Интеллектуальные системы управления, анализ данных

© 2025 г. Ю.С. ПОПКОВ, д-р техн. наук (popkov@isa.ru)
(Федеральный исследовательский центр
"Информатика и управление" РАН, Москва;
Институт проблем управления им. В.А. Трапезникова РАН, Москва),
А.Ю. ПОПКОВ, канд. техн. наук (apopkov@isa.ru),
Ю.А. ДУБНОВ (yury.dubnov@phystech.edu)
(Федеральный исследовательский центр
"Информатика и управление" РАН, Москва)

МЕТОДЫ РАНДОМИЗИРОВАННОГО МАШИННОГО ОБУЧЕНИЯ ДЛЯ ГЕНЕРАЦИИ АНСАМБЛЕЙ СЛУЧАЙНЫХ ДАННЫХ С ЗАДАННЫМИ ЧИСЛОВЫМИ ХАРАКТЕРИСТИКАМИ[1]

Рассматривается задача генерации случайных ансамблей данных с заданными числовыми характеристиками. Развивается метод ее решения, использующий процедуры рандомизированного машинного обучения, которые строятся на последовательности задач функционального энтропийно-линейного программирования. В качестве ограничений в них рассматриваются нормированные моменты. Задача генерации сводится к системе нелинейных уравнений с интегральными компонентами. Адаптируется разработанный авторами асимптотический аналитический метод преобразования указанных уравнений к системе уравнений с полиномиальной левой частью. Развитые аналитические методы применены для генерации случайных ансамблей данных, прогнозирующих динамику стоимости финансовых активов.

Ключевые слова: энтропия, рандомизация, нормированные моменты, случайные ансамбли, машинное обучение, степенные ряды, аналитические функции, многомерные интегралы, многомерные полиномы.

DOI: 10.31857/S0005231025070065, **EDN:** JSAVQB

1. Введение

Методы машинного обучения, использующиеся для решения большого количества практических и научно-практических задач, основаны на использовании данных с требуемыми свойствами. В контексте методов машинного обучения, основанных на статистическом подходе, требования к данным реализуются в виде их вероятностных и числовых характеристик [1]. Помимо

[1] Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ (проект № 075-15-2024-544).

традиционных задач классификации и прогнозирования [2–7], можно также отметить ряд таких областей, как, например, тестирование программного обеспечения [8–10] и контроля знаний [11–13], где данные либо служат для обучения моделей процессов, либо используются для статистического оценивания гипотез. Параллельно развивается так называемый «сценарный подход», в рамках которого для параметризованной модели процесса, чаще всего экспертно, составляются массивы параметров и генерируются соответствующие им ансамбли данных [14, 15]. В любом случае свойства данных должны учитываться как для корректного применения соответствующих теоретических методов и подходов, так и при практическом применении разработанных и обученных моделей.

Необходимо также отметить, что в современных условиях во многих областях недостатка в данных нет, так как они в избытке собираются и накапливаются в автоматическом режиме при функционировании информационных систем и различных технических устройств. Но в то же время по-прежнему существует проблема генерации необходимых данных (с требуемыми свойствами) в целях разработки, обучения и тестирования как методов, так и все тех же устройств. Конечно, в различных областях и задачах требуемые свойства могут быть формализованы различными способами. В данной работе такая формализация проводится в рамках вероятностной концепции, в частности, под данными понимаются выборки, обладающие подходящими функциями плотности распределениями вероятностей (ПРВ). При этом предполагается, что подходящие распределения можно тем или иным способом сэмплировать, т.е. трансформировать в соответствующие случайные последовательности.

Таким образом, генерирование подходящих данных основано на восстановлении с учетом заданных требований, оптимальных в принятом смысле функций ПРВ. Поскольку при формулировке указанных требований имеет место довольно большая неопределенность, то естественным критерием оптимизации ПРВ является информационная энтропия [16–20]. Однако этого недостаточно, и часто требуется учитывать какие-то дополнительные свойства функций ПРВ. Некоторые из них можно сформулировать в терминах числовых характеристик, а именно, моментов, семиинвариантов и др. Следовательно, задача генерирования желательных функций ПРВ сводится к условной максимизации функционала информационной энтропии. По своему формальному представлению эта задача близка к некоторым математическим моделям рандомизированного машинного обучения (РМО), изучаемым в [21, 22]. Определенные отличия связаны с системой ограничений в рассматриваемой задаче.

Настоящая работа направлена на дальнейшее развитие методов рандомизированного машинного обучения в следующих направлениях:

- рандомизированное обучение при дополнительных ограничениях моментного типа;

- адаптация аналитического метода решения нелинейных уравнений с интегральными компонентами к рассматриваемой задаче;
- исследование эффективности развитых методов на примере задачи прогнозирования стоимости финансовых активов.

Отметим, что задачи РМО и задачи генерирования желаемых распределений сводятся к решению существенно нелинейных уравнений, содержащих так называемые интегральные компоненты. Они представляют собой многомерные интегралы на простых множествах (параллелепипедах) с экспоненциальными подынтегральными функциями, параметризованные множителями Лагранжа. С использованием аналитических свойств экспоненциальных функций указанные интегралы аппроксимируются параметризованными интегралами от многомерных полиномов, вычисление которых выполняется аналитически.

В результате указанных трансформаций многомерных параметризованных интегралов система нелинейных уравнений аппроксимируется системой уравнений с полиномиальной левой частью. Для их решения применяется аналитический метод, основанный на представлении решения в виде абстрактного степенного ряда [23, 24].

С учетом изложенного, статья организована следующим образом. В разделе 2 формулируется общая задача, на решение которой направлен предлагаемый метод, в разделе 3 приводится описание подхода к ее решению и необходимый теоретический инструментарий, раздел 4 посвящен результатам решения задачи прогнозирования стоимости финансового актива, в разделе 5 обсуждаются особенности полученных результатов и направления дальнейших исследований, раздел 6 посвящен концентрированному обзору результатов статьи.

2. Постановка задачи

Рассмотрим случайную последовательность $u[n]$, где $n \in \mathcal{N} = \overline{1, N}$, для которой задана матрица числовых характеристик $Y_{(s\times N)}$, элементы которой характеризуют значения *нормированных моментов*[2] порядка $k = \overline{1, s}$ в точках наблюдения n:

$$Y_{(s\times N)} = \left\{ \left(\mathcal{M}\{u^k[n]\} \right)^{1/k} \right\} = \left\{ y^k[n] \,|\, k = \overline{1, s}, n = \overline{1, N} \right\}. \tag{1}$$

В частности, такого рода информация часто встречается в задачах прогнозирования ценообразования торгуемых финансовых инструментов [26–28].

Задачу генерации данных с заданными свойствами можно сформулировать следующим образом.

[2] Здесь в качестве числовых характеристик рассматриваются нормированные моменты, но также могут быть избраны семиинварианты или математические ожидания непрерывных функций от случайных последовательностей

В каждой точке n наблюдения требуется генерировать ансамбли (наборы) $\mathcal{Z}_n$ случайных последовательностей $z[n]$ с s нормированными моментами

$$m^{(k)}[n] = \left(\mathcal{M}\{z^k[n]\}\right)^{1/k}, \quad k = \overline{1,s}, \tag{2}$$

равными заданным нормированным моментам $y^k[n]$, определяемым (1).

Генератором ансамбля $\mathcal{Z}_n$ является *«вход–выход»-модель* или *«авто»-модель*[3]. В том и другом варианте параметры $\mathbf{a} \in \mathcal{A} \subset R^r$ в моделях случайные, интервальные:

$$\mathbf{a} \in \mathcal{A} \subset R^r, \quad \mathcal{A} = \left[\mathbf{a}^-, \mathbf{a}^+\right], \tag{3}$$

и характеризуются для каждого момента времени n функцией ПРВ $P_{(n)}(\mathbf{a})$, которая предполагается непрерывно дифференцируемой.

В зависимости от наличия априорной информации о происхождении данных используются либо статическая, либо динамическая «вход–выход»-модель.

Статическая модель – генератор случайных последовательностей $z[n]$ – характеризуется нелинейной дифференцируемой функцией φ с параметрами $\mathbf{a}$:

$$z[n|\mathbf{a}] = \varphi\left(\mathbf{x}[n] \,|\, \mathbf{a}\right), \quad n = \overline{1,N}, \tag{4}$$

где $\mathbf{x}[n] = \{x_1[n], \ldots, x_m[n]\}$ является входом модели, а z – выходом.

В дальнейшем символ «|» означает, что для каждого момента времени n реализуются случайные параметры $\mathbf{a}$ с ПРВ $P_{(n)}(\mathbf{a})$.

Динамическая модель характеризуется нелинейным непрерывным функционалом $\mathcal{B}$:

$$z[n|\mathbf{a}] = \mathcal{B}\left[\mathbf{x}[\tau],\, n-p \le \tau \le n \,|\, \mathbf{a}\right], \quad n = \overline{1,N}, \tag{5}$$

где p – «память» модели, определяющая количество прошлых значений входа, влияющих на выход в данный момент времени. Параметры $\mathbf{a}$ также случайные и интервальные, определяемые (3).

Сформулированная выше задача решается в два этапа. На первом этапе определяются оптимальные вероятностные характеристики случайных параметров, а именно, их функции ПРВ $P_{(n)}(\mathbf{a})$ для всех $n = \overline{1,N}$ при использовании статической модели (4) и для всех $n = \overline{1-p,N}$ при использовании динамической модели (5). Второй этап состоит в сэмплировании этих ПРВ, т.е. в трансформации их в соответствующие ансамбли $\mathcal{Z}_n$.

[3] Здесь будет использоваться «вход–выход»-модель, в примере — «авто»-модель, описываемая в виде разностных уравнений.

3. Материалы и методы

3.1. Оптимизация функций ПРВ параметров модели

Для решения задачи первого этапа воспользуемся методологией рандомизированного машинного обучения [22], в которой генерация указанных последовательностей осуществляется моделью типа «вход–выход» с оптимизированными по критерию информационной энтропии случайными параметрами.

Для удобства дальнейшего изложения введем следующие обозначения:

- вектор заданных нормированных моментов для каждого n

$$\mathbf{y}^{(n)} = \{y_1[n], \ldots, y_s[n]\}; \tag{6}$$

- вектор выхода модели для каждого n

$$\mathbf{z}^{(n)}(\mathbf{a}) = \{z_1[n \mid \mathbf{a}], \ldots, z_s[n \mid \mathbf{a}]\}; \tag{7}$$

- вектор нормированных моментов на выходе модели для каждого n

$$\mathbf{m}^{(n)} = \left\{ \int\limits_{\mathcal{A}} P_{(n)}(\mathbf{a})\, z[n \mid \mathbf{a}] d\mathbf{a}, \ldots, \left(\int\limits_{\mathcal{A}} P_{(n)}(\mathbf{a})\, z^s[n \mid \mathbf{a}] d\mathbf{a} \right)^{1/s} \right\}. \tag{8}$$

Согласно [22], задача определения оптимальных ПРВ $P_{(n)}(\mathbf{a})$ может быть сформулирована для каждого n в виде максимизации функционала информационной энтропии:

$$\mathcal{H}_{(n)}[P_{(n)}(\mathbf{a})] = -\int\limits_{\mathcal{A}} P_{(n)}(\mathbf{a}) \ln P_{(n)}(\mathbf{a}) d\mathbf{a} \Rightarrow \max \tag{9}$$

при ограничениях:

- нормировки

$$\int\limits_{\mathcal{A}} P_{(n)}(\mathbf{a})\, d\mathbf{a} = 1, \tag{10}$$

- балансов нормированных моментов выхода модели с данными

$$\mathbf{m}^{(n)} = \mathbf{y}^{(n)}. \tag{11}$$

Полагая, что функции ПРВ непрерывно-дифференцируемые, решение, параметризованное множителями Лагранжа $\boldsymbol{\lambda}^{(n)} = \left\{\lambda_1^{(n)}, \ldots, \lambda_s^{(n)}\right\}$, имеет вид (см. [22])

$$P^*_{(n)}(\mathbf{a}) = \frac{\exp\left(-\langle \boldsymbol{\lambda}^{(n)}, \mathbf{z}^{(n)}(\mathbf{a})\rangle\right)}{\mathbb{P}_{(n)}(\boldsymbol{\lambda}^{(n)})}, \tag{12}$$

где

$$\mathbb{P}_{(n)}(\boldsymbol{\lambda}^{(n)}) = \int\limits_{\mathcal{A}} \exp\left(-\langle \boldsymbol{\lambda}^{(n)}, \mathbf{z}^{(n)}(\mathbf{a})\rangle\right)\, d\mathbf{a}, \tag{13}$$

$\langle \bullet, \bullet \rangle$ обозначает скалярное произведение.

Множители Лагранжа $\boldsymbol{\lambda}^{(n)}$ определяются следующей системой из s уравнений, называемыми *балансовыми*:

$$\int\limits_{A} \exp\left(-\langle \boldsymbol{\lambda}^{(n)}, \mathbf{z}^{(n)}(\mathbf{a})\rangle\right)\, \left[\mathbf{z}^{(n)}(\mathbf{a}) - \mathbf{y}^{(n)}\right]\, d\mathbf{a} = \mathbf{0}. \tag{14}$$

Эта система уравнений определяет векторы множителей Лагранжа $\boldsymbol{\lambda}^{(n)}$ для каждого моменте n наблюдений из интервала $[1, N]$.

Задача второго этапа, а именно трансформация энтропийно-оптимальной функции ПРВ в соответствующую случайную последвательность, может быть реализована методами, изложенными в [25].

3.2. Аналитический метод решения балансовых уравнений

Из уравнения (14) следует, что для определения множителей Лагранжа требуется вычисление многомерных интегралов, а затем решение сформированных нелинейных уравнений. Разработанный аналитический метод позволяет объединить эти оба этапа.

В рассматриваемой задаче существуют некоторые полезные свойства, которыми можно воспользоваться для формирования приближенного аналитического метода. К ним относятся простая область определения интегралов, а именно, параллелепипед, и подынтегральные функции – экспоненты от непрерывных функций.

Заметим, что экспоненциальная функция аналитическая, а множители Лагранжа и выход модели ограничены. Поэтому имеет место полиномиальная аппроксимация степени q следующего вида:

$$\exp(-v_{(n)}) = \sum_{h=0}^{q} \frac{(-1)^h}{h!}\, v_{(n)}^h, \tag{15}$$

где

$$v_{(n)} = \langle \boldsymbol{\lambda}^{(n)}, \mathbf{z}^{(n)}(\mathbf{a})\rangle < \pm M < \infty. \tag{16}$$

Здесь [4]

$$(v_{(n)})^h = \sum_{i_j\geq 0;\,\sum_{j=1}^{s} i_j = h}^{s} A^{(h)}_{i_1,\ldots,i_h} (\lambda^{(n)}_1)^{i_1}\cdots(\lambda^{(n)}_s)^{i_h} \times \times (z^1[n\,|\,\mathbf{a}])^{i_1}\cdots(z^s[n\,|\,\mathbf{a}])^{i_h}, \tag{17}$$

$$A^{(h)}_{i_1,\ldots,i_h} = \frac{h!}{i_1!\cdots i_h!}. \tag{18}$$

Тогда с использованием этих обозначений система (14) примет следующий вид:

$$\sum_{h=0}^{q} \frac{(-1)^h}{h!} \sum_{(i_1,\ldots,i_h)=1}^{s} \lambda^{(n)}_{i_1}\cdots\lambda^{(n)}_{i_h}\, u^{(n)}_{i_1,\ldots,i_h;k} - v^k_n = 0, \tag{19}$$

где

$$u^{(n)}_{i_1,\ldots,i_h;k} = \int\limits_{\mathcal{A}} z^{i_1}[n\,|\,\mathbf{a}]\cdots z^{i_q}[n\,|\,\mathbf{a}]\left(z^k[n\,|\,\mathbf{a}] - y^{(k)}[n]\right) d\mathbf{a}, \tag{20}$$

$$i_1,\ldots,i_h,k = \overline{1,s}.$$

Здесь индексы $i_1,\ldots,i_h$ принимают значения в интервале $[1,s]$. Эта система содержит s переменных – множителей Лагранжа, и каждое уравнение представляет собой многомерный полином степени q.

Поскольку предполагается, что модель характеризуется непрерывной функцией (или функционалом), то она представима многомерным полиномом, и, следовательно, многомерные интегралы в (20) трансформируются в произведение одномерных интегралов, вычисляемых аналитически. Поэтому балансовая система в виде (19) имеет полиномиальную левую часть.

Рассмотрим семейство по параметру $\varepsilon \in [0,1]$ систем следующего вида:

$$-\sum_{i_1=1}^{s} \lambda^{(n)}_{i_1}\, u^{(n)}_{i_1,k} + \varepsilon \sum_{h=1}^{q} \frac{(-1)^h}{h!} \sum_{\substack{(i_1,\ldots,i_h)=1 \\ \sum_{j=1}^{h} i_j = h}}^{s} \lambda^{(n)}_{i_1}\cdots\lambda^{(n)}_{i_h}\, u^{(n)}_{i_2,\ldots,i_h;k} - v^k_n = 0. \tag{21}$$

При $\varepsilon = 0$ имеем так называемую *базовую* систему, линейную с квадратной матрицей $U^{(n)} = [u^n_{i_1,k},\,|\,(i_1,k) = \overline{1,s}]$:

$$U^{(n)}\boldsymbol{\lambda}^{(n)} = -\mathbf{v}_n, \tag{22}$$

где $\mathbf{v}_n = \{v^1_n,\ldots,v^s_n\}$.

[4] Выражение в правой части равенства (17) представляет собой полиномы Ньютона степени h [31].

Если $\det U^{(n)} \neq 0$, то множители Лагранжа, образующие *базовое* решение, равны

$$\boldsymbol{\lambda}_{(\bullet)}^{(n)} = -[U^{(n)}]^{-1}\mathbf{v}_n. \tag{23}$$

Представим решение системы (21) в виде абстрактного степенного ряда по параметру ε [23, 24]:

$$\lambda_{k,\star}^{(n)} = \lambda_{k,\bullet}^{(n)} + \varepsilon\lambda_{k,I}^{(n)} + \varepsilon^2\lambda_{k,II}^{(n)} + \cdots, \quad k = \overline{1,s}, \tag{24}$$

где $\lambda_{k,I}^{(n)}, \lambda_{k,II}^{(n)}, \ldots$ – I-ая, II-ая, ... коррекции базового решения.

Для последовательного определения коррекций, используя метод неопределенных коэффициентов [23], будем иметь вектор первой коррекции

$${}^*\boldsymbol{\lambda}_I^{(n)} = -[U^{(n)}]^{-1}\,\mathbf{b}_{(1)}^{(n)}(\boldsymbol{\lambda}_{(\bullet)}^{(n)}), \tag{25}$$

где

$$\mathbf{b}_{(I)}^{(n)}(\boldsymbol{\lambda}_{(\bullet)}^{(n)}) = \left\{ \frac{1}{2}\sum_{(i_1,i_2)=1}^{s} \lambda_{i_1,\bullet}^{(n)}\lambda_{i_2,\bullet}^{(n)}u_{i_1,i_2,1}^{(n)}, \ldots, \frac{1}{2}\sum_{(i_1,i_2)=1}^{s} \lambda_{i_1,\bullet}^{(n)}\lambda_{i_2,\bullet}^{(n)}u_{i_1,i_2,s}^{(n)} \right\}; \tag{26}$$

вектор второй коррекции

$${}^*\boldsymbol{\lambda}_{(II)}^{n} = -[U^{(n)}]^{-1}\,\mathbf{b}_{(II)}^{(n)}(\boldsymbol{\lambda}_{(\bullet)}^{n}, \boldsymbol{\lambda}_{(I)}^{n}), \tag{27}$$

где

$$\left\{ \frac{1}{2}\sum_{(i_1,i_2)=1}^{s} \lambda_{i_1,I}^{(n)}\lambda_{i_2,\bullet}^{(n)}u_{i_1,i_2,k}^{(n)} + \frac{1}{3!}\sum_{(i_1,i_2,i_3)=1}^{s} \lambda_{i_1,\bullet}^{(n)}\lambda_{i_2,\bullet}^{(n)}\lambda_{i_3,\bullet}^{(n)}u_{i_1,i_2,i_3,k}^{(n)} \right\}. \tag{28}$$

Таким образом, решение балансовой системы (21)представимо в виде

$$\boldsymbol{\lambda}_{\star}^{(n)} = \boldsymbol{\lambda}_{(\bullet)}^{(n)} + {}^*\boldsymbol{\lambda}_I^{(n)} + {}^*\boldsymbol{\lambda}_{II}^{(n)} + \cdots. \tag{29}$$

3.3. Модель ценообразования финансового актива

Рассмотрим процесс формирования стоимости финансового актива в процессе торговых сессий и применим для его прогнозирования развиваемый выше метод генерации случайных данных с заданными числовыми характеристиками.

Существует теоретический консенсус, что цена финансового инструмента в каждый момент времени является продуктом балансировки реального спроса и реального предложения. Однако трейдер ориентируется на ожидаемые спрос и предложение, которые могут существенно отличаться от реальных. В этих условиях важное значение приобретает прогноз стоимости при достаточно высоком уровне неопределенности.

Поэтому кажется естественной попытка максимимзации информационной энтропии как меры неопределенности на обучающих ретроспективных данных, содержащих значения средних и дисперсий стоимости.

3.3.1. Модель динамики стоимости

Рассмотрим автономную модель в виде линейного разностного уравнения порядка p со случайными параметрами интервального типа (интервалы $[d, w]$ всех параметров одинаковые):

$$C[t] = \sum_{i=1}^{p} a_i C[t-i], \quad t \in \mathcal{T}, \quad a_i \in \mathcal{A}_i = [d, w], \ \mathcal{A}^p = \prod_{i=1}^{p} \mathcal{A}_i. \tag{30}$$

Этой моделью будем пользоваться на этапах *обучения* $\mathcal{T}_{ln} = [t_0, t_0 + p]$ и *прогнозирования* $\mathcal{T}_{pr} = [t_0 + p + 1, t_0 + p + 1 + t_{pr}]$, где t_{pr} — одно- или двухдневные прогнозы (по аналогии могут быть рассмотрены и более длинные прогнозы).

Вероятностные свойства параметров характеризуются непрерывно дифференцированной функцией ПРВ $P_t(\mathbf{a})$.

Предполагается, что по результатам реальных торгов в каждой сессии t формируются два стоимостных индикаторов: средняя цена $m^*_C[t]$ и второй момент $D^*_C[t]$ цены в качестве характеристики ее волатильности:

$$D^*_C[t] = (V^*_C[t])^2 + (m^*_C[t])^2, \tag{31}$$

где $V^*_C[t]$ – оценка средне-квадратического отклонения цены, которая строится из величины максимального и минимального ее отклонения.

3.3.2. Данные

Пусть в реальных торгах на интервале $[t_0 - p, t_0 - 1]$ имеются данные об изменении средней стоимости:

$$m^*_C[t_0 - p], m^*_C[t_0 - p + 1], \ldots, m^*_C[t_0 - 1]$$

и оценке второго момента:

$$D^*_C[t_0 - p], D^*_C[t_0 - p + 1], \ldots, D^*_C[t_0 - 1].$$

Используя эти данные, получим значения стоимости на интервале обучения $\mathcal{T}_{ln}$, которые *генерируются моделью* (30) по данным на «прошлом» интервале $[t_0 - p, t_0 - 1]$:

$$\mathbb{C}[\mathbf{a} \,|\, t] = \sum_{i=1}^{p} a_i m^*_C[t-i], \quad t \in \mathcal{T}_{ln}. \tag{32}$$

3.3.3. Энтропийно-оптимальное оценивание функций ПРВ параметров модели

Для оптимизации функции ПРВ $P_t(\mathbf{a})$ воспользуемся методикой, развитой для рандомизированного машинного обучения [22] и изложенной в (9)–(11),

в рамках которой решаются следующие задачи для каждой торговой сессии $t \in \mathcal{T}_{ln}$:

$$\mathcal{H}_t[P_t(\mathbf{a})] = -\int\limits_{\mathcal{A}^p} P_t(\mathbf{a})\, ln P_t(\mathbf{a})\, d\mathbf{a} \Rightarrow \max_{P(\mathbf{a})} \tag{33}$$

при ограничениях:

$$\int\limits_{\mathcal{A}^p} P_t(\mathbf{a})\, d\mathbf{a} = 1, \tag{34}$$

$$\int\limits_{\mathcal{A}^p} P_t(\mathbf{a})\, \mathbb{C}[\mathbf{a}\,|\,t]\, d\mathbf{a} = m^*_C[t], \quad \int\limits_{\mathcal{A}} P_t(\mathbf{a}) \mathbb{C}^2[\mathbf{a}\,|\,t]\, d\mathbf{a} = D^*_C[t], \ t \in \mathcal{T}_{ln}, \tag{35}$$

где $\mathbb{C}_t[\mathbf{a}\,|\,t]$ определяются равенствами (32).

Задача (33)–(35) согласно (12), (13) имеет аналитическое решение:

$$\begin{aligned} P^*_t(\mathbf{a}) &= \frac{\exp\left(-\lambda^{(t)}_1 \mathbb{C}[\mathbf{a}\,|\,t] - \lambda^{(t)}_2 \mathbb{C}^2[\mathbf{a}\,|\,t]\right)}{\mathcal{P}_t(\lambda^{(t)}_1, \lambda^{(t)}_2)}, \\ \mathcal{P}^*_t(\lambda^{(t)}_1, \lambda^{(t)}_2) &= \int\limits_{\mathcal{A}^p} \exp\left(-\lambda^{(t)}_1 \mathbb{C}[\mathbf{a}\,|\,t] - \lambda^{(t)}_2 \mathbb{C}^2[\mathbf{a}\,|\,t]\right) d\mathbf{a}, \end{aligned} \qquad t \in \mathcal{T}_{ln}. \tag{36}$$

Множители Лагранжа $\lambda^{(t)}_1, \lambda^{(t)}_2$ определяются решением следующих двух балансовых уравнений:

$$\begin{aligned} &\int\limits_{\mathcal{A}^p} \exp\left(-\lambda^{(t)}_1 \mathbb{C}[\mathbf{a}\,|\,t] - \lambda^{(t)}_2 \mathbb{C}^2[\mathbf{a}\,|\,t]\right) \left(\mathbb{C}[\mathbf{a}\,|\,t-1] - m^*_C[t]\right) d\mathbf{a} = 0, \\ &\int\limits_{\mathcal{A}^p} \exp\left(-\lambda^{(t)}_1 \mathbb{C}[\mathbf{a}\,|\,t] - \lambda^{(t)}_2 \mathbb{C}^2[\mathbf{a}\,|\,t]\right) \left(\mathbb{C}^2[\mathbf{a}\,|\,t-1] - D^*_C[t]\right) d\mathbf{a} = 0, \end{aligned} \qquad t \in \mathcal{T}_{ln}. \tag{37}$$

Согласно развитому в [29] методу воспользуемся аппроксимациями экспоненты полиномом степени 2:

$$\begin{aligned} \exp(x) &\approx \left(1 - \lambda^{(t)}_1 \mathbb{C}[\mathbf{a}\,|\,t] + \frac{1}{2}\left(\lambda^{(t)}_1 \mathbb{C}[\mathbf{a}\,|\,t]\right)^2\right) \mathbb{C}[\mathbf{a}\,|\,t], \\ \exp(y) &\approx \left(1 - \lambda^{(t)}_2\, \mathbb{C}^2[\mathbf{a}\,|\,t] + \frac{1}{2}\left(\lambda^{(t)}_2\, \mathbb{C}^2[\mathbf{a}\,|\,t]\right)^2\right) \mathbb{C}^2[\mathbf{a}\,|\,t]. \end{aligned} \tag{38}$$

Используя аппроксимации (38), получим следующие формы балансовых уравнений:

$$\begin{aligned} &\lambda_1 B_1(t) + \lambda_2 B_2(t) + \lambda^2_1 B_3(t) + \lambda^2_2 B_4(t) + \lambda_1 \lambda_2 B_5(t) = B_0(t), \\ &\lambda_1 Z_1(t) + \lambda_2 Z_2(t) + \lambda^2_1 Z_3(t) + \lambda^2_2 Z_4(t) + \lambda_1 \lambda_2 Z_5(t) = Z_0(t), \\ &t \in \mathcal{T} = [t_0, t_0 + p]. \end{aligned} \tag{39}$$

В первом из этих уравнений коэффициенты имеют вид:

$$
\begin{gathered}
B_0(t) = A\, m_C^*[t] - I_{p,1}^{(t)}, \quad B_1(t) = m_C^*[t]\, I_{p,1}^{(t)} - I_{p,2}^{(t)}, \\
B_2(t) = m_C^*[t]\, I_{p,2}^{(t)} - I_{p,3}^{(t)}, \quad B_3(t) = -\frac{1}{2}\, B_2^{(2,3)}(t), \\
B_4(t) = -\frac{1}{2}\, \left(m_C^*[t]\, I_{p,4}^{(t)} - I_{p,5}^{(t)} \right) \quad B_5(t) = -\left(m_C^*[t]\, I_{p,3}^{(t)} - I_{p,4}^{(t)} \right),
\end{gathered} \tag{40}
$$

Коэффициенты во втором уравнении имеют вид:

$$
\begin{gathered}
Z_0(t) = A\, D_C^*[t] - I_{p,2}^{(t)}, \quad Z_1(t) = D_C^*[t]\, I_{p,1}^{(t)} - I_{p,3}^{(t)}, \\
Z_2(t) = D_C^*[t]\, I_{p,2}^{(t)} - I_{p,4}^{(t)}, \quad Z_3(t) = -\frac{1}{2}\, Z_2^{(2,4)}(t), r \\
Z_4(t) = -\frac{1}{2}\, \left(D_C^*[t]\, I_{p,4}^{(t)} - I_{p,6}^{(t)} \right), \quad Z_5(t) = -\left(D_C^*[t]\, I_{p,3}^{(t)} - I_{p,5}^{(t)} \right).
\end{gathered} \tag{41}
$$

В этих выражениях

$$
A = \int\limits_{\mathcal{A}} d\mathbf{a} = (w-d)^p,
$$

$$
I_{p,n}^{(t)}(k_1, \ldots, k_n) = \underbrace{\int\limits_d^w \cdots \int\limits_d^w}_{p} \mathbb{C}^n[\mathbf{a} \,|\, t]\, d\mathbf{a} =
$$

$$
\begin{aligned}
&= \sum_{k_j \geq 0;\, \sum_{j=1}^n k_j = n} \frac{n!}{k_1! \cdots k_n!} \left(\underbrace{\int\limits_d^w \cdots \int\limits_d^w}_{p} a_1^{k_1} \cdots a_p^{k_n}\, da_1 \cdots da_p \right) \times \\
&\times (m_C^*[t])^{k_1} \cdots (m_C^*[t-p])^{k_n}, \quad n = \overline{0,6}.
\end{aligned} \tag{42}
$$

4. Прогнозирование стоимости финансового актива

Экспериментальное прогнозирование предложенного метода будем проводить для задач *одно- и двухдневного прогнозирования* средней стоимости и дисперсии финансового актива – акций ПАО «Газпром» в течение 2020 г. на Московской бирже.

Рассматриваются 12 торговых сессий, каждая на начало месяца. Для удобства все вычисления проводятся в условных единицах, соответствующих 1/1000 исходной цены. Данные о стоимости приведены в табл. 1.

4.1. Обучение модели стоимости

Модель стоимости (30) имеет память $p = 2$, начальный момент обучения $t_0 = 3 \to$ март, интервалы для двух параметров, равные $d = -1$, $w = 2$.

Таблица 1. Стоимостные индикаторы акций ПАО «Газпром» в 2020 г.

Месяц	янв.	февр.	март	апр.	май	июнь
Дата t	1	2	3	4	5	6
Цена m^*_C	0,259	0,223	0,208	0,178	0,188	0,200
Макс $C^*_{\max}$	0,262	0,240	0,212	0,196	0,202	0,208
Мин $C^*_{\min}$	0,227	0,201	0,158	0,177	0,182	0,190
V^*_C	0,017	0,020	0,027	0,009	0,011	0,009
D^*_C	0,084	0,069	0,070	0,041	0,046	0,049
Месяц	июль	авг.	сент.	окт.	ноябрь	дек.
Дата t	7	8	9	10	11	12
Цена m^*_C	0,195	0,182	0,181	0,171	0,154	0,183
Макс $C^*_{\max}$	0,202	0,195	0,186	0,173	0,189	0,215
Мин $C^*_{\min}$	0,179	0,180	0,170	0,154	0,152	0,182
V^*_C	0,011	0,007	0,009	0,010	0,019	0,017
D^*_C	0,206	0,189	0,190	0,181	0,173	0,200

Обучение модели стоимости будем проводить на интервале $\mathcal{T} = [t_0, t_0 + 2] = [3, 5]$. Исторический период $\mathcal{I}_p = [t_0 - 2, t_0 - 1] = [1, 2]$.

Поскольку модель (32) содержит два параметра, то приобретает следующий вид:

$$\mathbb{C}[\mathbf{a} \,|\, t] = a_1\, m^*_C[t-1] + a_2\, m^*_C[t-2]. \tag{43}$$

Функции ПРВ в соответствующих торговых сессиях имеют вид (36) с моделью (43). Заметим, что энтропийно-оптимальные ПРВ параметров, генерируемые линейной моделью (43), отличны от нормального распределения.

Для определения значений множителей Лагранжа, т.е. решения балансовых уравнений, применялся приближенный аналитический метод, изложенный в разделе 2.2.

Энтропийно-оптимальные функции ПРВ (с приближенными до первой коррекции) для 3-й, 4-й и 5-й торговых сессий имеет вид:

$$P^*_3(\mathbf{a} \,|\, 1{,}068;\ -0{,}871) = 0{,}131\, \exp(-0{,}238a_1 - 0{,}277a_2 + \\ + 0{,}043a_1^2 + 0{,}058a_2^2 + 0{,}100a_1a_2); \tag{44}$$

$$P^*_4(\mathbf{a} \,|\, 0{,}958;\ 0{,}102) = 0{,}133\, \exp(-0{,}199a_1 - 0{,}214a_2 - \\ - 0{,}004a_1^2 - 0{,}005a_2^2 - 0{,}005a_1a_2); \tag{45}$$

$$P^*_5(\mathbf{a} \,|\, -1{,}994;\ 2{,}609) = 0{,}092\, (0{,}355a_1 + 0{,}415a_2 - \\ - 0{,}083a_1^2 - 0{,}112a_2^2 - 0{,}193a_1a_2). \tag{46}$$

На рис. 1–3 показаны их графики.

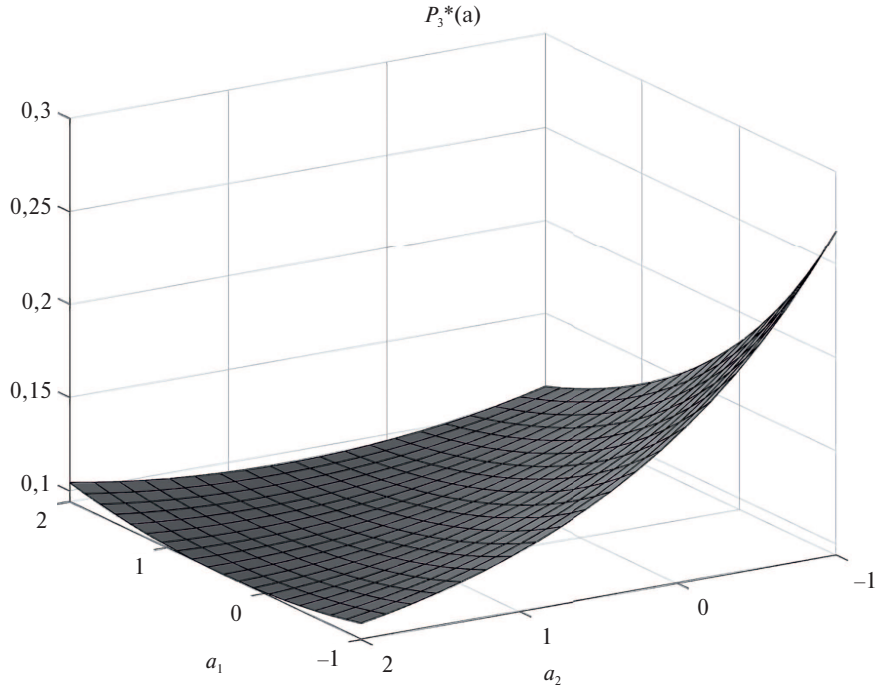

Рис. 1. Распределение $P_3^*(\mathbf{a})$.

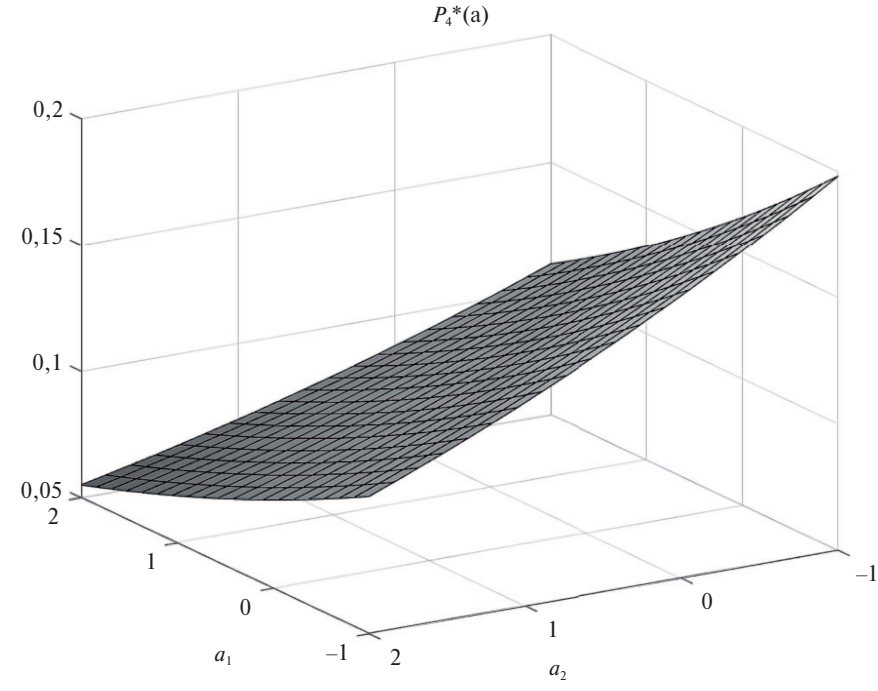

Рис. 2. Распределение $P_4^*(\mathbf{a})$.

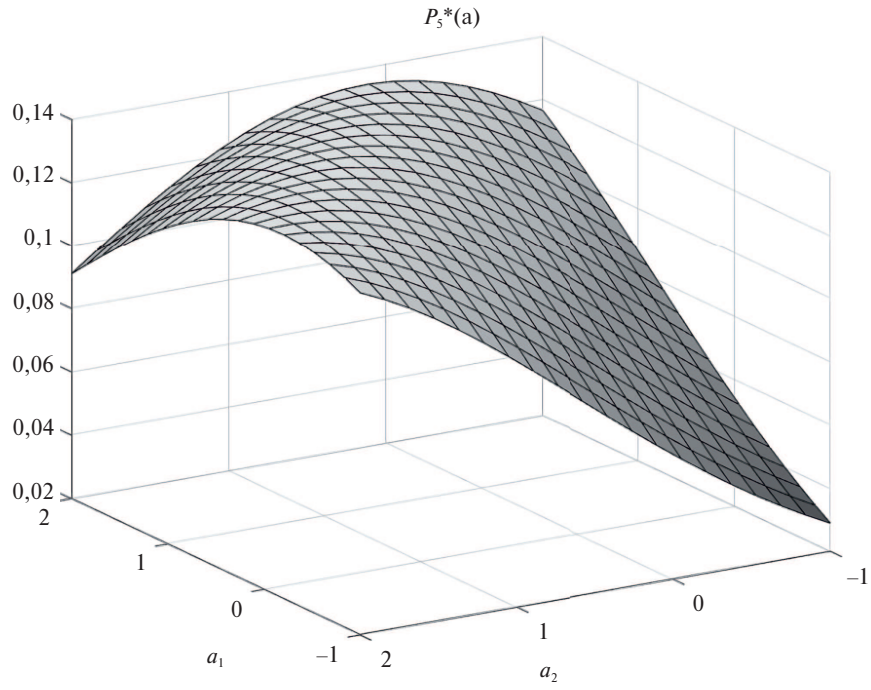

Рис. 3. Распределение $P_5^*(\mathbf{a})$.

4.2. Прогнозирование средней стоимости и среднего индикатора волатильности

Полученные выше энтропийно-оптимальные функции ПРВ $P_3^*(\mathbf{a})$, $P_4^*(\mathbf{a})$, $P_5^*(\mathbf{a})$ параметров модели (32) будем использовать для генерирования ансамблей данных, по которым вычисляются прогнозируемые значения $m_C[t]$ и $D_C[t]$ в торговых сессиях: апр. (4)–нояб. (11). Реализованные значения этих переменных известны (табл. 1), что позволяет оценить точность прогнозов при использовании различных стратегий прогнозирования.

4.2.1. Однодневные прогнозы $P_k^*(\mathbf{a}) \to (m_C[k+1], V_C[k+1])$

При однодневных прогнозах используется оптимальная ПРВ для k-ой торговой сессии и прогнозируются результы $(k+1)$-й сессии. Рассмотрим процедуру формирования прогноза $3 \to 4$. Для этого используется ПРВ $P_3^*(\mathbf{a})$ (44) и прогнозная модель (32), которая в данном примере приобретает следующий вид:

$$\mathbb{C}[\mathbf{a}\,|\,4] = a_1 m_C^*[3] + a_2 m_C^*[2]. \tag{47}$$

ПРВ $P_3^*(\mathbf{a})$ (44) трансформируем в случайную последовательность $\{a_1, a_2\}$. Генерируемый ансамбль содержит 1000 значений $\mathbb{C}[\mathbf{a}\,|\,4]$. Вычисляем $\bar{m}_C[4] = \tilde{\mathcal{M}}(\mathbb{C}[\mathbf{a}\,|\,4])$ и $\bar{\sigma}_C^2[4] = \tilde{\mathcal{M}}\{(\mathbb{C}[\mathbf{a}\,|\,4] - \bar{m}_C[4])^2\}$, где под оператором $\tilde{\mathcal{M}}\{\bullet\}$ понимается эмпирическое среднее.

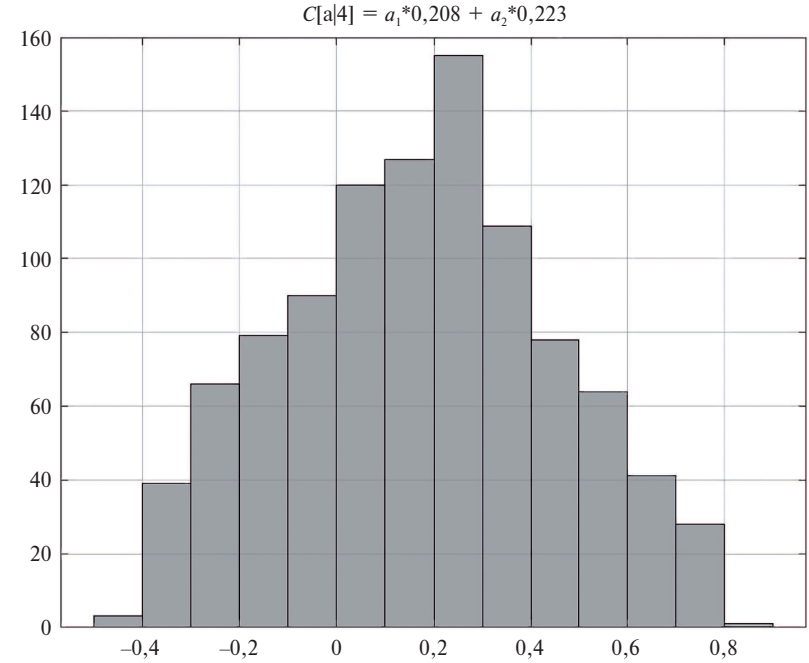

Рис. 4. Эмпирическая ПРВ $\mathbb{C}[\mathbf{a} \,|\, 4]$ (однодневный прогноз $3 \to 4$).

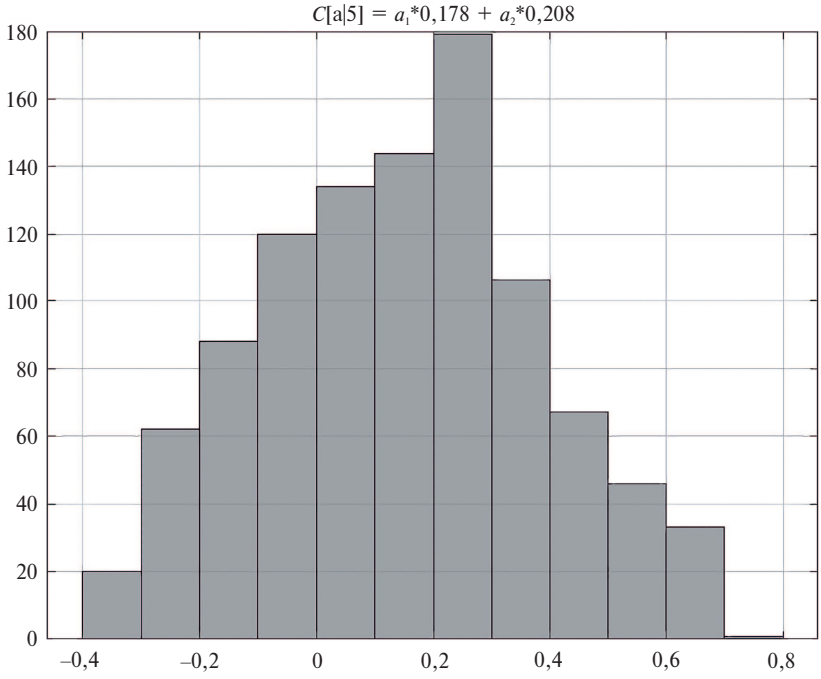

Рис. 5. Эмпирическая ПРВ $\mathbb{C}[\mathbf{a} \,|\, 5]$ (однодневный прогноз $4 \to 5$).

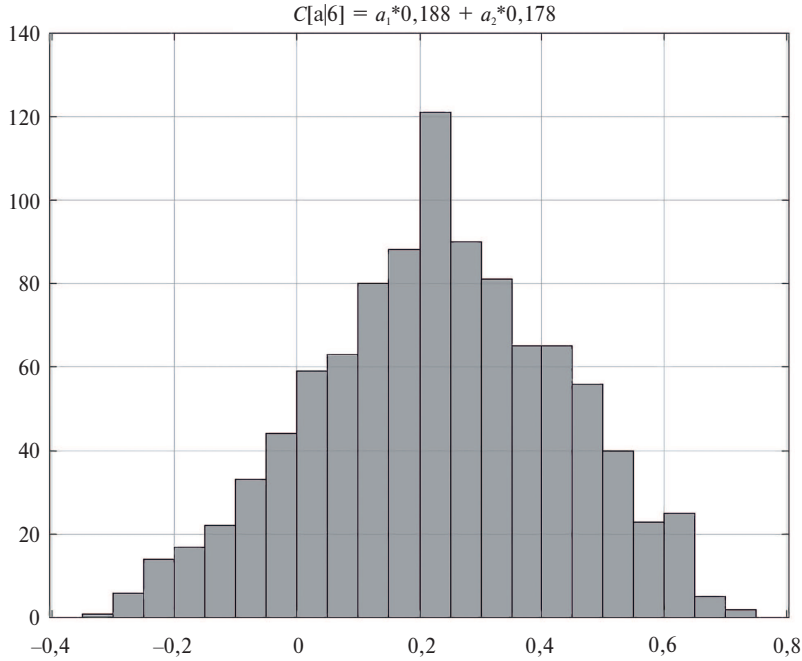

Рис. 6. Эмпирическая ПРВ $\mathbb{C}[\mathbf{a}\,|\,6]$ (однодневный прогноз $5 \to 6$).

Прогнозы $4 \to 5$ и $5 \to 6$ формируются аналогично. Результаты однодневных прогнозов и оценки их точности (в сравнении с реализованными значениями в торгах) приведены в табл. 2.

Таблица 2. Однодневные прогнозы

прогноз $\bullet$	$3 \to 4$	$4 \to 5$	$5 \to 6$
$\bar{m}_C[\bullet]$	0,175	0,145	0,229
$m^*_C[\bullet]$	0,178	0,188	0,200
$\bar{\sigma}^2_C[\bullet]$	0,076	0,056	0,041
$V^*_C[\bullet]$	0,041	0,046	0,049
$\lvert\delta_m[\bullet]\rvert$	0,003	0,043	0,029
$\lvert\delta_\sigma[\bullet]\rvert$	0,035	0,010	0,008

В этой таблице переменные

$$\delta_m[\bullet] = \bar{m}_C[\bullet] - m^*_C[\bullet], \tag{48}$$
$$\delta_\sigma[\bullet] = \bar{\sigma}^2_C[\bullet] - V^*_C[\bullet].$$

На рис. 4–6 показаны эмпирические ПРВ прогнозных значений стоимости в сессиях 4–6 при однодневных прогнозах.

«Интегральная» относительная ошибка *прогноза средней стоимости* при однодневных прогнозах имеет вид:

$$\Delta_m = \frac{\sqrt{\sum_{t=4}^{6} \delta_m^2[t]}}{\sqrt{\sum_{t=4}^{6} m_C^2[t]} + \sqrt{\sum_{t=4}^{6} (m_C^*[t])^2}} = 8\%. \tag{49}$$

«Интегральная» относительная ошибка *прогноза средней дисперсии* при однодневных прогнозах имеет вид:

$$\Delta_\sigma = \frac{\sqrt{\sum_{t=4}^{6} \delta_\sigma^2[t]}}{\sqrt{\sum_{t=4}^{6} \sigma_C^2[t]} + \sqrt{\sum_{t=4}^{6} (V_C^*[t])^2}} = 17\%. \tag{50}$$

4.2.2. Двухдневные прогнозы

$$P_k^*(\mathbf{a}) \to (m_C[k+1], V_C[k+1]), (m_C[k+2], V_C[k+2])$$

При двухдневных прогнозах используется оптимальная ПРВ для торговой сессии k и прогнозируется результат в сессии $k+2$.

Прогноз $3 \to 4, 5$, ПРВ $P_3^*(\mathbf{a})$ может быть реализован в варианте

$$\begin{aligned} \mathbb{C}[\mathbf{a} \,|\, 4] &= a_1\, m^*[3] + a_2\, m_C^*[2], \\ \mathbb{C}[\mathbf{a} \,|\, 5] &= a_1\, m^*[4] + a_2\, m_C^*[3]. \end{aligned} \tag{51}$$

и в варианте

$$\begin{aligned} &\mathbb{C}[\mathbf{a} \,|\, 4] = a_1\, m^*[3] + a_2\, m_C^*[2] = a_1\, 0{,}208 + a_2\, 0{,}223, \\ &\mathbb{C}[\mathbf{a} \,|\, 5] = a_1\, \bar{m}^*[4] + a_2\, m_C^*[3] = a_1\, \bar{m}^*[4] + a_2\, 0{,}208, \\ &\bar{m}^*[4] = \tilde{\mathcal{M}}\{\mathbb{C}[\mathbf{a} \,|\, 4]\}. \end{aligned} \tag{52}$$

Двухдневные прогнозы $4 \to 5, 6$ и $5 \to 6, 7$ производятся аналогично и их результаты и оценки точности приведены в табл. 3.

Таблица 3. Двухдневные прогнозы

$\bullet$	$3 \to 4$	$3 \to 5$	$4 \to 5$	$4 \to 6$	$5 \to 6$	$5 \to 7$
$\bar{m}_C[\bullet]$	0,175	0,185	0,145	0,206	0,229	0,200
$m_C^*[\bullet]$	0,178	0,188	0,188	0,200	0,200	0,195
$\bar{\sigma}_C^2[\bullet]$	0,056	0,060	0,056	0,012	0,041	0,024
$V_C^*[\bullet]$	0,041	0,051	0,046	0,009	0,049	0,011
$\lvert\delta_m[\bullet]\rvert$	0,003	0,033	0,043	0,078	0,029	0,065
$\lvert\delta_\sigma[\bullet]\rvert$	0,035	0,049	0,010	0,030	0,009	0,043

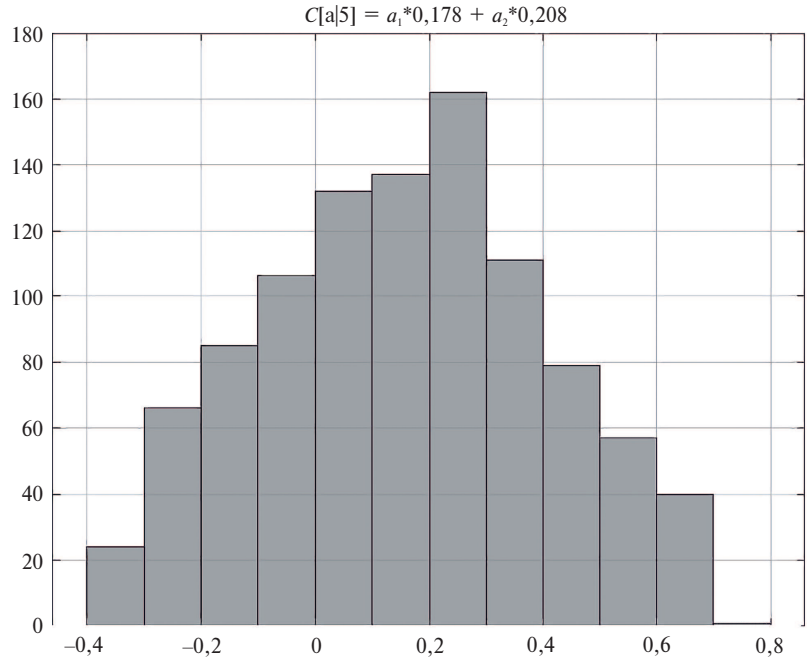

Рис. 7. Эмпирическая ПРВ $\mathbb{C}[\mathbf{a} \mid 5]$ (двухдневный прогноз $3 \to 4, 5$).

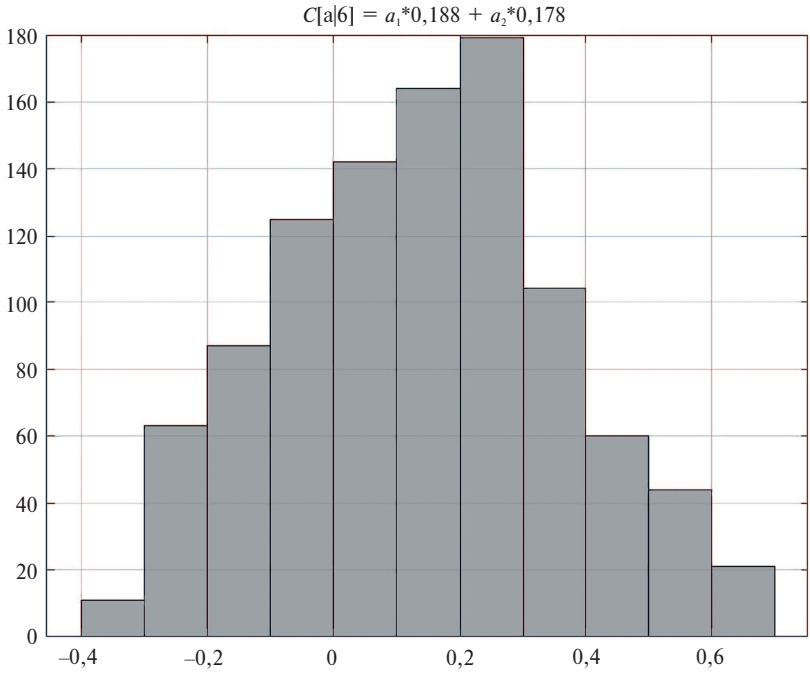

Рис. 8. Эмпирическая ПРВ $\mathbb{C}[\mathbf{a} \mid 6]$ (двухдневный прогноз $4 \to 5, 6$).

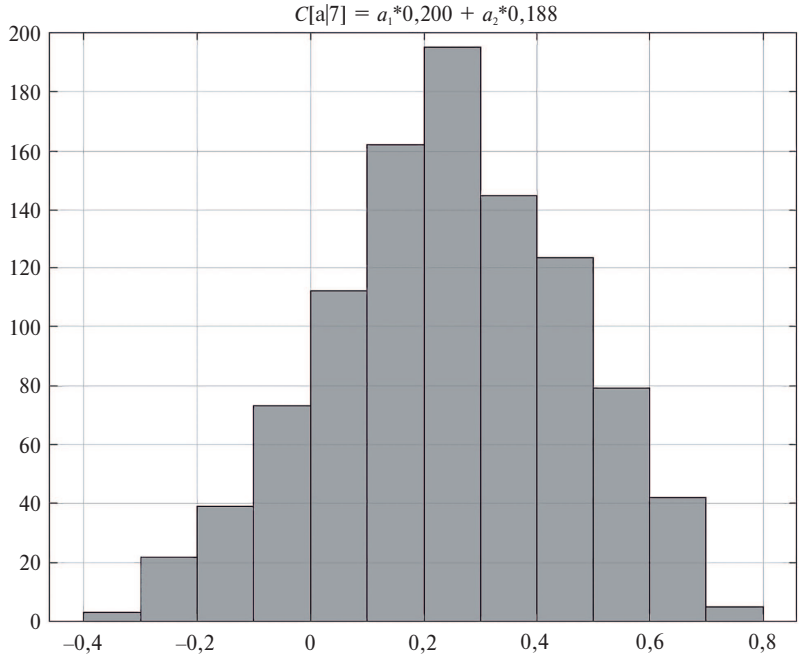

Рис. 9. Эмпирическая ПРВ $\mathbb{C}[\mathbf{a}\,|\,7]$ (двухдневный прогноз $5 \to 6,7$)

Интегральная относительная ошибка *прогноза средней стоимости* для группы двухдневных прогнозов имеет вид

$$\Delta_m = \frac{\sqrt{\sum_{t=5}^{7} \delta_m^2[t]}}{\sqrt{\sum_{t=5}^{7} m_C^2[t]} + \sqrt{\sum_{t=5}^{7} (m_C^*[t])^2}} = 7{,}2\%. \tag{53}$$

Интегральная относительная ошибка *прогноза средней волатильности* для группы двухдневных прогнозов имеет вид

$$\Delta_\sigma = \frac{\sqrt{\sum_{t=5}^{7} \delta_\sigma^2[t]}}{\sqrt{\sum_{t=5}^{7} \sigma_C^2[t]} + \sqrt{\sum_{t=5}^{7} (V_C^*[t])^2}} = 25\%. \tag{54}$$

На рис. 7–9 показаны эмпирические ПРВ прогнозных значений стоимости в сессиях 5–7 при двухдневных прогнозах.

5. Обсуждение

Проблема генерации подходящих данных для тестирования и прогнозирования является весьма популярной в современной компьютерной науке. В статье предлагается адаптация и развитие технологии рандомизированного машинного обучения для генерации ансамблей данных с заданными числовыми характеристиками.

В отличие от существующей технологии предлагается расширение, учитывающее моментные характеристики от 1-го до s-го порядка. Показано, что это приводит к функциям ПРВ не гаусового класса даже в случае линейной модели данных. Предлагаемое расширение, так же как и существующее, сводится к решению соответствующих балансовых уравнений, содержащих интегральные компоненты. В статье развивается приближенный аналитический метод их решения, основанный на использовании степенных рядов и методе неопределенных коэффициентов.

Он применяется в задаче прогнозирования стоимости финансового инструмента, результаты которой сравнивались с реализованными данными по одно- и двухдневным прогнозам. В рамках проведенных исследований обнаружилась вполне приемлемая точность приближенного решения в составе двух коррекций. Однако необходимо более глубокое изучение приближенного метода, как его теоретических аспектов, так и численного исследования.

6. Заключение

В статье развита теория и алгоритм генерации ансамблей тестовых данных с заданными свойствами в виде числовых характеристик, основанная на модификации структуры процедуры рандомизированного машинного обучения [22]. Известно, что ядром указанной процедуры являются балансовые уравнения относительно множителей Лагранжа, содержащие так называемые интегральные компоненты, т.е. многомерные интегралы от экспоненциальных подынтегральных функций.

Для решения этих уравнений адаптирован метод асимптотического *аналитического* решения, развитый в [29], который позволяет свести задачу многомерного интегрирования к сумме произведений одномерных интегралов от степенных функций.

Разработан метод рандомизированного прогнозирования и применен к построению одно- и двухдневных прогнозов средних стоимости и дисперсии биржевого финансового актива.

СПИСОК ЛИТЕРАТУРЫ

1. *Rubinstein R.Y., Kroese D.P.* Simulation and the Monte Carlo Method, 2016, John Wiley & Sons.
2. *Vapnik V.N.* Statistical Learning Theory, Wiley, 1998.
3. *Bishop C.M.* Pattern Recognition and Machine Learning, Springer, 2006.
4. *Hastie T., Tibshirant R., Friedman J.* The Elements of Statistical Learning, Springer, 2009.
5. *Vovk V., Shafer G.* Good Randomized sequential probability forecasting is always possible // J. Royal. Stat. Soc. B. 2005. V. 67. No. 5. P. 747–763.
6. *Hong T., Prinson P., Fan S., Zareijpour H., Triccoli A., Hyndman R.J.* Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond // Int. J. Forecast. 2016. V. 32. No. 3. P. 896–913.

7. *Zhang et al.* Stock pricepredictionvia discovering multy-frequency trading patterns // Proc. 23rd ACM SIGKDD international Conference on Knowlege Discovery and Data Mining. 2017. P. 2141–2149.
8. *Myers G.J.* The Art of Software Testing. John Wiley & Sons, 1979.
9. *Городецкий В.И., Грушицкий М.С., Хабалов А.В.* Многоагентные системы (обзор) // Новости искусственного интеллекта, 1998, № 2, С. 116.
10. *Patton R.* Software Testing, SAWS Publishers, 2005.
11. *Лысенков А.И., Бут Г.С., Диденко Д.А.* Система для разработки компьютерных тестов. http://www.bytic.ru/cue99m/cf7pvke.html, 2002.
12. *Мицель А.А., Погуда А.А.* Нейросетевой подход к задаче тестирования // Прикладная информатика, 2011. № 5 (35). С. 60–67.
13. *Заозерская Л.А., Платонова В.А.* Математические модели формирования оптимального комплекса структур тестов для контроля знаний // Омский научный вестник. 2012. № 3.
14. *Campi M.C., Garatti S., Prandini M.* The scenario approach for systems and control design // Ann. Rev. Control. 2009. V. 33. No. 2. P. 149–157.
15. *Chi Z., Liu Y., Turrini A., Zhang L., Jansen D.N.* A scenario approach for parametric Markov decision processes / In Principlies of Verification: Cycling the Probabilistic Landscape: Essay Dedicated to Joost-Pieter Katoen on the Occasiion of His 60th Birthday, Part II. Cham, Springer Nature Switzerland. 2024. P. 234–266.
16. *Boltzmann L.* Vorlesungen uber Gastheory. Leipzig, 1896, V. 1, J.A.Barth; 1898, V. 2, J.A.Barth.
17. *Jaynes E.T.* Information theory and statistical Mechanics // Phys. Rev. 1957. V. 104. No. 4. P. 620–630.
18. *Jaynes E.T.* Gibbs vs Boltzmann entropy // Amer. J. Phys. 1965. V. 33. P. 391–398.
19. *Rosenkrantz R.D., Jaynes E.T.* Paper on Probability, Statistics,and Statistical Physics. Kluwer Academic Pablishers, 1989.
20. *Jaynes E.T.* Probability theory: the logic of science. Cambridge Uni. Press, 2003.
21. *Попков Ю.С.* Асимптотическая эффективность оценок максимальной энтропии // Докл. АН. 2020. Т. 493. С. 100–103.
22. *Popkov Yu.S., Popkov A.Yu., Dubnov Yu.A.* Entropy Randomization in Machine Learning. CRC Press, 2023.
23. *Красносельский М.А., Вайникко Г.М., Забрейко П.П., Рутицкий Я.Б., Стеценко В.Я.* Приближенные решения операторных уравнений. М.: Наука, 1969.
24. *Малкин И.Г.* Некоторые задачи теории нелинейных колебаний. М.: УРСС, 2004.
25. *Darkhovsky, B.S., Popkov, Y.S., Popkov, A.Y., Aliev, A.S.* A Method of Generating Random Vectors with a Given Probability Density Function // Autom. Remote Control. 2018. V. 79. No. 9. P. 1569–1581. https://doi.org/10.1134/S0005117918090035
26. *Avellaneda M.* Minimum-relative-entropy calibration of asset-pricing models // Int. J. Theor. App. Finance. 1998. V. 1. No. 04. P. 447–472.
27. *Jackwerth J.C.* Recovering Risk Aversion from OptionPrices and Realized Returns // Rev. Financ. Stud. 2000. V. 11. No. 2. P. 437.
28. *Ant-Sahalia Y., Lo A.W.* Nonparametric Risk Management and Implied Risk Avarsion // J. Econom. 2000. V. 94. P. 4–5.

29. *Popkov Yu.S.* Analytic Method for Solving One Class of Nonlinear Equations // Doklady Mathematics. 2024. https://doi.org/10.1134/S1064562424601392
30. *Фихтенгольц Г.М.* Курс дифференциального и интегрального исчисления. М.: Физматгиз, 1962.
31. *Фелер В.* Введение в теорию вероятностей и ее приложения. М.: Мир, 1967.
32. *Соболь И.М.* Численные методы Монте-Карло. М.: Наука, 1973.
33. *Бахвалов Н.С., Жидков Н.П., Кобельков Г.М.* Численные методы. М.: Бином, 2003.

Статья представлена к публикации членом редколлегии О.Н. Граничиным.

Поступила в редакцию 23.02.2025

После доработки 29.04.2025

Принята к публикации 29.05.2025