

И.И.Еремин

ФЕЙЕРОВСКИЕ МЕТОДЫ СИЛЬНОЙ ОТДЕЛИМОСТИ ВЫПУКЛЫХ ПОЛИЭДРАЛЬНЫХ МНОЖЕСТВ¹

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 распознавания образов, включающая 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:

$$M := \{x \mid Ax \leq b\} \neq \emptyset, \quad N := \{x \mid Bx \leq d\} \neq \emptyset, \quad M \cap N = \emptyset. \quad (1.1)$$

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

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

Следующее утверждение было сформулировано автором в 1957 г., позднее оно было включено в монографию С.Н.Черникова: *Линейные неравенства*. М.: Наука, 1968, 488 с.

¹Работа выполнена при государственной поддержке ведущих научных школ (проект НШ-5595.2006.1) и финансовой поддержке РФФИ (проект 04-01-00108)

Утверждение. Система

$$\left. \begin{aligned} A^T u + B^T v &= 0, & [u, v] &\geq 0, \\ (b, u) + (d, v) &\leq -1 \end{aligned} \right\} \quad (1.2)$$

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

Если $[\bar{u}, \bar{v}]$ — некоторое решение системы (1.2), то гиперплоскость H при $c = A^T \bar{u}$ и $\alpha = 1/2$ (или при $c = -B^T \bar{v}$ 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 ([1], § 11), симплекс-методом или фейеровскими методами ([1], гл. VI).

В настоящей 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.

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

3°. Смешанный случай:

$$M := \{x \mid Ax \leq b\}, \quad N := \text{conv}\{q_j\}_1^s. \quad (1.3)$$

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

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

Если $\bar{x} \in M$ и $\bar{y} \in N$ являются \arg -точками задачи (1.3), т.е. $\text{opt}(1.4) = \|\bar{x} - \bar{y}\|$, 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

$$(x - \bar{x}, \bar{x} - \bar{y}) \leq 0 \quad \text{и} \quad (y - \bar{y}, \bar{x} - \bar{y}) \geq 0.$$

Гиперплоскость H_π , задаваемую уравнением $\left(x - \frac{\bar{x} + \bar{y}}{2}, \bar{x} - \bar{y}\right) = 0$, 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 $\bar{x} \in M$ и $\bar{y} \in N$, 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.

К задаче (1.4) сделаем такое замечание. Достижимость в (1.4), вообще говоря, необязательна, однако если хотя бы одно из множеств M и N ограничено, то достижимость имеет место, при этом π -слой не вырождается в гиперплоскость. Однако, если M и N — многогранники, задаваемые системами линейных неравенств, то факт достижимости верен.

В настоящей работе предлагается метод сведения (редукции) поставленных задач к некоторой конструктивно задаваемой системе линейных 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

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

Поставим ей в соответствие задачу

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

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

утверждением.

Прямая теорема двойственности (для задачи ВП) ([2], стр. 164). Пусть задача выпуклого программирования (2.1) разрешима в точке $\tilde{x} \in \text{Arg}(2.1)$ и $\exists p \in \mathbb{R}^n : [f_1(p), \dots, f_m(p)] < 0$ (условие Слейтера). Тогда существует вектор $\tilde{u} \geq 0$ такой, что $[\tilde{x}, \tilde{u}] \in \text{Arg}P^*$, при этом $\text{opt } P = \text{opt } P^*$, т.е. $f_0(\tilde{x}) = F(\tilde{x}, \tilde{u})$.

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

$$f_0(\bar{x}) \geq F(x', u'). \quad (2.3)$$

Следствие. Если в (2.3) выполняется равенство

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

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

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

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

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

Перепишем задачу (2.4) в эквивалентном виде:

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

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

$$S : f_j(x) \leq 0, j = 1, \dots, m, \nabla F_x(x, u) = 0, u \geq 0, f_0(x) \leq F(x, u). \quad (2.5)$$

Справедливо

Утверждение. Если \bar{x} и $\bar{u} \geq 0$ удовлетворяют системе S , то $\bar{x} \in \text{Arg } P$, $[\bar{x}, \bar{u}] \in \text{Arg } P^*$.

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

Приведем пример. Если P — задача ЛП: $\min \{(c, x) \mid Ax \leq b\}$, то симметрической задачей будет

$$S : Ax \leq b, A^T u = c, u \geq 0, (c, x) \leq (b, u).$$

Введем термин “ S -технология”. Смысл S -технологии состоит в редукции задачи оптимизации к системе неравенств через технику формирования симметрической системы S и ее преобразовании к виду, удобному для применения тех или иных процедур ее решения.

Изобразим схематически названную технологию

$$\begin{array}{c} P \\ \searrow \\ \downarrow \\ P^* \nearrow \end{array} S \xrightarrow{\sim} \tilde{S} \longrightarrow \mathcal{F}_{\tilde{M}}.$$

Прокомментируем эту схему:

P — задача ВП (2.1);

P^* — двойственная задача (2.2);

S — симметрическая задача (2.5);

\tilde{S} — эквивалентно преобразованная задача с множеством решений $\tilde{M} \neq \emptyset$, 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 ([1], гл. VI)

Базовым понятием, лежащим в основании фейеровских методов, является понятие M -фейеровского отображения $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ такого, что образ $T(x)$ элемента $x \notin M$ ближе к любой точке $y \in M$, при этом $T(y) = y, \forall y \in M$. Здесь M — выпуклое замкнутое множество из \mathbb{R}^n . Простейшим примером M -фейеровского отображения является оператор метрического проектирования на $M : \text{Pr}_M(x)$. 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. Приведем пример.

Пусть

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

т.е. M — множество решений конечной системы линейных неравенств. Положим $T(x) := \sum_{j=1}^m \alpha_j \text{Pr}_{M_j}(x)$, где $\alpha_j > 0, \sum_{j=1}^m \alpha_j = 1$. Отображение $T(x)$, синтезированное из элементарных проектирований, является непрерывным M -фейеровским, при этом $\{x_{k+1} = T(x_k)\}_{k=0}^{\infty} \rightarrow x' \in M$.

Методы фейеровского типа, соотнесенные к задачам решения систем линейных и выпуклых неравенств, а также 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 $\text{Fix } T =: M \neq \emptyset$. Здесь $\text{Fix } T$ — множество точек неподвижности отображения T .

О п р е д е л е н и е 3.1. Отображение $T \in \{X \rightarrow X\}$ называется *M-фейеровским*, если для него выполняются соотношения

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

О п р е д е л е н и е 3.2. Мнозначное отображение $T \in \{X \rightarrow 2^X\}$ называется *M-фейеровским*, если выполняются соотношения

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

Приведем пример такого отображения.

Пусть $M := \{x \mid f_j(x) \leq 0, j = 1, \dots, m\} \neq \emptyset$, $\{f_j(x)\}_1^m$ — выпуклые функции, $d(x) = \max_{(j)} f_j(x)$, $J(x) := \{j \mid d(x) = f_j(x)\}$. Положим

$$T(x) := \left\{ x - \lambda \frac{d^+(x)}{\|h\|^2} h \mid h \in \sum_{j \in J(x)} \alpha_j \partial f_j(x), \alpha_j \geq 0, \sum_{j \in J(x)} \alpha_j = 1 \right\}, \quad (3.1)$$

$0 < \lambda < 2$, $\partial f_j(x)$ — 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 выполнении следующей импликации

$$(\{x_k\} \rightarrow x', \{y_k\} \rightarrow y', y_k \in T(x_k)) \implies y' \in T(x').$$

Класс замкнутых *M-фейеровских* отображений будем обозначать $\overline{\mathcal{F}}_M$. Если отображение $T \in \mathcal{F}_M$ однозначно и замкнуто, то оно непрерывно.

О п р е д е л е н и е 3.3. Последовательность $\{x_k\} \subset \mathbb{R}^n$ называется *M-фейеровской*, если

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

О п р е д е л е н и е 3.4. Последовательность $\{x_k\} \subset \mathbb{R}^n$ назовем *почти M -фейеровской*, если

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

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

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

П р и м е ч а н и е к определению 3.4. Если из почти M -фейеровской последовательности убрать возможные повторения, то останется последовательность, именуемая *слабо M -фейеровской*, что соответствует следующему определению.

О п р е д е л е н и е 3.5. Последовательность $\{x_k\}$ называется *слабо M -фейеровской*, если

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

Заметим, что такие последовательности порождаются слабо M -фейеровскими отображениями T со следующим их определением:

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

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 необходимые для дальнейшего свойства и факты, относящиеся к M -фейеровским отображениям и M -фейеровским последовательностям.

1⁰. Если $\mathcal{F}_M \neq \emptyset$, то множество M автоматически выпукло и замкнуто.

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

3⁰. 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.

4⁰. Если $\{x_k\}$ — M -фейеровская последовательность и некоторая предельная ее точка x' принадлежит M , то $\{x_k\} \rightarrow x'$.

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

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

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

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

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

3.2. Базовые конструкции фейеровских отображений

Как уже отмечалось, оператор $\text{Pr}_M(x)$ метрического проектирования на выпуклое замкнутое множество является непрерывным M -фейеровским. Проектирование на простейшие выпуклые множества осуществляется весьма просто, к таким простым (элементарным) множествам можно отнести гиперплоскость, полупространство, параллелепипед, неотрицательный ортант, шар и др. Если имеется некоторая совокупность простых выпуклых множеств $M_j \subset \mathbb{R}^n$, $j = 1, \dots, m$ и соответствующая совокупность операторов проектирования на них $\text{Pr}_{M_j}(x)$, $j = 1, \dots, m$ (они являются M_j -фейеровскими), то из этих элементарных проектирований можно конструировать (разными способами) более сложные фейеровские отображения относительно множества $M := \bigcap_{j=1}^m M_j \neq \emptyset$.

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

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

Утверждение 3.3. Справедлива импликация:

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

здесь $\alpha_j > 0$, $\sum_{j=1}^m \alpha_j = 1$.

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

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

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 $d(x) := \max_{(j)} d_j(x)$ и $J(x) := \{j \mid d(x) = d_j(x)\}$ (эта ситуация копирует уже рассмотренные ситуации).

Введем отображение

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

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 , $j = 1, \dots, m$ 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 1.

Рассмотрим задачу проектирования элемента p на многогранник $M := \{x \mid Ax \leq b\}$, что можно записать в форме оптимизационной задачи

$$\min \left\{ \frac{1}{2} \|x - p\|^2 \mid Ax \leq b \right\}.$$

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

Имеем

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

Выпишем симметрическую задачу:

$$\left. \begin{array}{l} Ax \leq b; \\ x - p + A^T u = 0, \quad u \geq 0; \\ \underbrace{\frac{1}{2} \|x - p\|^2}_{f_0(x)} \leq \underbrace{\frac{1}{2} \|x - p\|^2 + (Ax - b, u)}_{F(x, u)}. \end{array} \right\} \quad (4.1)$$

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

$$-(AA^T)u \leq b - Ap, \quad u \geq 0, \quad \|A^T u\|^2 - (Ap - b, u) \leq 0,$$

или

$$\boxed{\begin{array}{l} u^T(AA^T)u - (Ap - b, u) \leq 0, \quad u \geq 0; \\ -(AA^T)u \leq b - Ap. \end{array}} \quad (4.2)$$

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

Пример 2.

Рассмотрим задачу квадратичного программирования:

$$P : \min \left\{ \frac{1}{2} (x - p)^T Q (x - p) + (c, x) \mid Ax \leq b, Bx = d \right\} \quad (4.3)$$

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

$$A(p - Q^{-1}(A^T u + B^T v + c)) \leq b, \quad (4.4)$$

$$B(p - Q^{-1}(A^T u + B^T v + c)) = d, \quad (4.5)$$

$$(A^T u + B^T v)^T Q^{-1}(A^T u + B^T v) + (b^T + (Q^{-1}c - p)^T A^T)u + \\ +(d^T + (Q^{-1}c - p)^T B^T)v \leq 0, \quad u \geq 0. \quad (4.6)$$

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 $[\bar{u}, \bar{v}]$ — некоторое решение системы (4.4)–(4.6), то решение задачи (4.3) отыскивается по формуле $\bar{x} = p - Q^{-1}(A^T \bar{u} + B^T \bar{v} + c)$.

5. Задача поиска расстояния между двумя многогранниками

Пусть

$$M := \{x \mid Ax \leq b\} \quad \text{и} \quad N := \{y \mid By \leq d\}.$$

Сформулируем задачу

$$\min \left\{ \frac{1}{2} \|x - y\|^2 \mid Ax \leq b, By \leq d \right\}. \quad (5.1)$$

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 сложный вопрос — о разделимости слоев наибольшей толщины, что эквивалентно, в конструктивном плане, задаче отыскания точек $\bar{x} \in M$ и $\bar{y} \in N$, реализующих

$$\min_{\substack{x \in M \\ y \in N}} \|x - y\|.$$

Итак, рассмотрим задачу (5.1). Запишем ее функцию Лагранжа:

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

Подсчитаем

$$\nabla_{x,y} F(x, y; u, v) = 0 : (x - y) + A^T u = 0, \quad y - x + B^T v = 0. \quad (5.2)$$

Для формирования симметрической задачи нужно сформировать неравенство

$$\frac{1}{2} \|x - y\|^2 \leq F(x, y; u, v) = \frac{1}{2} \|x - y\|^2 + (A^T u, x) - (b, u) + (B^T v, y) - (d, v),$$

что дает

$$0 \leq (A^T u, x) + (B^T v, y) - (b, u) - (d, v). \quad (5.3)$$

Подставляя в (5.3) $A^T u$ и $B^T v$ из (5.2), получаем $\|x - y\|^2 + (b, u) + (d, v) \leq 0$.

В итоге получаем систему

$$\left. \begin{aligned} Ax &\leq b, & By &\leq d; & (5.4)_1 \\ x - y &= -A^T u; & & & (5.4)_2 \\ x - y &= B^T v; & & & (5.4)_3 \\ \|x - y\|^2 + (b, u) + (d, v) &\leq 0; & & & (5.4)_4 \\ [u, v] &\geq 0. & & & (5.4)_5 \end{aligned} \right\} \quad (5.4)$$

Имея в виду (5.4)₂, величину $\|x - y\|^2$ ($= (x - y, x - y)$) можно заменить на $(A^T u, A^T u) = u^T (AA^T) u$. Так как $x = y - A^T u$, то подставляя это значение в (5.4)₁ и (5.4)₃, получим итоговую систему:

$$\boxed{\begin{aligned} A(y - A^T u) &\leq b, & By &\leq d, \\ A^T u + B^T v &= 0, \\ u^T (AA^T) u + (b, u) + (d, v) &\leq 0, \\ [u, v] &\geq 0. \end{aligned}} \quad (5.5)$$

Система (5.5) содержит подсистему линейных относительно $[y, u, v]$ неравенств и одно выпуклое квадратичное неравенство. Найдя ее решение $[\bar{y}, \bar{u}, \bar{v}]$,

по формуле $x = \bar{y} - A^T \bar{u}$ находим вектор \bar{x} , который в паре с \bar{y} дает точки из M и N , реализующие расстояние между множествами. По \bar{x} и \bar{y} строится сильно разделяющая M и N гиперплоскость $H_\pi := \left\{ x \mid \left(x - \frac{\bar{x} + \bar{y}}{2}, \bar{x} - \bar{y} \right) \right\} = 0$.

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

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

$$M = \left\{ x \in \mathbb{R}^n \mid x = \sum_{i=1}^k \alpha_i p_i, \alpha_i \geq 0, \sum_{i=1}^k \alpha_i = 1 \right\},$$

$$N = \left\{ y \in \mathbb{R}^n \mid y = \sum_{j=1}^s \beta_j q_j, \beta_j \geq 0, \sum_{j=1}^s \beta_j = 1 \right\}.$$

Точки p_i и q_j будем считать векторами-столбцами, равным образом векторы $\alpha := [\alpha_1, \dots, \alpha_k]^T$ и $\beta := [\beta_1, \dots, \beta_s]^T$. Образует 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

$$\min \left\{ \frac{1}{2} \|x - y\|^2 \mid x \in M, y \in N \right\}.$$

В аналитической записи эта задача будет иметь вид:

$$\min \left\{ \frac{1}{2} \|x - y\|^2 \mid x = A\alpha, y = B\beta, e_1^T \alpha = 1, e_2^T \beta = 1, \alpha \geq 0, \beta \geq 0 \right\}. \quad (6.1)$$

Решение сформированной задачи будет осуществляться по S -технологии. Ограничениям задачи (6.1) поставим в соответствии двойственные переменные:

$$\begin{aligned} A\alpha - x = 0 &\longrightarrow u_\alpha, & B\beta - y = 0 &\longrightarrow u_\beta; \\ -\alpha \leq 0 &\longrightarrow v_1, & -\beta \leq 0 &\longrightarrow v_2; & e_1^T \alpha - 1 &\longrightarrow t_1, & e_2^T \beta - 1 &\longrightarrow t_2. \end{aligned}$$

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

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

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

$$\left. \begin{aligned} \nabla_x F(\cdot) &= x - y - u_\alpha = 0, & (6.2)_1 \\ \nabla_y F(\cdot) &= y - x - u_\beta = 0, & (6.2)_2 \\ \nabla_\alpha F(\cdot) &= A^T u_\alpha - v_1 + t_1 e_1^T = 0, & (6.2)_3 \\ \nabla_\beta F(\cdot) &= B^T u_\beta - v_2 + t_2 e_2^T = 0; & (6.2)_4 \\ v_1 &\geq 0, \quad v_2 \geq 0. \end{aligned} \right\} (6.2)$$

Симметрическая система S имеет структуру:

$$\left. \begin{aligned} &\text{ограничения прямой задачи [6.1];} & (6.3)_1 \\ &\left. \begin{aligned} &\text{ограничения двойственной задачи,} \\ &\text{т.е. соотношения (6.2);} \end{aligned} \right] & (6.3)_2 \\ &\left. \begin{aligned} &\text{неравенство } \frac{1}{2} \|x - y\|^2 \leq F(\cdot), \\ &\text{что с учетом соотношений (6.2)}_3, (6.2)}_4 \text{ дает:} \\ &u_\alpha^T x + u_\beta^T y + t_1 + t_2 \leq 0. \end{aligned} \right] & (6.3)_3 \end{aligned} \right\} (6.3)$$

Умножив скалярно (6.2)₁ на x , (6.2)₂ — на y и сложив, получим

$$\|x - y\|^2 - u_\alpha^T x - u_\beta^T y = 0,$$

что позволяет последнее неравенство в (6.3)₃ записать в виде

$$\|x - y\|^2 + t_1 + t_2 \leq 0.$$

Запишем все ограничения, формирующие систему S :

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

Взяв из верхней строчки (6.4)₁ значения $x = A\alpha$, $y = B\beta$ и подставив в соотношения (6.4)₂, получили

$$A\alpha - B\beta = u_\alpha, \quad B\beta - A\alpha = u_\beta.$$

Далее, подставив u_α и u_β в (6.4)₃, в итоге получим систему

$$\boxed{\begin{array}{l} A^T(A\alpha - B\beta) + t_1 e_1 \geq 0, \\ B^T(B\beta - A\alpha) + t_2 e_2 \geq 0, \\ \|A\alpha - B\beta\|^2 + t_1 + t_2 \leq 0, \\ \sum_i \alpha_i = 1, \quad \sum_j \beta_j = 1, \\ \alpha \geq 0, \quad \beta \geq 0. \end{array}} \quad (6.5)$$

Если $[\bar{\alpha}, \bar{\beta}, \bar{t}_1, \bar{t}_2]$ — некоторое решение системы (6.5), то значения переменных x и y находятся согласно соотношений

$$\bar{x} = A\bar{\alpha}, \quad \bar{y} = B\bar{\beta},$$

а двойственные оценки u_α , u_β , v_1 и v_2 :

$$\bar{u}_\alpha = A\bar{\alpha} - B\bar{\beta}, \quad \bar{u}_\beta = -\bar{u}_\alpha; \quad \bar{v}_1 = A^T \bar{u}_\alpha + \bar{t}_1 e_1, \quad \bar{v}_2 = B^T \bar{u}_\beta + \bar{t}_2 e_2.$$

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

Пусть

$$M := \{x \mid Ax \leq b\}, \quad N := \text{conv}\{b_j\}_1^k \subset \mathbb{R}^n.$$

Положим $\beta^T := [\beta_1, \dots, \beta_k]$, $e^T = \underbrace{[1, \dots, 1]}_k$. Множество N можно представить так:

$$N = \{y \mid y = B\beta, e^T\beta = 1, \beta \geq 0\},$$

где $B = [b_1, \dots, b_k]$.

Запишем задачу

$$\min \left\{ \frac{1}{2} \|x - y\|^2 \mid Ax \leq b, y = B\beta, e^T\beta = 1, \beta \geq 0 \right\}. \quad (7.1)$$

Выпишем для ограничений задачи (7.1) двойственные переменные:

$$\left. \begin{array}{l} Ax \leq b \rightarrow u \geq 0, \quad B\beta - y = 0 \rightarrow u_\beta, \\ e^T\beta = 1 \rightarrow t, \quad -v \leq 0 \rightarrow v \geq 0 \end{array} \right\}$$

и функцию Лагранжа:

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

Составим аналог системы (2.5):

$$\left. \begin{array}{l} Ax \leq b, \quad B\beta - y = 0, \quad e^T\beta = 1, \quad \beta \geq 0 \\ \left. \begin{array}{l} \left. \begin{array}{l} \nabla_x F(\cdot) = x - y + A^T u = 0, \quad u \geq 0, \\ \nabla_y F(\cdot) = y - x - u_\beta = 0, \\ \nabla_\beta F(\cdot) = B^T u_\beta + te - v = 0, \quad v \geq 0 \\ (\sim B^T u_\beta + te = v \geq 0); \end{array} \right\} \quad (7.2)_1 \\ \left. \begin{array}{l} \left. \begin{array}{l} \nabla_x F(\cdot) = x - y + A^T u = 0, \quad u \geq 0, \\ \nabla_y F(\cdot) = y - x - u_\beta = 0, \\ \nabla_\beta F(\cdot) = B^T u_\beta + te - v = 0, \quad v \geq 0 \\ (\sim B^T u_\beta + te = v \geq 0); \end{array} \right\} \quad (7.2)_2 \\ \frac{1}{2} \|x - y\|^2 \leq \frac{1}{2} \|x - y\| + F_0(\cdot). \end{array} \right\} \quad (7.2)_3 \end{array} \right\} \quad (7.2)$$

Неравенство (7.2)₃ с учетом 1-го и 2-го соотношений из (7.2)₂ преобразуется в неравенство

$$u^T(AA^T)u + (b, u) + t \leq 0 \quad (7.2)'_3$$

(в этом преобразовании учитываются равенства $(Ax, u) = (A^T u, x) = (y - x, x)$ 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

$$\left. \begin{aligned} Ax \leq b, \quad B\beta - y = 0, \quad e^T\beta = 1, \quad \beta \geq 0; & \quad (7.3)_1 \\ x - y + A^T = 0, \quad u \geq 0; & \quad (7.3)_2 \\ y - x = u\beta; & \quad (7.3)_3 \\ B^T u\beta + te \geq 0; & \quad (7.3)_4 \\ u^T(AA^T)u + b^T u + t \leq 0. & \quad (7.3)_5 \end{aligned} \right\} \quad (7.3)$$

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), а именно. Взяв x из (7.3)₂, u_β — из (7.3)₃ и подставив в остальные соотношения, получим систему S в окончательном виде:

$$\boxed{\begin{aligned} A(y - A^T u) \leq b, \quad u \geq 0, \quad B\beta - y = 0, \quad e^T\beta = 1, \quad \beta \geq 0 \\ B^T(A^T u) + te \geq 0, \quad u^T(AA^T)u + b^T u + t \leq 0. \end{aligned}} \quad (7.4)$$

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 $[\bar{y}, \bar{\beta}, \bar{t}]$ системы (7.4), то значения остальных переменных находятся (точно или приближенно) согласно соотношениям: $\bar{x} = \bar{y} - A^T \bar{u}$, $\bar{u}_\beta = \bar{y} - \bar{x}$, $\bar{v} = B^T \bar{u}_\beta + \bar{t}e$ (все эти соотношения вытекают из системы (7.2)).

Сделаем такое замечание. 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. Еремин И.И. *Теория линейной оптимизации*. — Екатеринбург: изд-во "Екатеринбург", 1999. —312 с.
2. *Введение в нелинейное программирование* / Под ред. К.-Х. Эльстера. — М.: Наука, 1985. —264 с.

УДК 519.6

Фейеровские методы сильной отделимости выпуклых полиэдральных множеств

Еремин И.И. Известия вузов. Математика.

Предлагается метод редукции задач выпуклого квадратичного программирования (в разных постановках) к некоторой конструктивно задаваемой системе линейных и одного выпуклого неравенства, а для решения последней подключается фейеровский процесс. Редукция осуществляется путем формирования симметрической системы неравенств, являющейся результатом объединения ограничений прямой и двойственной задач и одного неравенства, связывающего их целевые функции. Далее эта система упрощается за счет гауссовских исключений. В целом сформулированный процедурный процесс в работе назван S -технологией.

По схеме S -технологии предложены методы решения общих задач квадратично-выпуклого программирования, а также задачи отделимости непересекающихся выпуклых полиэдров слоем максимальной толщины при разных заданиях этих полиэдров, а именно: при задании либо системами линейных неравенств, либо выпуклыми оболочками конечного числа точек. Рассмотрен и смешанный вариант. В статье указывается на области приложимости рассмотренных алгоритмов.

Библиогр. 2 назв. [Институт математики и механики УрО РАН]