

На правах рукописи

Матвеев Евгений Леонидович

ОПТИМИЗАЦИЯ КВАНТИЛЬНОГО КРИТЕРИЯ  
ПРИ ВЫПУКЛОЙ ЦЕЛЕВОЙ ФУНКЦИИ С ПОМОЩЬЮ  
СТОХАСТИЧЕСКОГО КВАЗИГРАДИЕНТНОГО АЛГОРИТМА

Специальность 05.13.01

Системный анализ, управление и обработка информации  
(авиационная и ракетно-космическая техника)

АВТОРЕФЕРАТ

диссертации на соискание ученой степени  
кандидата физико-математических наук

Москва, 2010

Работа выполнена на кафедре Теории вероятностей Московского авиационного института (государственного технического университета).

Научный руководитель: доктор физико-математических наук,  
профессор Кибзун Андрей Иванович

Официальные оппоненты: доктор физико-математических наук,  
профессор Матасов Александр Иванович  
кандидат физико-математических наук,  
доцент Руденко Евгений Александрович

Ведущая организация: Институт проблем управления им. В.А.  
Трапезникова РАН

Защита состоится “11” июня 2010 г. в 12 ч. 00 мин. на заседании Диссертационного совета Д 212.125.04 Московского авиационного института по адресу: 125993, Москва, А-80, ГСП-3, Волоколамское ш., 4, Ученый совет МАИ.

С диссертацией можно ознакомиться в библиотеке Московского авиационного института (государственного технического университета).

Отзыв на автореферат, заверенный гербовой печатью организации, просьба направлять по указанному адресу в двух экземплярах.

Автореферат разослан “\_\_\_” \_\_\_\_\_ 2010 г.

Ученый секретарь Диссертационного совета Д 212.125.04,  
кандидат физико-математических наук

М.В. Ротанина

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Объект исследования.** Диссертационная работа посвящена анализу и оптимизации конечномерных стохастических систем с квантильным критерием качества.

**Актуальность темы.** Поиск решения в практических задачах часто приходится вести в случае, когда некоторые исходные данные не являются детерминированными, а определяются случайными параметрами с известными законами распределения вероятностей. При детерминированном подходе к моделированию систем влияние случайных факторов не учитывается. Но с практической точки зрения в этом случае может быть потерян реализм присутствующих в задаче явлений. Например, при управлении летательными аппаратами необходимо решать задачу минимизации промаха, характеризующего точность попадания объекта в заданную область фазового пространства. В силу того, что на движение летательного аппарата оказывают воздействие случайные составляющие, задача сводится к минимизации функции со случайными параметрами. Одним из способов решения задач стохастического управления является сведение их к задачам стохастического программирования. Задачам стохастического программирования посвящены работы Дж. Бёржа, Р. Ветса, С. Волласа, Ю. М. Ермольева, П. Калла, Ю. С. Кана, А. И. Кибзуна, В. Купера, Р. Леша, К. Марти, А. В. Назина, Н. М. Новиковой, Б. Т. Поляка, А. Прекопы, Э. Райка, С. П. Урясьева, А. Чарнса, Д. Б. Юдина и многих других авторов.

Естественный, на первый взгляд, путь анализа стохастических систем (замена случайных параметров их средними значениями и нахождение оптимальных управлений полученных таким образом детерминированных задач) не всегда оправдан и может нарушить адекватность модели изучаемого явления. Решение детерминированной задачи с усредненными параметрами может не удовлетворять условиям задачи при различных реализациях случайных факторов. Учёт неопределённости в рамках детерминированного подхода зачастую приводит к чрезвычайно пессимистическим выводам, например, множество допустимых решений может быть пусто. Таким образом, простейшие пути учета случайного характера условий задачи - замена случайных переменных их средними значениями или переход к жесткой постановке - не всегда приводит к осмысленному решению задачи стохастического программирования. В качестве разумного критерия, по которому производится оптимизация, является функция квантили.

Функция квантили характеризует гарантированный при заданном уровне вероятности результат принимаемого решения. При таком подходе маловероятные, наихудшие сочетания параметров исключаются из рассмотрения за счёт того, что риск их появления ограничивается на некотором допустимом уровне. Таким образом, оптимизация стохастических систем с квантильным критерием качества является актуальной проблемой. Исследованием этой проблемы занимались известные российские и зарубежные ученые в области теории управления, такие как Бахшиян Б. Ц., Зубов В. И., Кац И. Я., Кибзун А. И., Колмановский В. Б., Красовский Н. Н., Кузьмин В. П., Лидов М. Л., Матасов А. И., Назиров Р. Р., Черноусько Ф. Л., Эльясберг П. Е., Юби Э., Ярошевский В. А. и др.

При анализе систем в присутствии случайных параметров по квантильному критерию качества возникают определённые сложности.

Во-первых, в большинстве постановок квантильной оптимизации найти аналитическое выражение для критерия при заданном уровне доверительной вероятности, как правило, невозможно. Это связано с неявным определением функции квантили. Кроме того, анализ усложняется нелинейностью квантильного критерия. Во-вторых, если количество опытных данных фиксировано, не понятно, как улучшить точность оценки квантили.

В силу указанных причин при разработке численных методов для вычисления квантили приходится использовать различные аппроксимации. Так как в большинстве случаев точное распределение неизвестно, то вместо значения квантили распределения используется её статистическая оценка. Проблемой статистического оценивания квантили в разное время занимались Галамбош Я., Azzalini A., David H. A., Falk M., Marron J. S., Nadaraya E. A., Nagaraja H. N., Parzen E., Reiss R. D., Shao Y., Sheather S. J., Tamm E., Xiang X., Yang S. и др.

Квантильный критерий относится к группе вероятностных критериев и впервые введен в рассмотрение С. Катаокой. Создание основ теории оптимизационных задач с таким критерием связано с именем эстонского математика Э. Райка и его учеников Р. Леппа, Э. Тамм и Э. Юби. Существенный вклад в развитие этой теории внесли С. Бойд, К. Марти, Домбровский В. В., Ермольев Ю. М., Кан Ю. С., Келли Дж., Кибзун А. И., Красильщиков М. Н., Малышев В. В., Норкин В. И., Панков А. Р., Роечко Н. В., Урясьев С. П. и др.

Задачи стохастического программирования, особенно при оптимизации по квантильному критерию, являются достаточно

сложными. Это объясняется в основном отсутствием аналитического вида квантильного критерия, а также сложностью построения конструктивных численных методов решения подобных задач.

Одним из наиболее эффективных путей решения задач стохастического программирования является использование стохастического квазиградиентного алгоритма. Построением и изучением свойств квазиградиентов занимались Гупал А. М., Ермольев Ю. М., Кан Ю. С., Кибзун А. И., Курбаковский В. Ю., Лепп Р., Михалевич В. С., Норкин В. И., Поляк Б. Т., Роечко Н. В., Третьяков Г., Урясьев С. П., Юби Э. и др.

Таким образом, проблема разработки эффективных методов поиска оптимальных стратегий в задачах стохастического программирования с использованием квантильного критерия качества является актуальной.

**Цель работы.** Целью работы является оптимизация квантильного критерия при выпуклой целевой функции с помощью стохастического квазиградиентного алгоритма. Для достижения данной цели предлагается обосновать стохастический квазиградиентный алгоритм минимизации функции квантили с использованием порядковых статистик, получить универсальное доказательство достаточных условий квазивогнутости вероятностной меры, построить децентрализованную оценку квантили и исследовать её свойства, построить оптимальную ядерную оценку квантили в классе финитных функций с моментными ограничениями, продемонстрировать работоспособность алгоритмов на примере решения задачи оптимизации площади взлётно-посадочной полосы.

**Методы исследования.** Для исследования теоретических проблем использовались современные методы стохастического программирования, теории вероятностей, математической статистики, нелинейного программирования, выпуклого анализа. Для исследования прикладных задач использовались методы компьютерного статистического моделирования.

**Достоверность результатов.** Достоверность результатов обеспечивается:

- 1) строгостью постановок и доказательств утверждений;
- 2) приведением численных расчетов, подтверждающих справедливость результатов;
- 3) рассмотрением конструктивных примеров, которые демонстрируют достоверность приведенных результатов.

**Научная новизна.** В работе получены новые результаты эффективного решения задач квантильной оптимизации, среди которых можно выделить следующие:

1) построена оптимальная ядерная оценка квантили в классе финитных функций с моментными ограничениями;

2) разработан алгоритм получения децентрализованной оценки квантили и найдено её асимптотическое распределение;

3) получено универсальное доказательство достаточных условий квазивогнутости и логарифмической вогнутости вероятностной меры;

4) разработан стохастический квазиградиентный алгоритм минимизации функции квантили и доказана его сходимость с вероятностью единица;

5) продемонстрирована работоспособность алгоритма на примере решения задачи оптимизации площади взлётно-посадочной полосы.

**Практическая значимость.** Диссертация обладает практической значимостью, поскольку полученные результаты позволяют быстро и эффективно решить прикладные задачи квантильной оптимизации.

**Апробация работы.** Результаты работы докладывались на следующих научных конференциях и симпозиумах: международные конференции “Системный анализ, управление и навигация” (Евпатория, 2007, 2008, 2009гг.), а также на научных семинарах в МАИ и ИПУ РАН.

Работа была поддержана грантами РФФИ №05-08-17963, №04-01-00758, №09-08-00369.

**Публикации.** Основные результаты диссертации опубликованы в четырёх статьях [1-4] в журналах из Перечня ВАК, а также в трудах научных конференций [5-6].

**Структура и объем диссертации.** Диссертация состоит из введения, четырех глав, заключения и списка литературы (100 источников). Объем диссертации включает 103 машинописных страниц, включая 7 рисунков.

## СОДЕРЖАНИЕ РАБОТЫ

**Во введении** обоснована актуальность исследуемой проблемы, сформулированы цель и задачи диссертационной работы, описана структура диссертации, перечислены полученные в ней результаты.

**В первой главе** рассматриваются различные оценки квантили. Пусть имеется априорная выборка  $X_i$ ,  $i = \overline{1, n}$  из независимых, одинаково распределенных случайных величин с плотностью

вероятности  $f_X(x)$ . Тогда ядерная оценка квантили, соответствующая ядру  $k(x)$ , определяется выражением

$$\tilde{X}_\alpha(h_n, k(\cdot)) \triangleq \sum_{i=1}^n X_{(i)} \int_{\frac{i-1}{n}}^{\frac{i}{n}} \frac{1}{h_n} k\left(\frac{t-\alpha}{h_n}\right) dt, \quad (1)$$

где  $h_n \rightarrow 0$  при  $n \rightarrow \infty$ , а  $X_{(i)}$  —  $i$ -й элемент вариационного ряда.

Для минимизации среднеквадратической ошибки ядерной оценки квантили сформирован следующий класс  $\Xi_m$  при  $m \geq 2$  финитных функций с моментными ограничениями.

$$\Xi_m \triangleq \left\{ k(x) : \text{supp}[k(x)] = [-1, 1], \int_{-1}^1 k(x) dx = 1, \int_{-1}^1 x^i k(x) dx = 0 \right. \\ \left. \text{для } i = \overline{1, m-1}, \left| \int_{-1}^1 x^m k(x) dx \right| < \infty \right\}.$$

Введено понятие оптимального ядра в классе  $\Xi_m$  с точки зрения минимума среднеквадратичного отклонения.

*Определение 1.1.* Назовем оптимальным в классе  $\Xi_m$  ядро  $k^*(\cdot) \in \Xi_m$ , если

а)  $\mathbf{M}[|\tilde{X}_\alpha(h_n, k^*(\cdot)) - x_\alpha|] \rightarrow 0, n \rightarrow \infty,$

б)  $\mathbf{M}[|\tilde{X}_\alpha(h_n, k^*(\cdot)) - x_\alpha|^2] \leq \mathbf{M}[|\tilde{X}_\alpha(h_n, \tilde{k}(\cdot)) - x_\alpha|^2]$  для всех  $\tilde{k}(\cdot) \in \Xi_m, n \rightarrow \infty.$

Установлен вид оптимального в смысле определения 1.1 ядра для ядерной оценки квантили.

*Теорема 1.1.* Пусть плотность вероятности  $f_X(x)$  дифференцируема в окрестности точки  $\alpha \in (0, 1)$  и  $f_X(x_\alpha) > 0$ . Тогда

а) оптимальное в классе  $\Xi_m$  ядро  $k^*(\cdot)$  будет иметь вид

$$k^*(x) \triangleq \frac{1}{2} - \frac{5}{4} R_2(x) = \frac{1}{2} - \frac{5}{8} (3x^2 - 1), \quad (2)$$

где  $R_2(x)$  — полином Лежандра 2-й степени;

б) смещение оценки определяется соотношением

$$\mathbf{M}[\tilde{X}_\alpha(h_n, k^*(\cdot)) - x_\alpha] = o(h_n^2) + o(n^{-1}); \quad (3)$$

в) среднеквадратичная погрешность определяется соотношением

$$\mathbf{M}[|\tilde{X}_\alpha(h_n, k^*(\cdot)) - x_\alpha|^2] = \frac{1}{n} \alpha(1-\alpha)(x'_\alpha)^2 - \frac{3}{56} h_n (x'_\alpha)^2 + o(n^{-1} h_n), \quad (4)$$

где  $\tilde{X}_\alpha(h_n, k^*(\cdot))$  — ядерная оценка типа (1),  $x'_\alpha \triangleq \frac{1}{f_X(x_\alpha)}$ .

Кроме того, приводится децентрализованная оценка, которая позволяет децентрализовать процесс нахождения квантили. Алгоритм базируется на использовании многопроцессорной архитектуры и позволяет находить значение квантили при достаточно общих ограничениях.

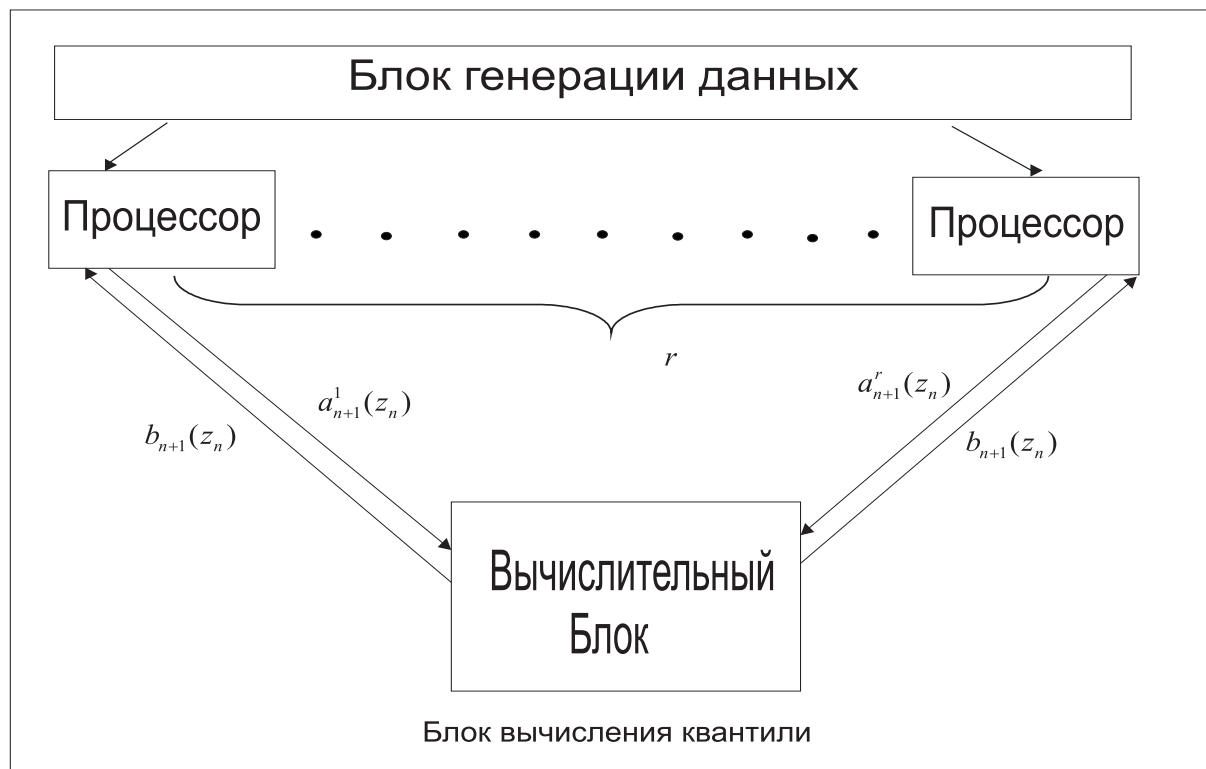


Рис. 1

Предположим, что в нашем распоряжении имеется вычислительная сеть (см. рис. 1), состоящая из блока генерации данных, генерирующего независимые наблюдения и передающего их для обработки процессорам, процессоров, которые накапливают информацию о выборке и вычислительного блока, обрабатывающего данные, поступающие от процессоров.

Тогда децентрализованный алгоритм оценивания квантили уровня  $\alpha \in (0, 1)$  выглядит следующим образом:

*А л г о р и т м 1.*

1.  $n = 0$ . Сформировать выборку  $x_1, x_2, \dots$  случайной величины  $X$  с плотностью вероятности  $f_X(x) > 0$ . Выбрать числовую последовательность  $\rho_n$ ,  $n = 1, 2, \dots$ , удовлетворяющую следующим условиям

$$\rho_n > 0, \sum_{n=0}^{\infty} \rho_n = \infty, \sum_{n=0}^{\infty} \rho_n^2 < \infty.$$



Выбрать константу  $\eta_r > 0$  и начальное значение квантили  $z_0$ , одинаковое для всех  $r$  процессоров.

2. Вычислить значения индикаторных функций  $a_{n+1}^i(z_n), i = 1, \dots, r$

$$a_{n+1}^i(z_n) \triangleq \begin{cases} 1, & \text{если } x_{r \cdot n + i} \leq z_n; \\ 0, & \text{иначе,} \end{cases} \quad (5)$$

3. Вычислить усреднённую величину

$$b_{n+1}(z_n) \triangleq \begin{cases} 1, & \text{если } \frac{1}{r} \sum_{i=1}^r a_{n+1}^i(z_n) \leq \alpha; \\ 0, & \text{иначе,} \end{cases} \quad (6)$$

4. Вычислить новую оценку квантили  $z_{n+1}$  следующим образом

$$z_{n+1} = z_n + \rho_n \eta_r (b_{n+1}(z_n) - \beta), \quad (7)$$

где

$$\beta \triangleq \sum_{i=0}^{[r\alpha]} C_r^i \alpha^i (1 - \alpha)^{r-i}. \quad (8)$$

5. Положить  $n = n + 1$  и перейти к пункту 2.

Другими словами, предположим, что блок генерации данных содержит независимые наблюдения случайной величины  $X$  с плотностью вероятностей  $f_X(x) > 0$ . На каждом шаге  $n = 0, 1, \dots$ , блок генерации данных передаёт  $(r \cdot n + i)$ -е наблюдение случайной величины  $X$  в каждый  $i$ -й процессор  $i \in \{1, \dots, r\}$ . Далее каждый процессор  $i = 1, 2, \dots, r$  передаёт 1 бит  $a_{n+1}^i(z_n)$ , вычисленный по определённому алгоритму, описанному ниже, в вычислительный блок. Вычислительный блок обрабатывает полученную информацию и передает 1 бит  $b_{n+1}(z_n)$ , вычисляемый определённым образом, назад каждому процессору в которых уточняется текущая оценка квантили  $z_n$ .

1. Каждый  $i$ -й процессор вычисляет значение индикаторной функции

$$a_{n+1}^i(z_n) \triangleq \begin{cases} 1, & \text{если } x_{n+1}(i) \leq z_n; \\ 0, & \text{иначе,} \end{cases} \quad (9)$$

используя реализацию  $x_{n+1}(i)$  случайной величины  $X_{n+1}$ , которую он получил из блока генерации данных, а затем передаёт её в вычислительный блок.

2. Вычислительный блок рассчитывает значение индикаторной функции по усреднённой величине, характеризующей полученную информацию от всех процессоров.

$$b_{n+1}(z_n) \triangleq \begin{cases} 1, & \text{если } \frac{1}{r} \sum_{i=1}^r a_{n+1}^i(z_n) \leq \alpha; \\ 0, & \text{иначе,} \end{cases} \quad (10)$$

и передаёт её обратно каждому процессору.

3. Каждый процессор вычисляет новую оценку квантили  $z_{n+1}$  по формуле

$$z_{n+1} = z_n + \rho_n \eta_r (b_{n+1}(z_n) - \beta), \quad (11)$$

где  $\beta$  определена в (8).

*Теорема 1.2. Рассмотрим последовательность  $Z_n$ , полученную с помощью описанного выше алгоритма. Тогда*

1) для произвольного начального состояния  $z_0$ :

$$Z_n \xrightarrow{n.н.} x_\alpha \text{ при } n \rightarrow \infty;$$

2) если константа  $\eta_r$  выбрана таким образом, что

$$\eta_r > \frac{1}{2f_X(x_\alpha) r C_{r-1}^{[r\alpha]} \alpha^{[r\alpha]} (1-\alpha)^{r-[r\alpha]-1}}, \quad (12)$$

то

$$\sqrt{n} (Z_n - x_\alpha) \xrightarrow{d} \mathcal{N} \left( 0, \frac{\eta_r^2 \beta (1-\beta)}{2\eta_r f_X(x_\alpha) r C_{r-1}^{[r\alpha]} \alpha^{[r\alpha]} (1-\alpha)^{r-[r\alpha]-1} - 1} \right), \quad (13)$$

где  $\beta$  определена в (8).

**Во второй главе** изучаются свойства выпуклости функции квантили. В основе результатов лежит понятие  $\gamma$ -вогнутой функции, которое является обобщением понятия выпуклой функции.

Пусть дана функция  $\Phi(u, x) : \mathbb{R}^m \times \mathbb{R}^s \rightarrow \mathbb{R}^1$ , измеримая по  $X$  и имеющая смысл функции потерь, которую необходимо минимизировать. Вектор  $u$  размерности  $m$  имеет смысл управления,  $X$  — случайный вектор с реализациями в  $\mathbb{R}^s$  и плотностью вероятности  $f_X(x)$ .

Функции вероятности и квантили имеют вид

$$P_\varphi(u) \triangleq \mathcal{P}\{X : \Phi(u, X) \leq \varphi\}, \quad \varphi_\alpha(u) \triangleq \min\{\varphi : P_\varphi(u) \geq \alpha\},$$

где  $\varphi \in \mathbb{R}^1$ ,  $\alpha \in (0, 1)$ , а  $\mathcal{P}(\cdot)$  – вероятностная мера на  $\mathbb{R}^s$ , определённая с помощью плотности  $f_X(x)$ .

Функция вероятности равна вероятности того, что потери при выбранной стратегии  $u$  не превысят заданный порог  $\varphi$ . Функция квантили показывает, что потери при выбранной стратегии  $u$  с вероятностью, не меньшей  $\alpha$ , не превысят значения  $\varphi_\alpha(u)$ .

Основным требованием, обеспечивающим вогнутость функции вероятности, является квазивогнутость вероятностной меры и выпуклость целевой функции. В диссертации используются следующие понятия.

*Определение 1.2. (Прекопа)* Вероятностная мера  $\mathcal{P}$  называется квазивогнутой на  $\mathbb{R}^s$ , если для любых двух выпуклых непустых множеств  $A, B \in \mathbb{R}^s$  и для любого  $\lambda \in (0, 1)$  выполняется неравенство

$$\mathcal{P}(\lambda A \oplus (1 - \lambda)B) \geq \min(\mathcal{P}(A), \mathcal{P}(B)), \quad (14)$$

где  $A \oplus B$  – обозначает сумму множеств  $A$  и  $B$  по Минковскому, т.е.  $A \oplus B \triangleq \{x + y : x \in A, y \in B\}$ .

*Определение 1.3. (Прекопа)* Вероятностная мера  $\mathcal{P}$  называется логарифмически вогнутой на  $\mathbb{R}^s$ , если для любых двух выпуклых непустых множеств  $A, B \in \mathbb{R}^s$  и для любого  $\lambda \in (0, 1)$  выполняется неравенство

$$\mathcal{P}(\lambda A \oplus (1 - \lambda)B) \geq \mathcal{P}^\lambda(A) \mathcal{P}^{1-\lambda}(B). \quad (15)$$

*Теорема 1.3. (Кан, Кибзун)* Если функция  $\Phi(u, x)$  выпукла на выпуклом множестве  $U \times \mathcal{X} \subset \mathbb{R}^m \times \mathbb{R}^s$  и вероятностная мера  $\mathcal{P}$  квазивогнута, то функция квантили  $\varphi_\alpha(u)$  является выпуклой на  $U$ .

Для получения достаточных условий для обеспечения квазивогнутости вероятностной меры используется понятие  $\gamma$ -вогнутой функции.

*Определение 1.4. (Норкин, Роечко)* Неотрицательная функция  $f(\cdot) : \mathbb{R}^s \rightarrow \mathbb{R}_+^1$  называется  $\gamma$ -вогнутой на выпуклом множестве  $A \subset \mathbb{R}^s$  если для любого  $\lambda \in [0, 1]$  и для любых  $x, y \in A$  выполняются неравенства

1. если  $\gamma > 0$ , то

$$f^\gamma(\lambda x + (1 - \lambda)y) \geq \lambda f^\gamma(x) + (1 - \lambda)f^\gamma(y);$$

2. если  $\gamma < 0$ , то

$$f^\gamma(\lambda x + (1 - \lambda)y) \leq \lambda f^\gamma(x) + (1 - \lambda)f^\gamma(y);$$

3. если  $\gamma = 0$ , то

$$f(\lambda x + (1 - \lambda)y) \geq f^\lambda(x)f^{1-\lambda}(y).$$

Приведено достаточное условие, обеспечивающее квазивогнутость вероятностной меры.

*Теорема 1.4.* Пусть существует  $\gamma \in [-\frac{1}{s}, \infty) \setminus \{0\}$  такое, что плотность вероятности  $f_X(x)$  с выпуклым носителем  $\mathcal{X} \subset \mathbb{R}^s$  оказывается  $\gamma$ -вогнутой на  $\mathcal{X}$ . Тогда вероятностная мера  $\mathcal{P}$ , порождённая плотностью  $f_X(x)$ , квазивогнута.

Доказательство этого утверждения было приведено у нескольких специалистов: Borell, Brascamp, Das Gupta, Prékopa, Rinott. Однако доказательства содержат ряд неточностей, о которых указано в диссертации. В диссертации приведено новое доказательство данного утверждения, лишённое этих недостатков. Доказательство носит достаточно универсальный характер и позволяет использовать его с небольшими модификациями для нахождения достаточных условий логарифмической вогнутости вероятностной меры.

*Теорема 1.5.* Пусть плотность вероятности  $f_X(x)$  логарифмически вогнута на  $\mathbb{R}^s$ . Тогда вероятностная мера  $\mathcal{P}$ , порождённая плотностью  $f_X(x)$ , является логарифмически вогнутой.

На основе теорем 1.3, 1.4 и 1.5 получены достаточные условия для выпуклости функции квантили.

*Теорема 1.6.* Если функция  $\Phi(u, x)$  выпукла на выпуклом множестве  $U \times \mathcal{X} \subset \mathbb{R}^m \times \mathbb{R}^s$  и существует  $\gamma \in [-\frac{1}{s}, \infty)$  такое, что плотность вероятности  $f_X(x)$  с выпуклым носителем  $\mathcal{X} \subset \mathbb{R}^s$  оказывается  $\gamma$ -вогнутой на  $\mathcal{X}$ , то функция квантили  $\varphi_\alpha(u)$  является выпуклой на  $U$ .

**Третья глава** посвящена стохастическому квазиградиентному алгоритму минимизации функции квантили.

Рассматривается задача стохастического программирования с использованием квантильного критерия

$$u_\alpha = \arg \min_{u \in U} \varphi_\alpha(u), \quad (16)$$

где  $U \subset \mathbb{R}^m$  — выпуклый компакт. Предположим, что точка минимума функции  $\varphi_\alpha(u)$  на множестве  $U$  единственна.

Пусть для любого  $u \in U \subset \mathbb{R}^m$  есть возможность моделировать серию независимых выборок  $\{\Phi(u, X_i)\}_{i=1}^{n_k}$  функции  $\Phi(u, X)$ . Обозначим через

$\{\Phi_{(i)}(u)\}_{i=1}^{n_k}$  вариационный ряд выборки, т.е. расположенные в порядке возрастания значения

$$\Phi_{(1)}(u) \leq \Phi_{(2)}(u) \leq \dots \leq \Phi_{(n_k)}(u).$$

На основе данного вариационного ряда сконструируем выборочную оценку квантили

$$\widehat{\varphi}_k(u) \triangleq \Phi_{([\![n_k \alpha]\!] + 1)}(u). \quad (17)$$

*Определение 1.5.* Назовем стохастическим квазиградиентом функции квантили случайный вектор

$$\xi_k(u, \delta_k) \triangleq \frac{1}{2\delta_k} \sum_{i=1}^m [\widehat{\varphi}_k(\tilde{u}_1, \dots, u_i + \delta_k, \dots, \tilde{u}_m) - \widehat{\varphi}_k(\tilde{u}_1, \dots, u_i - \delta_k, \dots, \tilde{u}_m)] e_i, \quad (18)$$

где  $\tilde{u}_i$  независимые равномерно распределенные случайные величины на отрезке  $[u_i - \delta_k, u_i + \delta_k]$ .

Рассмотрен стохастический квазиградиентный алгоритм минимизации функции квантили.

*А л г о р и т м 2.*

1.  $k = 0$ . Положить начальное значение  $u^{(0)} \sim \mathcal{R}(U)$ . Выбрать  $\varepsilon > 0$ , константу  $L$  таким образом, что почти наверное выполняется соотношение  $u_\alpha \in \{u : \|\xi_k(u, \delta_k)\| \leq L\}$  и числовую последовательность  $\rho_k$ .
2. Вычислить новое значение алгоритма следующим образом

$$u^{(k+1)} = \begin{cases} \Pi_U [u^{(k)} - \rho_k \xi_k(u^{(k)}, \delta_k)], & \|\xi_k(u^{(k)}, \delta_k)\| \leq L, \\ \tilde{u}, & \|\xi_k(u^{(k)}, \delta_k)\| > L, \end{cases} \quad (19)$$

где  $\tilde{u}$  имеет равномерное распределение на множестве  $U$ , а  $\Pi_U$  оператор рандомизированного проецирования на область  $U_1$ , т.е. отображение  $\Pi_U : \mathbb{R}^m \rightarrow U_1$ , задаваемое соотношением

$$\Pi_U(u) \triangleq \begin{cases} u, & u \in U, \\ \hat{u}, & u \notin U, \end{cases} \quad (20)$$

где случайная величина  $\hat{u}$  имеет равномерное распределение на шаре

$$B_\varepsilon \triangleq \left\{ u_1 \in U_1 : \left\| u_1 - \arg \min_{v \in U} \|u - v\| \right\| \leq \varepsilon \right\},$$

а  $U_1 \triangleq \{u \in \mathbb{R}^m : \exists v \in U \text{ такое, что } \|u - v\| \leq \varepsilon\}$ .

3. Положить  $k = k + 1$  и перейти к шагу 2.

Доказана сходимость данного алгоритма к минимальному значению функции квантили на выпуклом компакте  $U$ .

*Теорема 1.7. Пусть последовательность  $\{u^{(k)}\}$  получена в результате описанного выше алгоритма и выполнены следующие условия:*

- (i) функция квантили  $\varphi_\alpha(u)$  выпукла, удовлетворяет условию Липшица на выпуклом компакте  $U_1$  и почти наверное выполняется условие  $u_\alpha \in \{u : \|\xi_k(u, \delta_k)\| \leq L\}$ ;
- (ii) для каждого  $u \in \text{int}(U_1)$  функция вероятности  $P_\varphi(u)$  непрерывно дифференцируема по  $\varphi$  и  $\nabla_\varphi P_\varphi(u) > 0$  для всех  $\varphi \in \mathbb{R}^1$ ;
- (iii) последовательность  $\rho_k$  удовлетворяет условиям

$$\rho_k > 0, \quad \sum_{k=1}^{\infty} \rho_k = \infty, \quad \sum_{k=1}^{\infty} \rho_k^2 < \infty;$$

- (iv) последовательность  $\delta_k$  удовлетворяет условиям

$$\delta_k \rightarrow 0, \quad \sum_{k=1}^{\infty} \frac{1}{n_k \delta_k^2} < \infty, \quad \sum_{k=1}^{\infty} \frac{\rho_k}{\delta_k \sqrt{n_k}} < \infty;$$

- (v) существует  $\gamma > 0$ , удовлетворяющее условию  $\gamma > \max \left\{ \frac{1}{n_k - [n_k \alpha]}, \frac{1}{[n_k \alpha] + 1} \right\}$ , такое, что

$$\mathbf{M}[|\Phi(u, X)|^\gamma] < \infty \text{ для всех } u \in U_1;$$

- (vi)  $\text{mes} \{(u, x) \in U_1 \times \mathbb{R}^s : \Phi(u, x) = \varphi\} = 0$  для всех  $\varphi \in \mathbb{R}^1$ .

Тогда  $u^{(k)} \xrightarrow{n.k.} u_\alpha$ , где последовательность  $u^{(k)}$  образована алгоритмом (24), а  $u_\alpha$  удовлетворяет (16).

Также приведён стохастический квазиградиентный алгоритм минимизации функции квантили на основе децентрализованного алгоритма оценивания квантили, который приведён в первой главе диссертации.

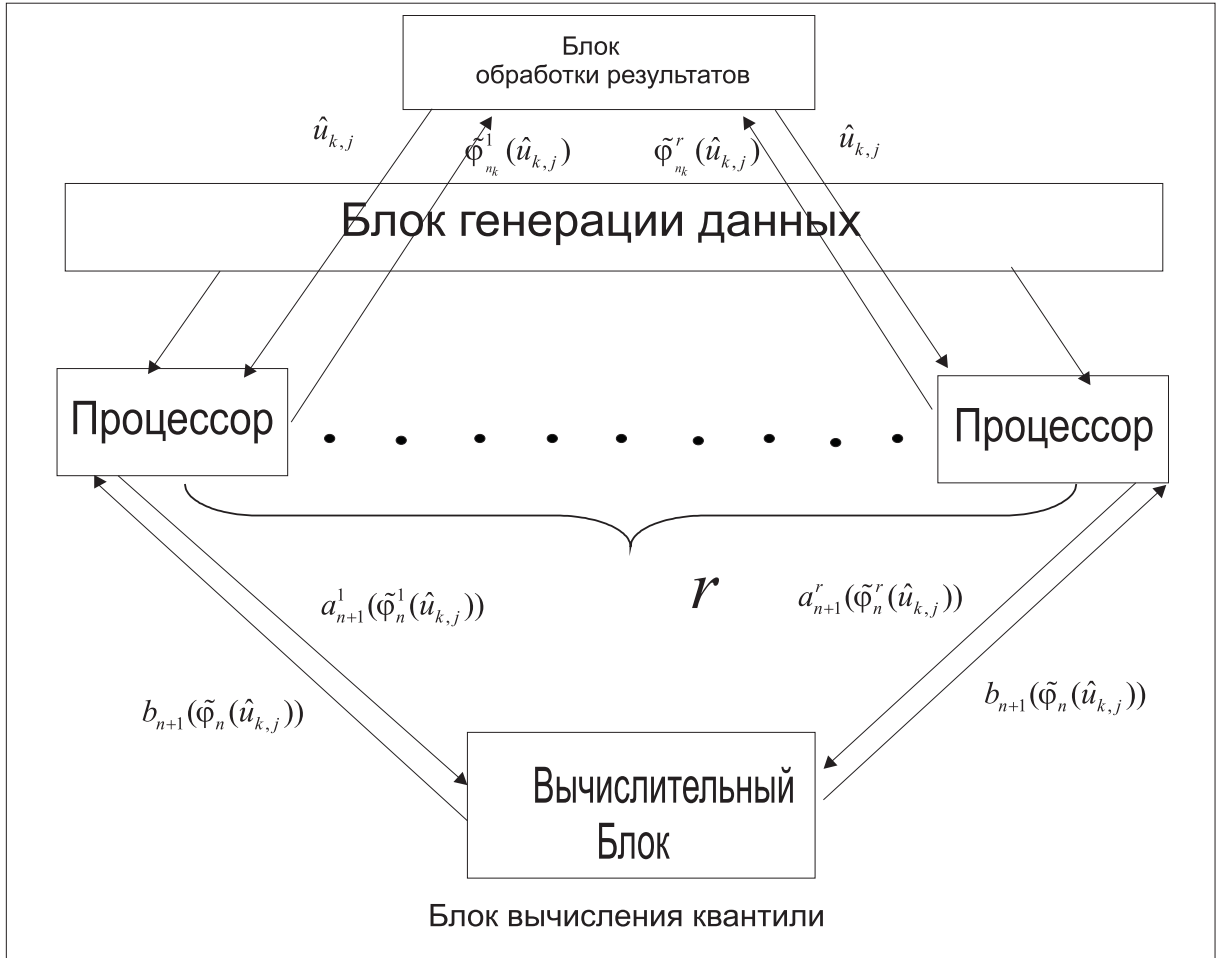


Рис. 2

Предположим, что в нашем распоряжении имеется вычислительная сеть (см. рис. 2) состоящая из двух основных блоков: блока обработки результатов и блока вычисления квантили, состоящего из  $r$  процессоров, блока генерации данных и вычислительного блока. В блоке обработки результатов происходит уточнение текущей оценки минимума функции квантили, а в блоке вычисления квантили по текущему значению  $u^{(k)}$  происходит вычисление новой оценки квантили, согласно алгоритму, описанному в первой главе диссертации.

Стохастический квазиградиентный алгоритм минимизации функции квантили при помощи децентрализованной оценки квантили выглядит следующим образом.

*А л г о р и т м 3.*

1.  $k = 0$ . Положить начальное значение  $u^{(0)} \sim \mathcal{R}(U)$ . Выбрать  $\varepsilon > 0$ , константу  $L$  таким образом, что почти наверное выполняется соотношение  $u_\alpha \in \{u : \|\xi_k(u, \delta_k)\| \leq L\}$  и числовые последовательности  $\rho_k > 0, n_k > 0, \delta_k > 0$  таким образом, чтобы

ВЫПОЛНЯЛИСЬ СООТНОШЕНИЯ

$$\sum_{k=1}^{\infty} \rho_k = \infty, \sum_{k=1}^{\infty} \rho_k^2 < \infty, \delta_k \rightarrow 0, \sum_{k=1}^{\infty} \frac{1}{n_k \delta_k^2} < \infty, \sum_{k=1}^{\infty} \frac{\rho_k}{\delta_k \sqrt{n_k}} < \infty.$$

2. Положить  $\widehat{u}^{(k)} = u_{1,j}^{(k)}$ , где  $u_{1,j}^{(k)} \triangleq \text{col} \left( \tilde{u}_1^{(k)}, \dots, u_j^{(k)} + \delta_k, \dots, \tilde{u}_r^{(k)} \right)$ , а  $\tilde{u}_l^{(k)}$  независимые равномерно распределенные величины на отрезке  $[u_l^{(k)} - \delta_k, u_l^{(k)} + \delta_k]$ ,  $l = \overline{1, m}$ ,  $l \neq j$ .

3. Сформировать выборку  $x_1, x_2, \dots$  случайной величины  $X$  с плотностью вероятности  $f_X(x) > 0$ . Выбрать константу  $\eta_r > 0$  и начальное значение квантили  $\tilde{\varphi}_0^i(\widehat{u}^{(k)})$ , одинаковое для всех  $r$  процессоров.

4. Вычислить значения индикаторных функций  $a_{n+1}^i(\tilde{\varphi}_n^i(\widehat{u}^{(k)}))$ ,  $i = 1, \dots, r$

$$a_{n+1}^i(\tilde{\varphi}_n^i(\widehat{u}^{(k)})) \triangleq \begin{cases} 1, & \text{если } \Phi(\widehat{u}^{(k)}, x_{r \cdot n + i}) \leq \tilde{\varphi}_n^i(\widehat{u}^{(k)}); \\ 0, & \text{иначе,} \end{cases} \quad (21)$$

5. Вычислить усреднённую величину

$$b_{n+1}(\tilde{\varphi}_n(\widehat{u}^{(k)})) \triangleq \begin{cases} 1, & \text{если } \frac{1}{r} \sum_{i=1}^r a_{n+1}^i(\tilde{\varphi}_n^i(\widehat{u}^{(k)})) \leq \alpha; \\ 0, & \text{иначе,} \end{cases} \quad (22)$$

6. Вычислить новую оценку квантили  $\tilde{\varphi}_{n+1}^i(\widehat{u}^{(k)})$  следующим образом

$$\tilde{\varphi}_{n+1}^i(\widehat{u}^{(k)}) = \tilde{\varphi}_n^i(\widehat{u}^{(k)}) + \rho_n \eta_r \left( b_{n+1}(\tilde{\varphi}_n(\widehat{u}^{(k)})) - \beta \right), \quad (23)$$

где  $\beta$  определена в (8).

7. Если  $n < n_k$ , то положить  $n = n + 1$  и перейти к пункту 4. В противном случае перейти к пункту 8.

8. Вычислить

$$\tilde{\varphi}_{n_k}(\widehat{u}^{(k)}) = \frac{1}{r} \sum_{i=1}^r \tilde{\varphi}_{n_k}^i(\widehat{u}^{(k)}).$$

9. Вычислить

$$u_{2,j}^{(k)} \triangleq \text{col} \left( \tilde{u}_1^{(k)}, \dots, u_j^{(k)} - \delta_k, \dots, \tilde{u}_r^{(k)} \right),$$



где  $\tilde{u}_l^{(k)}$  независимые равномерно распределенные случайные величины на отрезке  $[u_j^{(k)} - \delta_k, u_j^{(k)} + \delta_k]$ ,  $l = \overline{1, m}, l \neq j$ . После этого шаги 3 – 8 повторить для  $\hat{u}^{(k)} = u_{2,j}^{(k)}$ .

10. Вычислить

$$\xi_k(u^{(k)}, \delta_k) = \frac{1}{2\delta_k} \sum_{j=1}^m [\tilde{\varphi}_{n_k}(u_{1,j}^{(k)}) - \tilde{\varphi}_{n_k}(u_{2,j}^{(k)})] e_j.$$

11. Вычислить новое значение алгоритма следующим образом

$$u^{(k+1)} = \begin{cases} \Pi_U [u^{(k)} - \rho_k \xi_k(u^{(k)}, \delta_k)], & \|\xi_k(u^{(k)}, \delta_k)\| \leq L, \\ \tilde{u}, & \|\xi_k(u^{(k)}, \delta_k)\| > L, \end{cases} \quad (24)$$

где  $\tilde{u}$  имеет равномерное распределение на множестве  $U$ , а  $\Pi_U$  оператор рандомизированного проецирования на область  $U_1$ , т.е. отображение  $\Pi_U : \mathbb{R}^m \rightarrow U_1$ , задаваемое соотношением

$$\Pi_U(u) \triangleq \begin{cases} u, & u \in U, \\ \hat{u}, & u \notin U, \end{cases} \quad (25)$$

где случайная величина  $\hat{u}$  имеет равномерное распределение на шаре

$$B_\varepsilon \triangleq \left\{ u_1 \in U_1 : \left\| u_1 - \arg \min_{v \in U} \|u - v\| \right\| \leq \varepsilon \right\},$$

а  $U_1 \triangleq \{u \in \mathbb{R}^m : \exists v \in U \text{ такое, что } \|u - v\| \leq \varepsilon\}$ .

12. Положить  $k = k + 1$  и перейти к шагу 2.

Глава завершается рассмотрением примера оптимизации площади взлётно-посадочной полосы. Данная задача рассматривалась ранее в работах Кана, Кибзуна, Курбаковского, Малышева.

Как показано у Кана, задача оптимизации площади эквивалентна следующей задаче квантильной оптимизации.

$$\varphi_\alpha(u_1, u_2) \rightarrow \min_{u_1, u_2 > 0}, \quad (26)$$

где  $\varphi_\alpha(u_1, u_2)$  — функция квантили уровня  $\alpha$  для функции потерь  $\Phi(u_1, u_2, X, Z)$ , где

$$\Phi(u_1, u_2, X, Z) \triangleq \max \{ \Phi_1(u_1, u_2, X), \Phi(u_1, u_2, X), \Phi_3(u_2, Z) \}, \quad (27)$$

а функции  $\Phi_1(u_1, u_2, X)$ ,  $\Phi_2(u_1, u_2, X)$ ,  $\Phi_3(u_2, Z)$  определяются следующим образом.

$$\Phi_1(u_1, u_2, X) \triangleq \frac{-X(1 + u_1)/u_1 + l_0}{\sqrt{u_2}}, \quad (28)$$

$$\Phi_2(u_1, u_2, X) \triangleq \frac{X(1 + u_1) + l_0}{\sqrt{u_2}}, \quad (29)$$

$$\Phi_3(u_2, Z) \triangleq 2|Z|\sqrt{u_2}. \quad (30)$$

Оптимизация функции квантили проводилась двумя различными способами – квазиградиентным алгоритмом на основе выборочной оценки квантили (выборочный) и квазиградиентным алгоритмом на основе децентрализованной оценки квантили (параллельный). Значения стратегий и соответствующих им критериев сравниваются с аналитически гарантированным значением, полученным в работе Кана (аналитический).

$\alpha$	Метод	$u_1$	$u_2$	Площадь, км <sup>2</sup>
0.99	выборочный	1,9189	27,014	0,173
	параллельный	1,9210	27,098	0,169
	аналитический	1	23.15	0.192
0.999	выборочный	1,424	22,043	0,2348
	параллельный	1,416	22,132	0,2361
	аналитический	1	20.12	0.25
0.9999	выборочный	2,132	21,35	0,308
	параллельный	2,150	21,236	0,310
	аналитический	1	18.32	0.304

На основании полученных данных можно сделать вывод, что оценка, полученная при использовании стохастического квазиградиентного алгоритма на основе децентрализованной оценки квантили, при малых объемах выборки ведет себя более устойчиво чем на основе выборочной, а при возрастании объема выборки они сближаются. Также можно сказать, что при достаточно высоком уровне вероятности  $\alpha$  (больше 0,9999) оценка, полученная с помощью аналитического метода не сильно отличается от точного решения (первые два алгоритма).

### ОСНОВНЫЕ РЕЗУЛЬТАТЫ, ВЫНОСИМЫЕ НА ЗАЩИТУ

1) Построена оптимальная ядерная оценка квантили в классе финитных функций с моментными ограничениями [1,5].

2) Разработан алгоритм получения децентрализованной оценки квантили и найдено её асимптотическое распределение [2].

3) Получено универсальное доказательство достаточных условий квазивогнутости и логарифмической вогнутости вероятностной меры [3].

4) Разработан стохастический квазиградиентный алгоритм минимизации функции квантили и доказана его сходимость с вероятностью единица [6].

5) Продемонстрирована работоспособность алгоритма на примере решения задачи оптимизации площади взлётно-посадочной полосы.

#### **Публикации в журналах из перечня ВАК**

1) Кибзун А. И., Матвеев Е. Л. Оптимизация функции квантили на основе ядерных оценок // Автоматика и Телемеханика, 2007, № 1, С. 187 – 189.

2) Кибзун А. И., Матвеев Е. Л. Алгоритм распараллеливания процесса оптимизации функции квантили // Вестник МАИ, 2008, Т.15, №2, С. 51 – 59.

3) Кибзун А. И., Матвеев Е. Л. Достаточные условия квазивогнутости функции вероятности // Автоматика и Телемеханика, 2010, № 3, С. 54 – 71.

4) Кибзун А. И., Матвеев Е. Л. Стохастический квазиградиентный алгоритм минимизации функции квантили // Автоматика и Телемеханика, 2010, № 6, С. 64 – 78.

#### **Публикации в других изданиях**

5) Кибзун А. И., Матвеев Е. Л. Квантильная оптимизация на основе ядерных оценок // Проектирование и изготовление аэрокосмических аппаратов М.: Изд-во МАИ, 2006, С. 276 – 283.

6) Кибзун А. И., Матвеев Е. Л. Квазиградиентный алгоритм минимизации функции квантили // Тезисы межд. конф. “Системный анализ, управление и навигация”, Евпатория, 2009, С. 96 – 97.