ИЗВЕСТИЯ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

2005 МАТЕМАТИКА № 12 (523)

УДК 519.854

А.С. СТРЕКАЛОВСКИЙ, Т.В. ЯКОВЛЕВА

МОДИФИКАЦИЯ МЕТОДА РОЗЕНА В ОБРАТНО-ВЫПУКЛОЙ ЗАДАЧЕ

1. Введение

Рассмотрим задачу математического программирования

$$h(x) \downarrow \min, \quad x \in S, \quad g(x) \geq 0, \tag{P}$$

где $h(\cdot)$, $g(\cdot)$ — выпуклые на $\mathbb{R}^n$ функции, S — выпуклое множество из $\mathbb{R}^n$, $g(\cdot)$ — непрерывно дифференцируемая на выпуклом открытом множестве Ω из $\mathbb{R}^n$ функция, $S \subset \Omega$, $\{x \in \mathbb{R}^n \mid g(x) = 0\} \subset \Omega$. Пусть

$$\mathcal{V}(P) \triangleq h_* := \inf(h, D) \triangleq \inf\{h(x) \mid x \in S,\ g(x) \geq 0\} > -\infty. \tag{1}$$

Специфика *задачи* (P) определяется обратно-выпуклым ограничением $g(x) \geq 0$, которое задает дополнение выпуклого открытого множества $\{x \in \mathbb{R}^n \mid g(x) < 0\}$. Естественно предполагать, что $g(x_*) = 0$, где x_* — решение задачи (P), т.к. если это предположение не выполнено, то решение релаксированной задачи

$$h(x) \downarrow \min, \quad x \in S, \tag{PW}$$

которая проще, чем (P), является решением задачи (P).

С помощью функции $\mathcal{V}(P)$, определенной в (1), можно предложить и другую форму условия того, что обратно-выпуклое ограничение активно на решениях задачи (P):

$$\exists v \in S\ g(v) < 0 : h(v) < h_* \triangleq \mathcal{V}(P). \tag{2}$$

Одним из следствий условия (2) является то, что при решении задачи (PW) каким-нибудь из методов оптимизации [1] с произвольной начальной точкой $x^0 \in S$, $g(x^0) > 0$, будет непременно найдена точка $\widetilde{x}$ такая, что $g(\widetilde{x}) = 0$, $h(v) < h(\widetilde{x}) < h(x^0)$. Эта процедура является спуском на поверхность уровня $g(x) = 0$.

Один из методов локального поиска связан с условиями глобальной оптимальности [2] и заключается в отыскании точки на поверхности уровня $g(x) = 0$, а также в последовательном решении линеаризованных задач (которые, как нетрудно видеть, являются задачами выпуклого программирования):

$$(LQ(u, \beta)): \ \langle \nabla g(u), x \rangle \uparrow \max, \quad x \in S, \quad h(x) \leq \beta, \tag{3}$$

где u и β — параметры задачи (3) [2].

Этот метод в [2] назван специальным методом локального поиска (СМЛП). С его помощью получается (приближенно) стационарная точка для ассоциированной задачи выпуклой максимизации (при некотором β) [3]

$$(Q(\beta)): \ g(x) \uparrow \max, \quad x \in S, \quad h(x) \leq \beta,$$

а не для исходной задачи (P).

В данной статье используется подход Розена [4], основанный на линеаризации функции $g(\cdot)$, задающей обратно-выпуклое ограничение; описывается метод отыскания локального решения, являющийся модификацией метода Розена, исследуется его сходимость; описываются результаты вычислительного эксперимента.

2. Модифицированный метод Розена

Розен в [4] предложил метод, идея которого заключается в последовательном решении линеаризованных (в точке u) задач

$$(PL(u)): \ h(x) \downarrow \min, \quad x \in S, \quad \langle \nabla g(u), x - u \rangle + g(u) \geq 0, \tag{4}$$

так что каждая точка x^{s+1} является (точным) решением линеаризованной (в точке x^s) задачи $(PL(x^s))$ вида (4).

Заметим, что целевая функция $h(\cdot)$ задачи (P) остается целевой и в задаче (4), в отличие от специального метода локального поиска [2], где функция $h(\cdot)$ задает ограничение в задаче (3).

В работах [4], [5] была изучена сходимость метода Розена в предположении о компактности допустимого множества задачи (4) и точном решении этой линеаризованной задачи. Далее будет рассмотрена модификация метода Розена и изучена ее сходимость без предположения о компактности допустимого множества и с приближенным решением вспомогательной задачи (4).

Пусть в задаче (P) выполнено (при некотором $\varkappa > 0$) условие невырожденности

$$\forall y \in S \ \ g(y) = 0 \ \ \exists d = d(y) \in S: \ \langle \nabla g(y), d - y \rangle \geq \varkappa > 0, \tag{5}$$

а также условие (2).

Представим метод локального поиска для задачи (P), состоящий из следующих двух процедур.

Процедура 1, начиная с точки $y \in S$, $g(y) \geq 0$, строит точку $x = x(y)$ со свойствами $x \in S$, $g(x) = 0$, $h(x) \leq h(y)$, т. е.

— если $g(y) = 0$, то полагаем $x(y) = y$;

— если же $g(y) > 0$, то в силу выпуклости функции $g(\cdot)$ найдется $\mu \in]0, 1[$ такое, что $g(x_\mu) = 0$, где $x_\mu = \mu v + (1 - \mu)y$, поскольку $g(y) > 0 > g(v)$, где v — вектор из условия (2). Тогда полагаем $x(y) = x_\mu$. При этом в силу выпуклости $h(\cdot)$ справедлива цепочка неравенств

$$h(x(y)) = h(x_\mu) \leq \mu h(v) + (1 - \mu)h(y) < \mu h_* + (1 - \mu)h(y) \leq h(y).$$

Процедура 2 по допустимой точке $\widetilde{x} \in S$, $g(\widetilde{x}) = 0$, строит последовательность $\{u^r\}$ такую, что $(r = 0, 1, 2, \dots)$

$$u^r \in S, \quad g(u^r) \geq 0, \quad h(u^r) \leq h(\widetilde{x}), \quad r = 0, 1, 2, \dots, \tag{6}$$

где $u^0 := \widetilde{x}$. Последовательность $\{u^r\}$ формируется по следующему правилу. Если известна точка u^r, $r \geq 0$, удовлетворяющая (6), то следующая точка u^{r+1} строится как приближенное решение задачи выпуклого программирования

$$(PL_r): \ h(x) \downarrow \min, \quad x \in S, \quad g(u^r) + \langle \nabla g(u^r), x - u^r \rangle \geq 0. \tag{7}$$

Другими словами, справедливо неравенство

$$h(u^{r+1}) \leq \inf\{h(x) \mid x \in S, \ g(u^r) + \langle \nabla g(u^r), x - u^r \rangle \geq 0\} + \delta_r, \tag{8}$$

где $\{\delta_r\}$ — числовая последовательность такая, что

$$\delta_r > 0, \quad r = 0, 1, 2, \dots, \quad \sum_{r=1}^{\infty} \delta_r < +\infty. \tag{9}$$

Обозначим допустимое множество задачи (7) через

$$D_r := D(u^r) \triangleq \{x \in S \mid g(u^r) + \langle \nabla g(u^r), x - u^r \rangle \geq 0\}.$$

Пусть $u_0 \in S$ — некоторая точка.

Предложение. i) *Предположим, что выполнено условие* (1), *а функции* $h(\cdot)$ *и* $g(\cdot)$ *выпуклы и непрерывно дифференцируемы на* Ω. *Тогда последовательность* $\{u^r\}$, *генерируемая процедурой* 2, *удовлетворяет условию*

$$\lim_{r\to\infty}[\inf\{h(x) - h(u^{r+1}) \mid x \in D_r\}] = 0. \tag{10}$$

ii) *При этом, если множество* $F_0 = \left\{x \in S \mid h(x) \le h(u_0) + \sum\limits_{r=1}^{\infty} \delta_r\right\}$ *ограничено, то всякая предельная точка* u^* *последовательности* $\{u^r\}$ *является решением задачи*

$$h(x) \downarrow \min, \quad x \in D(u^*). \tag{11}$$

iii) *В случае замкнутости множества* S *эта предельная точка* u^* *является нормально стационарной в задаче* (P).

Доказательство. i) Из неравенства (8) получаем

$$\begin{aligned} -\delta_r \le \inf\{h(x) \mid x \in D_r\} - h(u^{r+1}) = \\ = \inf\{h(x) - h(u^r) \mid x \in D_r\} + h(u^r) - h(u^{r+1}) \le h(u^r) - h(u^{r+1}). \end{aligned} \tag{12}$$

Отсюда имеем $h(u^{r+1}) - \delta_r \le h(u^r)$, так что числовая последовательность $\{h(u^r)\}$ является почти монотонно убывающей. Принимая во внимание (9) и ограниченность снизу функции $h(\cdot)$ на допустимом множестве задачи (P), по лемме 2 ([1], с. 87, § 2.6) заключаем, что существует конечный предел числовой последовательности $\{h(u^r)\}$: $\lim\limits_{r\to\infty} h(u^r) = h_*$. При переходе к пределу в соотношениях (12) получаем (10).

ii) Согласно (8) $u^{r+1} \in D_r \subset S$ и $h(u^{r+1}) \le h(u^r) + \delta_r$, следовательно, $\{u^r\} \subset F_0$, а F_0 ограничено. Поэтому с точностью до подпоследовательности можно считать, что $u^r \to u^*$.

Далее покажем полунепрерывность сверху [6] в точке $u^* \in S$ многозначного отображения $u^r \to D(u^r)$, а именно, что $x^* \in D_*$, если $u^r \to u^*$, $x^r \in D_r$, $x^r \to x^*$.

Поскольку $x^r \in D_r$, то это означает $x^r \in S$ и

$$g(u^r) + \langle \nabla g(u^r), x^r - u^r\rangle \ge 0. \tag{13}$$

В силу непрерывной дифференцируемости функции $g(\cdot)$ и непрерывности скалярного произведения при переходе к пределу в (13) при $r \to \infty$ получаем

$$g(u^*) + \langle \nabla g(u^*), x^* - u^*\rangle \ge 0.$$

Это означает принадлежность точки x^* множеству D_*, а, следовательно, и полунепрерывность сверху в точке $u^* \in S$ многозначного отображения $u \to D(u)$.

Тот факт, что предельная точка u^* является решением задачи (11), следует из (10) по непрерывности функции $h(\cdot)$ и полунепрерывности сверху многозначного отображения $u^r \to D(u^r)$.

iii) Поскольку u^* является решением задачи (11), то найдутся множители $\lambda_0, \lambda \ge 0$, $\lambda_0 + \lambda > 0$, при которых

$$\begin{gathered} \langle \lambda_0 \nabla h(u^*) - \lambda \nabla g(u^*), x - u^*\rangle \ge 0 \;\; \forall x \in S, \\ \lambda g(u^*) = 0. \end{gathered} \tag{14}$$

Обратим внимание на то, что (14) является системой оптимальности для задачи (P).

Покажем, что множители λ_0 и λ строго положительны в (14).

Пусть $\lambda_0 = 0$. Тогда $\lambda > 0$, и из системы (14) следует

$$\langle \nabla g(u^*), x - u^*\rangle \le 0 \;\; \forall x \in S, \quad g(u^*) = 0,$$

что противоречит условию (5). Значит, $\lambda_0 > 0$.

Пусть теперь $\lambda = 0$. Тогда $\lambda_0 > 0$, и из системы (14) получаем

$$\langle \nabla h(u^*), x - u^* \rangle \geq 0 \ \ \forall x \in S.$$

Отсюда в силу выпуклости функции $h(\cdot)$ имеем

$$0 \leq \langle \nabla h(u^*), x - u^* \rangle \leq h(x) - h(u^*) \ \ \forall x \in S,$$

или $h(u^*) \leq h(x)$ $\forall x \in S$, что противоречит условию (2). Значит, $\lambda > 0$. $\square$

Теперь рассмотрим вопрос о том, как с помощью последовательности $\{u^r\}$, где $u_0 = \widetilde{x}$, определить точку $y(\widetilde{x})$. Для этого зададим число $\varepsilon > 0$. Тогда, если $\delta_r \leq \varepsilon$ и справедливо неравенство

$$h(u^r) - h(u^{r+1}) \leq \varepsilon, \tag{15}$$

то полагаем $y = y(\widetilde{x}, \varepsilon) = u^{r+1}$, поскольку в этом случае из (12) вытекает

$$\inf\{h(x) - h(u^{r+1}) \mid x \in D_r\} \leq \varepsilon,$$

т. е. u^{r+1} является ε-решением линеаризованной задачи (7).

Объединяя процедуры 1 и 2 в один метод, запишем его в форме алгоритма.

Пусть задана числовая последовательность $\{\varepsilon_s\}$, $\varepsilon_s > 0$, $s = 0, 1, 2, \ldots$, $\varepsilon_s \downarrow \infty$.

Модифицированный метод Розена (ММР)

Шаг 0 (инициализация). Положить $s := 0$, $x^s := x_0$.

Шаг 1 (процедура 2). Начиная с точки x^s, построить точку $y^s = y(x^s, \varepsilon_s)$, являющуюся ε_s-решением линеаризованной задачи (7), т. е.

$$h(x) + \varepsilon_s \geq h(y^s) \ \ \forall x \in S : \ \langle \nabla g(y^s), x - y^s \rangle + g(y^s) \geq 0.$$

Шаг 2 (критерий останова). Если $g(y^s) \leq 0$, то stop.

Шаг 3 (процедура 1). С помощью y^s построить точку $u := x(y^s)$,

$$u \in S, \ \ g(u) = 0, \ \ h(u) < h(y^s) \leq h(x^s).$$

Шаг 4. Положить $x^{s+1} := u$, $s := s + 1$ и вернуться к шагу 1. $\square$

Лемма. *Модифицированный метод Розена*

i) *либо конечен, причем $g(y^N) = 0$, где N — итерация останова;*

ii) *либо генерирует две последовательности $\{x^s\}$ и $\{y^s\}$ со свойствами*

$$x^s \in S, \ \ g(x^s) = 0, \ \ y^s \in S, \ \ g(y^s) > 0, \tag{16}$$

$$h(x^{s+1}) < h(y^s) \leq h(x^s). \tag{17}$$

При этом справедливо

$$\lim_{s\to\infty} h(x^s) = \lim_{s\to\infty} h(y^s). \tag{18}$$

Доказательство. i) Согласно критерию останова (шаг 3), если N — итерация останова, то $g(y^N) \leq 0$. Поскольку для каждой точки y^{r+1}, $r = 0, 1, 2, \ldots$, выполнено

$$0 \leq g(y^r) + \langle \nabla g(y^r), y^{r+1} - y^r \rangle \leq g(y^r) + g(y^{r+1}) - g(y^r) = g(y^{r+1}), \ \ r = 0, 1, 2, \ldots,$$

то и $g(y^s) \geq 0$, $s = 0, 1, 2, \ldots$ Поэтому $g(y^N) = 0$.

ii) Если процесс счета бесконечен, то благодаря выпуклости $h(\cdot)$ и по построению $x^{s+1} = \mu_s v + (1 - \mu_s) y^s$ получаем (17). А поскольку

$$-\infty < h(v) < h(x^{s+1}) < h(y^s) \leq h(x^s),$$

то существует предел $\lim h(x^s) \geq h(v)$, откуда получаем (18). $\square$

Теорема. *Пусть функция $h(\cdot)$ и множество S выпуклы, а множество*

$$F = \{x \in S \mid h(x) \le h(x_0)\}$$

ограничено, и, кроме того, справедливо условие (2).

Тогда модифицированный метод Розена

i) *при конечном числе итераций получает точку $y^N \in S$, $g(y^N) = 0$, которая является ε_N-решением линеаризованной задачи* (7), *где N — номер итерации останова;*

ii) *при бесконечном числе итераций для последовательностей $\{y^s\}$ и $\{x^s\}$, обладающих свойствами* (16)–(18), *справедливы условия*

$$0 = g(x^s) = \lim_{s\to\infty} g(y^s), \quad y_* = \lim_{s\to\infty} x^s = \lim_{s\to\infty} y^s \tag{19}$$

для некоторого $y_* \in \mathbb{R}^n$, $g(y_*) = 0$.

Кроме того, точка y_* является решением линеаризованной задачи

$$\begin{gathered}(PL_*): \ h(x) \downarrow \min, \quad x \in S, \quad g(y_*) + \langle \nabla g(y_*), x - y_* \rangle \ge 0, \\ \text{т. е. } h(x) \ge h(y_*) \ \ \forall x \in S: \ \langle \nabla g(y_*), x - y_* \rangle + g(y_*) \ge 0,\end{gathered} \tag{20}$$

и нормально стационарной в задаче (P).

Доказательство. Утверждение i) вытекает из леммы и (16).

ii) Из (17), (18) и выпуклости функции $h(\cdot)$ следует

$$h(x^{s+1}) \le \mu_s[h(v) - h(y^s)] + h(y^s)$$

или

$$0 \le \mu_s[h_* - h(v)] \le \mu_s[h(y^s) - h(v)] \le h(y^s) - h(x^{s+1}),$$

где $h_* \triangleq \mathcal{V}(P)$, поэтому из (18), поскольку $h(y^s) \ge h_* > h(v)$, получаем

$$\lim_{s\to\infty} \mu_s = 0, \quad \mu_s \in]0,1[.$$

По построению $\{x^s\} \subset F$, а F ограничено. Поэтому с точностью до подпоследовательности можно считать, что $x^s \to x_*$.

Далее, поскольку $x^{s+1} = y^s + \mu_s(v - y^s)$, то

$$\|y^s - x_*\| = \|x^{s+1} + \mu_s(y^s - v) - x_*\| \le \|x^{s+1} - x_*\| + \mu_s\|y^s - v\| \le \|x^{s+1} - x_*\| + \mu_s K$$

для некоторого $K > 0$ такого, что $\|y^s - v\| \le K$. Эта оценка справедлива, т. к. $y^s \in F_0$, $s = 0,1,2,\ldots$ Поэтому в силу равенств $g(x^s) = 0$, $s = 0,1,2,\ldots,$ и непрерывности $g(\cdot)$ имеем $0 = g(x^s) = \lim_{s\to\infty} g(y^s) = 0$, тем самым условия (19) доказаны.

Наконец, по построению (см. (17)) для каждого $y^s = y(x^s, \varepsilon_s)$ согласно (12) и (15) справедливо неравенство $\inf\{h(x) - h(y^s) \mid x \in D_s\} \le \varepsilon_s$, откуда при $s \to \infty$ по непрерывности функции $h(\cdot)$ и непрерывности сверху для каждого $y^s \in S$ многозначного отображения $y^s \to D(y^s) \equiv D_s$ следует (20).

Нормальная стационарность точки y_* доказывается так же, как в предложении выше (см. также предложение 5.2.1 из [2], с. 194).

3. Численный эксперимент

Сравнение эффективности специального метода локального поиска (СМЛП) [2] и модифицированного метода Розена (ММР) проведено на серии тестовых задач, предоставленных K. Moshirvaziri (www.csulb.edu/~moshir/Test_Problems):

$$\begin{gathered} h(x) = \langle c, x\rangle \downarrow \min, \quad x \in S = \{x \mid Ax \le b,\ x \ge 0\}, \\ g_l(x) \ge 0, \end{gathered} \tag{P$_l$}$$

где $A \in \mathcal{M}_{m\times n}$, $b \in \mathbb{R}^m$, $c, x \in \mathbb{R}^n$, а $g_l(\cdot)$ выпукла и дифференцируема на $\mathbb{R}^n$.

В качестве примера функции $g_l(x)$, $l = 1, 2$, выбирались в виде

$$\begin{aligned} g_1(x) &\triangleq \|x\|^2 - \langle r, x\rangle - \gamma, \\ g_2(x) &\triangleq \langle x, Qx\rangle - \langle r, x\rangle - \gamma, \end{aligned} \tag{21}$$

где $Q \in \mathcal{M}_{n\times n}$, $Q = Q^T > 0$, $r \in \mathbb{R}^n$ и $\gamma \in \mathbb{R}$.

Поскольку целевая функция задачи (P_l) линейна, то $(LQ(u,\beta))$ и $(PL(u))$ оказались задачами линейного программирования.

Согласно теоремам сходимости каждого метода в качестве критерия останова было выбрано выполнение неравенства $g(y^s) \le \tau$. При расчетах задавалось $\tau = 10^{-4}$.

Как выяснилось, ММР работает быстрее, но СМЛП в некоторых случаях находит лучшую критическую точку. Превосходство СМЛП может объясняться, по-видимому, тем, что градиенты функции $g(\cdot)$ в некоторых критических точках, найденных алгоритмами локального поиска, оказались близки по направлению к собственным векторам, соответствующим минимальным собственным значениям плохо обусловленных матриц Q, задающих функции (21). Этот факт не влияет на качество решения линеаризованных двойственных к (P_2) задач $(LQ(u,\beta))$, что является преимуществом СМЛП.

По результатам численного эксперимента можно сделать заключение о том, что при решении задач с линейной целевой функцией и квадратичным обратно-выпуклым ограничением в качестве метода локального поиска целесообразно использовать комбинацию двух методов. Сначала находить критическую точку посредством ММР, а затем пытаться ее улучшить с помощью СМЛП [7].

Литература

1. Васильев Ф.П. *Методы оптимизации.* – М.: Факториал Пресс, 2002. – 823 с.
2. Стрекаловский А.С. *Элементы невыпуклой оптимизации.* – Новосибирск: Наука, 2003. – 356 с.
3. Horst R., Tuy H. *Global optimization (deterministic approach).* – Berlin: Springer-Verlag, 1993. – 698 p.
4. Rosen J.B. *Iterative solution of nonlinear optimal control problems* // J. SIAM Control. – 1966. – V. 4. – № 1. – P. 223–244.
5. Meyer R.R. *The validity of a family of optimization methods* // J. SIAM Control. – 1970. — V. 8. – № 1. – P. 41–54.
6. Демьянов В.Ф., Васильев Л.В. *Недифференцируемая оптимизация.* – М.: Наука, 1981. – 383 с.
7. Стрекаловский А.С., Яковлева Т.В. *О локальном и глобальном поиске в невыпуклых задачах оптимизации* // Автоматика и телемеханика. – 2004. – № 3. – С. 23–34.

Институт динамики систем
и теории управления
Сибирского отделения
Российской академии наук

Поступила
27.05.2004