Н. А. С о к о л о в (Москва, ЦЭМИ РАН). **О точности решения вспомогательных задач ЛП и КП в методе уровней.**

Метод уровней [1], [2] был разработан как эффективный метод оракульного типа для решения задач оптимизации и отыскания седловых точек. Он получил дальнейшее развитие в работах [3]–[7], в которых рассмотрены варианты оракула более сложного вида, приводящие к обобщенным методам уровней.

1. Суть *базового* метода уровней такова. Пусть имеется многогранник $Q \subset \mathbf{E}$ ($\mathbf{E}$ — конечномерное пространство), содержащий некоторый непустой выпуклый компакт Q^*. Нашей целью является нахождение какой-нибудь точки из Q^* либо приближение к множеству Q^* с некоторой точностью. Каждой точке $z \in Q$ приписано непустое множество ответов (откликов) оракула $\mathbf{B}(z) = \{(b(z), \beta(z)\}$, ответом оракула считается пара $(b(z), \beta(z))$, где $b(z)$ — вектор, $\beta(z)$ — число. Множество $\mathbf{B}(z)$ содержит набор $b(z) = 0$, $\beta(z) = 0$ лишь для точек $z \in Q^*$. Для остальных наборов $\|b(z)\| \neq 0$ и каждое полупространство точек $y \in E$, для которых выполнено неравенство

$$\langle b(z), y - z\rangle + \beta(z) \leqslant 0, \tag{1}$$

содержит множество Q^*.

В методе уровней при $z \in Q \setminus Q^*$ оракул выдает один или несколько наборов векторов $b = b(z)$, $\|b(z)\| \neq 0$, и неотрицательных чисел $\beta = \beta(z)$, $(b, \beta) \in \mathbf{B}(z)$, таких, что для каждого набора выполнено (1), если же оракул определяет $b(\overline{z}) = 0$ и $\beta(\overline{z}) = 0$ для некоторой точки $\overline{z} \in Q$, то заведомо $\overline{z} \in Q^*$.

Опишем вариант метода уровней, использующий введенный оракул. Зададим множества $S_0 = \varnothing$, $\Lambda_0 = \varnothing$, $Q_0 = Q$ и число $n_0 = 0$. Выберем положительно определенную симметричную матрицу H и произвольную точку $z_1 \in Q$.

Пусть уже построено k ($k \geqslant 1$) точек $z_i \in Q$, $i = 1, 2, \dots, k$, определено множество S_{k-1}, $|S_{k-1}| = n_{k-1}$. С точкой z_k свяжем m_k ($m_k \geqslant 1$) откликов оракула, т. е. зададим m_k наборов векторов $b_j = b_j(z_k)$ и чисел $\beta_{j,k} = \beta_{j,k}(z_k)$, $j \in J_k = \{n_{k-1} + 1, \dots, n_{k-1} + m_k\}$, и для каждого набора выполняется условие (1). Положим $S_k = S_{k-1} \cup J_k$, где $|J_k| = m_k$, $|S_k| = n_k = n_{k-1} + m_k$. Далее будем предполагать, что $\|b_j\| \neq 0$ для всех $j \in J_k$ (иначе точка $z_k \in Q^*$), и что для каждого $k \geqslant 1$ найдется хотя бы один индекс $\widehat{\jmath} \in J_k$, для которого $\beta_{\widehat{\jmath},k} \geqslant 0$.

Будем предполагать также, что многогранник

$$Q_k \equiv \{z \in Q: \ \ \langle b_j, z - z_i\rangle + \beta_{j,k} \leqslant 0, \ \ j \in J_i, \ \ i = 1, \dots, k\},$$

где числа $\beta_{j,k} \geqslant \beta_{j,k-1}$ для всех $j \in J_i$, $i = 1, \dots, k-1$, является непустым множеством и содержит Q^*.

Поиск следующей точки z_{k+1} в методе уровней начинается с решения задачи линейного программирования

$$\mathcal{A}_k: \quad t \to \max, \quad z \in Q, \quad \langle b_j, z - z_i\rangle + \beta_{j,k} + t \leqslant 0, \qquad j \in J_i, \quad i = 1, \dots, k, \tag{2}$$

относительно переменных z и t. Обозначим Δ_k ее оптимальное значение, $\Delta_k \geqslant 0$. В силу своего определения числа Δ_k не возрастают с ростом k. При помощи Δ_k производится оценка сверху меры близости приближенного решения исходной оптимизационной задачи к множеству Q^*.

Если Δ_k «достаточно мало», то процесс прекращается. В противном случае к множеству Λ_{k-1} добавляем $m_k \geqslant 1$ чисел $\lambda_j \in (0, 1)$, $j \in J_k$, полагаем $\Lambda_k = \Lambda_{k-1} \cup \{\lambda_j,\ j \in J_k\}$, и (непустой) многогранник $Q_k(\Lambda_k, \Delta_k)$ определяется как

$$\{z \in Q: \quad \langle b_j, z - z_i\rangle + \beta_{j,k} + \lambda_j \Delta_k \leqslant 0, \qquad j \in J_i, \quad i = 1, \dots, k\}. \tag{3}$$

В заключение k-й итерации, новая точка z_{k+1} находится как проекция точки z_k на многогранник $Q_k(\Lambda_k, \Delta_k)$, т. е. решается задача квадратичного программирования

$$\mathcal{B}_k: \quad \langle H(z - z_k), z - z_k \rangle \to \min, \quad z \in Q_k(\Lambda_k, \Delta_k). \tag{4}$$

Далее будем предполагать, что $0 < \underline{\lambda} \leqslant \lambda_j \leqslant \overline{\lambda} < 1$ для всех $\lambda_j \in \Lambda_k$ и $\|b_j\| \leqslant \overline{B}$ для всех $j \in S_k$, $k \geqslant 1$, где ответы оракула задаются системой неравенств (1). Пусть d — диаметр Q, $\mu = h_{\max}/h_{min}$ — число обусловленности матрицы H.

Лемма 1. *Для исследуемого варианта метода уровней последовательность* $\{\Delta_k\}$, *которую вырабатывает метод, удовлетворяет неравенству* (*при* $\sigma = 1$)

$$\Delta_k \leqslant \mathcal{Q}\, k^{-1/2}, \quad k = 1, 2, \ldots, \quad \text{где} \quad \mathcal{Q} = \frac{\overline{B} d \mu^{1/2}}{\underline{\lambda}(1 - \overline{\lambda}^2)^{1/2}}\, \sigma. \tag{5}$$

Лемма 1 в [6] доказана в случае, когда при определении множества $Q_k(\Lambda_k, \Delta_k)$ и, следовательно, новой точки z_{k+1}, вместо оптимального значения Δ_k задачи $\mathcal{A}_k$ можно использовать приближенное значение $\widehat{\Delta}_k$ $(0 \leqslant \widehat{\Delta}_k \leqslant \Delta_k)$ этой задачи. Более того, и квадратичную задачу $\mathcal{B}_k$ можно решать приближенно.

Пусть $z_{k+1} \in G$ — приближенное решение квадратичной задачи $\mathcal{B}_k$, т. е. такая точка из $Q_k(\Lambda_k, \Delta_k)$, для которой

$$\begin{aligned} \langle H(z_k - \overline{z}_{k+1}), z_k - \overline{z}_{k+1} \rangle &\leqslant \langle H(z_k - z_{k+1}), z_k - z_{k+1} \rangle \\ &\leqslant \langle H(z_k - \overline{z}_{k+1}), z_k - \overline{z}_{k+1} \rangle (1 + \tau_k^2), \end{aligned} \tag{6}$$

где τ_k^2 $(k = 1, 2, \ldots)$ — относительная погрешность решения квадратичной задачи $\mathcal{B}_k$, $\overline{z}_{k+1}$ — ее точное решение.

Применив разложение $H = \overline{H}^T \overline{H}$, можно получить $\|\overline{H}(z_{k+1} - \overline{z}_{k+1})\| \leqslant \tau_k \|\overline{H}(z_k - z_{k+1})\|$, и τ_k можно считать относительной погрешностью решения квадратичной задачи $\mathcal{B}_k$ (в метрике $\|\cdot\|_{\overline{H}}$).

Лемма 2. *Пусть выполнены все предположения леммы* 1, *задача* $\mathcal{B}_k$ *решается приближенно с погрешностью* τ_k, $k \geqslant 1$, *а также выполнены условия* (6), *и пусть* $S = \sum_{k=1}^{\infty} (2\tau_k + \tau_k^2)$, *где*

$$\sum_{k=1}^{\infty} \tau_k < \infty. \tag{7}$$

Тогда последовательность $\{\Delta_k\}$ (*при точном или приближенныом решении задач* $\mathcal{A}_k$), *вырабатываемая методом уровней, удовлетворяет условию* (5) *при* $\sigma = 1 + S$. *Если условие* (7) *заменить на условие*

$$0 \leqslant \tau_k \leqslant \mathbf{C} \|\overline{H}(z_k - z_{k+1})\|, \quad 0 \leqslant \mathbf{C} < \overline{\mathbf{C}} \equiv \frac{\sqrt{2} - 1}{d\sqrt{h_{\max}}},$$

то последовательность $\{\Delta_k\}$ (*при точном или приближенном решении задач* $\mathcal{A}_k$), *вырабатываемая методом уровней, удовлетворяет условию* (5) *при* $\sigma = 1/(1 - \widetilde{S})$, *где* $\widetilde{S} = 2\mathbf{C} d\sqrt{h_{\max}} + \mathbf{C}^2 d^2 h_{\max}$, $0 \leqslant \widetilde{S} < 1$.

2. Рассмотрим теперь вариант оракула, соответствующий *обобщенному* методу уровней. Пусть в многограннике Q содержится выпуклый компакт $\overline{Q}$, который, в свою очередь, содержит Q^* (множества Q и Q^* определены ранее). Множества откликов оракула J_k и S_k разделим на два непересекающихся подмножества: $J_k = J_k' \cup J_k''$, $J_k' \cap J_k'' = \varnothing$, $S_k' = \cup_{1 \leqslant i \leqslant k} J_k'$, $S_k'' = \cup_{1 \leqslant i \leqslant k} J_k''$, $k = 1, 2, \ldots$. Это разделение производится по принципу принадлежности точки z_k множеству $\overline{Q}$. К первому типу относим точки $z_k \in \operatorname{int} \overline{Q}$, для них любой отклик оракула гарантирует лишь то, что в полупространстве, задаваемом линейным неравенством (1), содержится множество

Q^*. Ко второму типу отнесем точки $z_k \in Q \setminus \overline{Q}$, для них любой отклик оракула гарантирует, что полупространство, задаваемое неравенством (1), отделяет z_k от $\overline{Q}$. Для граничных точек $z_k \in \partial\overline{Q}$ оракул может выдавать отклики как первого, так и второго типов.

При формулировке леммы 3 *не* предполагается, что точки $\{z_i\}$, $1 \leqslant i \leqslant k$, построены с помощью решения задачи квадратичной задачи $\mathcal{B}_k$. Пусть z_i, $1 \leqslant i \leqslant k$ — *произвольный* набор точек из Q, $z_i \neq z_j$ при $i \neq j$, для каждой точки z_i множество ответов оракула J_i разбито на два непересекающихся подмножества J_i' и J_i'', $1 \leqslant i \leqslant k$, и пусть Δ_k — оптимальное значение задачи линейного программирования $\mathcal{A}_k$.

Лемма 3 [3]. *Пусть множество $\overline{Q}$ содержит шар радиуса $r > 0$. Тогда, если $\Delta_k \leqslant \underline{B}r$, где $0 < \underline{B} = \inf_{j \in S_k'',\, k \geqslant 1} \|b_j\| >$ то множество $S_k' \equiv \cup_{1 \leqslant i \leqslant k} J_i' \neq \varnothing$ и*

$$\Delta_k = \max_{z \in Q} \min_{1 \leqslant i \leqslant k} \begin{cases} \min\limits_{j \in J_i'} \{\langle b_j, z_i - z\rangle - \beta_{j,k}\}, & J_i' \neq \varnothing, \\ +\infty, & J_i' = \varnothing. \end{cases}$$

3. Пусть функция $f(z)$ определена на выпуклом компакте $\overline{Q} \subset Q \subset \mathbf{E}$ ($\overline{Q}$, Q и $\mathbf{E}$ заданы ранее). Пусть $f(z)$ выпукла и удовлетворяет на $\overline{Q}$ условию Липшица с константой L. Обозначим $\partial f(z_k) = \{l(z_k)\} \neq \varnothing$ субдифференциал функции в точке z_k, Q^* — (непустое) множество решений оптимизационной задачи $f(z) \to \min_{z \in \overline{Q}}$. В качестве добавленных отсечений в точке z_k оракул обычно выдает набор(ы) $(b(z_k), \beta(z_k))$ вида (см. [6])

$$b_j = b_j(z_k) = \begin{cases} \dfrac{1}{n_{jk}} l_j(z_k), & j \in J_k', \\ a_j(z_k), & j \in J_k'', \end{cases} \quad \beta_{j_k} = \beta_{jk}(z_k) = \begin{cases} \dfrac{1}{n_{jk}} \rho(f(z_k) - f_k^*), & j \in J_k', \\ \alpha_j(z_k), & j \in J_k'', \end{cases} \tag{8}$$

где $l_j(z_k) \in \partial f(z_k)$, $n_{jk} \equiv 1$ либо $n_{jk} = \|l_j(z_k)\|$ (ненормированный либо нормированный метод уровней), $\rho = 1$ или $\rho = 0$ («с глубоким отсечением» или без него); $f_k^* = \min_{1 \leqslant i \leqslant k:\, z_i \in \overline{Q}} f(z_i)$ — «рекорд» на k-й итерации (вначале он равен $+\infty$), $f^* = \min_{z \in \overline{Q}} f(z)$; обычно для отсечений $j \in J_k''$ выбираются $\|a_j(z_k)\| = 1$, $\alpha_j(z_k) \geqslant 0$.

Лемма 4. *Пусть отклики оракула задаются формулой* (8) *и выполнены условия леммы* 3. *Тогда если $\Delta_k < r$, то $J_k' \neq \varnothing$, причем для меры близости z_k к Q^* верна оценка*

$$f_k^* - f^* \leqslant \mathcal{C}\Delta_k, \quad \text{где} \quad \mathcal{C} = \begin{cases} L(d-r)/r + 1 & \text{для нормированного варианта}, \\ Ld/r & \text{для ненормированного варианта}. \end{cases}$$

Теорема. *Пусть выполнены предположения лемм 1–4, $\widehat{\Delta}_k$ — приближение к оптимальному значению Δ_k задачи $\mathcal{A}_k$, отыскиваемое с погрешностью $0 \leqslant \delta = \Delta_k - \widehat{\Delta}_k < r$, и пусть $\widehat{\Delta}_k < r - \delta$. Тогда $f_k^* - f^* \leqslant \mathcal{C}(\mathcal{Q}k^{-1/2} + \delta)$ при $k > \mathcal{Q}^2/(r-\delta)^2$.*

Для нахождения седловых точек также могут быть получены результаты, аналогичные изложенным для оптимизационных задач (см. [7]).

Работа выполнена при финансовой поддержке РФФИ, проекты № 09–01–00156 и № 09–06–00218.

СПИСОК ЛИТЕРАТУРЫ

1. *Lemaréchal C.*, *Nemirovskii A.*, *Nesterov Yu.* New Variants of Bundle Methods. — Math. Progr., 1995, v. 69, p. 111–147.
2. *Гольштейн Е. Г.*, *Немировский А. С.*, *Нестеров Ю. Е.* Метод уровней, его обобщения и приложения. — Экономика и матем. методы, 1995, т. 31, № 3, с. 164–180.
3. *Бэр К.*, *Гольштейн Е. Г.*, *Соколов Н. А.* Об использовании метода уровней для минимизации выпуклых функций, не все значения которых конечны. — Экономика и матем. методы, 2000, т. 36, № 4, с. 95–107.

4. *Бэр К.*, *Гольштейн Е. Г.*, *Соколов Н. А.* Метод отыскания седловой функции, область определения которой содержится в многограннике. — Экономика и матем. методы, 2001, т. 37, № 3, с. 97–105.
5. *Гольштейн Е. Г.* Обобщенный седловой вариант метода уровней. — Журнал вычислит. матем. и матем. физ., 2001, т. 41, № 8, с. 1139–1147.
6. *Соколов Н. А.* Новые варианты обобщенного метода уровней для минимизации выпуклой недифференцируемой функции, не все значения которой конечны. — Журнал вычислит. матем. и матем. физики, 2007, т. 47, № 12, с. 2037–2054.
7. *Соколов Н. А.* Новые модификации обобщенного седлового варианта метода уровней. — Журнал вычислит. матем. и матем. физики, 2009, т. 49, № 1, с. 26–50.