МЕТОДЫ ОПТИМИЗАЦИИ

"МНОГОШАГОВЫЙ" МЕТОД ЭЛЛИПСОИДОВ

Немировский А.С., Ненахов Э.И.

(Москва, Киев)

Описывается новый метод негладкой выпуклой минимизации, соединяющий идеи метода эллипсоидов [1, 2] и субградиентного спуска. Гарантированная оценка скорости сходимости для предлагаемого алгоритма такая же, как и для обычного метода эллипсоидов, но новый метод представляется более гибким, и можно ожидать, что его поведение на реальных задачах при разумной реализации будет лучше, чем у прототипа.

1. **Формулировка задачи. Предварительные замечания.** Рассмотрим задачу выпуклого программирования в следующей постановке

$$f(x) \to \min |x \in K \in R^n. \tag{1}$$

где K – выпуклое замкнутое ограниченное тело в R^n, а f – непрерывная выпуклая функция, определенная на K (все последующее без труда распространяется и на задачу выпуклого программирования с функциональными ограничениями).

Будем предполагать, что в нашем распоряжении имеется некоторая процедура Ω, которая по входу $x \in R^n$ сообщает, лежит ли x в $\mathrm{int}K$, и, кроме того, в случае $x \in \mathrm{int}K$ сообщает значение $f(x)$ и опорный к f в x функционал $f'(x)$, а при $x \notin \mathrm{int}K$ – функционал, разделяющий x и $\mathrm{int}K$ (т.е. пару, образованную ненулевым вектором $e \in R^n$ и неотрицательным a, такую, что $\langle e, y\rangle \leqslant \langle e, x\rangle - a$ при всех $y \in K$).

Назовем локализатором для (f, K) любую пару (u, W), в которой $u \in K$, а W – выпуклый компакт в R^n, такие, что

$$\{x \in K \mid f(x) \leqslant f(u)\} \subset W. \tag{2}$$

Нетрудно показать [1], что погрешность u-компоненты любого локализатора в качестве решения (1) может быть оценена через n-мерный объем W-компоненты (который будем называть объемом локализатора). Именно для любого локализатора (u, W) справедливо неравенство

$$\epsilon_f(u) \equiv (f(u) - \min_K f)/(\max_K f - \min_K f) \leqslant \{|W|/|K|\}^{1/n}, \tag{3}$$

где $|\cdot|$ – n-мерный объем.

В силу (3) для решения (1) достаточно научиться строить последовательности локализаторов с по возможности быстро убывающими объемами. В предлагаемом методе (как и в обычном методе эллипсоидов) W-компонента каждого локализатора всегда эллипсоид, т.е. множество вида

$$W = \{z + Bu \mid \|u\|_2 \leqslant 1\},$$

определяемое центром z и $n \times n$-матрицей B положительным детерминантом.

Начиная с этого места в определение локализатора включается требование эллипсоидальности W.

Заметим, что если u_0 – произвольная точка K, а W_0 – какой-нибудь содержащий K эллипсоид, то пара (u_0, W_0) является локализатором. Таким образом, дело сводится к построению процедуры π преобразования имеющегося локализатора (u, W) в новый – $(u', W') = \pi(u, W)$ – меньшего объема. Процедуру такого рода целесообразно характеризовать следующими тремя величинами:

гарантированным сокращением $\nu(\pi)$ объема

$$\nu(\pi) = \inf\{n^{-1}\ln(|W|/|W'|)|(u', W') = \pi(u, W)\}$$

(inf берется по всем выпуклым задачам (1) данной размерности n и всем локализаторам для таких задач);

числом $\Omega(\pi)$ обращений к процедуре Ω в ходе преобразования локализатора;

числом $A(\pi)$ арифметических операций по обработке ответов процедуры Ω.

Назовем π допустимой процедурой, если $\nu(\pi) > 0$. Из сказанного ясно, что при любом $\epsilon \in (0, 1)$ отыскание ϵ-решения (т.е. такой точки $u \in K$, что $\epsilon_f(u) \leqslant \epsilon$) всякой выпуклой задачи (1) с n переменными на базе допустимой процедуры π "стоит" не более

$$N_\pi(\epsilon) = \Omega(\pi)]\,\nu^{-1}(\pi)\ln(\beta/\epsilon)[\tag{4}$$

обращений к Ω и, сверх того, не более

$$M_\pi(\epsilon) = A(\pi)]\,\nu^{-1}(\pi)\ln(\beta/\epsilon)[\tag{5}$$

арифметических операций; величина

$$\beta = \{|W_0|/|K|\}^{1/n} \tag{6}$$

"отвечает" за качество исходного локализатора. Здесь и далее $]t[$ – наименьшее целое, не меньшее t.

Для примера укажем, что обычный метод эллипсоидов основан на процедуре π_1, у которой

$$\nu(\pi_1) = O(1)n^{-2}, \quad \Omega(\pi_1) = 1, \quad A(\pi_1) = O(1)n^2 \tag{7}$$

(здесь и далее 0(1) – положительные абсолютные постоянные).

Цель настоящей работы – построить семейство процедур π_k, $k = 1, \ldots, n$, у которых

$$\nu(\pi_k) = O(1)kn^{-2}, \quad \Omega(\pi_k) = k, \quad A(\pi_k) = O(1)kn^2. \tag{8}$$

Таким образом, наши новые процедуры с теоретической точки зрения не имеют никаких преимуществ перед исходным методом эллипсоидов. Следует, однако, иметь в виду, что теоретические оценки наихудшего случая типа приведенных выше далеко не исчерпывающим образом описывают метод: многие методы на практике обычно демонстрируют гораздо более быструю скорость сходимости, чем можно ожидать на основании теоретической оценки эффективности. Метод эллипсоидов таким свойством не обладает (что объясняется чрезвычайной "жесткостью" схемы метода, в которой просто нет места для ускоряющих процедур). Процедуры π_k при "больших" k представляются в этом плане, как увидим, более гибкими.

2. **Основные неравенства.** Обозначим через (u, W) подлежащий преобразованию локализатор, и пусть $W = (z + Bu|\,\|u_2\| \leqslant 1)$. Сделав замену переменных $x = z + Bu$, придем к следующей ситуации. Имеются выпуклое ограниченное замкнутое тело Q в R^n и определенная на Q непрерывная выпуклая функция g; пара (g, Q) находится в тех же условиях наблюдения, что и исходная пара (f, K) (процедуру, распознающую принадлежность точки телу Q и вычисляющую соответственно значение и субградиент g в данной точке или гиперплоскость, отделяющую эту точку от $\operatorname{int} Q$, по-прежнему будем называть Ω). Кроме того, дано некоторое $\nu \in Q$, такое, что $(\nu, V = \{x|\,\|x\|_2 \leqslant 1\})$ – локализатор для (g, Q). Наша задача – преобразовать этот локализатор в локализатор меньшего объема. Сформируем с этой целью последовательно точки $x_1, \ldots, x_p, p \leqslant k$, и, построив x_i, обратимся к Ω с этим входом. В результате будут получены некоторые линейные функционалы h_i и числа a_i, причем при данном i окажется либо

$$x_i \in \operatorname{int} Q, \quad h_i = g'(x_i), \quad a_i = g(x_i),$$

либо

$$x_i \notin \operatorname{int} Q, \quad a_i \geqslant 0, \quad h_i \neq 0, \quad \langle h_i, y\rangle \leqslant \langle h_i, x_i\rangle - a_i \forall y \in Q.$$

Будем считать, что и в первом случае $h_i \neq 0$ (иначе найдено точное решение задачи). Пусть I – множество значений i, отвечающих первой альтернативе, ν' – лучшая (с минимальным значением g) среди точек $\nu, \{x_i \mid i \in I\}$. В силу выпуклости g и того факта, что (ν, V) – локализатор для (g, Q), имеем

$$Q' \equiv \{x \in Q, \ g(x) \leqslant g(\nu')\} \subset A_1 \cap \ldots \cap A_p,$$
$$A_i = \{x \mid \|x\|_2 \leqslant 1, \ \langle e_i, x\rangle \leqslant b_i \equiv \langle e_i, x_i\rangle - \alpha_i\}, \tag{9}$$

где обозначено

$$e_i = h_i/\|h_i\|_2; \quad \alpha_i = \begin{cases} a_i/\|h_i\|_2, & i \notin I, \\ (a_i - g(\nu'))/\|h_i\|_2, & i \in I \end{cases}$$

((9) устанавливается непосредственной проверкой). Отметим, что $\alpha_i \geqslant 0$ (определение Ω при $i \notin I$ и определение ν' при $i \in I$). Поскольку $\|e_i\|_2 = 1$, линейная форма $\langle e_i, x\rangle$ на A_i не меньше -1 (и не больше b_i по определению A_i). Стало быть, при любом $c_i \geqslant b_i$ все точки множества A_i удовлетворяют квадратичному неравенству

$$(\langle e_i, x\rangle + 1)(\langle e_i, x\rangle - c_i) \leqslant 0. \tag{10}$$

Умножая i-е из этих неравенств на неотрицательное λ_i, суммируя по i и добавляя выполненное на Q' неравенство $\|x\|_2^2 \leqslant 1$, мы приходим к следующему результату.

Лемма. *Каковы бы ни были точки x_i, скаляры $c_i \geqslant b_i$ и неотрицательные λ_i, $1 \leqslant i \leqslant p$, все точки множества Q' удовлетворяют квадратичному неравенству*

$$\langle S(\lambda)x, x\rangle - 2\langle s(c, \lambda), x\rangle \leqslant \sigma(c, \lambda), \tag{11}$$

где

$$S(\lambda) = I + \sum_{i=1}^{p} \lambda_i e_i e_i^T; \quad 2s(c, \lambda) = \sum_{i=1}^{p} \lambda_i (c_i - 1) e_i;$$
$$\sigma(c, \lambda) = 1 + \sum_{i=1}^{p} \lambda_i c_i. \tag{12}$$

Заметим, что (11) задает эллипсоид $V'(c, \lambda)$ с центром в точке

$$x(c, \lambda) = S^{-1}(\lambda) s(c, \lambda),$$

причем $(\nu, V'(c, \lambda))$ – локализатор для (g, Q), такой, что

$$\gamma^2(c, \lambda) \equiv \{|V|/|V'(c, \lambda)|\}^2 = \rho^{-n}(c, \lambda) \det(S(\lambda)),$$
$$\rho(c, \lambda) = \sigma(c, \lambda) + \langle S^{-1}(\lambda) s(c, \lambda), s(c, \lambda)\rangle. \tag{13}$$

Теперь наша цель – выбрать $\{x_i\}$, $\{c_i\}$ и $\{\lambda_i\}$ так, чтобы сделать $\gamma^2(c, \lambda)$ по возможности большим. Заметим, что если $p = 1$ и $x_1 = 0$, то $\gamma(c, \lambda)$ допускает максимизацию в явном виде, и результирующая процедура – обычный метод эллипсоидов (в версии "с глубоким отсечением" при оптимальном выборе c_1 ($c_1 = b_1$) или в базовой версии ($c_1 = 0$)).

В случае $p > 1$ ограничимся максимизацией нижней оценки γ. Прежде всего, очевидно, матрица $S(\lambda) - I$ неотрицательно определена, так что замена члена $\langle S^{-1}(\lambda) s(c, \lambda), s(c, \lambda)\rangle$ в выражении для $\rho(c, \lambda)$ на $\|s(c, \lambda)\|_2^2$ не уменьшает ρ и, следовательно, не увеличивает γ. Далее, ввиду $\|e_i\|_2^2 = 1$ получаем $Tr\{S(\lambda) - 1\} = \sum_{i=1}^{p} \lambda_i$, так что собственные числа матрицы $S(\lambda)$ имеют вид $1 + \theta_j$, $\theta_j \geqslant 0$, $\sum_{j=1}^{n} \theta_j = \sum_{i=1}^{p} \lambda_i$, поэтому $\det(S(\lambda)) \geqslant 1 + \sum_{i=1}^{p} \lambda_i$. Итак,

$$\gamma^2(c, \lambda) \geqslant \gamma_*^2(c, \lambda) = \rho_*^{-n}(c, \lambda)(1 + \sum_{i=1}^{p} \lambda_i), \quad \rho_*(c, \lambda) = \sigma(c, \lambda) + \|s(c, \lambda)\|_2^2. \tag{14}$$

3. **Конструкция процедуры** π_k. Рассмотрим следующий, самый прямолинейный способ действий. Зададимся положительными θ и μ и натуральными $k \leqslant n$, и пусть x_i определены рекуррентным соотношением

$$x_1 = 0, \quad x_i = x_{i-1} - \theta e_{i-1}; \tag{15}$$

отметим, что (15) – просто субградиентный спуск.

Процесс (15) будем вести до момента, когда впервые окажется

$$\|x_i\|_2 > \mu n^{-1/2}, \tag{16}$$

если это случится при некотором $i \leqslant k$; в противном случае оборвем процесс при $i = k+1$ (так что число обращений к Ω не превзойдет k). Пусть $p+1$ – максимальное значение i, полученное в ходе процесса (15). Возможны два случая

$$\|x_{p+1}\|_2 \leqslant \mu n^{-1/2} \tag{17}$$

или

$$\|x_{p+1}\|_2 > \mu n^{-1/2}. \tag{17'}$$

А. Начнем с рассмотрения (17'). Прежде всего имеем $p = k$. Кроме того, ввиду (15) $\|x_i\|_2^2 = \|x_{i-1}\|_2^2 - 2\theta\langle e_{i-1}, x_{i-1}\rangle + \theta^2$, что дает

$$\sum_{i=1}^{k} \langle e_i, x_i\rangle = (2\theta)^{-1}(k\theta^2 - \|x_{k+1}\|_2^2). \tag{18}$$

Стало быть, при выборе

$$c_i = \langle e_i, x_i\rangle \tag{19}$$

(допустимом неравенством $c_i \geqslant b_i$ ввиду $\alpha_i \geqslant 0$) и

$$\lambda_i \equiv \pi > 0 \tag{20}$$

получим

$$\sigma(c, \lambda) = 1 + \pi(2\theta)^{-1}(k\theta^2 - \|x_{k+1}\|_2^2) \tag{21}$$

и $2s(c, \lambda) = \pi\left(\sum_{i=1}^{k} c_i e_i + \sum_{i=1}^{k} (-e_i)\right)$. Вторая сумма в последнем выражении, учитывая (15), есть $\theta^{-1} x_{p+1}$. Следовательно,

$$\|s(c, \lambda)\|_2^2 \leqslant (\pi^2/2)\{\|\sum_{i=1}^{k} c_i e_i\|_2^2 + \theta^{-2}\|x_{k+1}\|_2^2\}.$$

В силу $c_i = \langle e_i, x_i\rangle$ и $\|e_i\|_2 = 1$ имеем $\|c_i e_i\|_2 \leqslant \|x_i\|_2$, что при интересующих нас i не превосходит $\mu n^{-1/2}$. Итак, $\|s(c, \lambda)\|_2^2 \leqslant (\pi^2/2)\{k^2\mu^2 n^{-1} + \theta^{-2}\|x_{p+1}\|_2^2\}$, что дает

$$\rho_*(c, \lambda) \leqslant 1 + \pi(2\theta)^{-1}(k\theta^2 - \|x_{k+1}\|_2^2) + (\pi^2/2)\{k^2\mu^2 n^{-1} + \theta^{-2}\|x_{k+1}\|_2^2\}. \tag{22}$$

Чтобы упростить выкладки, примем

$$\pi = \theta, \quad \mu = 1, \tag{23}$$

тогда (22), с учетом $k \leqslant n$, превратится в

$$\rho_*(c, \lambda) \leqslant 1 + \pi^2 k. \tag{24}$$

Следовательно, в условиях (19), (20), (23)

$$\gamma_*^2(c, \lambda) \geqslant (1 + \pi^2 k)^{-n}(1 + k\pi). \tag{25}$$

Остается выбрать в качестве π величину, максимизирующую правую часть (25), т.е. взять

$$\pi = \pi^* = \pi^*(k, n) \equiv (n + (n^2 + (2n-1)k)^{1/2})^{-1}. \tag{26}$$

Проверим, что при этом выборе π будет $\gamma_*(c, \lambda) \geqslant 1 + 0(1)n^{-1}k$. Действительно, правая часть (25) при $\pi = \pi^*$ не меньше, чем при $\pi = 1/(2n)$, когда соответствующее выражение равно $r(k/(2n))$, $r(\tau) = (1+\tau)(1+\tau/(2n))^{-n}$. Имеем

$$r'(\tau) = (1+\tau/(2n))^{-n-1}(1-1/(2n)) \geqslant O(1), \quad 0 \leqslant \tau \leqslant ½,$$

что и дает

$$r(k/(2n)) \geqslant 1 + O(1)k/n.$$

Б. Теперь посмотрим, что происходит в случае (17′). Пусть $x \in Q' = \{x \in Q \mid g(x) \leqslant g(v')\}$, где, напомним, v' – лучшая по значению g среди точек множества, составленного из v и всех лежащих в Q элементов последовательности $x_1, \ldots, x_p$. Как мы видели (см. (9)),

$$\langle e_i, x_i - x\rangle \geqslant 0, \quad 1 \leqslant i \leqslant p. \tag{27}$$

Как и в п. А, имеем $-\sum_{i=1}^{p} e_i = \theta^{-1}x_{p+1}$, $\sum_{i=1}^{p}\langle e_i, x_i\rangle = (2\theta)^{-1}(p\theta^2 - \|x_{p+1}\|_2^2)$, что дает $-\theta^{-1}\langle x_{p+1}, x\rangle \leqslant (2\theta)^{-1}(p\theta^2 - \|x_{p+1}\|_2^2)$, или

$$x \in Q' \Rightarrow \langle x_{p+1}, x\rangle \geqslant 2^{-1}(\|x_{p+1}\|_2^2 - p\theta^2). \tag{28}$$

В условиях (19), (20), (23), (26) $p\theta^2 \leqslant k(\pi^*)^2 \leqslant (2n)^{-1}$, тогда как в случае (17′) $\|x_{p+1}\|_2 \geqslant n^{-½}$. Следовательно, при

$$e = x_{p+1}/\|x_{p+1}\|_2 \tag{29}$$

(28) дает

$$x \in Q' \Rightarrow \langle e, x\rangle \geqslant \nu \equiv 4^{-1}n^{-½}. \tag{30}$$

Последнее означает, что Q' содержится в евклидовом шаре V' радиуса $(1-\nu^2)^{½}$ с центром в точке νe. В силу определения Q' пара (v', V') является локализатором для (g, Q), причем по определению v

$$|V|/|V'| \geqslant 1 + O(1) \geqslant 1 + O(1)k/n. \tag{31}$$

Таким образом, в случае (17′) новый локализатор для (g, Q) можно получить совсем просто, не обращаясь к схеме п. 3.

Суммируем вышеизложенное в следующем предложении.

Теорема. *Имея локализатор вида (u, W) для пары (f, K), задавшись некоторым $k \leqslant n$ и выполнив не более чем k шагов процесса* (15) *с параметрами θ и μ, выбранными по* (19), (20), (23), (26) (*процесс применяется к паре (g, Q), получающейся из (f, k) описанной в* п. 3 *аффинной заменой координат*), *можно построить новый локализатор (u', W') для (f, K), обеспечив неравенство*

$$|W|/|W'| \geqslant \exp\{O(1)k/n\}. \tag{32}$$

Замечание. Полученная процедура π_k очевидным образом удовлетворяет первым двум соотношениям из (8); легко видеть, что при естественной реализации она также удовлетворяет и третьему из соотношений (8).

4. **Обсуждение.** Идея, лежащая в основе процедур π_k, весьма проста. Обычный метод эллипсоидов представляет собой простейший метод отсечений: каждое новое обращение к Ω позволяет отсечь от текущего локализатора-эллипсоида "половину", которую затем приходится заключать в новый эллипсоид, чтобы поддержать геометрическую стабильность форм локализаторов и гарантировать тем самым умеренную ($O(n^2)$) арифметическую стоимость итерации. В наших нынешних процедурах новый локализатор строится после многих (например, порядка n) отсечений. Укажем, что идея использовать несколько отсечений в схеме метода эллипсоидов не нова; она обсуждалась в [3], где, однако, не содержится сколько-нибудь общего результата (в [3] рассмотрено несколько частных случаев расположения отсекающих гиперплоскостей, для которых удается указать "хороший" новый локализатор). Общее утверждение типа сформулированного в (8), насколько нам известно, является новым.

Укажем, еще, что намеченная выше реализация указанной идеи не претендует на окончательность. Даже оставаясь в рамках схемы п. 3, можно оптимизировать выбор параметров π, θ, μ (наш выбор мотивирован стремлением к простоте выкладок, но никак не сутью дела). Далее, построив x_i, скажем, с помощью процесса (15), можно попытаться оптимизировать объем нового локализатора за счет более рационального выбора c_i и λ_i. Даже грубая оценка степени уменьшения объема локализатора, даваемая соотношением (14), позволяет пойти дальше, чем это было продемонстрировано. В самом деле, величина $\gamma_*(c, \lambda)$, которую мы стремимся сделать по возможности большей за счет выбора $c = \{c_1, \ldots, c_p\}$ и $\lambda = \{\lambda_1, \ldots, \lambda_p\}$, на самом деле зависит лишь от $q_1(\lambda) = \sum_{i=1}^{p} \lambda_i$, $q_2(\lambda, c) = \sum_{i=1}^{p} \lambda_i c_i$ и $q_3(\lambda) = \| \sum_{i=1}^{p} \lambda_i e_i \|_2$, и можем попытаться ее оптимизировать многими способами (скажем, выбрать вначале c_i и λ_i так же, как и выше, подсчитать для этого выбора значения трех указанных величин – ξ_1, ξ_2, ξ_3 – и решить задачу минимизации q_3 выбором неотрицательных λ (при тех же c) при ограничениях $q_1(\lambda) \geqslant \xi_1$, $q_2(\lambda, c) \leqslant \xi_2$). Таким образом, представляется, что "многошаговые" версии метода эллипсоидов более гибкие, чем исходная, и предоставляют некоторые возможности для применения разного рода ускоряющих процедур.

ЛИТЕРАТУРА

1. *Юдин Д.Б., Немировский А.С.* Информационная сложность и эффективные методы решения выпуклых экстремальных задач // Экономика и мат. методы. 1976. т. XII. Вып. 2.
2. *Шор Н.З.* Метод отсечения с растяжением пространства для решения задач выпуклого программирования // Кибернетика. 1977. Вып. 1.
3. *Шор Н.З., Гершович В.И.* Метод эллипсоидов, его обобщения и приложения // Кибернетика, 1982. Вып. 5.

Поступила в редакцию
30 X 1990