On detecting predictability of one-sided sequences

Nikolai Dokuchaev

August 18, 2017

Abstract

We consider predicting and predictability for discrete time processes in a pathwise setting where only a sole one-sided semi-infinite sequences of past values is available, and where statistical properties of the ensemble of the paths are unknown. We obtain some sufficient conditions of predictability in the terms of degeneracy of frequency characteristics representing one-sided modifications of the Z-transform. This characteristics are defined for one-sided semi-infinite sequences representing of the past values. Some predictors are discussed.

Keywords: One-sided sequences, predictability, forecasting, frequency analysis, Z-transform.

1 Introduction

The paper considers a problem of characterization of predictable processes in discrete time setting. It is known that this predictability is related to the degeneracy of their frequency characteristics. For example, for stochastic stationary discrete time processes, there is a criterion of predictability given by the classical Szegö-Kolmogorov Theorem. This theorem says that the optimal prediction error is zero if the spectral density ϕ is such that

$$\int_{-\pi}^{\pi} \log \phi \left(e^{i\omega} \right) d\omega = -\infty; \tag{1}$$

see [13, 24, 25] and recent literature reviews in [1, 23]. This means that a stochastic stationary process is predictable if its spectral density is vanishing with a certain rate at a point of the unit circle $\mathbb{T} = \{z \in \mathbf{C} : |z| = 1\}$. In particular, this holds if the spectral density vanishes on an arc of the unit circle, i.e. the process is bandlimited. This result was expanded on more general stochastic processes featuring spectral densities; see e.g. [3].

To apply the theory based on the spectral density, one has to collect sufficient quantity of statistical data and estimate the spectral density of the underlying processes; this density is used to construct the predictors. In practice, there are situations where statistical data are insufficient or absent. Therefore, there is a need in methods for pathwise setting oriented on data inputs being considered as sole sequences rather than members of an ensemble with a known probability distribution. This task is quite challenging. First of all, it is not obvious how to distinct predictable sequences from non-predictable. In a more general framework, there is not yet a comprehensive criterion of randomness of a sequence, and it is unclear how to detect potential randomness of a sole sequence. This problem and some related problems were intensively studied in the framework of the concept of intrinsic randomness and the problem of distinguishability of random sequences; see the references in [15, 12]. In particular, the approach from Borel (2009) [2], Mises (1919) [18], and Church (1940) [4], was based on limits of the sampling proportions of zeros in the binary sequences and subsequences; Kolmogorov (1965) [14] and Loveland (1966) [16] developed a different concept of the algorithmic randomness and compressibility; Schnorr (1971) [22] suggested an approach based on predicability and martingale properties. This paper studies predictability of sequences which is an aspect of non-randomness. It has to be clarified that it is only a particular aspect; for example, predictability is not equivalent to a possibility to recover missing values without error that can also be associated with non-randomness.

In [5, 6, 8, 9, 10, 11, 17, 19, 20], predictability was readdressed in the pathwise deterministic setting oriented on sole sequences without known statistics for similar sequences. The corresponding criteria of predictability were formulated in the terms of degeneracy of the standard two-sided Z-transform on $\mathbb{T}$, which reminded degeneracy of the spectral density in (1). This similarity seemed to be unexpected since the spectral density and Z-transform are quite different by their nature, despite the fact that they both are regarded as frequency characteristics and are used in similar frameworks involving transfer functions. However, Z-transform and spectral density share a common feature: their calculations require information about the values of the underlying sequences for arbitrarily remote future times. For Z-transform, this information includes the entire future path; for the spectral density, this information is contained in the stationarity assumption which is a probabilistic type assumptions prescribing the future evolution of the ensemble members. This setting is reasonable if there are sufficient historical observations of similar processes.

If we have to consider a sole sequence rather than a member of an ensemble of paths with known statistics, then it could be unreasonable to hypothesize about the remote future. In this situation, it is more convenient to represent data flow as one-sided sequences $\{x(t)\}_{t=0,-1,-2,..}$, with the past members of diminishing significance as $t \to -\infty$. As was mentioned above, predictability are often studied using the notion of bandlimitness or its relaxed versions such as (1) for the spectral densities. Z-transform for two-sided sequences is a symmetrical transformation around the current instant t, sym-

metrical with relation to the passed time and the future time. Unfortunately, the notion of classical bandlimitness based on the standard Z-transform cannot be applied directly to one-sided semi-infinite sequences extended by zeros into two-sided sequences, since the corresponding one-sided Z-transforms will not vanish on an arc of the unit circle. This cannot be fixed via moving the center of symmetry of Z-transform into the middle of the sample. Therefore, it is still unknown how to expand the notion of bandlimitness on the one-sided sequences $\{x(t)\}_{t=0,-1,-2,\dots,-\infty}$. This paper represents an attempt to attack this problem; as far as we know, this is the first attempt in the existing literature.

We suggest some special new frequency characteristics for one-sided semi-infinite sequences representing some modifications of Z-transforms. A degeneracy of these characteristics allows to detect predictability of certain classes of one-sided sequences. In addition, we found that the predictors introduced in [9] for sequences with degenerate standard two-sided Z-transforms can be also used for one-sided semi-infinite sequences given that the are predictable, i.e. given that some degeneracy conditions of these new special frequency characteristics are satisfied.

2   Some definitions

We denote by $\mathbf{Z}$ the set of all integers.

For $\tau \in \mathbf{Z} \cup \{+\infty\}$ and $\theta < \tau$, we denote by $\ell_r(\theta, \tau)$ a Banach space of sequences $x = \{x(t)\}_{\theta-1<t<\tau+1} \subset \mathbf{C}$, with the norm $\|x\|_{\ell_r(\theta,\tau)} = \left(\sum_{t=\theta}^{\tau} |x(t)|^r\right)^{1/r} < +\infty$ for $r \in [1,\infty)$ or $\|x\|_{\ell_\infty(\theta,\tau)} = \sup_{t:\ \theta-1<t<\tau+1} |x(t)| < +\infty$ for $r = +\infty$; the cases where $\theta = -\infty$ or $\tau = +\infty$ are not excluded. As usual, we assume that all sequences with the finite norm of this kind are included in the corresponding space.

For brevity, we will use the notations $\ell_r = \ell_r(-\infty, \infty)$, and $\ell_r^- = \ell_r(-\infty, 0)$.

Let $\mathbb{T} \triangleq \{z \in \mathbf{C} :\ |z| = 1\}$.

For $x \in \ell_2$, we denote by $X = \mathcal{Z}x$ the Z-transform

$$X(z) = \sum_{t=-\infty}^{\infty} x(t)z^{-t},$$

defined for $z \in \mathbf{C}$ such that the series converge. If $x \in \ell_1$, then X is continuous on $\mathbb{T}$, and the corresponding series converge uniformly on $\mathbb{T}$. If $x \in \ell_2$, then $X|_{\mathbb{T}}$ is defined as an element of $L_2(\mathbb{T})$, and the corresponding series converge on $\mathbb{T}$ in the norm $L_2(\mathbb{T})$.

The inverse $x = \mathcal{Z}^{-1}X$ is defined as

$$x(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X\left(e^{i\omega}\right) e^{i\omega t} d\omega, \quad t \in \mathbf{Z}.$$

In addition, we will be using the one-sided Z-transform defined as $\mathcal{Z}^- x = \sum_{t=-\infty}^{0} z^{-t}x(t)$. Again, for any $x \in \ell_2^-$, this transform converges on $\mathbb{T}$ in the norm $L_2(\mathbb{T})$.

Let $D \triangleq \{z \in \mathbf{C} : |z| < 1\}$ and $D^c \triangleq \{z \in \mathbf{C} : |z| > 1\}$. For $r \in [1, +\infty]$, let $H^r(D)$ be the Hardy space of functions that are holomorphic on D, and let $H^r(D^c)$ be the Hardy space of functions that are holomorphic on D^c including the point at infinity; see e.g. [21].

3   The main result

It appears that the notion of bandlimitness of degeneracy of the frequency characteristics is non-applicable to the one-sided Z-transforms since $X^- = \mathcal{Z}^- x^- \in H^2(D)$ for $x \in \ell_2^-$ and $x^-(t) = x(t)\mathbb{I}_{t\leq 0}$, and hence $\int_{-\pi}^{\pi} \log |X^- (e^{i\omega})| \, d\omega > -\infty$; see e.g. Theorem 17.17 from [21].

We suggest below modifications of one-sided Z-transform oriented on detecting a degeneracy of the frequency characteristics for one-sided sequences.

Definition 1 *For* $x \in \ell_2^-$*, we introduce transforms* $\xi_1 = \Xi_1 x \in L_2([0, \pi], \mathbf{C})$ *and* $\xi_2 = \Xi_2 x \in L_2([0, \pi], \mathbf{C}) \times \mathbf{C}$ *defined as*

$$\xi_1(\omega) = 2 \sum_{t=-\infty}^{-1} \cos(\omega t) x(t) + x(0),$$

$$\xi_2(\omega) = (\xi_2'(\omega), \xi_2'') = \left(2 \sum_{t=-\infty}^{-1} \sin(-\omega t) x(t), \; x(0) \right),$$

where $\omega \in [0, \pi]$.

Remark 1 *For* $x \in \ell_1^-$*, the corresponding functions* $\xi_k(\omega)$ *are continuous in* $\omega \in [0, \pi]$*. In a general case* $x \in \ell_2^-$*, the convergence of the corresponding series for* ξ_1 *and* ξ_2' *is in* $L_2([0, \pi], \mathbf{C})$.

Remark 2   *(i) It follows from the definitions that* $\mathcal{Z}^- x = \xi_1 + i\xi_2'$.

(ii) If $x \in \ell_2$ *is real valued then* $\xi_1(\omega) = \mathrm{Re}\, X (e^{i\omega})$ *and* $\xi_2'(\omega) = \mathrm{Im}\, X (e^{i\omega})$ *for* $\xi_k = \Xi_k x^-$ *and* $X = \mathcal{Z} x$, $\omega \in (-\pi, \pi]$*, where* $x_- \in \ell_2^-$ *is such that* $x^-(t) = x(t)\mathbb{I}_{t\leq 0}$*. In addition, if* x *is real and even then* $\xi_1(\omega) = \mathrm{Re}\, X (e^{i\omega}) = X (e^{i\omega})$*; if* x *is real and odd then* $\xi_2'(\omega) = \mathrm{Im}\, X (e^{i\omega}) = X (e^{i\omega})$.

Lemma 1 *The mappings* $\Xi_1 : \ell_2^- \to L_2([0, \pi], \mathbf{C})$ *and* $\Xi_2 : \ell_2^- \to L_2([0, \pi], \mathbf{C}) \times \mathbf{C}$ *are continuous bijections, and the corresponding inverse mappings are also continuous bijections such that* $x_k =$

$\Xi_k^{-1}\xi_k$ *are defined as*

$$x_1(t) = \frac{1}{\pi}\int_0^\pi \xi_1(\omega)\cos(\omega t)d\omega,$$
$$t = 0, -1, -2, ...,$$
$$x_2(t) = \frac{1}{\pi}\int_0^\pi \xi_2'(\omega)\sin(-\omega t)d\omega,, \quad t = -1, -2, ...,$$
$$x_2(0) = \xi_2'',$$
$$where \quad \xi_2 = (\xi_2'(\omega), \xi_2''). \tag{2}$$

The following corollary follows immediately from Lemma 1 and shows that our new transforms, despite being related to $\mathcal{Z}^-$ as is shown in Remark 2(i), still allows to use the notion of bandlimitiness.

Corollary 1 *For any open subset $I \subset [0, \pi]$, there exist non-zero sequences $x \in \ell_2^-$ such that either $\xi_1|_I = 0$ or $\xi_2'|_I = 0$, where $\xi_k = \Xi_k x$ and $\xi_2 = (\xi_2'(\omega), \xi_2'')$.*

Types of predicability of classes one-sided sequences

Definition 2 *Let $\widehat{\mathcal{K}}$ be the class of functions $\widehat{k} : \mathbf{Z} \to \mathbf{R}$ such that $\widehat{k}(t) = 0$ for $t < 0$ and such that $\widehat{K}(\cdot) = \mathcal{Z}\widehat{k} \in H^\infty(D^c)$.*

Definition 3 *Let $\mathcal{Y} \subset \ell_r^-$ be a class of one-sided sequences, $r \in [1, +\infty]$.*

(i) *We say that this class is unilaterally ℓ_r^--predictable if there exists a sequence $\{\widehat{k}_m(\cdot)\}_{m=1}^{+\infty} \subset \widehat{\mathcal{K}}$ such that*

$$\|\widehat{x}_m(t-1) - x(t)\|_{\ell_r(-\infty,-1)} \to 0 \quad as \quad m \to +\infty$$
$$\forall x \in \mathcal{Y}.$$

Here $\widehat{x}_m(t) \triangleq \sum_{s=-\infty}^{t} \widehat{k}_m(t-s)x(s)$.

(ii) *We say that the class $\mathcal{Y}$ is uniformly unilaterally ℓ_r^--predictable if, for any $\varepsilon > 0$, there exists $\widehat{k}(\cdot) \in \widehat{\mathcal{K}}$ such that*

$$\|\widehat{x}(t-1) - x(t)\|_{\ell_r(-\infty,-1)} \le \varepsilon \quad \forall x \in \mathcal{Y}.$$

Here $\widehat{x}(t) \triangleq \sum_{s=-\infty}^{t} \widehat{k}(t-s)x(s)$.

In particular, Definition 3 implies that $x(t)$ is uniquely defined by the values $\{x(s)\}_{s \le t-1}$ for all $t \le -1$, and can be recovered from these values using some kernels from $\widehat{K}$.

Some predictable classes

Let some $q > 1$ be given. For $c > 0$ and $\omega \in (-\pi, \pi]$, set

$$h(\omega, c) = \exp \frac{c}{[(\cos(\omega) + 1)^2 + \sin^2(\omega)]^{q/2}}.$$

Let $\mathcal{W}(c)$ be the class of all sequences $x \in \ell_2^-$ such that, for each $x \in \ell_2^-$, at least one of the following conditions holds: either there exist $a \in \mathbf{C}$ such that

$$\operatorname*{ess\,sup}_{\omega \in [0,\pi]} |\xi_1(\omega) - a| h(\omega, c) < +\infty,$$

or

$$\operatorname*{ess\,sup}_{\omega \in [0,\pi]} |\xi_2'(\omega)| h(\omega, c) < +\infty.$$

Here $\xi_k = \Xi_k x$, $\xi_2 = (\xi_2'(\omega), \xi_2'')$.

Note that $h(\omega, c) \to +\infty$ as $\omega \to \pm\pi$ and that, for $x \in \mathcal{W}(c)$, either $(\xi_1(\omega) - a)$ or $\xi_2''(\omega)$ vanishes with a sufficient rate of decay as $\omega \to \pi$.

Remark 3 *By Corollary 1, $\mathcal{W}(c) \neq \emptyset$ for any $c > 0$.*

Let $\mathcal{W} = \cup_{c>0} \mathcal{W}(c)$.

Further, for $c > 0$ and $d > 0$, let $\mathcal{V}(c, d)$ be a class of processes $x \in \mathcal{W}(c)$ such that there exists $a \in \mathbf{C}$ such that

$$\min \left(\operatorname*{ess\,sup}_{\omega \in [0,\pi]} |(\xi_1(\omega) - a) h(\omega, c)|, \operatorname*{ess\,sup}_{\omega \in [0,\pi]} |\xi_2''(\omega) h(\omega, c)| \right) \leq d, \tag{3}$$

where $\xi_k = \Xi_k x$ and $\xi_2 = (\xi_2'(\omega), \xi_2'')$.

Theorem 1 *Let either $r = 2$ or $r = +\infty$.*

(i) *The class $\mathcal{W}$ is unilaterally ℓ_r^--predictable in the sense of Definition 3(i).*

(ii) *For any given $c > 0$ and $d > 0$, the class $\mathcal{V}(c, d)$ is uniformly unilaterally ℓ_r-predictable in the sense of Definition 3(ii).*

(iii) *Prediction required in (i) and (ii) can be achieved with kernels introduced in [9]. These kernels were defined as $\widehat{k}(\cdot) = \widehat{k}(\cdot, \gamma) = \mathcal{Z}^{-1} \widehat{K}$, where*

$$\widehat{K}(z) = z \left(1 - \exp \left[-\frac{\gamma}{z + 1 - \gamma^{2\mu/(1-q)}} \right] \right). \tag{4}$$

Here $\mu > 1$, $q > 1$, and $\gamma > 0$, are parameters; the prediction error vanishes as $\gamma \to +\infty$.

The predicting kernels (4) were introduced in [9] for sequences with degenerate two-sided Z-transforms. They represent a discrete time version of the continuous time processes predicting kernels from [6]. Some numerical experiments for these predictors are described in [11].

Remark 4 *The novelty of Theorem 1 comparing with [9] is that the predicability criteria in this theorem were formulated in the terms of degeneracy of frequency characteristics defined by one-sided semi-infinite sequences; on the other hand, the predicability criteria in [9] were formulated in the terms of degeneracy of the standard two-sided Z-transform defined by two-sided infinite sequences.*

Remark 5 *Theorem 1 describes classes $W(c)$ and $\mathcal{V}(c,d)$ of predictable sequences, and gives a method of prediction of $x(t)$ based on the observations for $\{x(s)\}_{s \leq t-1}$, for $t \leq 0$ and for sequences belonging to a predictable class. Consider, for example, a scenario where $x(0)$ has to be predicted. In this case, predictability means that an agent can find, using observations $\{x(s)\}_{s \leq -1}$, an unique $x(0)$ such that the extended sequence $\{x(s)\}_{s \leq 0}$ belongs to this predicable class.*

4 Proofs

Proof of Lemma 1. Let $x \in \ell_2^-$ be fixed. We define $\widetilde{x}_k \in \ell_2$, $k = 1, 2$, such that $x_k(t) = x(t)$ for $t \leq 0$ and $\widetilde{x}_1(t) = x(-t)$, $\widetilde{x}_2(t) = -x(-t)$ for $t > 0$. For $\xi_k = \Xi_k x$ and $X_k = \mathcal{Z}\widetilde{x}_k$, it follows from the definitions that

$$X_1\left(e^{i\omega}\right) = \xi_1(|\omega|), \quad \omega \in (-\pi, \pi], \tag{5}$$

and, for $\xi_2 = (\xi_2'(\omega), \xi_2'')$,

$$X_2\left(e^{i\omega}\right) = i\xi_2'(\omega) + \xi_2'', \quad \omega \in [0, \pi],$$
$$X_2\left(e^{i\omega}\right) = -i\xi_2'(-\omega) + \xi_2'', \quad \omega \in [-\pi, 0). \tag{6}$$

Since $\widetilde{x}_k \in \ell_2$, it follows that $X_k\left(e^{i\omega}\right) \in L_2([-\pi, \pi], \mathbf{C})$. Hence $\xi_k = \Xi_k x \in L_2([0, \pi], \mathbf{C})$, and the mappings $\Xi_1 : \ell_2^- \to L_2([0, \pi], \mathbf{C})$ and $\Xi_2 : \ell_2^- \to L_2([0, \pi], \mathbf{C}) \times \mathbf{C}$ are continuous.

Further, let $\xi_1 \in L_2([0, \pi], \mathbf{C})$ and $\xi_2 = (\xi_2'(\omega), \xi_2'') \in L_2([0, \pi], \mathbf{C}) \times \mathbf{C}$ be fixed. Let us define mappings $X_k : \mathbb{T} \to \mathbf{C}$ by (5) and (6). Let $\widetilde{x}_k = \mathcal{Z}^{-1} X_k$, i.e.

$$\widetilde{x}_k(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X_k\left(e^{i\omega}\right) e^{i\omega t} d\omega, \quad t = 0, \pm 1, \pm 2, \dots.$$

It can be verified immediately that $\widetilde{x}_1(t) = \widetilde{x}_1(-t)$, $\widetilde{x}_2(t) = -\widetilde{x}_2(-t)$, $t > 0$, and that (2) holds for $x_k \stackrel{\triangle}{=} \widehat{x}_k|_{\{t \leq 0\}}$. It follows that $\Xi_k x_k = \xi_k$. Hence the mappings $\Xi_1 : \ell_2^- \to L_2([0, \pi], \mathbf{C})$ and $\Xi_2 : \ell_2^- \to L_2([0, \pi], \mathbf{C}) \times \mathbf{C}$ are bijections. It is known that an inverse of a continuous bijection between Banach spaces is also continuous. This completes the proof of Lemma 1. $\square$

Proof of Theorem 1. Let $x \in \mathcal{W}$ and $\xi_k = \Xi_k x$, and let $\widetilde{x}_k$ and X_k be such as defined in the proof of Lemma 1 above; in particular, we assume that (5)-(6) hold.

Let us prove statement (i). Let $x \in \mathcal{W}(c)$ be given for some $c > 0$. If x is such that the condition for $\xi_1(\omega)$ in the definition of $W(c)$ holds, we set $k = 1$ and memorize the corresponding $a \in \mathbf{C}$. Otherwise, we set $k = 2$ and select $a = x(0)$; this happens if the definition of $\mathcal{W}(c)$ holds for x with the condition for $\xi_2''(\omega)$ only. Consider a process $x^a(t)$ such that $x^a(t) = x(t)$ for $t < 0$ and $x^a(0) = x(0) - a$. Let $\xi_k^a = \Xi_k x^a$. By the definitions, it follows that $\mathrm{ess\,sup}_{\omega \in [0,\pi]} |\zeta_k^a(\omega)| h(\omega, c) < +\infty$, where $\zeta_1^a = \xi_1$ and $\zeta_2^a = \xi_2'$, for $\xi_2 = (\xi_2'(\omega), \xi_2'')$.

Let $\widetilde{x}_k^a(t) \in \ell_r$ be defined similarly to $\widetilde{x}_k$ in the proof of Lemma 1, with x replaced by x^a. By the definitions, $\widetilde{x}_k^a$ belongs to the class $\mathcal{X} \subset \ell_r$ used in Theorem 1 (i) [9]. By this theorem, the class $\mathcal{X}$ is ℓ_r-predictable in the sense of Definition 2(i) [9]. We have that $\widetilde{x}(t) = x_k^a(t)$ for $x \in \mathcal{W}$, $t < 0$ and $x(0) = x_k^a(0) + a$ for $x \in \mathcal{W}$. Since a is defined by the observations of the one-sided sequence $\{x(t)\}_{t \leq 0}$, it follows that the predicability of $\mathcal{X}$ in the sense of Definition 2(i) [9] implies predictability of $\mathcal{W}$ in the sense of Definition 3(i).

Let us prove statement (ii) following the same steps. Let $x \in \mathcal{V}(c, d)$. It suffices to consider $d = 1$ only. If the definition of $\mathcal{V}(c, 1)$ holds for this x with the minimum in (3) achieved for $\xi_1(\omega)$, we set $k = 1$ and select $a \in \mathbf{C}$ to be the corresponding a. Otherwise, we set $k = 2$ and $a = x(0)$; this happens if the minimum in (3) is achieved for $\xi_2''(\omega)$ only.

Let x^a, $\xi_k^a = \Xi_k x^a$, and $\widetilde{x}_k^a(t) \in \ell_r$, be defined as in the proof for the statement (i) above.

By the definitions, $\widetilde{x}_k^a$ belong to the class $\mathcal{U}(c)$ defined in Theorem 1 (ii) [9]. The class $\mathcal{U}(c) \subset \ell_r$ is uniformly ℓ_r-predictable in the sense of Definition 2(ii) [9]. By the definitions, $x(t) = \widetilde{x}_k^a(t)$ for $t < 0$, and $x(0) = \widetilde{x}_k^a(0) + a$ for $x \in \mathcal{V}(c, 1)$. Again, since a is defined by the observations of the one-sided sequence $\{x(t)\}_{t \leq 0}$, it follows that the predicability for $\mathcal{U}(c_0)$ in the sense of Definition 2(ii) [9] implies predictability in the sense of Definition 3(ii) for $\mathcal{V}(c, 1)$. This completes the proof of Theorem 1. $\square$

5 Discussion and future developments

(i) The present paper considers one-sided sequences with certain one-sided frequency characteristics vanishing at a point of the unit circle. Similar problems arise for a setting where one-sided frequency characteristics are supposed to vanish on an arc of the unit circle, in the spirit of the classical bandlimitness. Again, it is unclear how to extend the notion of bandlimitness on one-sided sequences since the classical one-sided Z-transforms cannot vanish on an arc of the unit circle. Since the extensions on $t > 0$ are uniquely defined by the observations $\{x(t)\}_{t \leq 0}$ for predictable two-sided band-limited sequences [8], it follows that there are two types of one-sided semi-infinite sequences $\{x(t)\}_{t \leq 0}$: the ones representing the history of band-limited processes,

and the ones representing non-band-limited processes. The following problem arises: *Assume that the values $\{x(t)\}_{t\leq 0}$ are observed. How to determine if there exists a sequence $\{x(t)\}_{t>0}$ such that the sequence $\{x(t)\}_{t\in\mathbf{Z}}$ is band-limited?* We leave this for the future research; so far, the present papers gives only sufficient conditions of a related property described via degeneracy of certain one-sided frequency characteristics. A different approach to this problem was suggested in [7, 10]

(ii) It is possible to consider real valued sequences, with small modification of the definitions; more precisely, the spaces ℓ_r have to be defined as spaces of real valued sequences. In this case, ξ_1, ξ_2', and ξ_2'' will be real-valued, and a in the definition of $W(c)$ has to be selected among real numbers.

(iii) A straightforward modification leads to similar versions of the transforms Ξ_k for sequences from $\ell_r(0, +\infty)$.

Acknowledgment

References

[1] N. H. Bingham. (2012). Szegö's theorem and its probabilistic descendants. *Probability Surveys* 9, 287-324.

[2] E. Borel. (1909). Les probabilites denombrables et leurs applications arithmetique. *Rend. Circ. Mat. Palermo* 27, 247–271.

[3] S. Cambanis and A.R. Soltani. (1984). Prediction of stable processes: spectral and moving average representations. *Z. Wahrsch. Verw. Gebiete* 66, no. 4, 593–612.

[4] A. Church. (1940). On the concept of random sequence. *Bull. Amer.Math. Soc.* 46, 254–260.

[5] N. Dokuchaev. (2008). The predictability of band-limited, high-frequency, and mixed processes in the presence of ideal low-pass filters. *Journal of Physics A: Mathematical and Theoretical* **41**, No 38, 382002. (7pp).

[6] N. Dokuchaev. (2010). Predictability on finite horizon for processes with exponential decrease of energy on higher frequencies. *Signal Processing* **90**, Iss. 2, 696–701.

[7] N. Dokuchaev. (2012). On one sided band-limitness and extrapolation of sequences. Working paper. ArXiv: https://arxiv.org/abs/1208.3278

[8] N. Dokuchaev. (2012). On predictors for band-limited and high-frequency time series, *Signal Processing* **92**, iss. 10, pp. 2571–2575.

[9] N. Dokuchaev. (2012). Predictors for discrete time processes with energy decay on higher frequencies. *IEEE Transactions on Signal Processing* **60**, No. 11, 6027-6030.

[10] N. Dokuchaev. (2013). Forecasting for discrete time processes based on causal band-limited approximation. ICORES 2013. In: *Proc. 2nd International Conference on Operations Research and Enterprise Systems. Barcelona, Spain. 16-18 February, 2013.* Ed. B. Vitoriano and F.Valente, pp. 81-85.

[11] N. Dokuchaev. (2016). Near-ideal causal smoothing filters for the real sequences. *Signal Processing* **118**, iss. 1, pp. 285-293.

[12] R. Downey.(2004). Some Recent Progress in Algorithmic Randomness. In: *"Mathematical foundations of computer science". Lect. Not .Comp. Sci* 3153, 42–83.

[13] A.N. Kolmogorov. (1941). Interpolation and extrapolation of stationary stochastic series. *Izv. Akad. Nauk SSSR Ser. Mat.,* 5:1, 3–14.

[14] A.N. Kolmogorov. (1965). Three approaches to the quantitative definition of information. *Problems of Information and Transmission*, 1:1, 1–7.

[15] M. Li, P.B.M. Vitanyi. (1993). *An introduction to Kolmogorov complexity and its applications*, Springer.

[16] D.M. Loveland. (1966). A new interpretation of von Mises' concept of random sequence. *Z. Math. Logik Grundlagen Math* **12**, 279–294.

[17] A.G. Miamee, and M. Pourahmadi. (1988). Best approximations in $L^p(d\mu)$ and prediction problems of Szegö, Kolmogorov, Yaglom, and Nakazi. *J. London Math. Soc.* 38, no. 1, 133–145.

[18] R. von Mises. (1919). Grundlagen der Wahrscheinlichkeitsrechnung. *Math. Z.* 5, 52–99.

[19] T. Nakazi, Two problems in prediction theory.(1984). *Studia Math.* 78, no. 1, 7–14.

[20] T. Nakazi and K. Takahashi. (1980). Prediction n units of time ahead. *Proc. Amer. Math. Soc.* 80, no. 4, 658–659.

[21] W. Rudin. (1987). *Real and Complex Analysis*. 3rd ed. Boston: McGraw-Hill, 1987.

[22] C.P. Schnorr. (1971). A unified approach to the definition of a random sequence. Mathematical Systems Theory 5 (3): 246–258.

[23] B. Simon. (2011). Szegö's Theorem and Its Descendants. Spectral Theory for L^2 Perturbations of Orthogonal Polynomials. *M.B. Porter Lectures.* Princeton University Press, Princeton.

[24] G. Szegö. (1920). Beiträge zur Theorie der Toeplitzschen Formen. *Math. Z.* 6, 1920, 167–202. Part II: *Math. Z.* 9, 1921, 167–190.

[25] S. Verblunsky. (1936). On positive harmonic functions (second paper). *Proc. London Math. Soc.* 40, 290–320.