А. И. Кибзун, Распараллеливание алгоритмов оптимизации функции квантили, *Автомат. и телемех.*, 2007, выпуск 5, 59–70

Автоматика и телемеханика, № 5, 2007

PACS 2.50.-r, 02.60.Pn

© 2007 г. **А.И. КИБЗУН, д-р физ.-мат. наук**
(Московский авиационный институт)

РАСПАРАЛЛЕЛИВАНИЕ АЛГОРИТМОВ ОПТИМИЗАЦИИ ФУНКЦИИ КВАНТИЛИ[1]

Исследуется задача оптимизации для функции потерь, выпуклой отдельно по стратегии и по случайному вектору. Для решения задачи используется доверительный метод, с помощью которого получается верхняя оценка функции квантили. Рассматриваются два способа, с помощью которых можно получить искомую верхнюю оценку. Оба способа позволяют распараллелить процесс вычисления оценки, а задача сводится к решению набора задач выпуклого программирования.

1. Введение

В многочисленных задачах экономики и авиационной техники возникает функция квантили [1, 2], которая характеризует гарантированные с заданной вероятностью потери. Например, такими критериями являются гарантированная с заданной вероятностью терминальная точность приведения некоторого объекта в цель и потери некоторого инвестора, непревышение которых гарантируется с заданной вероятностью, и др. Функция квантили с математической точки зрения оказывается довольно сложным объектом в отличие, скажем, от математического ожидания. Например, если функция потерь выпукла по стратегии, то из этого не следует выпуклость функции квантили, в то время как критерий в виде математического ожидания в этом случае оказывается выпуклым [3]. Поэтому трудно предложить эффективные вычислительные процедуры для решения задачи квантильной оптимизации. Для преодоления этой проблемы был предложен в [2] доверительный метод, сводящий задачу квантильной оптимизации к некоторой минимаксной. Этот метод, по сути, является методом регуляризации исходной стохастической задачи, некорректной, так как небольшие погрешности при вычислении вероятности приводят к значительным изменениям значения функции квантили.

Основные свойства функции квантили, а также разнообразные методы ее оптимизации, в том числе доверительный метод, приведены в [1]. В основе изложенных в [1] алгоритмов квантильной оптимизации лежит построение стохастического квазиградиента вида [3], а сами алгоритмы являются, по сути, рекуррентными, очень медленно сходящимися. Как известно, рекуррентные алгоритмы нельзя в принципе распараллелить, т.е. ускорить процесс решения, используя современную вычислительную технику.

В данной статье предлагаются два алгоритма получения верхних оценок для оптимального значения функции квантили, которые удается распараллелить. При

[1] Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (гранты № 04-01-00758, 05-08-17963).

этом рассматривается случай, когда функция потерь выпукла отдельно по стратегии и по случайному вектору, который имеет нормальное распределение. При этом исходно невыпуклая задача квантильной оптимизации сводится к решению набора задач выпуклого программирования, которые можно решать независимо друг от друга.

2. Постановка задачи

Пусть задана функция потерь $\Phi(u,x) : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^1$, определенная для каждой стратегии $u \in U \subset \mathbb{R}^m$ и каждой реализации $x \in \mathbb{R}^n$ случайного вектора X с известной функцией распределения $F_X(x)$.

Рассмотрим функцию вероятности $P_\varphi(u) : U \to [0,1]$, равную вероятности такого события, при котором потери $\Phi(u,X)$ не превысят допустимый порог $\varphi \in \mathbb{R}^1$:

$$P_\varphi(u) \triangleq \mathcal{P}\{X : \Phi(u,X) \leqslant \varphi\}. \tag{1}$$

Рассмотрим также функцию квантили $\Phi_\alpha(u) : U \to \mathbb{R}^1$, определяющую максимальные потери, непревышение которых гарантируется с вероятностью $\alpha \in (0,1)$:

$$\Phi_\alpha(u) \triangleq \min\{\varphi : P_\varphi(u) \geqslant \alpha\}. \tag{2}$$

Пусть задача оптимизации заключается в минимизации функции квантили

$$u_\alpha = \arg\min_{u \in U} \Phi_\alpha(u) \tag{3}$$

в предположении существования оптимальной стратегии u_α.

Свойства непрерывности и полунепрерывности функций $P_\varphi(u), \Phi_\alpha(u)$, а также условия существования решений задачи (3) установлены в [1].

Для решения задачи (3) можно воспользоваться стохастическим квазиградиентным алгоритмом, введенным в [3] для минимизации средних потерь, следующего вида.

Пусть X_i, $i = \overline{1,N}$, – выборка объема N случайного вектора $X \sim F_X(x)$. По ней сформируем выборку $\Phi(u, X_i)$, $i = \overline{1,N}$, для фиксированной стратегии $u \in U$. Построим вариационный ряд для полученной выборки

$$\Phi_N^{(1)}(u) \leqslant \Phi_N^{(2)}(u) \leqslant \ldots \leqslant \Phi_N^{(N)}(u).$$

Определим выборочную оценку функции квантили

$$\widehat{\Phi}_N(u) \triangleq \Phi_N^{[\alpha N]+1}, \tag{4}$$

где $[\cdot]$ – означает целую часть.

Рассмотрим случайную последовательность

$$u_{k+1} = \Pi_U(u_k - \rho_k \xi_k), \quad k = 1, 2, \ldots, \tag{5}$$

где Π_U – оператор проецирования на множество U, ρ_k – длина рабочего шага алгоритма, ξ_k – стохастический квазиградиент:

$$\xi_k \triangleq \frac{1}{2\delta_k} \sum_{j=1}^{m} \left[\widehat{\Phi}_{N_k}(\widetilde{u}_{k_1}, \ldots, u_{k_j} + \delta_k, \ldots, \widetilde{u}_{k_m}) - \widehat{\Phi}_{N_k}(\widetilde{u}_{k_1}, \ldots, u_{k_j} - \delta_k, \ldots, \widetilde{u}_{k_m})\right] e_j,$$

где N_k – объем выборки на k-м шаге алгоритма, δ_k – длина пробного шага, e_j – единичный орт, направленный вдоль j-й оси, $\widetilde{u}_{k_j}$, $j = \overline{1,m}$, – независимые равномерно распределенные на $[u_{k_j} - \delta_k, u_{k_j} + \delta_k]$ случайные величины. Справедливо следующее утверждение [1].

Т е о р е м а 1. *Пусть выполнены следующие условия:*

(i) *функция квантили* $\Phi_\alpha(u)$ *выпукла на выпуклом компакте* U*;*

(ii) *функция квантили непрерывно дифференцируема на* U*;*

(iii) *последовательность* ρ_k *удовлетворяет условиям:*

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

(iv) *последовательности* δ_k *и* N_k *удовлетворяют условиям*

$$\delta_k \to 0, \quad N_k \to \infty, \quad \frac{k^{(1+\varepsilon)}}{\delta_k^2 N_k} \to 0 \quad \text{при} \quad k \to \infty \quad \text{при некотором} \quad \varepsilon > 0.$$

Тогда $u_k \xrightarrow{п.н.} u_\alpha$.

З а м е ч а н и е 1. Условия (i) и (ii) труднопроверяемы. Но даже при их выполнении алгоритм (5) крайне медленно сходится, так как должны выполняться условия (iii) и (iv), накладываемые на последовательности ρ_k, δ_k, N_k. Более того, стохастический квазиградиентный алгоритм (5) в принципе нельзя ускорить, используя какую-либо технику распараллеливания вычислений, так как он рекуррентный. Например если выбрать две начальные точки для алгоритма (5), то получатся две независимые задачи. Но так как в обеих задачах минимум один и тот же, то продолжительность работы такой пары алгоритмов будет определяться наиболее медленно сходящейся последовательностью в связи с тем, что заранее нельзя отдать предпочтение какому-либо алгоритму из двух. Рассмотрим другие способы решения задачи (3), позволяющие распараллелить процесс поиска оптимальной стратегии u_α.

3. Доверительный метод

В [1] установлено, что задача (3) эквивалентна следующей минимаксной

$$\Phi_\alpha(u_\alpha) = \min_{S \in \mathcal{F}_\alpha} \psi(S, u^S), \tag{6}$$

где $\mathcal{F}_\alpha \triangleq \{S \in \mathcal{F}, \mathcal{P}(S) \geqslant \alpha\}$ – семейство доверительных множеств, $\mathcal{F}$ – σ-алгебра борелевских множеств,

$$\psi(S, u) \triangleq \sup_{x \in S} \Phi(u, x), \tag{7}$$

$$u^S = \arg\min_{u \in U} \psi(S, u). \tag{8}$$

Сведение задачи квантильной оптимизации (3) к эквивалентной минимаксной (6)–(8) называется в [1] *доверительным методом*. Заметим, что минимум в (6) достигается на оптимальном доверительном множестве

$$S_\alpha \triangleq \{x : \Phi(u_\alpha, x) \leqslant \Phi_\alpha(u_\alpha)\},$$

которое, правда, заранее неизвестно.

Важно, что для любого доверительного множества $S \in \mathcal{F}_\alpha$ можно получить минимаксную оценку сверху для неизвестного оптимального значения функции квантили

$$\Phi_\alpha(u_\alpha) \leqslant \psi(S, u^S) \ \forall S \in \mathcal{F}_\alpha. \tag{9}$$

З а м е ч а н и е 2. Задача (8) записана при предположении, что ее решение u^S существует. Но доверительный метод справедлив и в более общем случае, когда не существуют u_α и u^S, а вместо задач (3) и (8) рассматриваются последовательности стратегий для поиска нижних граней соответствующих функций.

Перебирая различные $S \in \mathcal{F}_\alpha$ можно улучшать верхнюю оценку для $\Phi_\alpha(u_\alpha)$. Степень близости верхней оценки $\psi(S, u^S)$ к $\Phi_\alpha(u_\alpha)$ можно определить, если известна нижняя оценка для $\Phi_\alpha(u_\alpha)$. С этой целью рассмотрим частный случай.

Пусть случайный вектор X имеет нормальное распределение $\mathcal{N}(0, I)$, где I – единичная матрица, и функция потерь $\Phi(u, x)$ выпукла по $x \in \mathbb{R}^n$ для всех $u \in U$.

Рассмотрим два шара B_ρ и B_R с центрами в начале координат и с радиусами ρ и R, удовлетворяющими условиям

$$\mathcal{P}(B_R) = \alpha, \mathcal{P}(X_1 \leqslant \rho) = \alpha, \tag{10}$$

где X_1 – первая компонента вектора $X \sim \mathcal{N}(0, I)$. Таким образом, ρ – квантиль уровня α стандартного нормального распределения $\mathcal{N}(0, 1)$, а B_R – доверительный шар, у которого радиус R определяется путем решения трансцендентного уравнения [1]

$$\frac{1}{2^{(n-2)/2}\Gamma(n/2)} \int\limits_0^R t^{n-1} \exp(-t^2/2) dt = \alpha,$$

где $\Gamma(\cdot)$ – гамма функция.

В [1] содержится следующее утверждение.

Т е о р е м а 2. Если $\Phi(u, x)$ выпукла на $\mathbb{R}^n$ для всех $u \in U$, а $X \sim \mathcal{N}(0, I)$, то справедливы следующие оценки

$$\psi(B_\rho, u^\rho) \leqslant \Phi_\alpha(u_\alpha) \leqslant \psi(B_R, u^R), \tag{11}$$

где u^ρ и u^R – решения минимаксной задачи (8) для множеств B_ρ и B_R соответственно, причем

$$\psi(B_R, u^R) - \psi(B_\rho, u^\rho) \to 0 \quad \textit{при} \quad \alpha \to 1.$$

Таким образом, по величине $\psi(B_R, u^R) - \psi(B_\rho, u^\rho)$ можно судить о точности верхней оценки $\psi(B_R, u^R)$. Более того, разница между этой верхней оценкой и минимальным значением функции квантили тем меньше, чем ближе α к 1.

З а м е ч а н и е 3. Отметим, что приведенные нижняя и верхняя оценки для $\Phi_\alpha(u_\alpha)$ оказываются достигаемыми. Например, пусть

$$\Phi(u, x) = b^{\mathrm{T}}(u)x + c(u),$$

где $b(u)$ – некоторая n-мерная векторная функция, а $c(u)$ – скалярная функция. Тогда легко показать, что

$$\psi(B_\rho, u) \triangleq \max_{x \in B_\rho} \Phi(u, x) = \|b(u)\|\rho + c(u),$$

так как максимум рассматриваемой функции $\Phi(u, x)$ на B_ρ достигается в точке

$$x^* = \frac{b(u)}{\|b(u)\|}\rho.$$

Кроме того, согласно [1] в данном случае можно найти явное выражение для функции квантили

$$\Phi_\alpha(u) = \|b(u)\|\rho + c(u).$$

Тогда нижняя оценка $\psi(B_\rho, u^\rho)$ совпадет с оптимальным значением функции квантили

$$\psi(B_\rho, u^\rho) = \Phi_\alpha(u_\alpha), \quad u^\rho = u_\alpha.$$

Рассмотрим теперь функцию потерь вида

$$\Phi(u, x) = b(u)g(\|x\|),$$

где $g(\cdot) : [0, \infty) \to [0, \infty)$ – строго возрастающая функция, а $b(u) > 0$ – некоторая строго положительная функция. Тогда оптимальным доверительным множеством S_α будет шар B_R, так как

$$S_\alpha = \{x : b(u)g(\|x\|) \leqslant \Phi_\alpha(u)\} = \left\{\|x\| \leqslant g^{-1}\left(\frac{\Phi_\alpha(u)}{b(u)}\right)\right\} = B_R,$$

где из условия нормировки найдено

$$R = g^{-1}\left(\frac{\Phi_\alpha(u)}{b(u)}\right).$$

Значит, согласно (6)

$$\Phi_\alpha(u) = \psi(S_\alpha, u) = \psi(B_R, u).$$

Поэтому верхняя оценка $\psi(B_R, u^R)$ совпадает с оптимальным значением функции квантили

$$\Phi_\alpha(u_\alpha) = \psi(B_R, u^R), \quad u^R = u_\alpha.$$

4. Алгоритм улучшения верхней оценки квантили

Пусть функция потерь $\Phi(u, x)$ выпукла отдельно по $u \in U$ и по $x \in \mathbb{R}^n$. Примером такой функции может служить

$$\Phi(u, x) = \max_{j=\overline{1,J}} \sum_{i=1}^{I} \varphi_{ij}(u) g_{ij}(x), \tag{12}$$

где все $\varphi_{ij}(u)$, $i = \overline{1, I}$, $j = \overline{1, J}$, выпуклы по $u \in U$ и неотрицательны и все $g_{ij}(x)$, $i = \overline{1, I}$, $j = \overline{1, J}$, выпуклы по $x \in \mathbb{R}^n$ и неотрицательны. В частности, такими свойствами обладает кусочно-линейная функция

$$\Phi(u, x) = \max_{j=\overline{1,J}} \left\{a_j^{\mathrm{T}} u + b_j^{\mathrm{T}} x + c_j\right\}, \tag{13}$$

где a_j – вектор размерности m, а b_j – вектор размерности n. Данная функция является выпуклой одновременно по $u \in U$ и по $x \in \mathbb{R}^n$, причем в данном случае элементы векторов a_j и b_j могут быть отрицательными.

Рассмотрим минимаксную задачу

$$(14)\qquad u^r = \arg\min_{u\in U} \psi(B_r, u),$$

$$(15)\qquad \psi(B_r, u) \triangleq \max_{x\in B_r} \Phi(u, x)$$

для шара B_r с произвольным радиусом $r \in [\rho, R]$.

Так как по предположению $\Phi(u, x)$ выпукла по $u \in U$ для каждого $x \in \mathbb{R}^n$, то (14) является задачей выпуклого программирования. В частном случае, когда $\Phi(u, x)$ – кусочно-линейная вида (13) и U – выпуклый многогранник, задача (14) сводится согласно [4] к задаче линейного программирования, так как функция максимума $\psi(B_\rho, u)$ в этом случае равна

$$(16)\qquad \psi(B_r, u) = \max_{j=\overline{i,J}} \left\{ a_j^{\mathrm{T}} u + \|b_j\| r + c_j \right\}.$$

Сформулируем следующее утверждение.

Теорема 3. Если $\Phi(u, x)$ выпукла по $x \in \mathbb{R}^1$ для всех $u \in U$, $X \sim \mathcal{N}(0, I)$ и решение u_α задачи (3) существует, то найдется такой радиус $r_\alpha \in [\rho, R]$, что

$$(17)\qquad \psi(B_\rho, u^\rho) \leqslant \psi(B_{r_\alpha}, u^{r_\alpha}) \leqslant \psi(B_{r_\alpha}, u_\alpha) = \Phi_\alpha(u_\alpha) \leqslant \Phi_\alpha(u^{r_\alpha})$$

и

$$(18)\qquad \Phi_\alpha(u_\alpha) \leqslant \Phi_\alpha(u^R) \leqslant \psi(B_R, u^R).$$

Доказательство приведено в Приложении.

Рассмотрим вспомогательную задачу

$$(19)\qquad r(\alpha) = \arg\min_{r\in[\rho,R]} \Phi_\alpha(u^r),$$

где u^r – стратегия, оптимальная для критерия $\psi(B_r, u)$ с шаром радиуса $r \in [\rho, R]$.

В общем случае $r(\alpha) \neq r_\alpha$, которое фигурирует в теореме 1. Но очевидно, что

$$(20)\qquad \Phi_\alpha(u_\alpha) \leqslant \Phi_\alpha(u^{r(\alpha)}) \leqslant \Phi_\alpha(u^R)$$

и

$$(21)\qquad \Phi_\alpha(u_\alpha) \leqslant \Phi_\alpha(u^{r(\alpha)}) \leqslant \Phi_\alpha(u^\rho),$$

так как $r(\alpha) \in [\rho, R]$.

Таким образом, оценка $\Phi_\alpha(u^{r(\alpha)})$ является улучшенной по сравнению с $\psi(B_R, u^R)$ верхней оценкой для $\Phi_\alpha(u_\alpha)$. Заметим, что аналогичная процедура для случая кусочно-линейной функции потерь рассмотрена в [4].

5. Распараллеливание процедуры вычисления верхней оценки функции квантили

Рассмотрим два способа распараллеливания процедуры вычисления верхней оценки для $\Phi_\alpha(u_\alpha)$ для случая, когда $\Phi(u, x)$ отдельно выпукла по $u \in U$ и по $x \in \mathbb{R}^n$, а $X \sim \mathcal{N}(0, I)$.

Способ 1. Рассмотрим набор доверительных множеств: B_R – шар с радиусом R и семейство n-мерных прямоугольных параллелепипедов

$$(22)\qquad \Pi(a, b) \triangleq \{a_i \leqslant x_i \leqslant b_i, i = \overline{1, n}\},$$

где векторы $a = \mathrm{col}\,(a_1, \ldots, a_n), b = \mathrm{col}\,(b_1, \ldots, b_n)$ связаны между собой условием $\mathcal{P}(\Pi(a, b)) = \alpha$.

Так как $X \sim \mathcal{N}(0, I)$, то получаем

$$\prod_{i=1}^{n}[F(b_i) - F(a_i)] = \alpha, \tag{23}$$

где $F(\cdot)$ – функция Лапласа. Пусть C – ортогональная матрица размерности $n \times n$, т.е. $C^{\mathrm{T}}C = I$. Так как $X \sim \mathcal{N}(0, I)$, то после поворота параллелепипедов $\Pi(a, b)$ с помощью матрицы C его мера не изменится, т.е.

$$\mathcal{P}(\Pi(C, a, b)) = \alpha, \tag{24}$$

где

$$\Pi(C, a, b) = \{x : a \leqslant Cx \leqslant b\}. \tag{25}$$

Построенное семейство доверительных множеств принадлежит классу $\mathcal{F}_\alpha$ всех доверительных множеств. Поэтому согласно (6) получаем верхнюю оценку для $\Phi_\alpha(u_\alpha)$:

$$\Phi_\alpha(u_\alpha) \leqslant \psi_\alpha, \tag{26}$$

где

$$\psi_\alpha \triangleq \min\{\psi(B_R, u^R), \overline{\psi}_\alpha\}, \tag{27}$$

$$\overline{\psi}_\alpha = \min_{a,b,C} \psi(\Pi(C, a, b), u(C, a, b)),\ \mathcal{P}(\Pi(C, a, b)) = \alpha,\ C^{\mathrm{T}}C = 1, \tag{28}$$

где

$$u(C, a, b) = \arg\min_{u \in U} \psi(\Pi(C, a, b), u). \tag{29}$$

Замечание 4. Подчеркнем, что верхняя оценка $\psi(\Pi(C, a, b), u(C, a, b))$ может совпадать с оптимальным значением функции квантили. Например, пусть

$$\Phi(u, x) = \max_{i=\overline{1,n}} |x_i - u_i|, \quad \text{где} \quad m = n.$$

Тогда нетрудно установить из соображений симметрии, что

$$u_\alpha = 0 \quad \text{и} \quad S_\alpha = \{x : |x_i| \leqslant \varphi_\alpha\},$$

где $\varphi_\alpha = \Phi_\alpha(u_\alpha)$ определяется из условия

$$\mathcal{P}(S_\alpha) = [2F(\varphi_\alpha) - 1]^n = \alpha.$$

Таким образом, оптимальное доверительное множество является n-мерным доверительным кубом, т.е. принадлежит рассматриваемому семейству прямоугольных параллелепипедов, $S_\alpha = \Pi(-\varphi_\alpha, \varphi_\alpha)$. Тогда минимаксная задача

$$u^{S_\alpha} = \arg\min_{u \in U}\left[\max_{x \in S_\alpha} \max_{i=\overline{1,n}} |x_i - u_i|\right]$$

имеет следующее решение

$$u^{S_\alpha} = 0, \psi(S_\alpha, \quad u^{S_\alpha}) = \varphi_\alpha,$$

которое совпадает с решением задачи квантильной оптимизации для рассматриваемой функции потерь.

Рассмотрим отдельно две подзадачи, возникающие при получении оценки (26). В частности,

$$u^R = \arg\min_{u\in U} \psi(B_R, u), \tag{30}$$

$$\psi(B_R, u) = \max_{x\in B_R} \Phi(u, x). \tag{31}$$

Если $\Phi(u, x)$ – кусочно-линейная по x, т.е.

$$\Phi(u, x) = \max_{j=\overline{1,J}} \left\{c_j(u) + b_j^{\mathrm{T}}(u)x\right\}, \tag{32}$$

то задача (31) имеет аналитическое решение (см. (16))

$$\psi(B_R, u) = \max_{j=\overline{1,J}} \{c_j(u) + \|b_j(u)\|R\}. \tag{33}$$

В противном случае, задачу (30) можно решать, например, методом случайного поиска. С этой целью сформируем выборку $\{Y^i\}_{i=1}^N$ для случайного вектора Y, равномерно распределенного на поверхности ∂B_R шара B_R. Для этого можно использовать согласно [1] выборку случайного вектора $X \sim \mathcal{N}(0, I)$ с последующим проецированием на поверхность ∂B_R. Нетрудно убедиться, исходя из соображений симметрии, что выборка, получаемая таким образом на поверхности шара, будет обладать нужными свойствами. Рассмотрим две экстремальные порядковые статистики в вариационном ряде

$$\Phi_N^{(1)}(u) \leqslant \ldots \leqslant \Phi_N^{(N)}(u), \tag{34}$$

построенные по выборке $\{\Phi(u, Y^i)\}_{i=1}^N$. Найдем оценку функции максимума $\psi(B_R, u)$ вида

$$\widehat{\psi}_N(u) = \Phi_N^{(N)}(u) + \frac{n-1}{2}\left[\Phi_N^{(N)}(u) - \Phi_N^{(N-1)}(u)\right], \tag{35}$$

которая согласно [5] является асимптотически несмещенной и, более того,

$$\mathbf{M}\left[\widehat{\psi}_N(u)\right] - \psi(B_R, u) = o(1/N). \tag{36}$$

Далее, так как по предположению $\Phi(u, x)$ выпукла не только по $x \in \mathbb{R}^n$, но и по $u \in U$, то функция $\psi(B_R, u)$ согласно (31) будет по-прежнему выпуклой по $u \in$ $\in U$ и задача (30) оказывается задачей выпуклого программирования, для решения которой можно использовать стандартные методы выпуклого программирования [6].

Исследуем теперь задачу (29). Так как функция $\Phi(u, x)$ выпукла по $x \in \mathbb{R}^n$ для всех $u \in U$, то согласно [1] максимум функции $\Phi(u, x)$ достигается в одной из 2^n вершин n-мерного прямоугольного параллелепипеда $\Pi(C, a, b)$. В связи с выпуклостью $\Phi(u, x)$ по $u \in U$ для всех $x \in \mathbb{R}^n$ полученная функция $\psi(\Pi(C, a, b), u)$ является также выпуклой по $u \in U$, а задача (29) сводится к задаче выпуклого программирования, для решения которой можно использовать стандартные методы [6].

Пусть (C_k, a_k, b_k), $k = \overline{1, K}$, – некоторый набор матриц C_k и векторов a_k и b_k, удовлетворяющих условиям (28).

Таким образом, решая отдельно и независимо задачи (30) и (29) для набора (C_k, a_k, b_k), $k = \overline{1, K}$, можно распараллелить процесс поиска верхней оценки ψ_α для $\Phi_\alpha(u_\alpha)$. После решения этих задач в отдельности объединяем полученные оценки и получаем

$$\Phi_\alpha(u_\alpha) \leqslant \ \widehat{\psi}_\alpha = \min\left\{\psi(B_R, u^R), \widetilde{\psi}_\alpha\right\}, \tag{37}$$

где

$$(38)\qquad \widetilde{\psi}_\alpha = \min_{k=\overline{1,K}} \psi(\Pi(C_k, a_k, b_k), u(C_k, a_k, b_k)),$$

$$(39)\qquad u(C_k, a_k, b_k) = \arg\min_{u\in U} \psi(\Pi(C_k, a_k, b_k), u).$$

Напомним, что каждая из возникающих задач является задачей выпуклого программирования. В то же время нельзя сказать что-либо определенное о выпуклых свойствах функции квантили $\Phi_\alpha(u)$ для рассматриваемого случая. Иными словами, в данном случае нельзя воспользоваться алгоритмом (5) для решения задачи квантильной оптимизации (см. теорему 1).

Способ 2. Воспользуемся теоремой 2 для построения 2-го способа распараллеливания процесса поиска верхней оценки для $\Phi_\alpha(u_\alpha)$.

С этой целью рассмотрим задачи (14) и (15) для произвольного $r \in [\rho, R]$. Как отмечалось выше, если функция потерь имеет кусочно-линейный вид (32), то задача (15) согласно (32) аналитически решается

$$(40)\qquad \psi(B_r, u) = \max_{j=\overline{1,J}} \{c_j(u) + \|b_j(u)\| r\}.$$

В случае произвольно выпуклой функции $\Phi(u, x)$ по $x \in \mathbb{R}^n$ можно воспользоваться оценкой (35) для оценивания $\psi(B_r, u)$. Так как функция $\Phi(u, x)$ является также выпуклой по $u \in U$, то функция максимума $\psi(B_r, u)$ будет выпуклой по $u \in U$, а задача (14) – задачей выпуклого программирования, для решения которой можно использовать стандартные методы [6].

Далее, для найденной стратегии u^r вычисляем квантиль $\Phi_\alpha(u^r)$, используя для этого выборочную оценку (4), т.е.

$$(41)\qquad \widehat{\Phi}_N(u^r) = \Phi_N^{[\alpha N]+1}(u^r),$$

где $\Phi_N^{[\alpha N]+1}(u^r)$ – порядковая статистика с номером $[\alpha N] + 1$ для выборки $\{\Phi(u^r, X_i)\}_{i=1}^N$. Теперь решим задачу (19), используя для этого вариацию метода дихотомии, который обладает робастным свойством в отличие от многих других методов одномерной оптимизации, в которых пробные точки могут располагаться близко друг от друга. Робастность метода дихотомии в данном случае важна, так как значение функции квантили $\Phi_\alpha(u^r)$ вычисляется неточно (будут погрешности при вычислении u^r и при использовании выборочной оценки (41)).

Разобьем отрезок $[\rho, R]$ на 8 частей и для 9 точек радиуса r шара B_r:

$$(42)\qquad r_1 = \rho, \quad r_9 = R, \quad r_i = \rho + (1/8)^{i-1}(R - \rho), \quad i = \overline{2,8},$$

решим задачи (14), (15), а затем вычислим квантиль $\Phi_\alpha(u^{r_i})$, $i = \overline{1,9}$. Выберем такой номер I из набора $i = \overline{1,9}$, при котором значение $\Phi_\alpha(u^{r_I})$ минимально. Оставим из начального набора по две точки справа от r_I и слева от r_I, т.е. для $i = I-2, I-1, I+1, I+2$. Произведем дальнейшее разбиение полученных четырех отрезков пополам с помощью новых четырех точек. Далее процедура повторяется.

Таким образом, вначале необходимо решить задачу выпуклого программирования (14) отдельно 9 раз для точек (42), а затем на каждом шаге итерации – 4 раза для 4-х новых точек. Это значит, что при k итерациях алгоритма необходимо решить $9 + 4^k$ задач выпуклого программирования (14). При этом интервал неопределенности будет равен $(1/2)^{k+2}[R - \rho]$.

Процесс поиска $r(\alpha)$, т.е. получения верхней оценки $\Phi_\alpha(u^{r(\alpha)})$ для $\Phi_\alpha(u_\alpha)$, можно в 4 раза ускорить, если решать задачу выпуклого программирования (14) для четырех новых точек $r \in [\rho, R]$ независимо.

6. Пример

Рассмотрим функцию потерь вида (12)

$$\Phi(u, x) = g(x)\varphi(u), \tag{43}$$

где $g(x)$ – выпуклая по $x \in \mathbb{R}^n$ и неотрицательная функция, а $\varphi(u)$ – выпуклая по $u \in U$ функция, такая, что $\varphi(u) > 0$. Данная функция является выпуклой отдельно по $x \in \mathbb{R}^n$ и по $u \in U$. Пусть по-прежнему $X \sim \mathcal{N}(0, I)$. В данном случае согласно [1] можно найти детерминированный эквивалент для функции квантили

$$\Phi_\alpha(u) = g_\alpha \varphi(u), \tag{44}$$

где g_α – квантиль уровня α распределения случайной величины $g(X)$, $X \sim \mathcal{N}(0, I)$. Очевидно, что

$$u_\alpha = \arg\min_{u \in U} \varphi(u). \tag{45}$$

Рассмотрим второй способ получения верхней оценки для $\Phi_\alpha(u_\alpha)$. С этой целью решим задачу (14), (15) для произвольного $r \in [\rho, R]$. В данном случае

$$\psi(B_r, u) = \varphi(u) \max_{x \in B_r} g(x). \tag{46}$$

Поэтому

$$u^r = \arg\min_{u \in U} \varphi(u),$$

т.е. $u^r = u_\alpha$ для любого $r \in [\rho, R]$, и верхняя оценка совпадает с оптимальным значением функции квантили

$$\Phi_\alpha(u^r) = \Phi_\alpha(u_\alpha). \tag{47}$$

Данный пример показывает, что алгоритм, соответствующий способу 2, в рассматриваемом случае позволяет получить точное решение квантильной задачи за одну итерацию. В более общем случае можно рассчитывать на быструю сходимость алгоритма, так как длина начального интервала неопределенности $[\rho, R]$ невелика, а метод дихотомии имеет высокую скорость сходимости.

7. Заключение

Рассмотрена задача квантильной оптимизации, когда функция потерь является выпуклой отдельно по стратегии и по случайному вектору, который имеет нормальное распределение. В этом случае функция квантили может оказаться, вообще говоря, невыпуклой. Поэтому, известные стохастические квазиградиентные алгоритмы становятся здесь неприменимыми. В статье предложены два способа получения верхней оценки для оптимального значения функции квантили, которые основаны на доверительном методе, сводящем исходную стохастическую задачу к некоторой минимаксной. В первом способе производится параметризации семейства доверительных прямоугольных параллелепипедов, с помощью чего задача поиска верхней оценки сводится к решению набора задач выпуклого программирования, которые можно решать независимо, т.е. параллельно. Второй способ состоит в решении набора минимаксных задач на шарах с переменным радиусом, ограниченным снизу радиусом ядра вероятностной меры, а сверху – радиусом доверительного шара.

В результате получается уточненная верхняя оценка искомой квантили, а процесс решения сводится к решению набора задач выпуклого программирования, которые можно решать так же, как и в первом способе, независимо друг от друга.

Благодарность. Автор выражает благодарность И.И. Еремину за высказанную им идею заменить исходную задачу на набор задач, позволившую распараллелить алгоритмы квантильной оптимизации.

ПРИЛОЖЕНИЕ

Доказательство теоремы 1. Рассмотрим оптимальное доверительное множество

$$S_\alpha \triangleq \{x : \Phi(u_\alpha, x) \leqslant \Phi_\alpha(u_\alpha)\}, \tag{П.1}$$

для которого $\psi(S_\alpha, u_\alpha) = \Phi_\alpha(u_\alpha)$. Данное множество выпукло, так как по предположению $\Phi(u, x)$ выпукла по $x \in \mathbb{R}^n$ для всех $u \in U$. Согласно [1] $S_\alpha \supset B_\rho$. Пусть B_{r_α} – шар радиуса r_α, вписанный в S_α. Очевидно, что $r_\alpha \geqslant \rho$. Но так как S_α и B_R имеют одну и ту же меру α, а $B_{r_\alpha} \subset S_\alpha$, то $r_\alpha \leqslant R$. Значит, $B_\rho \subset B_{r_\alpha} \subset B_R$ и $r \in [\rho, R]$. Так как $B_\rho \subset B_{r_\alpha}$, то

$$\psi(B_\rho, u) \triangleq \max_{x \in B_\rho} \Phi(u, x) \leqslant \psi(B_{r_\alpha}, u) \triangleq \max_{x \in B_{r_\alpha}} \Phi(u, x). \tag{П.2}$$

Поэтому

$$\psi(B_\rho, u^\rho) = \min_{u \in U} \psi(B_\rho, u) \leqslant \psi(B_{r_\alpha}, u^{r_\alpha}) = \min_{u \in U} \psi(B_{r_\alpha}, u).$$

Далее, так как стратегия u_α, оптимальная для $\Phi_\alpha(u)$, не является в общем случае оптимальной для $\psi(B_{r_\alpha}, u)$, то

$$\psi(B_{r_\alpha}, u^{r_\alpha}) \leqslant \psi(B_{r_\alpha}, u_\alpha). \tag{П.3}$$

Но B_{r_α} – шар, вписанный в S_α, поэтому

$$\Phi_\alpha(u_\alpha) = \psi(S_\alpha, u_\alpha) = \psi(B_{r_\alpha}, u_\alpha) \triangleq \max_{x \in B_{r_\alpha}} \Phi(u_\alpha, x). \tag{П.4}$$

Так как стратегия u^{r_α} не обязательно является оптимальной для $\Phi_\alpha(u)$, то

$$\Phi_\alpha(u_\alpha) \leqslant \Phi_\alpha(u^{r_\alpha}). \tag{П.5}$$

Таким образом, получаем цепочку неравенств

$$\psi(B_\rho, u^\rho) \leqslant \psi(B_{r_\alpha}, u^{r_\alpha}) \leqslant \psi(B_{r_\alpha}, u_\alpha) = \Phi_\alpha(u_\alpha) \leqslant \Phi_\alpha(u^{r_\alpha}). \tag{П.6}$$

Теперь рассмотрим доверительный шар B_R с мерой α. Для любого $u \in U$ имеем

$$\mathcal{P}\{\Phi(u, X) \leqslant \psi(B_R, u)\} \geqslant \mathcal{P}(B_R) = \alpha. \tag{П.7}$$

Поэтому

$$\Phi_\alpha(u) \leqslant \psi(B_R, u) \ \forall u \in U. \tag{П.8}$$

В частности,

$$\Phi_\alpha(u^R) \leqslant \psi(B_R, u^R), \tag{П.9}$$

где u^R – стратегия, оптимальная для критерия $\psi(B_R, u)$. Но стратегия u^R в общем случае не является оптимальной для $\Phi_\alpha(u)$, т.е.

$$(\text{П}.10) \qquad \Phi_\alpha(u_\alpha) \leqslant \Phi_\alpha(u^R).$$

Объединяя найденные неравенства, получим

$$(\text{П}.11) \qquad \Phi_\alpha(u_\alpha) \leqslant \Phi_\alpha(u^R) \leqslant \psi(B_R, u^R).$$

Теорема доказана.

СПИСОК ЛИТЕРАТУРЫ

1. *Kibzun A.I., Kan Yu.S.* Stochastic Programming Problems with Probability and Quantile Functions. Chichester: John Wiley, 1996.
2. *Малышев В.В., Кибзун А.И.* Анализ и синтез высокоточного управления ЛА. М.: Машиностроение, 1987.
3. *Ермольев Ю.М.* Методы стохастического программирования. М.: Наука, 1976.
4. *Кибзун А.И., Наумов А.В.* Гарантирующий алгоритм решения задачи квантильной оптимизации // Космич. исслед. 1995. Т. 33. № 2. С. 160–165.
5. *Kibzun A.I., Kurbakovskiy V.Yu.* Guaranteeing Approach to Solving Quantile Optimization Problems // Ann. Oper. Res. 1991. V. 30. P. 81–94.
6. *Поляк Б.Т.* Введение в оптимизацию. М.: Наука, 1983.

Статья представлена к публикации членом редколлегии В.М. Вишневским.

Поступила в редакцию 18.12.2006