_n) = 0. \tag{4.5}$$

Since the spectral density $f(\lambda)$ is positive on

$$W^{-1}(E_f) = \{\lambda \in \Lambda : e^{i\lambda} \in E_f\},$$

the measure μ is absolutely continuous with respect measure μ_f. Taking into account that the measure μ is also finite, in view of (4.5), we conclude that

$$\lim_{n\to\infty} \mu(e_n) = 0, \quad n \in J_2. \tag{4.6}$$

For $n \in J_2$ define the sets $E_n = E_f \backslash e_n$ and observe that in view of (4.3), for $z \in E_n$ we have

$$|p_n(z,f)| \leqslant \sqrt{\sigma_n(f) \cdot \widetilde{m}_n(E_f)}. \tag{4.7}$$

Let $\varepsilon > 0$ be an arbitrary number satisfying

$$\frac{1}{(1+\varepsilon)^2} \geqslant 1-\rho, \tag{4.8}$$

and let $\delta := \delta(E_f, \varepsilon)$ be chosen according to Lemma 3.4. Then, in view of (4.6), for large enough $n \in J_2$, we have

$$\mu(E_f \backslash E_n) = \mu(e_n) < \delta(E_f, \varepsilon).$$

Therefore, in view of (4.7) and Lemma 3.4, we can write

$$\widetilde{m}_n(E_f) = \max_{z \in E_f} |\widetilde{T}_n(z, E_f)| \leqslant \max_{z \in E_f} |p_n(z, f)|$$

$$\leq (1+\varepsilon)^n \max_{z \in E_n} |p_n(z, f)| \leqslant (1+\varepsilon)^n \sqrt{\sigma_n(f) \cdot \widetilde{m}_n(E_f)}, \quad n \in J_2,$$

implying that

$$\sigma_n(f) \geqslant \frac{\widetilde{m}_n(E_f)}{(1+\varepsilon)^{2n}}, \quad n \in J_2.$$

Therefore, letting n tend to infinity, and taking into account the inequality (4.8), we obtain

$$\liminf_{n \to \infty} \sqrt[n]{\sigma_n(f)} \geqslant \widetilde{\tau}(E_f)(1-\rho). \tag{4.9}$$

Putting together (4.4) and (4.9), and taking into account arbitrariness of ρ, we obtain (4.2).

A combination of (4.1) and (4.2) implies (2.1), and thus completes the proof of the theorem. □

Proof of Theorem 2.2. Both assertions (a) and (b) of the theorem we infer from Theorem 2.1. To prove the assertion (a), observe first that since the spectral density $f(\lambda)$ is positive almost everywhere in some vicinity of zero, one can assume that $E_f \supset \Gamma_\alpha$ for some $0 < \alpha < \pi$, where $\Gamma_\alpha = \{e^{i\lambda},\ |\lambda| \leq \alpha\}$. Denoting by $f_\alpha(\lambda)$ the contraction of $f(\lambda)$ on the set $W^{-1}(\Gamma_\alpha)$:

$$f_\alpha(\lambda) = \begin{cases} f(\lambda) & \text{if} \quad |\lambda| \leq \alpha \\ 0 & \text{if} \quad |\lambda| > \alpha, \end{cases} \tag{4.10}$$

and taking into account the obvious inequality $f(\lambda) \geq f_\alpha(\lambda)$ and (3.3), we conclude that $\sigma_n(f) \geq \sigma_n(f_\alpha)$. Hence, in view of Theorem 2.1 and equality (3.17), we obtain

$$\liminf_{n \to \infty} \sqrt[n]{\sigma_n(f)} \geq \lim_{n \to \infty} \sigma_n(f_\alpha) = \widetilde{\tau}(\Gamma_\alpha) = 1. \tag{4.11}$$

Combining the relations (4.1) and (4.11) we obtain (2.2), and the assertion (a) of the theorem follows.

Now we proceed to prove the assertion (b) of the theorem. In this case, without loss of generality, we can assume that

$$E_f = \Gamma'_\alpha = \{e^{i\lambda},\ \alpha \leq |\lambda| \leq \pi\},$$

and hence, the equality (2.1) is fulfilled. So, to complete the proof, it remains to show that

$$\widetilde{\tau}(E_f) = \widetilde{\tau}(\Gamma'_\alpha) \leqslant \cos\frac{\alpha}{2}. \tag{4.12}$$

To prove (4.12), consider the following polynomial of degree n:

$$q_n(z) = \left[\frac{z+1}{2}\right]^n,$$

and observe that

$$q_n(z) \in \mathcal{Q}_n(1) \quad \text{and} \quad |q_n(e^{i\lambda})| = \left(\cos\frac{\lambda}{2}\right)^n.$$

Next, according to the definition of minimizing polynomial $\widetilde{T}_n(z)$, with $E_f = \Gamma'_\alpha$ we have

$$\widetilde{m}_n(E_f) = \max_{z\in E_f} |\widetilde{T}_n(z, E_f)| \leqslant \max_{z\in E_f} |q_n(z)| = \max_{|\lambda|\geq\alpha} |q_n(e^{i\lambda})| = \left(\cos\frac{\alpha}{2}\right)^n.$$

Taking the root of order n and passing to the limit as $n\to\infty$, we obtain (4.12). Finally, from (2.1) and (4.12) we obtain (2.3), and the assertion (b) of the theorem follows. □

Список литературы

[1] R. K. Adenstedt, "On large-sample estimation for the mean of a stationary random sequence", Ann. Stat., **2**, 1095 – 1107 (1974).

[2] R. K. Adenstedt, B. Eisenberg, "Linear estimation of regression coefficients", Quart. Appl. Math., **32** (3), 317 – 327 (1974).

[3] N. M. Babayan, "On the asymptotic behavior of prediction error", J. of Soviet Mathematics **27** (6), 3170 – 3181 (1984).

[4] N. M. Babayan, "On asymptotic behavior of the prediction error in the singular case", Theory Probab. Appl **29** (1), 147 – 150 (1985).

[5] J. Beran, "Recent developments in location estimation and regression for long-memory processes; New directions in time series analysis, part II, (D. Brillinger et al. eds.) **46**, 1 – 10 (1993).

[6] J. Beran, H. Künsch, "Location estimates for processes with long range dependence", Research Report No. 40, Seminar für Statistik, Eidenössische Technische Hochschule, Zürich (1985).

[7] J. Beran, Y. Feng, S. Ghosh, R. Kulik, Long-Memory Processes: Probabilistic Properties and Statistical Methods, Springer-Verlag, Berlin (2013).

[8] Ya. L. Geronimus, "On certain asymptotic properties of polynomials", Mat. Sb. **23 (65)** (1), 77 – 88 (1948).

[9] G. M. Goluzin, Geometric Theory of Functions of a Complex Variable. Providence, Amer. Math. Soc. (1969).

[10] U. Grenander, "Stochastic processes and statistical inference", Ark. Mat. **1** (17), 195 – 277 (1950).

[11] U. Grenander, "On Toeplitz forms and stationary processes", Ark. Mat. **1** (37), 555 – 571 (1952).

[12] U. Grenander, "On the estimation of regression coefficients in the case of an autocorrelated disturbance", Ann. Math. Stat. **25**, 252 – 272 (1954).

[13] U. Grenander, M. Rosenblatt, Statistical Analysis of Stationary Time Series. New York, Wiley (1957).

[14] U. Grenander, G. Szegö, Toeplitz Forms and Their Applications, University of California Press, Berkeley (1958).

[15] S. Mazurkievicz, "Un theoreme sur les polynomes", Ann. Soc. Polon. Math. **18**, 113 – 118 (1945).
[16] E. M. Nikishin, V. N. Sorokin, Rational Approximations and Orthogonality, Translations of Mathematical Monographs, **92**, American Mathematical Society, Providence, Rhode Island (1991).
[17] M. Rosenblatt, Some Purely Deterministic Processes, J. of Math. and Mech. **6** (6), 801 – 810 (1957).
[18] A. Samarov, M. S. Taqqu, "On the efficiency of the sample mean in long-memory noise", J. Time Series Analysis **9**, 191 – 200 (1988).
[19] R. A. Vitale, "An asymptotically efficient estimate in time series analysis", Quart. Appl. Math. **30**, 421–440 (1973).
[20] Y. Yajima, "On estimation of a regression model with long-memory stationary errors", Ann. Stat. **16**, 791 – 807 (1988).

Поступила 7 ноября 2018

После доработки 7 ноября 2018

Принята к публикации 25 апреля 2019