ИЗВЕСТИЯ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

2006 МАТЕМАТИКА № 12 (535)

УДК 519.677

И.И. ЕРЕМИН

ФЕЙЕРОВСКИЕ МЕТОДЫ СИЛЬНОЙ ОТДЕЛИМОСТИ ВЫПУКЛЫХ ПОЛИЭДРАЛЬНЫХ МНОЖЕСТВ

1. Введение

В математике фундаментальную роль играют теоремы отделимости непересекающихся множеств, в частности, выпуклых (в разных пространствах). При этом важны как теоремы существования, так и алгоритмы, реализующие отделимость. Имеется много задач как теоретического характера, так и прикладного, когда необходимо конструктивно обеспечивать разделимость. Одна из областей, в которых алгоритмическая сторона дела играет важнейшую роль, — это распознавание образов, включающее задачи дискриминации, классификации и много других.

Приведем пример конструктивной отделимости множеств $M \subset \mathbb{R}^n$ и $N \subset \mathbb{R}^n$, заданных системами линейных неравенств:

$$M := \{x \mid Ax \le b\} \ne \emptyset, \quad N := \{x \mid Bx \le d\} \ne \emptyset, \quad M \cap N = \emptyset. \tag{1.1}$$

Нужно построить строго их разделяющую гиперплоскость $H := \{x \mid (c,x) - \alpha = 0\}$, т. е. такую, что

$$(c,x) - \alpha < 0 \ \ \forall x \in M, \quad (c,x) - \alpha > 0 \ \ \forall x \in N.$$

Следующее утверждение было сформулировано автором в 1957 г., позднее оно было включено в [1].

Утверждение 1.1. *Система*

$$\begin{gathered} A^T u + B^T v = 0, \quad [u,v] \ge 0, \\ (b,u) + (d,v) \le -1 \end{gathered} \tag{1.2}$$

совместна тогда и только тогда, когда $M \cap N = \emptyset$.

Если $[\overline{u}, \overline{v}]$ *— некоторое решение системы* (1.2), *то гиперплоскость* H *при* $c = A^T\overline{u}$ *и* $\alpha = 1/2$ (*или при* $c = -B^T\overline{v}$ *и* $\alpha = -1/2$) *строго разделяет множества* M *и* N.

Заметим, что нахождение хотя бы одного решения системы (1.2) можно осуществить разными методами, например, методом Фурье ([2], § 11), симплекс-методом или фейеровскими методами ([2], гл. VI).

В данной работе рассматриваются задачи *сильной* отделимости множеств M и N для трех форм их задания:

1° M и N задаются согласно (1.1),

2° M и N задаются выпуклыми оболочками двух конечных совокупностей точек $\{p_i\}_1^k$ и $\{q_j\}_1^s$, т. е.

$$M := \operatorname{conv}\{p_i\}_1^k, \quad N := \operatorname{conv}\{q_j\}_1^s,$$

Работа выполнена при государственной поддержке ведущих научных школ (проект НШ-5595.2006.1) и финансовой поддержке Российского фонда фундаментальных исследований (проект 04-01-00108).

3° смешанный случай:

$$M := \{x \mid Ax \le b\}, \quad N := \mathrm{conv}\{q_j\}_1^s. \tag{1.3}$$

В названии данной статьи употреблен термин *сильной отделимости*, это означает *отделимость слоем наибольшей толщины* (*π-слоем*). Сильная отделимость, по существу, эквивалентна задаче отыскания расстояния между M и N в смысле метрики

$$\rho(M,N) := \min\{\|x-y\| \mid x \in M,\ y \in N\}. \tag{1.4}$$

Если $\overline{x} \in M$ и $\overline{y} \in N$ являются arg-точками задачи (1.3), т. е. $\mathrm{opt}(1.4) = \|\overline{x}-\overline{y}\|$, то слоем наибольшей толщины, разделяющим множества M и N, будет $P := \{x \mid x \in P_1 \cap P_2\}$, где P_1 и P_2 — полупространства, задаваемые линейными неравенствами $(x-\overline{x},\ \overline{x}-\overline{y}) \le 0$ и $(y-\overline{y},\ \overline{x}-\overline{y}) \ge 0$.

Гиперплоскость H_π, задаваемую уравнением $(x - \frac{\overline{x}+\overline{y}}{2},\ \overline{x}-\overline{y}) = 0$, будем называть *сильно разделяющей* множества M и N гиперплоскостью.

Если множества M и N достаточно просты (в смысле простоты операции проектирования точек на них), то, применяя к M и N известный метод *последовательного проектирования*, получим две подпоследовательности, сходящиеся к точкам $\overline{x} \in M$ и $\overline{y} \in N$, которые и будут arg-точками задачи (1.4). Но если M и N — произвольные многогранники, то поскольку сама операция проектирования на них далеко не элементарна, то метод последовательного проектирования не может быть признан эффективным.

Заметим, что достижимость в (1.4), вообще говоря, необязательна, но если хотя бы одно из множеств M и N ограничено, то достижимость имеет место, при этом π-слой не вырождается в гиперплоскость. В случае, когда M и N — многогранники, задаваемые системами линейных неравенств, факт достижимости верен.

В данной работе предлагается метод сведения (редукции) поставленных задач к некоторой конструктивно задаваемой системе линейных неравенств и одного выпуклого неравенства, а для ее решения подключается фейеровский процесс. В основе такой редукции лежит теория двойственности в математическом программировании.

2. Математический аппарат

2.1. Двойственность в выпуклом программировании (ВП). Запишем задачу

$$P := \min\{f_0(x) \mid f_j(x) \le 0,\ j = 1,\dots,m\}. \tag{2.1}$$

Поставим ей в соответствие называемую *двойственной* задачу

$$P^* := \max\{F(x,u) \ \mid\ \nabla_x F(x,u) = 0,\ u \ge 0\}, \tag{2.2}$$

где $F(x,u) := f_0(x) + \sum_{j=1}^{m} u_j f_j(x)$ — функция Лагранжа, соответствующая задаче (2.1). Задачи P и P^* связаны следующим утверждением.

Прямая теорема двойственности (для задачи ВП) ([3], с. 164). *Пусть задача выпуклого программирования* (2.1) *разрешима в точке* $\widetilde{x} \in \mathrm{Arg}(2.1)$ *и* $\exists p \in \mathbb{R}^n : [f_1(p),\dots,f_m(p)] < 0$ (*условие Слейтера*). *Тогда существует вектор* $\widetilde{u} \ge 0$ *такой, что* $[\widetilde{x},\widetilde{u}] \in \mathrm{Arg}\,P^*$, *при этом* $\mathrm{opt}\,P = \mathrm{opt}\,P^*$, *т. е.* $f_0(\widetilde{x}) = F(\widetilde{x},\widetilde{u})$.

Легко доказывается свойство: *если* $\overline{x}$ *и* $[x',u']$ *допустимы относительно задач* P *и* P^*, *то*

$$f_0(\overline{x}) \ge F(x',u'). \tag{2.3}$$

Следствие. Если в (2.3) выполняется равенство

$$f_0(\overline{x}) = F(x',u'),$$

то $\overline{x} \in \mathrm{Arg}\,P$ и $[x',u'] \in \mathrm{Arg}\,P^*$ (равным образом $x' \in \mathrm{Arg}\,P$).

Сделаем замечание к вопросу о дифференцируемости функций, формирующих задачу (2.1). Этого условия во многих случаях можно избежать. Пусть, например, имеем задачу

$$\min\{f_0(x) + r\overline{f}(x) \mid f_i(x) \le 0,\ i \in I \subset \overline{1,m}\}, \tag{2.4}$$

где $r > 0$, $\overline{f}(x) := \max\limits_{i \in \overline{1,m} \setminus I} f_j^+(x)$. Такие задачи возникают, например, в методах точных штрафных функций. Функция $\overline{f}(x)$ не является, вообще говоря, дифференцируемой.

Перепишем задачу (2.4) в эквивалентном виде

$$\min\{f_0(x) + rt \mid f_j(x) \le t,\ j \in \overline{1,m} \setminus I,\ f_i(x) \le 0,\ i \in I,\ t \ge 0\}.$$

Для этой задачи нужные условия дифференцируемости выполняются (естественно, в предположении дифференцируемости функций $\{f_j(x)\}_1^m$), поэтому к модифицированной ситуации применимо правило формирования двойственной задачи по правилу (2.2).

2.2. Симметрическая задача и S-технологии. Теорема двойственности в п. 2.1 и ее следствие подсказывают выделение в качестве самостоятельного объекта систему

$$S : f_j(x) \le 0,\ j = 1, \ldots, m,\ \bigtriangledown F_x(x,u) = 0,\ u \ge 0,\ f_0(x) \le F(x,u). \tag{2.5}$$

Справедливо

Утверждение 2.1. *Если $\overline{x}$ и $\overline{u} \ge 0$ удовлетворяют системе S, то $\overline{x} \in \operatorname{Arg} P$, $[\overline{x}, \overline{u}] \in \operatorname{Arg} P^*$.*

Следовательно, задача P сводится к системе неравенств. Решив ее, тем самым находим как решение $\widetilde{x}$ задачи P, так и решение $[\widetilde{x}, \widetilde{u}]$ задачи P^*.

Пример 2.1. Если P — задача ЛП: $\min\{(c,x) \mid Ax \le b\}$, то симметрической будет задача

$$S:\ Ax \le b,\quad A^T u = c,\quad u \ge 0,\quad (c,x) \le (b,u).$$

Введем термин "*S-технология*". Смысл S-технологии состоит в редукции задачи оптимизации к системе неравенств через формирование симметрической системы S и ее преобразование к виду, удобному для применения тех или иных процедур ее решения.

Изобразим схематически названную технологию

$$\begin{tikzcd}
P \arrow[dd] \arrow[dr] & & & \\
 & S \arrow[r, "\sim"] & \widetilde{S} \arrow[r] & \mathcal{F}_{\widetilde{M}}. \\
P^* \arrow[ur] & & &
\end{tikzcd}$$

Прокомментируем эту схему: P — задача выпуклого программирования (2.1), P^* — двойственная задача (2.2), S — симметрическая задача (2.5), $\widetilde{S}$ — эквивалентно преобразованная задача с множеством решений $\widetilde{M} \ne \emptyset$, для которого можно построить непрерывное M-фейеровское отображение (или замкнутое); если такое отображение построено, то итерационный процесс, им порождаемый, будет сходиться к решению системы S, следовательно, будет давать как решение задачи P, так и задачи P^*; $\mathcal{F}_{\widetilde{M}}$ — класс $\widetilde{M}$-фейеровских отображений.

3. Фейеровские процессы ([2], гл. VI)

Базовым понятием, лежащим в основании фейеровских методов, является понятие *M-фейеровского отображения* $T : \mathbb{R}^n \to \mathbb{R}^n$ такого, что образ $T(x)$ элемента $x \notin M$ ближе к любой точке $y \in M$, при этом $T(y) = y \quad \forall\, y \in M$. Здесь M — выпуклое замкнутое множество из $\mathbb{R}^n$. Простейшим примером M-фейеровского отображения является оператор метрического проектирования на $M : \mathrm{Pr}_M(x)$. Поэтому на M-фейеровское отображение можно смотреть как на некоторое обобщение операции метрического проектирования. Обозначим класс M-фейеровских отображений через $\mathcal{F}_M$. В прикладных аспектах M-фейеровские отображения конструируются, как из элементарных кирпичиков, из простых проектирований на простые множества: гиперплоскость, полупространство, параллелепипед, неотрицательный ортант, линейное многообразие, задаваемое конечной системой линейных уравнений, и др.

Так, например, при

$$M := \{x \mid l_j(x) := (a_j, x) - b_j \le 0,\ j = 1, \dots, m\} \ne \emptyset$$

отображение $T(x) := \sum\limits_{j=1}^{m} \alpha_j \mathrm{Pr}_{M_j}(x)$, где $\alpha_j > 0$, $\sum\limits_{j=1}^{m} \alpha_j = 1$, синтезированное из элементарных проектирований, является непрерывным M-фейеровским, при этом $\{x_{k+1} = T(x_k)\}_{k=0}^{\infty} \to x' \in M$.

Методы фейеровского типа, соотнесенные к задачам решения систем линейных и выпуклых неравенств, а также к задачам линейного и выпуклого программирования, развиты достаточно хорошо. Ниже остановимся кратко на основных свойствах и теоремах, относящихся к теории фейеровских операторов и порождаемых ими процессах, сходящихся к решению той или иной поставленной задачи.

3.1. Исходные определения, понятия и утверждения. Пусть T — однозначное отображение из X в X и $\mathrm{Fix}\, T =: M \ne \emptyset$. Здесь $\mathrm{Fix}\, T$ — множество неподвижных точек отображения T.

Определение 3.1. Отображение $T \in \{X \to X\}$ называется *M-фейеровским*, если для него выполняются соотношения

$$T(y) = y, \quad \|T(x) - y\| < \|x - y\| \ \ \forall\, y \in M,\ \ \forall\, x \notin M.$$

Определение 3.2. Многозначное отображение $T \in \{X \to 2^X\}$ называется *M-фейеровским*, если выполняются соотношения

$$T(y) = y, \quad \|z - y\| < \|x - y\| \ \ \forall\, y \in M,\ \ \forall\, z \in T(x),\ \ \forall\, x \notin M.$$

Пример 3.1. Пусть $M := \{x \mid f_j(x) \le 0,\ j = 1, \dots, m\} \ne \emptyset$, $\{f_j(x)\}_1^m$ — выпуклые функции, $d(x) = \max\limits_{(j)} f_j(x)$, $J(x) := \{j \mid d(x) = f_j(x)\}$. Положим

$$T(x) := \Big\{x - \lambda \frac{d^+(x)}{\|h\|^2} h \ \Big|\ h \in \sum_{j \in J(x)} \alpha_j \partial f_j(x),\ \alpha_j \ge 0,\ \sum_{j \in J(x)} \alpha_j = 1\Big\}, \tag{3.1}$$

$0 < \lambda < 2$, $\partial f_j(x)$ — субдифференциал выпуклой функции $f_j(x)$. Это отображение удовлетворяет определению 3.2.

Обозначим через $\mathcal{F}_M$ класс M-фейеровских (как однозначных, так и многозначных) отображений. В дальнейшем будем использовать понятие замкнутости отображения $T \in \mathcal{F}_M$, которое состоит в выполнимости импликации $(\{x_k\} \to x',\ \{y_k\} \to y',\ y_k \in T(x_k)) \Longrightarrow y' \in T(x')$. Класс замкнутых M-фейеровских отображений будем обозначать $\overline{\mathcal{F}}_M$. Если отображение $T \in \mathcal{F}_M$ однозначно и замкнуто, то оно непрерывно.

Определение 3.3. Последовательность $\{x_k\} \subset \mathbb{R}^n$ называется *M-фейеровской*, если

$$\{x_k\} \cap M = \emptyset, \quad \|x_{k+1} - y\| < \|x_k - y\| \ \ \forall\, y \in M,\ \ \forall\, k = 0, 1, \dots$$

Определение 3.4. Последовательность $\{x_k\} \subset \mathbb{R}^n$ назовем *почти M-фейеровской*, если

$$\{x_k\} \cap M = \emptyset, \quad \|x_{k+1} - y\| \le \|x_k - y\| \ \ \forall\, y \in M, \ \ \forall\, k = 0, 1, \dots,$$

причем эта последовательность содержит M-фейеровскую подпоследовательность, т. е. $\exists \{x_{j_k}\} \subset \{x_k\}$:

$$\|x_{j_{k+1}} - y\| < \|x_{j_k} - y\| \ \ \forall\, y \in M, \ \ \forall\, k = 0, 1, \dots$$

Примечание к определению 3.4. Если из почти M-фейеровской последовательности убрать возможные повторения, то останется последовательность, именуемая *слабо M-фейеровской*, что соответствует следующему определению.

Определение 3.5. Последовательность $\{x_k\}$ называется *слабо M-фейеровской*, если

$$\{x_k\} \cap M = \emptyset, \quad \|x_{k+1} - y\| \le \|x_k - y\| \ \ \forall\, y \in M, \quad x_{k+1} \ne x_k \ \ \forall\, k.$$

Заметим, что такие последовательности порождаются слабо M-фейеровскими отображениями T:

$$T(y) = y, \quad \|T(x) - y\| \le \|x - y\| \ \ \forall\, y \in M, \ \ \forall\, x \notin M, \quad T(x) \ne x.$$

Легко проверить свойство: *если $T(x)$ слабо M-фейеровское, то $T_\alpha(x) := \alpha T(x) + (1-\alpha)x$ является M-фейеровским при $\alpha \in (0,1)$.*

Ниже будем полагать $X = \mathbb{R}^n$.

Перечислим некоторые свойства и факты, относящиеся к M-фейеровским отображениям и M-фейеровским последовательностям.

1^0. Если $\mathcal{F}_M \ne \emptyset$, то множество M автоматически выпукло и замкнуто.

2^0. Если M выпукло и телесно, то любая M-фейеровская последовательность $\{x_k\}$ сходится к некоторому $x' \in \mathbb{R}^n$.

3^0. Если $T \in \mathcal{F}_M$ и последовательность $\{x_k\}$ порождена процессом

$$\{x_{k+1} \in T(x_k)\}_{k=0}^{\infty}, \tag{3.2}$$

причем $\{x_k\} \cap M = \emptyset$, то $\{x_k\}$ M-фейеровская.

4^0. Если $\{x_k\}$ — M-фейеровская последовательность и некоторая предельная ее точка x' принадлежит M, то $\{x_k\} \to x'$.

Утверждение 3.1. *Если $T \in \mathcal{F}_M$ и $T(x)$ непрерывно, то*

$$\{x_{k+1} = T(x_k)\}_{k=0}^{\infty} \to x' \in M.$$

Утверждение 3.2. *Если $T \in \overline{\mathcal{F}}_M$, то*

$$\{x_{k+1} \in T(x_k)\}_{k=0}^{\infty} \to x' \in M.$$

Поскольку отображение (3.1) является замкнутым M-фейеровским, то в силу утверждения 3.2 процесс типа (3.2) сходится к решению совместной системы выпуклых неравенств $f_j(x) \le 0$, $j = 1, \dots, m$.

3.2. Базовые конструкции фейеровских отображений. Как уже отмечалось, оператор $\mathrm{Pr}_M(x)$ метрического проектирования на выпуклое замкнутое множество является непрерывным M-фейеровским. Если имеется некоторая совокупность простых выпуклых множеств $M_j \subset \mathbb{R}^n$, $j = 1, \dots, m$, и соответствующая совокупность операторов проектирования на них $\mathrm{Pr}_{M_j}(x)$, $j = 1, \dots, m$ (они являются M_j-фейеровскими), то из этих элементарных проектирований можно конструировать (разными способами) более сложные фейеровские отображения относительно множества $M := \bigcap\limits_{j=1}^{m} M_j \ne \emptyset$.

Приведем базовые конструкции такого рода, однако они будут даны ниже с более общих позиций.

Итак, пусть $T_j \in \mathcal{F}_{M_j}$, $j = 1, \ldots, m$, и $M := \bigcap\limits_{j=1}^{m} M_j \neq \emptyset$.

Утверждение 3.3. *Справедлива импликация*

$$(T_j \in \mathcal{F}_{M_j}, \quad j = 1, \ldots, m) \implies T^{\alpha}(x) := \sum_{j=1}^{m} \alpha_j T_j(x) \in \mathcal{F}_M,$$

где $\alpha_j > 0$, $\sum\limits_{j=1}^{m} \alpha_j = 1$.

Утверждение 3.4. *Если* $\sigma := (j_1, \ldots, j_m)$ *— любое упорядочение индексов* $j = 1, \ldots, m$, *то*

$$T^{(\sigma)}(x) := T_{j_1}\, T_{j_2}\, \ldots\, T_{j_m}(x) \in \mathcal{F}_M.$$

Примечание 3.1. Если в этих утверждениях $\{T_j(x)\}_1^m$ непрерывны, то $T^{(\alpha)}(x)$ и $T^{(\sigma)}(x)$ непрерывны; если же они замкнуты, то $T^{\alpha}(x)$ и $T^{(\sigma)}(x)$ также замкнуты. Следовательно, к последним можно применить утверждения 3.1 и 3.2 о сходимости процессов, ими порождаемых.

Рассмотрим еще одну конструкцию отображения $T \in \mathcal{F}_M$, получаемого из $\{T_j\}_1^m$. Пусть для множеств M_j заданы функции $d_j(x)$, обладающие свойством $d_j(y) = 0$ для $y \in M_j$ и $d_j(x) > 0$ для $x \notin M_j$. Будем считать эти функции выпуклыми. Введем $d(x) := \max\limits_{(j)} d_j(x)$ и $J(x) := \{j \mid d(x) = d_j(x)\}$ (это копирует уже рассмотренные ситуации).

Введем отображение

$$T(x) := \mathrm{conv}\{T_{j_x}(x) \mid j_x \in J(x)\}; \tag{3.3}$$

здесь conv — символ выпуклой оболочки множеств, стоящих в скобках { }.

Утверждение 3.5. *Отображение* (3.3) *является* M*-фейеровским, при этом если* $\{T_j(x)\}_1^m$ *замкнуты, то отображение* $T(x)$ *замкнуто.*

Примечание 3.2. В утверждениях 3.3–3.5 все множества M_j, $j = 1, \ldots, m$, могут совпадать, при этом содержательность этих утверждений не утрачивается. Более того, эти частные утверждения подчеркивают в явном виде тот важный факт, что из всякой конечной совокупности M-фейеровских отображений можно строить новые отображения из $\mathcal{F}_M$ способами, зафиксированными в утверждениях 3.3–3.5. Из этих утверждений следует, что класс отображений $\mathcal{F}_M$ замкнут относительно выпуклой суммы любой конечной совокупности отображений из $\mathcal{F}_M$ и относительно любой их суперпозиции, т. е. $\mathcal{F}_M$ является своего рода выпуклой алгеброй.

4. Примеры применения *S*-технологии

Пример 4.1. Задачу проектирования элемента p на многогранник $M := \{x \mid Ax \le b\}$ можно записать в форме оптимизационной задачи

$$\min\{\tfrac{1}{2}\|x - p\|^2 \mid Ax \le b\}.$$

Положим $L(x, u) := \frac{1}{2}\|x - p\|^2 + (u, Ax - b)$ $(= \frac{1}{2}\|x - p\| + (A^T u, x) - (b, u))$. Имеем

$$\bigtriangledown_x L(x, u) = (x - p) + A^T u.$$

Выпишем симметрическую задачу

$$\begin{aligned} &Ax \le b; \\ &x - p + A^T u = 0, \quad u \ge 0; \\ &\underbrace{\tfrac{1}{2}\|x - p\|^2}_{f_0(x)} \le \underbrace{\tfrac{1}{2}\|x - p\|^2 + (Ax - b, u)}_{F(x,u)}. \end{aligned} \tag{4.1}$$

После элементарных преобразований и исключения переменной x, получим эквивалент системы (4.1)

$$\begin{aligned} u^T(AA^T)u - (Ap - b, u) \le 0, \quad u \ge 0;\\ -(AA^T)u \le b - Ap. \end{aligned} \tag{4.2}$$

Система (4.2) состоит из системы линейных относительно u неравенств (вторая строка) и одного квадратичного выпуклого неравенства (первая строка). Найдя решение $\overline{u}$ системы (4.2) тем или иным методом, например, фейеровским, по формуле $x = p - A^T\overline{u}$ находим $\overline{x}$, при этом вектор $\overline{x}$ будет проекцией точки p на многогранник $M = \{x \mid Ax \le b\}$.

Пример 4.2. Рассмотрим задачу квадратичного программирования

$$P : \min\{\tfrac{1}{2}(x-p)^TQ(x-p) + (c, x) \mid Ax \le b,\ Bx = d\} \tag{4.3}$$

в предположении, что Q — положительно-определенная матрица. В рамках схемы получения симметрической задачи, соответствующей задаче (2.1), приходим к следующему ее виду:

$$A(p - Q^{-1}(A^Tu + B^Tv + c)) \le b, \tag{4.4}$$

$$B(p - Q^{-1}(A^Tu + B^Tv + c)) = d, \tag{4.5}$$

$$(A^Tu + B^Tv)^TQ^{-1}(A^Tu + B^Tv) + (b^T + (Q^{-1}c - p)^TA^T)u + (d^T + (Q^{-1}c - p)^TB^T)v \le 0, \quad u \ge 0. \tag{4.6}$$

В этой системе подсистемы (4.4) и (4.5) являются линейными относительно$[u, v]$, а неравенство (4.6) выпуклое относительно $[u, v]$, так что к этой системе может быть применен тот или иной вариант фейеровского метода. Если $[\overline{u}, \overline{v}]$ — некоторое решение системы (4.4)–(4.6), то решение задачи (4.3) отыскивается по формуле $\overline{x} = p - Q^{-1}(A^T\overline{u} + B^T\overline{v} + c)$.

5. Задача поиска расстояния между двумя многогранниками

Пусть $M := \{x \mid Ax \le b\}$, $N := \{y \mid By \le d\}$. Сформулируем задачу

$$\min\{\tfrac{1}{2}\|x - y\|^2 \mid Ax \le b,\ By \le d\}. \tag{5.1}$$

Цель: свести ее к решению некоторой системы выпуклых неравенств (точнее, к системе линейных неравенств и одному выпукло-квадратичному неравенству).

Сформулированная задача является важной в связи с актуальностью вопросов сильной разделимости гиперплоскостью непересекающихся выпуклых множеств (пусть многогранников) в дискриминантном анализе. Вопросы разделимости многогранников хорошо разработаны. Более сложным является вопрос о разделимости слоем наибольшей толщины, что эквивалентно, в конструктивном плане, задаче отыскания точек $\overline{x} \in M$ и $\overline{y} \in N$, реализующих $\min\limits_{\substack{x\in M\\ y\in N}} \|x - y\|$.

Итак, для задачи (5.1) запишем функцию Лагранжа

$$F(x, y; u, v) := \tfrac{1}{2}\|x - y\|^2 + (Ax - b, u) + (By - d, v), \quad [u, v] \ge 0.$$

Подсчитаем

$$\bigtriangledown_{x,y} F(x, y; u, v) = 0 : (x - y) + A^Tu = 0, \qquad y - x + B^Tv = 0. \tag{5.2}$$

Для постановки симметрической задачи нужно сформировать неравенство

$$\tfrac{1}{2}\|x - y\|^2 \le F(x, y; u, v) = \tfrac{1}{2}\|x - y\| + (A^Tu, x) - (b, u) + (B^Tv, y) - (d, v),$$

что дает

$$0 \le (A^Tu, x) + (B^Tv, y) - (b, u) - (d, v). \tag{5.3}$$

Подставляя в (5.3) A^Tu и B^Tv из (5.2), получаем $\|x-y\|^2+(b,u)+(d,v)\le 0$, а в итоге — систему

$$Ax \le b, \quad By \le d; \tag{$5.4)_1$}$$
$$x - y = -A^Tu; \tag{$5.4)_2$}$$
$$x - y = B^Tv; \tag{$5.4)_3$}$$
$$\|x-y\|^2 + (b,u) + (d,v) \le 0; \tag{$5.4)_4$}$$
$$[u,v] \ge 0. \tag{$5.4)_5$}$$

Подставляя $(5.4)_2$ в $(5.4)_1$, $(5.4)_3$ и $(5.4)_4$, получим итоговую систему

$$\begin{gathered} A(y - A^Tu) \le b, \quad By \le d, \\ A^Tu + B^Tv = 0, \\ u^T(AA^T)u + (b,u) + (d,v) \le 0, \\ [u,v] \ge 0. \end{gathered} \tag{5.5}$$

Система (5.5) содержит подсистему линейных относительно $[y,u,v]$ неравенств и одно выпуклое квадратичное неравенство. Найдя ее решение $[\overline{y},\overline{u},\overline{v}]$, по формуле $x = \overline{y} - A^T\overline{u}$ находим вектор $\overline{x}$, который в паре с $\overline{y}$ дает точки из M и N, реализующие расстояние между множествами. По $\overline{x}$ и $\overline{y}$ строится сильно разделяющая M и N гиперплоскость $H_\pi := \{x \mid (x - \frac{\overline{x}+\overline{y}}{2}, \overline{x}-\overline{y})\} = 0$.

6. Сильная отделимость многогранников, заданных выпуклыми оболочками двух совокупностей конечного числа точек

Пусть заданы множества $M := \operatorname{conv}\{p_i\}_{i=1}^k$ и $N := \operatorname{conv}\{q_i\}_{j=1}^s$, т. е.

$$M = \Big\{x \in \mathbb{R}^n \;\Big|\; x = \sum_{i=1}^k \alpha_i p_i,\ \alpha_i \ge 0,\ \sum_{i=1}^k \alpha_i = 1\Big\},$$
$$N = \Big\{y \in \mathbb{R}^n \;\Big|\; y = \sum_{j=1}^s \beta_j q_j,\ \beta_j \ge 0,\ \sum_{j=1}^s \beta_j = 1\Big\}.$$

Точки p_i и q_j будем считать векторами-столбцами, равным образом векторы $\alpha := [\alpha_1,\dots,\alpha_k]^T$ и $\beta := [\beta_1,\dots,\beta_s]^T$. Образуем матрицы $A := [p_1,\dots,p_k]$ и $B := [q_1,\dots,q_s]$. Положим $e_1 = \underbrace{[1,\dots,1]^T}_{k}$, $e_2 = \underbrace{[1,\dots,1]^T}_{s}$. Сформулируем задачу

$$\min\{\tfrac12\|x-y\|^2 \mid x \in M,\ y \in N\}.$$

В аналитической записи эта задача будет иметь вид

$$\min\{\tfrac12\|x-y\|^2 \mid x = A\alpha,\ y = B\beta,\ e_1^T\alpha = 1,\ e_2^T\beta = 1,\ \alpha \ge 0,\ \beta \ge 0\}. \tag{6.1}$$

Решение сформулированной задачи будет осуществляться по S-технологии. Ограничениям задачи (6.1) поставим в соответствие двойственные переменные

$$A\alpha - x = 0 \longrightarrow u_\alpha, \quad B\beta - y = 0 \longrightarrow u_\beta;$$
$$-\alpha \le 0 \longrightarrow v_1, \quad -\beta \le 0 \longrightarrow v_2; \quad e_1^T\alpha - 1 \longrightarrow t_1, \quad e_2^t\beta - 1 \longrightarrow t_2.$$

В символах u_α и u_β буквы α и β — просто метки. Функция Лагранжа, соотнесенная задаче (6.1), запишется в виде

$$\begin{aligned}F(x,y,\alpha,\beta;u_\alpha,u_\beta;v_1,v_2;t_1,t_2) &:= \tfrac{1}{2}\|x-y\|^2 + u_\alpha^T(A\alpha - x) + \\ &+ u_\beta^T(B\beta - y) - v_1^T\alpha - v_2^T\beta + t_1(e_1^T\alpha - 1) + t_2(e_2^T\beta - 1) \equiv \\ &\equiv \frac{1}{2}\|x-y\|^2 + (A^Tu_\alpha - v_1 + t_1e_1^T)^T\alpha + (B^Tu_\beta - v_2 + t_2e_2^T)^T\beta - u_\alpha^Tx - u_\beta^Ty - t_1 - t_2.\end{aligned}$$

Выпишем градиенты функции $F(\cdot)$ по переменным x, y, α и β с приравниванием их нулю:

$$\begin{aligned}\nabla_x F(\cdot) &= x - y - u_\alpha = 0, && (6.2)_1\\ \nabla_y F(\cdot) &= y - x - u_\beta = 0, && (6.2)_2\\ \nabla_\alpha F(\cdot) &= A^Tu_\alpha - v_1 + t_1e_1^T = 0, && (6.2)_3\\ \nabla_\beta F(\cdot) &= B^Tu_\beta - v_2 + t_2e_2^T = 0, && (6.2)_4\\ v_1 \geq 0&, \quad v_2 \geq 0.\end{aligned}$$

Симметрическая система S имеет структуру

ограничения прямой задачи (6.1); $(6.3)_1$

ограничения двойственной задачи, т. е. соотношения (6.2); $(6.3)_2$

неравенство $\frac{1}{2}\|x-y\|^2 \leq F(\cdot)$, что с учетом
соотношений $(6.2)_3$, $(6.2)_4$ дает $u_\alpha^Tx + u_\beta^Ty + t_1 + t_2 \leq 0$. $(6.3)_3$

Умножив скалярно $(6.2)_1$ на x, $(6.2)_2$ — на y и сложив, получим $\|x-y\|^2 - u_\alpha^Tx - u_\beta^Ty = 0$, что позволяет последнее из неравенств $(6.3)_3$ записать в виде $\|x-y\|^2 + t_1 + t_2 \leq 0$.

Запишем все ограничения, формирующие систему S:

$$\begin{gathered}A\alpha = x, \quad B\beta = y,\\ \sum_i \alpha_i = 1, \quad \sum_j \beta_j = 1, \quad \alpha_i \geq 0, \quad \beta_j \geq 0; \qquad (6.4)_1\\ x - y = u_\alpha, \quad y - x = u_\beta,\\ A^Tu_\alpha = v_1 \geq 0, \quad B^Tu_\beta = v_2 \geq 0; \qquad (6.4)_2\\ \|x-y\|^2 + t_1 + t_2 \leq 0. \qquad (6.4)_3\end{gathered}$$

Из $(6.4)_1$ и $(6.4)_2$ следует $A\alpha - B\beta = u_\alpha$, $B\beta - A_\alpha = u_\beta$. Далее, подставив u_α и u_β в $(6.4)_3$, в итоге получим систему

$$\begin{gathered}A^T(A\alpha - B\beta) + t_1e_1 \geq 0,\\ B^T(B\beta - A\alpha) + t_2e_2 \geq 0,\\ \|A\alpha - B\beta\|^2 + t_1 + t_2 \leq 0,\\ \sum_i \alpha_i = 1, \quad \sum_j \beta_j = 1,\\ \alpha \geq 0, \quad \beta \geq 0.\end{gathered} \tag{6.5}$$

Если $[\overline{\alpha}, \overline{\beta}, \overline{t}_1, \overline{t}_2]$ — некоторое решение системы (6.5), то значения переменных x и y находятся согласно соотношениям $\overline{x} = A\overline{\alpha}$, $\overline{y} = B\overline{\beta}$, а двойственные оценки u_α, u_β, v_1 и v_2 — по формулам

$$\overline{u}_\alpha = A\overline{\alpha} - B\overline{\beta}, \quad \overline{u}_\beta = -\overline{u}_\alpha; \quad \overline{v}_1 = A^T\overline{u}_\alpha + \overline{t}_1e_1, \quad \overline{v}_2 = B^T\overline{u}_\beta + \overline{t}_2e_2.$$

7. Сильная отделимость двух многогранников, заданных прямо-двойственным способом

Пусть $M := \{x \mid Ax \le b\}$, $N := \mathrm{conv}\{b_j\}_1^k \subset \mathbb{R}^n$. Положим $\beta^T := [\beta_1, \dots, \beta_k]$, $e^T = [\underbrace{1, \dots, 1}_{k}]$. Множество N представим в виде $N = \{y \mid y = B\beta,\, e^T\beta = 1,\, \beta \ge 0\}$, где $B = [b_1, \dots, b_k]$.

Запишем задачу

$$\min\left\{\tfrac{1}{2}\|x - y\|^2 \mid Ax \le b,\ y = B\beta,\ e^T\beta = 1,\ \beta \ge 0\right\}. \tag{7.1}$$

Выпишем для ограничений задачи (7.1) двойственные переменные

$$\begin{gathered} Ax \le b \to u \ge 0, \quad B\beta - y = 0 \to u_\beta, \\ e^T\beta = 1 \to t, \quad -v \le 0 \to v \ge 0 \end{gathered}$$

и функцию Лагранжа

$$\begin{aligned} F(x, y, \beta; u, u_\beta, t, v) := \tfrac{1}{2}\|x - y\|^2 + (Ax - b, u) + (B\beta - y, u_\beta) + t[e^t\beta - 1] - v^T\beta \equiv \\ \equiv \tfrac{1}{2}\|x - y\|^2 + \underbrace{(A^T u, x) - (y, u_\beta) + (B^T u_\beta + te - v)_\beta^T - (b, u) - t}_{F_0(\cdot)}. \end{aligned}$$

Составим аналог системы (2.5)

$$Ax \le b, \quad B\beta - y = 0, \quad e^T\beta = 1, \quad \beta \ge 0 \ \ \text{(это ограничения задачи (7.1));} \tag*{$(7.2)_1$}$$

$$\begin{gathered} \nabla_x F(\cdot) = x - y + A^T u = 0, \quad u \ge 0, \\ \nabla_y F(\cdot) = y - x - u_\beta = 0, \ \ \nabla_\beta F(\cdot) = B^T u_\beta + te - v = 0, \quad v \ge 0 \ \ (\sim B^T u_\beta + te = v \ge 0); \end{gathered} \tag*{$(7.2)_2$}$$

$$\frac{1}{2}\|x - y\|^2 \le \frac{1}{2}\|x - y\| + F_0(\cdot). \tag*{$(7.2)_3$}$$

Неравенство $(7.2)_3$ с учетом 1-го и 2-го соотношений из $(7.2)_2$ преобразуется в неравенство

$$u^T(AA^T)u + (b, u) + t \le 0, \tag*{$(7.2)_3'$}$$

в этом преобразовании учитываются равенства

$$(Ax, u) = (A^T u, x) = (y - x, x) \ \text{ и } \ (y, u_\beta) = (y, y - x).$$

Таким образом, система S (аналог системы (2.5) применительно к задаче (7.1)) есть объединение систем $(7.2)_1$, $(7.2)_2$ и выпуклого неравенства $(7.2)_3'$.

Выпишем ее отдельно

$$Ax \le b, \quad B\beta - y = 0, \quad e^T\beta = 1, \quad \beta \ge 0; \tag*{$(7.3)_1$}$$

$$x - y + A^T = 0, \quad u \ge 0; \tag*{$(7.3)_2$}$$

$$y - x = u_\beta; \tag*{$(7.3)_3$}$$

$$B^T u_\beta + te \ge 0; \tag*{$(7.3)_4$}$$

$$u^T(AA^T)u + b^T u + t \le 0. \tag*{$(7.3)_5$}$$

Переменными в системе (7.3) являются x, y, β, а также u, u_β и t (двойственные переменные). Часть переменных из этой совокупности можно исключить (по методу Гаусса). Именно, взяв x из

$(7.3)_2$, u_β — из $(7.3)_3$ и подставив в остальные соотношения, получим систему S в окончательном виде

$$\begin{gathered} A(y - A^T u) \le b, \quad u \ge 0, \quad B\beta - y = 0, \quad e^T\beta = 1, \quad \beta \ge 0, \\ B^T(A^T u) + te \ge 0, \quad u^T(AA^T)u + b^T u + t \le 0\,. \end{gathered} \tag{7.4}$$

Система (7.4) содержит линейные неравенства и уравнения, а также одно выпуклое неравенство. К такой системе приложимы те или иные варианты фейеровских методов. Если найдено (точно или приближенно) некоторое решение $[\overline{y}, \overline{\beta}, \overline{\beta}, \overline{t}\,]$ системы (7.4), то значения остальных переменных находятся (точно или приближенно) согласно соотношениям $\overline{x} = \overline{y} - A^T\overline{u}$, $\overline{u}_\beta = \overline{y} - \overline{x}$, $\overline{v} = B^T\overline{u}_\beta + \overline{t}e$ (все эти соотношения вытекают из системы (7.2)).

Заметим, что все соотношения в системе (7.4) записаны в матричном виде. Это позволяет строить методы ее решения в матричных процедурах, что приложимо и к фейеровским методам.

Литература

1. Черников С.Н. *Линейные неравенства.* – М.: Наука, 1968. – 488 с.
2. Еремин И.И. *Теория линейной оптимизации.* – Екатеринбург: Изд-во “Екатеринбург”, 1999. – 312 с.
3. *Введение в нелинейное программирование* / Под ред. К.-Х. Эльстера. – М.: Наука, 1985. – 264 с.

Институт математики и механики
Уральского отделения
Российской академии наук

Поступила
17.02.2006