На правах рукописи

Лазутченко Алексей Николаевич

Робастное управление в случайных средах с различными дисперсиями применительно к параллельной обработке данных

05.13.18 – Математическое моделирование, численные методы и комплексы программ

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико-математических наук

Петрозаводск – 2017

Работа выполнена в Новгородском государственном университете имени Ярослава Мудрого.

Научный руководитель:	**Колногоров Александр Валерианович**, доктор физико-математических наук, доцент
Официальные оппоненты:	**Кириллов Александр Николаевич**, доктор физико-математических наук, доцент, ведущий научный сотрудник лаборатории информационных компьютерных технологий Института прикладных математических исследований Карельского научного центра РАН
	Золотухин Игорь Владимирович, кандидат физико-математических наук, старший научный сотрудник лаборатории оптики океана и атмосферы Санкт-Петербургского филиала Института океанологии имени П. П. Ширшова РАН
Ведущая организация:	Федеральный исследовательский центр «Информатика и управление» РАН

Защита состоится 7 апреля 2017 года в 11 часов на заседании диссертационного совета Д 212.190.03 при ФГБОУ ВО «Петрозаводский государственный университет», по адресу: 185910, г. Петрозаводск, пр. Ленина, 33

С диссертацией можно ознакомиться в библиотеке Петрозаводского государственного университета и на сайте http://www.petrsu.ru.

Автореферат разослан «____» ____________ 2017 г.

Ученый секретарь
диссертационного совета *Воронов Роман Владимирович*

Общая характеристика работы

Актуальность темы исследования. Задачи управления в условиях неопределенности в последние годы получают всё бо́льшую актуальность в различных областях человеческой деятельности: медицина, web-поиск, интернет-реклама, компьютерные сети и др. Рассматриваемую в данной работе задачу можно описать следующим образом. Имеются два альтернативных действия, каждое применение которых ведет к получению некоторого дохода. Доходы являются случайными, их распределения фиксированы и зависят только от текущих применяемых действий, но неизвестны лицу, принимающему решения (ЛПР). Наиболее известным приложением является задача о лечении пациентов, когда имеются два альтернативных лекарства с фиксированными, но неизвестными эффективностями. Требуется организовать лечение таким образом, чтобы математическое ожидание количества выздоровевших в результате лечения пациентов было максимальным. Здесь можно отметить работы «Асимптотически эффективные адаптивные правила распределения» Т. Лая, Г. Роббинса (Lai T. L., Robbins H. Asymptotically Efficient Adaptive Allocation Rules, 1985) и «Модели многоруких бандитов для оптимального управления клиническими испытаниями: достоинства и недостатки» С. Виллар (Villar S. S. et al. Multi-armed bandit models for the optimal design of clinical trials: benefits and challenges, 2015). Другое известное приложение — оптимизация обработки данных двумя альтернативными методами с фиксированными, но априори неизвестными эффективностями. Такая задача рассматривалась, например, в работе «Современный байесовский взгляд на многоруких бандитов» С. Скотта (Scott S. L. A modern Bayesian look at the multi-armed bandit, 2010). Отметим, что при решении этих задач может использоваться параллельная обработка. В задачах лечения пациентов количество этапов обработки является небольшим, обычно два, редко три–четыре. В задачах же обработки данных допускается большое число этапов, например, 30–50. Так как данные при параллельной обработке суммируются, то в силу центральной предельной теоремы для управления используются нормально распределенные значения доходов. В зависимости от рассматриваемой постановки задачи дисперсии этих доходов могут иметь попарно одинаковые или различные значения.

Наиболее известными постановками цели управления являются минимаксная и байесовская. Байесовская постановка рассматривается, например, в работах «Последовательное управление по неполным данным» Э. Пресмана, И. Сонина, 1982 г. и «Проблемы бандитов. Последовательное распределение экспериментов» Д. Берри, Б. Фриштеда (Berry D. A., Fristedt B. Bandit problems. Sequential Allocation of Experiments, 1985). Достоинством байесовского подхода является то, что он позволяет для любого априорного распределения найти байесовские стратегии и риск численными методами. Недостатком подхода является отсутствие ясных критериев выбора априорного распределения. Достоинством минимаксного подхода является его робастность, а недостатком — сложность нахождения робастных алгоритмов, поскольку прямые методы нахождения их отсутствуют. Примерами робастных алгоритмов являются алгоритм на основе метода стохастической аппроксимации («Адаптивный выбор вариантов» А. Назина, А. Позняка, 1986 г.) и алгоритм зеркального спуска («Оценки для стохастического многорукого бандита» А. Юдицкого, А. Назина, А. Цыбакова, Н. Ваятиса (Juditsky A. B., Nazin A. V., Tsybakov A. B., Vayatis N. Gap-free Bounds for Stochastic Multi-Armed Bandit, 2008) и «Об эффективности одного метода рандомизации зеркального спуска в задачах online оптимизации» А. Гасникова, Ю. Нестерова, В. Спокойного, 2015 г.). Отметим, что версии этих алгоритмов, допускающие параллельную обработку, насколько нам известно, не разрабатывались.

Байесовский и минимаксный подход объединяет основная теорема теории игр, согласно которой минимаксные стратегии и риск ищутся как байесовские, соответствующие наихудшему априорному распределению. Эту теорему можно использовать, если такое распределение удается характеризовать.

Таким образом, актуальной является разработка робастных (минимаксных) алгоритмов, с помощью которых осуществляется параллельная многоэтапная обработка данных. Математически эта задача описывается как управление в случайной среде с нормально распределенными доходами, характеризуемыми различными дисперсиями. В качестве основного инструмента для нахождения минимаксных стратегии и риска используется основная теорема теории игр.

Цели и задачи диссертационной работы. Целью диссертационной работы является нахождение минимаксных стратегий и рисков в средах с бинарными и нормально распределенными доходами, характеризуемыми различными дисперсиями.

Для достижения поставленных целей были решены следующие задачи:

1. Разработка математической модели оптимизации групповой обработки данных, если для обработки используются два альтернативных метода с фиксированными, но неизвестными эффективностями.

2. Исследование методами математического моделирования одно- и двухпороговых стратегий, обеспечивающих робастное управление в случайной среде, характеризуемое неулучшаемым по порядку величины максимальным значением функции потерь.

3. Разработка алгоритмов, позволяющих искать минимаксные стратегии и риски численными методами как байесовские, соответствующие наихудшему априорному распределению параметров среды с различными дисперсиями.

4. Разработка комплекса программ, который позволяет численными методами находить минимаксные стратегии и риски.

Научная новизна. Научная новизна работы заключается в разработке модели групповой обработки данных, которая математически описывается как задача об управлении в случайной среде с доходами, имеющими нормальное распределение с произвольными дисперсиями. Известные стратегии групповой обработки допускали малое число этапов обработки: обычно два, редко три-четыре. Предложенные в работе стратегии допускают бо́льшее число этапов, например 30–50. Такое число этапов позволяет вести обработку практически без потери качества, то есть без увеличения минимаксного риска. Нахождение минимаксных стратегий и рисков выполнено численными методами с использованием свойства инвариантности и основной теоремы теории игр, то есть минимаксные стратегии и риск ищутся как байесовские, соответствующие наихудшему априорному распределению.

Практическая значимость. Практические результаты, полученные в диссертации, могут быть использованы в различных задачах о лечении пациентов, для параллельной обработки данных и нахождения оптимальных стратегий управления. Полученные результаты могут применяться в системах, предварительное тестирование которых либо невозможно, либо нецелесообразно в силу различных причин.

Методология и методы исследования. В диссертации используются: теория вероятностей, теория игр, адаптивное управление. Для численного моделирования и программной реализации разработанных алгоритмов применяются методы вычислительной математики и прикладного программирования.

Положения, выносимые на защиту: на защиту выносятся следующие положения:

1. Модель параллельной обработки данных, которая описывается управлением в случайных средах с нормально распределенными доходами с произвольными дисперсиями.

2. Численный метод определения оптимальных параметров одно- и двухпороговых стратегий управления.

3. Численный метод нахождения минимаксных рисков и стратегий как байесовских, соответствующих наихудшему априорному распределению, для случайных сред с различными дисперсиями.

4. Комплекс программ для моделирования двухальтернативных случайных сред с бинарными и нормально распределенными доходами с различными дисперсиями, разработанный для обработки численными методами данных, получения и применения стратегий управлений и решения вспомогательных задач, возникших при исследований.

Степень достоверности и апробация результатов. Данные расчетов, полученных в работе, соответствуют результатам проведенных численных моделирований. Кроме того, для ряда полученных результатов установлена согласованность с уже известными, полученными другими авторами.

Основные результаты диссертации докладывались на следующих конференциях:

1. XVI, XVII, XVIII, XIX, XXIII научные конференции преподавателей, аспирантов и студентов НовГУ, 2009–2016 гг., Великий Новгород.

2. XIV Всероссийский симпозиум по прикладной и промышленной математике, 29 сентября — 5 октября 2013 г., Великий Новгород.

3. 57-я международная научная конференция МФТИ «Актуальные проблемы фундаментальных и прикладных наук в современном информационном обществе», 24 — 29 ноября 2014 г., Долгопрудный, Московская обл.

4. 58-я научная конференция МФТИ с международным участием, 23 — 28 ноября 2015 г., Долгопрудный, Московская обл.

5. IX Международная Петрозаводская конференция «Вероятностные методы в дискретной математике», 30 мая — 3 июня 2016 г., Петрозаводск, Россия.

А также на семинарах кафедры прикладной математики и информатики Новгородского государственного университета имени Ярослава Мудрого.

Результаты исследований внедрены в учебный процесс Новгородского государственного университета имени Ярослава Мудрого в рамках дисциплины «Модели искусственного интеллекта».

Публикации. Материалы диссертации опубликованы в 11 печатных работах, из них 4 статьи в рецензируемых журналах, рекомендованных ВАК [1–4], 3 статьи в сборниках трудов конференций [5–7] и 4 тезиса докладов [8–11].

Комплексы программ:

Свидетельство о государственной регистрации программы для ЭВМ «Оптимизация параметров пороговой стратегии управления в случайной среде» № 2012614942, зарегистрирована в Реестре программ для ЭВМ 1 июня 2012 г.

Свидетельство о государственной регистрации программы для ЭВМ «Моделирование случайных сред с одно- и двухпороговой стратегиями управления в случайной среде» № 2013617255, зарегистрирована в Реестре программ для ЭВМ 06 августа 2013 г.

Свидетельство о государственной регистрации программы для ЭВМ «Расчет байесовских рисков, потерь и стратегий управления для случайных сред с нормально распределенными доходами с произвольными дисперсиями» № 2016619416, зарегистрирована в Реестре программ для ЭВМ 18 августа 2016 г.

Структура и объем диссертации. Диссертация состоит из введения, 4 глав, заключения и списка литературы. Общий объем диссертации 96 страниц, включая 24 рисунка и 3 таблицы. Список литературы включает 64 наименования.

Содержание работы

Во Введении приведены общие сведения по исследуемой проблеме, обоснована актуальность диссертационной работы, сформулирована цель и задачи работы, аргументирована научная новизна исследований, показана практическая значимость полученных результатов, представлены выносимые на защиту положения, описана структура диссертации.

В первой главе рассмотрен краткий обзор литературы по исследуемой тематике, который приводит к постановке задачи о робастном параллельном управлении в случайной среде. Рассмотрены: автоматный и идентификационный подходы, метод зеркального спуска, алгоритм UCB1 (верхний доверительный предел), параллельное управление, а также минимаксный и байесовский подходы.

Одним из ключевых рассматриваемых в работе понятий является двухальтернативная случайная среда — это управляемый случайный процесс ξ_t, значения которого зависят только от выбираемых в текущие моменты времени действий $\ell = 1, 2$ и которые могут интерпретироваться как доходы. Здесь $t = 1, \ldots, T$, T — полное время управления средой (горизонт управления). В нашей постановке задачи горизонт управления известен и фиксирован.

Классические случайные среды имеют бинарное распределение доходов, т. е. случайный процесс ξ_t принимает только значения 1 и 0 (которые можно рассматривать как «успех» и «неудача») и зависит только от выбираемых в текущий момент времени действий. В рассматриваемой постановке задачи распределение доходов фиксировано, но ЛПР не знает, какое действие является более «доходным».

Таким образом, бинарную случайную среду можно описать следующим образом:

$$P(\xi_t = 1|y_t = \ell) = p_\ell, P(\xi_t = 0|y_t = \ell) = q_\ell,$$
$$p_\ell + q_\ell = 1, \quad \ell = 1,2, \quad t = 1,\dots,T.$$

При этом среда задается вектором вероятностей $\theta = (p_1, p_2)$.

Цель ЛПР — получить как можно бóльший ожидаемый выигрыш. Для этого он может использовать некоторые последовательности решений, которые в совокупности составляют его стратегию. Стратегия — это функция предыстории процесса, которая указывает, какое действие следует применять на очередном шаге, используя всю накопленную информацию о количестве использовании действий и накопленных в результате доходах. Мы будем использовать обозначение σ для стратегии и Σ — для множества стратегий.

В процессе управления возникают потери дохода вследствие того, что ЛПР не располагает полной информацией о параметрах случайной среды. Поэтому появляется понятие «функция потерь», мы будем использовать обозначение $L_T(\sigma, \theta)$, где σ — используемая стратегия, $\theta = (p_1, p_2)$ — векторный параметр случайной среды, p_1,p_2 — вероятности успешного исхода испытаний на действиях $\ell = 1, 2$ соответственно.

В случае, когда ЛПР известен параметр θ, наилучшей стратегией для него является такая, которая применяет только то действие, которому соответствует большая из величин p_1, p_2; при этом максимальный полный доход в этом случае равен $\max\limits_{\ell=1,2} p_\ell \cdot T$. В противном случае возникают потери, равные:

$$L_T(\sigma, \theta) = \max_{\ell=1,2} p_\ell \cdot T - \mathbb{E}_{\sigma,\theta}\left(\sum_{t=1}^{T} \xi_t\right).$$

Здесь $\mathbb{E}_{\sigma,\theta}$ представляет собой математическое ожидание потерь полного дохода при выбранной стратегии σ и заданном параметре среды θ.

Модификация рассматриваемой в настоящей работе задачи состоит в том, что доходы за применение действий имеют нормальное распределение, плотность которого равна

$$f_{D_\ell}(x|m_\ell) = \frac{1}{\sqrt{2\pi D_\ell}} e^{-\frac{(x-m_\ell)^2}{2D_\ell}}.$$

Такая случайная среда описывается вектором математических ожиданий $\theta = (m_1, m_2)$, который также неизвестен ЛПР, D_ℓ — дисперсии на действиях ℓ, а ограничения на множество допустимых значений параметра θ имеют вид:

$$\Theta = \{m_\ell : |m_1 - m_2| \leqslant 2c,\ 0 < c < \infty, \ell = 1, 2\}.$$

Условие $|m_1 - m_2| \leqslant 2c$ требуется для ограниченности функции потерь на Θ. Функция потерь для этого случая принимает вид

$$L_T(\sigma, \theta) = \max_{\ell=1,2} m_\ell \cdot T - \mathbb{E}_{\sigma,\theta}\left(\sum_{t=1}^{T} \xi_t\right).$$

Поясним, почему имеет смысл рассматривать такую постановку задачи. Дело в том, что использование нормального распределения тесно связано с возможностью параллельной обработки данных. Предположим, что нам необходимо обработать объем данных $T = M \cdot K$, результат обработки которых имеет бинарное распределение с вероятностями p_ℓ, где p_ℓ близки к p, $0 < p_\ell < 1$, $\ell = 1, 2$. Разделим все данные на M групп по K данных. Для всех K данных в одной группе мы можем применить параллельно один и тот же способ обработки. При этом к данным с номерами $t = (m-1)K + 1, \dots, mK$ применяется одно и то же действие, а затем формируется приведенное значение процесса ξ'_m:

$$\xi'_m = (DK)^{-1/2} \sum_{t=(m-1)K+1}^{mK} \xi_t, \quad m = 1, \dots, M,$$

где $D = p(1-p)$. Поскольку обрабатываемые данные не зависят друг от друга, то в силу центральной предельной теоремы распределение ξ'_m будет близким к нормальному.

Существуют различные подходы к решению рассматриваемой задачи. В данной работе рассматриваются два таких подхода, минимаксный и байесовский. При минимаксном подходе цель управления состоит в минимизации максимальных потерь на множестве параметров среды Θ по множеству стратегий Σ. При этом величина

$$R_T^M(\Theta) = \inf_{\Sigma} \sup_{\Theta} L_T(\sigma, \theta)$$

называется минимаксным риском. Достоинство этого подхода состоит в том, что он обладает свойством робастности. Оно заключается в следующем: использование минимаксной стратегии σ^M гарантирует, что при любых допустимых параметрах среды максимальные потери будут ограничены величиной минимаксного риска $R_T^M(\Theta)$. Недостатком минимаксного подхода является отсутствие простого прямого способа нахождения минимаксных риска и стратегии.

Другой рассматриваемый в работе подход — байесовский. Если обозначить априорное распределение на множестве параметров среды Θ как $\lambda(\theta)$, то байесовский риск можно искать по формуле

$$R_T^B(\lambda(\theta)) = \min_{\{\sigma\}} \int\limits_{\Theta} L_T(\sigma, \theta)\lambda(\theta)\mathrm{d}\theta,$$

соответствующая стратегия σ^B — байесовская. Плюсом подхода является возможность получения рисков и стратегий численными методами, минусом — необходимость задания априорного распределения.

В работе показано, что наибольшие потери дохода возникают в случае близких математических ожиданий, т.е. $m_1 - m_2 = xT^{-1/2}$, $0 < x < C$. В этом случае потери имеют порядок $\sqrt{T}$. При удаленных математических ожиданиях ($m_1 - m_2 > \Delta m$, $\Delta m > 0$) потери имеют порядок $\ln T$.

Во второй главе рассматриваются пороговые и двухпороговые стратегии управления применительно к случайным средам с бинарными и нормально распределенными доходами.

Пороговая стратегия σ используется для достижения минимаксной цели управления и заключается в следующем. ЛПР применяет действия среды $\ell = 1, 2$ по очереди, накапливая доходы $X_T^\ell(t)$ на каждом действии соответственно, где t — текущее время управления, $t = 1, \ldots, T$, T — горизонт управления. После каждых поочередных применений действий вычисляется абсолютная разность между текущими накопленными доходами $|X_T^1(t) - X_T^2(t)|$. Процесс продолжается до тех пор, пока не истечет время управления или эта величина в некоторый момент времени t^* не превысит порога $\alpha(DT)^{1/2}$, т.е.

$$|X_T^1(t^*) - X_T^2(t^*)| > \alpha(DT)^{1/2}.$$

где α — пороговая константа, $D = 0{,}25$ — максимальная дисперсия одношагового дохода для бинарных случайных сред. Коэффициент перед α рассматривается, исходя из свойства инвариантности функции потерь. Если $X_T^1(t^*) > X_T^2(t^*)$, то на заключительном этапе управления для оставшихся

$t = t^* + 1, \ldots, T$ применяем действие $\ell = 1$; если же $X_T^1(t^*) < X_T^2(t^*)$, то применяем $\ell = 2$. При этом не исключается, что в результате управления может быть исключено и лучшее действие, особенно при близких параметрах среды. После этого вычисляется суммарный накопленный доход

$$X_T = X_T^1 + X_T^2 + X_T^3,$$

где $X_T^1 = X_T^1(t^*)$, $X_T^2 = X_T^2(t^*)$ — доходы на действиях $\ell = 1, 2$, X_T^3 — доход, накопленный за время управления $t^* + 1, \ldots, T$, а затем математическое ожидание суммарного дохода $M(X_T)$. Далее вычисляются потери дохода, равные разности между теоретически максимально возможным значением дохода и математическим ожиданием дохода, полученным в результате управления:

$$L_T(\sigma, \theta) = \max_{\ell=1,2} p_\ell \cdot T - M(X_T).$$

В работе используется свойство инвариантности функции потерь, согласно которому для рассматриваемой пороговой стратегии σ и параметра $\theta = (0{,}5, 0{,}5 - \beta(D/T)^{1/2})$ имеет место приближенное равенство

$$(DT)^{-\frac{1}{2}} L_T(\sigma, \theta) \sim L(\alpha, \beta),$$

если T достаточно велико. Мы используем $T = 10\,000$. Задача сводится к отысканию значения $\min_\alpha \max_\beta L(\alpha, \beta)$ и соответствующих этому минимаксу параметров α и β.

Рассмотренная среда моделируется в программе TwoArmed для заданных диапазонов α и β, в результате была вычислена величина минимаксного риска $\min_\alpha \max_\beta L(\alpha, \beta)$, равная 0,748 при $\alpha = 0{,}57$, $\beta = 3{,}8$. Соответствующая таким параметрам стратегия σ^M выглядит следующим образом. При управлении действия среды применяются по очереди, накапливая доходы X_T^1 и X_T^2. Если в некоторый момент времени t^* величина $|X_T^1 - X_T^2|$ превышает порог $\alpha(DT)^{1/2} = 28{,}5$, то действие, показавшее к этому времени лучшую доходность, применяется оставшееся время $t = t^* + 1, \ldots, T$.

Пример графика функции потерь $L_T(\alpha, \beta)$ представлен на рисунке (1):

Аналогичная задача решена для случайной среды с нормально распределенными доходами с единичными дисперсиями $D_1 = D_2 = 1$, для которой $\min_\alpha \max_\beta L(\alpha, \beta) = 0{,}761$ при $\alpha = 0{,}55$, $\beta = 4$. В качестве горизонта

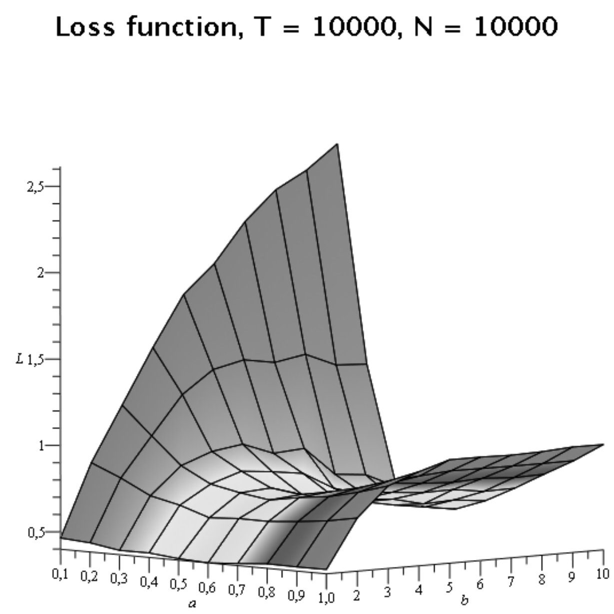

Рис. 1. График функции потерь для случайной среды с бинарно распределенными доходами

управления рассматривалось значение $T = 100$, соответствующие разбиению $T = M \cdot K = 10\,000$ данных на $M = 100$ групп по $K = 100$ данных.

Также в работе проверено, что найденная для бинарных сред стратегия σ^M остается оптимальной для сред с дисперсиями, удовлетворяющих ограничениям $D_0 < D_1 = D_2 < 1$. При моделировании рассматривалась $D_0 = 0{,}09$.

Далее рассмотрен вопрос оптимальности используемого алгоритма в вычислительном плане. Приведено и обосновано два метода оптимизации, сокращающие время расчетов без потери точности результатов. Первый метод позволяет не моделировать заключительный этап управления, а сразу вычислять математическое ожидание дохода на этом этапе. Идея второго метода основана на том, что накопление доходов X^1 и X^2 для текущего параметра среды β можно проводить не для конкретного порога α, а для целого набора порогов $\alpha_1 < \alpha_2 < \ldots < \alpha_I$, I — количество рассматриваемых порогов. Это приводит к значительному выигрышу в плане временных затрат без потери качества управления.

После этого мы представляем модификацию пороговой стратегии, двухпороговую стратегию управления. Идея состоит в следующем. Найденная пороговая стратегия σ^M является оптимальной для рассматриваемых параметров β. Однако при больших β эффективности действий различаются значительно, поэтому лучшее действие определяется достаточно быстро. В таких случаях нет необходимости рассматривать слишком большое значение порога. Проблема в том, что если уменьшить порог, то увеличатся потери в случае близких вероятностей, так как увеличится вероятность выбора неоптимального действия. Это приведет к росту общих потерь на заключительном этапе, когда применяется только одно действие.

Компромиссное решение состоит в том, что вводится дополнительный порог $\alpha_{\mathrm{s}} < \alpha_{\mathrm{fix}}$, действующий до определенного времени T_{s}. Если вероятности выигрыша различаются значительно, то оптимальное действие определяется по достижении порога α_{s}. Если же вероятности близки, то на определение оптимального действия требуется больше времени, и, соответственно, бóльший порог α_{fix}.

Цель управления задается в байесовской формулировке:

$$\int\limits_0^{\beta_0} L_T(\alpha, \beta, \alpha_{\mathrm{s}}, T_{\mathrm{s}}) f(\beta) \mathrm{d}\beta \to \min_{\alpha_{\mathrm{s}}, T_{\mathrm{s}}}. \tag{1}$$

с дополнительными ограничениями:

$$\begin{aligned} &L_T(\alpha, \beta, \alpha_{\mathrm{s}}, T_{\mathrm{s}}) \leqslant \mathfrak{L} + \varepsilon, \\ &\varepsilon > 0, \end{aligned}$$

где $\mathfrak{L} = \min\limits_{\alpha} \max\limits_{\beta} L_T(\alpha, \beta)$, а $f(\beta)$ имеет равномерное распределение.

Двухпороговая стратегия позволяет значительно снижать средние потери на рассматриваемом множестве параметров среды $\beta \in [0; \beta_0]$ по сравнению со случаем фиксированного порога, для которого средние потери можно вычислить по формуле (1), с учетом того, что $\alpha = \alpha_{\mathrm{fix}}$, $\alpha_{\mathrm{s}} = 0$, $T_{\mathrm{s}} = 0$. Результаты моделирования показывают, что для бинарных случайных сред средние потери, вычисленные с помощью двухпороговой стратегии для $\beta_0 = 30$, уменьшаются на 28,52% по сравнению с потерями при фиксированном пороге. Для этого случая найдена оптимальная двухпороговая стратегия при $a_{\mathrm{s}} = 0{,}3$, $T_{\mathrm{s}} = 4\,000$. Для $\beta_0 = 30$ и $T = 10\,000$ стратегия выглядит следующим образом: применять действия по очереди, накапливая доходы X_1, X_2. Если величина $|X_1 - X_2|$ превышает первый порог $\alpha_{\mathrm{s}}(DT)^{1/2} = 15$ в момент времени $t^* < T_{\mathrm{s}}$, то оптимальным является действие, показавшее лучшую доходность, после чего применять его оставшееся время $t = t^* + 1, \ldots, T$. Если же абсолютная разность превышает первый порог, но при $t > T_{\mathrm{s}}$, либо не превышает вовсе, то моделирование продолжать для второго порога $\alpha_{\mathrm{fix}}(DT)^{1/2} = 28{,}5$.

Результаты моделирования для случайных сред с нормально распределенными доходами показывают, что средние потери, вычисленные с помощью двухпороговой стратегии для $\beta_0 = 40$, уменьшаются на 21,7% по сравнению с потерями при фиксированном пороге. Для этого случая также найдена оптимальная двухпороговая стратегия при $a_{\mathrm{s}} = 0{,}34$, $T_{\mathrm{s}} = 55$.

В третьей главе рассмотрена задача об оптимальном управлении в случайной среде, использующая основную теорему теории игр. Минимаксные стратегия и риск ищутся как байесовские, соответствующие наихудшему априорному распределению. В работе рассмотрен случай, когда дисперсии доходов на действиях могут принимать различные значения из множества $D_0 \leqslant D \leqslant 1$, $D_0 > 0$. Поясним, почему имеет смысл рассматривать такую постановку задачи. В случае, когда бинарные доходы имеют близкие математические ожидания, при которых достигается максимум потерь, они имеют и близкие дисперсии. Если же мы оптимизируем, например, не доходы, а время обработки, дисперсии могут различаться достаточно сильно, их значения оцениваются на начальных этапах управления. Поэтому задача нахождения байесовских рисков и стратегий управления для случайных сред с произвольными дисперсиями является актуальной.

Итак, пусть λ — плотность апостериорного распределения. Обозначим через $\widetilde{R}_T^B(\lambda)$ байесовский риск, рассчитанный для доходов с дисперсиями $D_0 \leqslant D \leqslant 1$, $D_0 > 0$, относительно апостериорного распределения с плотностью λ. Тогда справедлива следующая теорема.

Теорема 1. *Байесовский риск $\widetilde{R}_T^B(\lambda)$, рассчитанный для доходов с дисперсиями $D_0 \leqslant D \leqslant 1$, $D_0 > 0$, связан с байесовским риском $R_T^B(\widetilde{\lambda})$, рассчитанным для доходов с $D_1 = D_2 = 1$ следующим соотношением:*

$$\widetilde{R}_T^B(\lambda) = \sqrt{D} R_T^B(\widetilde{\lambda}), \tag{2}$$

где $\widetilde{\lambda} = \widetilde{\lambda}(m_1, m_2) = \lambda(\widetilde{m}_1, \widetilde{m}_2)/D$, $\widetilde{m}_\ell = m_\ell/\sqrt{D}$, $\ell = 1, 2$.

Теорема 1 устанавливает связь между рисками, вычисленными для случая единичных дисперсий и для случая различных попарно равных на действиях дисперсий.

Приведем результаты численной оптимизации, учитывая (2). При этом предполагается, что плотность априорного распределения сосредоточена в двух точках $\widetilde{v} = \pm\widetilde{d}T^{-1/2}$ с вероятностями 0,5, где $\widetilde{d} = d/\sqrt{D}$, $d = |m_1 - m_2|$, m_1, m_2 — математические ожидания доходов на действиях $\ell = 1, 2$: Вычисления проводились с помощью программы Bayes. Наихудшее распределение соответствует максимуму приведенного байесовского риска $r_T(\widetilde{d}) = T^{1/2}\widetilde{R}_T^B$. Для найденной в результате вычисления этого риска стратегии для различных дисперсий вычисляются приведенные потери $l_T(\widetilde{d}) = T^{1/2}\widetilde{L}_T$ (рисунок 2):

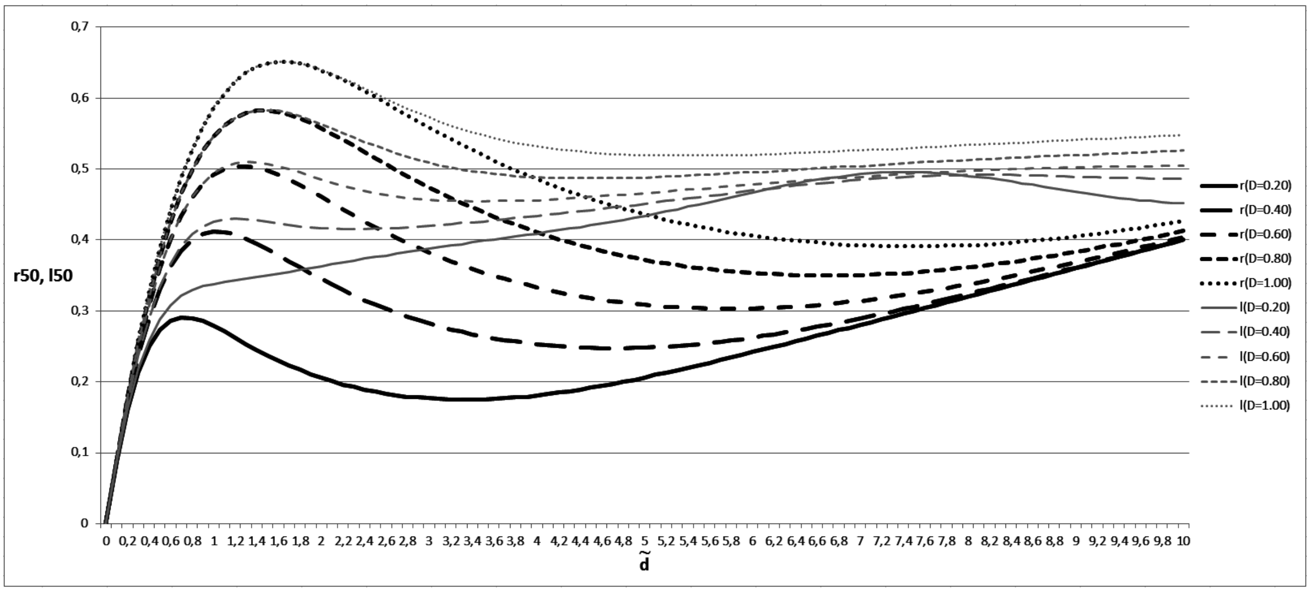

Рис. 2. Риски и потери при $0{,}2 \leqslant D \leqslant 1$

В результате вычислений были найдены оптимальные стратегии для $D =$ 0,2, 0,4, 0,6, 0,8, 1.

Далее рассматривается постановка задачи, при которой дисперсии на действиях могут принимать попарно различные значения из множества $D_0 < D \leqslant 1$, $D_0 > 0$. Получены рекуррентные уравнения, позволяющие находить байесовские риски методами динамического программирования. Если положить $m_1 = u + v$, $m_2 = u - v$, $\theta = (u + v, u - v)$, $\Theta = \{\theta : |v| \leqslant C\}$, то справедливы теоремы.

Теорема 2. *Не уменьшающая байесовский риск плотность распределения при $a \to \infty$ может быть выбрана в виде*

$$\nu_a(u, v) = \kappa_a(u)\rho(v), \tag{3}$$

где $\kappa_a(u)$ — постоянная плотность на отрезке $|u| \leqslant a$, $a \to \infty$.

Теорема 3. *Пусть $\nu_a(u,v)$ выбрана из условия (3) и $a \to \infty$. Тогда*

$$R(Z,t_1,t_2) = \min(R^{(1)}(Z,t_1,t_2), R^{(2)}(Z,t_1,t_2)),$$

где $R^{(1)}(Z,t_1,t_2) = R^{(2)}(Z,t_1,t_2) = 0$ при $t_1 + t_2 = T$ и

$$\begin{aligned} R^{(1)}(Z,t_1,t_2) &= \int\limits_{-C}^{0} 2|v| g(Z,v,t_1,t_2)\rho(v)\mathrm{d}v + \\ &+ \frac{1}{t_2}\int\limits_{-\infty}^{+\infty} R(Z+z,t_1+1,t_2) h^{(1)}\left(\frac{Z-t_1 z}{t_2}, t_1\right)\mathrm{d}z, \\ R^{(2)}(Z,t_1,t_2) &= \int\limits_{0}^{C} 2v g(Z,-v,t_1,t_2)\rho(v)\mathrm{d}v + \\ &+ \frac{1}{t_1}\int\limits_{-\infty}^{+\infty} R(Z+z,t_1,t_2+1) h^{(2)}\left(\frac{Z-t_2 z}{t_1}, t_2\right)\mathrm{d}z \end{aligned} \tag{4}$$

при $2 < t_1 + t_2 < T, t_1 \geqslant 1, t_2 \geqslant 1$. Здесь

$$\begin{aligned} g(Z,v,t_1,t_2) &= \frac{1}{(2\pi D_1 t_1 D_2 t_2 (t_1/D_1 + t_2/D_2))^{1/2}} \times \\ &\times \exp\left(-\frac{(Z + 2v t_1 t_2)^2}{2 D_1 t_1 D_2 t_2 (t_1/D_1 + t_2/D_2)}\right), \end{aligned} \tag{5}$$

$$h^{(\ell)}(z,t_\ell) = \left(\frac{t_\ell + 1}{2\pi D_\ell t_\ell}\right)^{1/2} \exp\left(-\frac{z^2}{2 D_\ell t_\ell (t_\ell + 1)}\right). \tag{6}$$

При этом байесовский риск вычисляется по формуле:

$$\lim_{a\to\infty} R_T^B(\rho(v)) = L(\rho(v)) + \int\limits_{-\infty}^{\infty} R(z,1,1)\mathrm{d}z. \tag{7}$$

Нахождение байесовских рисков для попарно различных дисперсий на действиях осуществлялось по формулам (4)–(7) с учетом того, что наихудшая априорная плотность имеет вид (3). На рисунке 3 приведены графики значений функции рисков $r_T(d)$, вычисленные при $T = 50$, $0 \leqslant d \leqslant 10$ с шагом 0,1 для некоторых дисперсий.

В целях самопроверки моделирование было проведено для попарно одинаковых дисперсий $D_1 = D_2 = 1$ и $D_1 = D_2 = 0{,}6$. Эти значения соответствуют

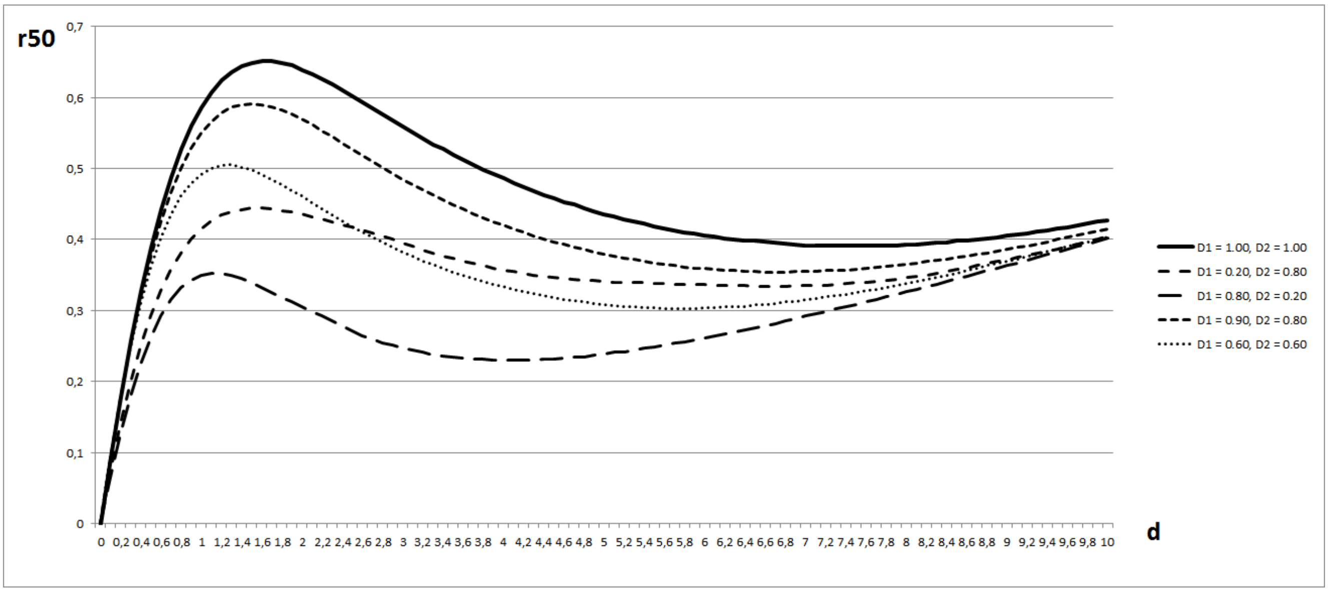

Рис. 3. Риски при некоторых дисперсиях

полученным ранее. По графикам видно, что чем меньше дисперсии доходов, тем меньше потери полного дохода.

Также в работе было проверено, что стратегия, найденная для случая попарно равных различных на действиях дисперсий, остается оптимальной и для случая попарно различных дисперсий.

В четвертой главе описывается программный комплекс, состоящий из программ TwoArmed и Bayes, с помощью которых осуществлялись моделирование всех рассматриваемых в диссертации процессов и вычисление необходимых результатов. Программы были написаны на языке C++ с использованием библиотек графического интерфейса Qt и MFC.

Заключение. В работе представлены результаты исследования задачи параллельной обработки данных в двухальтернативных случайных средах с бинарными и нормально распределенными доходами, имеющими произвольные дисперсии. Показано, что нормально распределенные доходы возникают при использовании параллельной обработки бинарных данных.

Один из разделов работы посвящен пороговой и двухпороговой стратегиям управления, реализующим минимаксный подход к рассматриваемой задаче, преимуществом которого является робастность. С помощью пороговой стратегии и моделирования методом Монте-Карло для бинарных случайных сред с были найдены величины минимаксного риска, соответствующие этим величинам параметры среды, а также оптимальные стратегии управления.

Аналогичные величины были получены для случайной среды с нормально распределенными доходами и единичными дисперсиями. Показано, что соответствующая этому случаю оптимальная стратегия остается оптимальной и для сред с произвольными дисперсиями. Результаты моделирований представлены в графическом и табличном видах.

Также в работе представлены результаты исследований, базирующиеся на основной теореме теории игр. А именно, минимаксные стратегии и риски ищутся как байесовские, соответствующие наихудшему априорному распределению. Установлен вид этого распределения. Использован байесовский подход к задаче, который позволяет вычислять потери методами динамического программирования. Получены рекуррентные уравнения для вычисления байесовских рисков, потерь и оптимальных стратегий управления. Произведено численное моделирование этих стратегий и представлены результаты для сред с произвольными дисперсиями.

Разработан комплекс программ, состоящий из программ TwoArmed и Bayes, который позволяет:

1. Моделировать случайные среды с бинарными и нормально распределенными доходами с различными дисперсиями.

2. Вычислять минимаксные и байесовские риски, потери и стратегии управления в таких средах.

3. Выводить результаты моделирований в формате для последующей визуализации.

Полученные в диссертации результаты могут быть развиты в следующих направлениях: модификация случайной среды для случая трех и более действий, стратегии управления с бо́льшим числом порогов, оптимизация размеров групп данных при параллельной обработке.

Список публикаций в изданиях, рекомендованных ВАК РФ

1. Лазутченко А. Н. Использование двухпороговой стратегии управления в бинарной случайной среде // Современные проблемы науки и образования. 2013. № 3. С. 430.

2. Лазутченко А. Н. Использование двухпороговой стратегии управления в случайной среде с нормально распределенными доходами // Современные проблемы науки и образования. 2014. № 2. С. 666.

3. Лазутченко А. Н. Минимаксная стратегия управления для класса гауссовских случайных сред с различными дисперсиями // Вестник Новгородского государственного университета им. Ярослава Мудрого. 2015. № 3–2 (86). С. 25–28.

4. Лазутченко А. Н. О робастном управлении в случайной среде, характеризуемой доходами с различными дисперсиями // Труды Карельского научного центра Российской академии наук. 2015. № 10. С. 107–113.

Список публикаций в других изданиях

5. Лазутченко А. Н., Колногоров А. В. О робастном управлении в случайной среде, характеризуемой нормальным распределением доходов с различными дисперсиями // Труды 57-й научной конференции МФТИ. Управление и прикладная математика. Т. 1. С. 114–116.

6. Лазутченко А. Н., Колногоров А. В. Минимаксный подход к решению одной из задач о двуруком бандите в случайной среде с нормально распределенными доходами // Труды 58-й научной конференции МФТИ; URL: http://conf58.mipt.ru/static/reports_pdf/214.pdf (дата обращения: 04.09.2016.

7. Лазутченко А. Н. Использование двухпороговой стратегии управления в

бинарной случайной среде // Обозрение прикладной и промышленной математики. 2013. Т. 20. вып. 4. с. 557.

8. Лазутченко А. Н. Задача о целесообразном управлении в случайной среде // Тезисы докладов аспирантов, соискателей, студентов. Часть 3. Великий Новгород. 2009. с. 38.

9. Лазутченко А. Н. Задача о целесообразном управлении в случайной среде // Тезисы докладов аспирантов, соискателей, студентов. Часть 3. Великий Новгород. 2010. С. 39–40.

10. Лазутченко А. Н. Задача о целесообразном поведении в случайной среде // Материалы докладов аспирантов, соискателей, студентов. Часть 3. Великий Новгород. 2011. С. 59–62.

11. Лазутченко А. Н. Оптимизация параллельного управления в случайной среде с нормально распределенными доходами и различными дисперсиями // IX Международная Петрозаводская конференция «Вероятностные методы в дискретной математике», 30 мая — 3 июня 2016 г., Петрозаводск, Россия.