Исследование устойчивости параллельного алгоритма решения задачи сильной отделимости на базе фейеровских отображений*

А.В. Ершова, И.М. Соколинская

Южно-Уральский национальный исследовательский университет

В работе рассматривается задача сильной отделимости, заключающаяся в разделении двух выпуклых непересекающихся многогранников слоем наибольшей толщины. Описывается параллельный алгоритм решения задачи сильной отделимости на базе фейеровских отображений, допускающий эффективную реализацию на многопроцессорных системах с массовым параллелизмом. Вводится понятие устойчиво фейеровского отображения. Доказывается теорема, определяющая условия, при которых фейеровское отображение будет устойчиво фейеровским.

1. Введение

Задача разделения двух выпуклых непересекающихся многогранников имеет большое значение теоретического и прикладного характера в распознавании образов, включающем задачи дискриминации, классификации и др. Задача сильной отделимости может быть решена в ходе итерационного процесса, использующего операцию проектирования. Однако на практике применение этого метода существенно ограничивается тем, что далеко не всегда удается построить конструктивную формулу для вычисления проекции точки на выпуклое множество. Поэтому целесообразно заменить операцию проектирования последовательностью фейеровских отображений [1], строящих некоторую *псевдопроекцию*. Фейеровские методы относятся к итерационным методам проекционного типа, применяемым для решения систем линейных неравенств и задач линейного программирования. Их конструкция, то есть конструкция соответствующих фейеровских операторов, базируется на той или иной суперпозиции элементарных проектирований, а именно – проектирований на полупространства.

Среди современных методов классификации и распознавания образов одно из ведущих мест занимает метод опорных векторов, известный также как метод обобщенного портрета или машины опорных векторов (Support Vector Machines, SVM) [2]. Системы, разработанные на его основе, успешно решают задачи в таких областях, как биоинформатика, машинное зрение, категоризация текстов, распознавание рукописных символов и др. Однако метод опорных векторов оказывается неэффективным в случае изменяющихся исходных данных. Алгоритмы разделения многогранников на основе фейеровских отображений обладают тем преимуществом по сравнению с этим и другими известными методами, что они применимы к нестационарным задачам [3], т.е. к задачам, в которых исходные данные могут меняться в процессе решения задачи.

Для итерационных алгоритмов, решающих нестационарные задачи особенно важным является вопрос их устойчивости, так как динамическое изменение входных данных и погрешности вычислений могут привести к тому, что алгоритм выдаст неверный результат, либо итерационный процесс вообще перестанет сходиться.

В данной работе исследуется вопрос устойчивости параллельного алгоритма решения задачи сильной отделимости на базе фейеровских отображений, предложенного в работе [4]. Далее описывается математическая модель сильной отделимости с использованием фейеровских отображений. Приводится итерационный алгоритм **F** решения задачи сильной отделимости, основанный на построении псевдопроекций. Рассматривается масштабируемый алгоритм **S** построения псевдопроекции, позволяющий реализовать параллельную версию

* Работа выполнена при поддержке РФФИ, грант № 12-01-00452-а.

алгоритма **F**. Вводится понятие устойчиво фейеровского отображения. Доказывается теорема, определяющая условия, при которых фейеровское отображение будет устойчиво фейеровским.

2. Математическая модель

Пусть даны два выпуклых непересекающихся многогранника $M \subset \mathbf{R}^n$ и $N \subset \mathbf{R}^n$, заданные системами линейных неравенств:

$$\begin{gathered} M = \{x \mid Ax \le b\} \ne \varnothing; \\ N = \{x \mid Bx \le d\} \ne \varnothing; \\ M \cap N = \varnothing. \end{gathered} \tag{1}$$

Задача сильной отделимости – это задача нахождения слоя наибольшей толщины (π-слоя), разделяющего M и N. Сильная отделимость, по существу, эквивалентна задаче отыскания расстояния между M и N в смысле метрики

$$\rho(M,N) = \min\{\|x-y\| \mid x \in M,\ y \in N\}. \tag{2}$$

Если $\bar{x} \in M$ и $\bar{y} \in N$ являются arg-точками задачи (2), то есть $\rho(M,N) = \|\bar{x} - \bar{y}\|$, то слоем наибольшей толщины, разделяющим множества M и N, является $P := \{x \mid x \in P_1 \cap P_2\}$, где P_1 и P_2 – полупространства, задаваемые линейными неравенствами

$$(x-\bar{x}, \bar{x}-\bar{y}) \le 0 \ \text{ и } \ (y-\bar{y}, \bar{x}-\bar{y}) \ge 0.$$

Задача сильной отделимости может быть решена с помощью известного метода *последовательного проектирования* [1]. Если множества M и N достаточно просты в смысле простоты реализации операции проектирования точек на них, то алгоритм на базе последовательного проектирования может быть использован на практике. Но если M и N – произвольные многогранники, то такой алгоритм не может быть признан эффективным, так как не известен универсальный конструктивный метод построения проекции точки на многогранник. Ситуацию можно исправить, если вместо операции проектирования использовать фейеровские отображения.

Дадим определение фейеровского отображения. Пусть $\varphi \in \{\mathbf{R}^n \to \mathbf{R}^n\}$. Отображение φ называется *M-фейеровским*, если выполняются следующие два условия:

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

Сконструируем M-фейеровское отображение, следуя работе [5]. Представим систему линейных неравенств, задающих многогранник M, в следующем виде:

$$Ax \le b : l_j(x) = (a_j, x) - b_j \le 0,\ j = 1,\dots,m, \tag{3}$$

где $a_j \ne 0$ для любого j. Определим $l_j^+(x)$ следующим образом:

$$l_j^+(x) = \max\{l_j(x), 0\},\ j = 1,\dots,m. \tag{4}$$

Тогда отображение вида

$$\varphi(x) = x - \sum_{j=1}^{m} \alpha_j \lambda_j \frac{l_j^+(x)}{\|a_j\|^2} a_j \tag{5}$$

будет M-фейеровским для любой системы положительных коэффициентов $\{\alpha_j > 0\}$, $j = 1,\dots,m$, таких, что $\sum_{j=1}^{m} \alpha_j = 1$ и коэффициентов релаксации $0 < \lambda_j < 2$. Аналогичным образом сконструируем N-фейеровское отображение ψ. Используя отображения φ и ψ, мы можем построить следующий алгоритм **F**, решающий задачу сильной отделимости с использованием фейеровских отображений.

Алгоритм F. Пусть задано произвольное начальное приближение $z_0 \in \mathbf{R}^n$. Зафиксируем положительное вещественное число ε. Алгоритм состоит из следующих шагов:

Шаг 0. $k := 0$.

Шаг 1. $x_{k+1} := \lim_{u\to\infty} \varphi^u(z_k)$.

Шаг 2. $y_{k+1} := \lim_{u\to\infty} \psi^u(z_k)$.

Шаг 3. $z_{k+1} := \dfrac{x_{k+1} + y_{k+1}}{2}$.

Шаг 4. $k := k + 1$.

Шаг 5. Если $\min\{\|x_{k+1} - x_k\|, \|y_{k+1} - y_k\|\} \geq \varepsilon$, перейти на шаг 1.

Шаг 6. Стоп.

Алгоритм **F** был исследован в работе [6]. Проведенные вычислительные эксперименты на модельных и случайных задачах подтвердили его эффективность. Однако для больших размерностей работа алгоритма **F** требовала значительного времени. Например, при размерности задачи n = 512 время счета на одном процессорном ядре составило 16 часов, а при размерности задачи $n = 1024$ – 608 часов (более 25 дней). В связи с этим возникла необходимость разработки параллельной версии этого алгоритма для многопроцессорных систем с массовым параллелизмом. Очевидно, что в алгоритме **F** ресурсоемкими являются шаги 1 и 2. На каждом из этих шагов реализуется *последовательный* фейеровский процесс, в результате которого мы получаем *псевдопроекцию* точки на многогранник. Исследования показали, что подобные фейеровские процессы не допускают эффективного распараллеливания на большом количестве процессорных узлов (предел масштабируемости в экспериментах не превышал 8-16 узлов). В следующем разделе описывается масштабируемый алгоритм построения псевдопроекции точки на многогранник с использованием фейеровских отображений, предложенный в работе [4].

3. Масштабируемый алгоритм S построения псевдопроекции

В основе масштабируемого алгоритма построения псевдопроекции на многогранник лежит метод разбиения пространства на подпространства. Для каждого подпространства организуется независимый фейеровский процесс. Через каждые s шагов результаты, полученные на подпространствах, соединяются в один вектор, который и является очередным приближением. Если расстояние между соседними приближениями меньше заданного положительного числа ε, то полученный вектор принимается в качестве псевдопроекции. В противном случае вычисления продолжаются.

Дадим формальное описание алгоритма построения псевдопроекции на выпуклый многогранник, допускающего эффективное распараллеливание на большом количестве процессорных узлов. Введем следующие обозначения. Для произвольного линейного подпространства $\mathrm{P} \subset \mathrm{R}^n$ через $\pi_\mathrm{P}(x)$ будем обозначать ортогональную проекцию $x \in \mathbb{R}^n$ на линейное подпространство P. Везде далее линейное подпространство будем называть просто подпространством. Через $\rho(\mathrm{P}, x) := \min_{p\in\mathrm{P}} \|p - x\|$ будем обозначать расстояние от точки x до подпространства P. Пусть линейное многообразие L получается из P сдвигом на некоторый вектор z: $\mathrm{L} = \mathrm{P} + z$. Через $\pi_\mathrm{L}(x)$ обозначим ортогональную проекцию $x \in \mathbb{R}^n$ на линейное многообразие L:

$$\pi_\mathrm{L}(x) = \pi_\mathrm{P}(x) + z. \tag{6}$$

Алгоритм S. Пусть задано однозначное непрерывное M-фейеровское отображение $\varphi \in \{\mathrm{R}^n \to \mathrm{R}^n\}$, M – выпукло и замкнуто. Зададим разбиение пространства R^n в прямую сумму ортогональных подпространств: $\mathbf{R}^n = \mathrm{P}_1 \oplus \ldots \oplus \mathrm{P}_r$, $\mathrm{P}_i \perp \mathrm{P}_j$ при $i \neq j$. Для каждого подпространства P_i $(i = 1, \ldots, r)$ построим линейное многообразие L_i следующим образом. Пусть $\overline{x}^i \in \operatorname{Arg}\min\limits_{x \in M} \rho(\mathrm{P}_i, x)$. Положим $\overline{z}^i = \pi_{\mathrm{P}_i^{\perp}}(\overline{x}^i) \in \mathrm{P}_i^{\perp}$. Здесь $\mathrm{P}_i^{\perp}$ обозначает ортогональное дополнение к подпространству P_i. Построим линейное многообразие L_i путем сдвига подпространства P_i на вектор $\overline{z}^i$:

$$\mathrm{L}_i = \mathrm{P}_i + \overline{z}^i. \tag{7}$$

Для каждого $i \in \{1, \ldots, r\}$ определим отображение $\varphi_i \in \{\mathrm{R}^n \to \mathrm{L}_i\}$:

$$\varphi_i(x) = \pi_{\mathrm{L}_i}\left(\varphi\left(\pi_{\mathrm{L}_i}(x)\right)\right). \tag{8}$$

Зафиксируем некоторое натуральное число s и положительное вещественное число ε. Положим $x_0 = \mathbf{0} \in \square^n$. Алгоритм состоит из следующих шагов:

Шаг 0. $k := 0$.

Шаг 1. $x_{k+1} := \sum\limits_{i=1}^{r}\left(\varphi_i^s\left(\pi_{\mathrm{L}_i}(x_k)\right) - \overline{z}^i\right)$.

Шаг 2. $k := k + 1$.

Шаг 3. Если $\|x_{k+1} - x_k\| \geq \varepsilon$ & $d_M(x_{k+1}) \geq \varepsilon$, перейти на шаг 1.

Шаг 4. Стоп.

Для выхода из итерационного процесса в алгоритме **S** на шаге 3 используется критерий завершения, включающий в себя функцию невязки d_M, определяющую степень близости точки x к многограннику M. В качестве такой функции в нашей реализации используется следующая функция

$$d_M(x) = \sum_{j=1}^{m} \max\left\{\left\langle a_j, x\right\rangle - b_j, 0\right\}.$$

На основе описанного подхода на языке программирования C++ была разработана параллельная программа, решающая задачу сильной отделимости многогранников для произвольных входных данных. Для организации обменов данными между процессами была использована система параллельного программирования MPI. Исходные тексты программ доступны в Интернет по адресу http://life.susu.ru/discr/. Проведенные вычислительные эксперименты на высокопроизводительном кластере подтвердили эффективность предложенного подхода [4]. В следующем разделе доказывается теорема, из которой следует устойчивость описанного алгоритма по отношению к динамически меняющимся исходным данным задачи.

4. Теорема об устойчиво фейеровском отображении

Пусть задана система линейных неравенств в пространстве R^n:

$$Ax \leq b. \tag{9}$$

Пусть $y = [A, b]$ – *информационный вектор*, задающий все параметры системы (9), $y \in \mathrm{R}^{nm+m}$. Обозначим через M_y многогранник решений системы (9), определяемой информационным вектором y. Имеем $M_y \subset \mathrm{R}^n$.

Лемма. Пусть $\overline{y} \in \mathrm{R}^{nm+m}$ – информационный вектор, задающий устойчиво совместную систему [7] неравенств

$$\overline{A}x \leq \overline{b}. \tag{10}$$

Тогда существует некоторая окрестность V точки $\overline{y}$, такая, что любая точка $\tilde{y} \in V$ также определяет устойчиво совместную систему неравенств

$$\tilde{A}x \le \tilde{b}\,, \tag{11}$$
где $\tilde{y} = \left[\tilde{A}, \tilde{b}\right]$.

Доказательство. Устойчивая совместность системы (10) означает, что существует $\overline{x} \in \mathrm{R}^n$ такой, что $\overline{A}\overline{x} < \overline{b}$, или, что то же самое, $\overline{A}\overline{x} - \overline{b} < 0$. Поскольку вектор-функция $A\overline{x} - b$ непрерывна по $y = \left[A, b\right]$, то для всех точек $\tilde{y} = \left[\tilde{A}, \tilde{b}\right]$ некоторой окрестности V точки $\overline{y}$ также будет выполняться неравенство $\tilde{A}\overline{x} - \tilde{b} < 0$, что равносильно $\tilde{A}\overline{x} < \tilde{b}$. А это, в свою очередь, означает, что для всех точек этой окрестности система (11) будет устойчиво совместной, что и требовалось доказать. Лемма доказана.

Определение. Пусть задано отображение $\varphi : \mathrm{R}^{nm+m} \times \mathrm{R}^n \to \mathrm{R}^n$ двух аргументов $y \in \mathrm{R}^{nm+m}$ и $x \in \mathrm{R}^n$. Обозначим через φ_y отображение из R^n в R^n, которое получается из φ, путем фиксации аргумента y. Отображение φ является *устойчиво фейеровским* относительно точки $\overline{y} \in \mathrm{R}^{nm+m}$, если $\varphi_{\overline{y}} \in \mathbf{F}_{M_{\overline{y}}}$ и существует окрестность $V \subset \mathrm{R}^{nm+m}$ точки $\overline{y}$ такая, что для любого $\tilde{y} \in V$ имеем $\varphi_{\tilde{y}} \in \mathbf{F}_{M_{\tilde{y}}}$.

Теорема. Пусть отображение $\varphi : \mathrm{R}^{nm+m} \times \mathrm{R}^n \to \mathrm{R}^n$ двух аргументов $y \in \mathrm{R}^{nm+m}$ и $x \in \mathrm{R}^n$ является непрерывным по x и y[1]. Пусть система (10), определяемая информационным вектором $\overline{y}$, является устойчиво совместной и $\varphi_{\overline{y}} \in \mathbf{F}_{M_{\overline{y}}}$. Тогда отображение φ является *устойчиво фейеровским* относительно точки $\overline{y} \in \mathrm{R}^{nm+m}$.

Доказательство. Доказательство проведем от противного. Предположим, что для любой окрестности V точки $\overline{y}$ существует
$$\tilde{y} \in V \tag{12}$$
такой, что $\varphi_{\tilde{y}} \notin \mathbf{F}_{M_{\tilde{y}}}$. В соответствии с определением фейеровского отображения это означает, что не выполняется одно из следующих условий:

1) $\varphi_{\tilde{y}}(z) = z, \qquad \forall z \in M_{\tilde{y}}$;
2) $\left\|\varphi_{\tilde{y}}(x) - z\right\| < \left\|x - z\right\|, \qquad \forall z \in M_{\tilde{y}}, \qquad \forall x \notin M_{\tilde{y}}$.

В силу свойства 39.4 [5] при доказательстве мы можем ограничиться $\operatorname{int} M_{\tilde{y}}$, представляющим подмножество внутренних точек множества $M_{\tilde{y}}$. В силу леммы 1 мы можем выбрать окрестность V таким образом, чтобы для всех $\tilde{y} \in V$ получалась устойчиво совместная система, то есть $\operatorname{int} M_{\tilde{y}} \ne \varnothing$.

Предположим сначала, что *не выполняется условие 1)*. Это означает, что существует точка $\tilde{z} \in \operatorname{int} M_{\tilde{y}}$ такая, что
$$\varphi_{\tilde{y}}(\tilde{z}) \ne \tilde{z}\,. \tag{13}$$

Так как $\tilde{z}$ является внутренней точкой многогранника $M_{\tilde{y}}$, то $\tilde{A}\tilde{z} - \tilde{b} < 0$. Поскольку вектор-функция $A\tilde{z} - b$ непрерывна относительно $y = [A, b]$, то существует окрестность $\tilde{V}$ точки $\tilde{y}$ такая, что для любой точки $\tilde{y}' \in \tilde{V}$ имеем $\tilde{A}'\tilde{z} - \tilde{b}' < 0$. Выберем окрестность V таким образом, чтобы выполнялось вложение $V \subset \tilde{V}$. Тогда получается, что $\tilde{z}$ также является внутренней точкой многогранника $M_{\overline{y}}$. Поскольку по условию теоремы $\varphi_{\overline{y}} \in \mathbf{F}_{M_{\overline{y}}}$, отсюда следует, что
$$\varphi_{\overline{y}}(\tilde{z}) = \tilde{z}\,. \tag{14}$$

---
[1] Отображение $\varphi(y, x)$ *непрерывно* по x в точке $\overline{x}$, если при фиксированном y для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что отображение φ определено в δ-окрестности точки $\overline{x}$ и $\left\|x - \overline{x}\right\| < \delta \Rightarrow \left\|\varphi\left(y, x\right) - \varphi\left(y, \overline{x}\right)\right\| < \varepsilon$. Непрерывность по y определяется аналогичным образом.

Так как отображение φ является непрерывным по y, то из (13) следует, что существует окрестность $\tilde{V}'$ точки $\tilde{y}$ такая, что для любой точки $\tilde{y}' \in \tilde{V}'$ имеем $\varphi_{\tilde{y}'}(\tilde{z}) \neq \tilde{z}$. Выберем окрестность V таким образом, чтобы выполнялось вложение $V \subset \tilde{V}'$. Тогда получается, что $\bar{y} \in \tilde{V}'$, откуда следует $\varphi_{\bar{y}}(\tilde{z}) \neq \tilde{z}$. Получили противоречие с (14). Таким образом, мы доказали, что условие 1) выполняется.

Теперь предположим, что *не выполняется условие 2)*. Это означает, что существует точка $\tilde{z} \in \operatorname{int} M_{\tilde{y}}$ и точка $\tilde{x} \notin M_{\tilde{y}}$ такие, что

$$\left\|\varphi_{\tilde{y}}(\tilde{x}) - \tilde{z}\right\| \geq \left\|\tilde{x} - \tilde{z}\right\|. \tag{15}$$

Так как $\tilde{z}$ является внутренней точкой многогранника $M_{\tilde{y}}$, то $\tilde{A}\tilde{z} - \tilde{b} < 0$. Поскольку вектор-функция $A\tilde{z} - b$ непрерывна относительно $y = [A, b]$, то существует окрестность $\tilde{V}$ точки $\tilde{y}$ такая, что для любой точки $\tilde{y}' \in \tilde{V}$ имеем $\tilde{A}'\tilde{z} - \tilde{b}' < 0$. Выберем окрестность V таким образом, чтобы выполнялось вложение $V \subset \tilde{V}$. Тогда получается, что $\tilde{z}$ является внутренней точкой многогранника $M_{\bar{y}}$. Поскольку по условию теоремы $\varphi_{\bar{y}} \in \mathbf{F}_{M_{\bar{y}}}$, отсюда следует, что

$$\left\|\varphi_{\bar{y}}(\tilde{x}) - \tilde{z}\right\| < \left\|\tilde{x} - \tilde{z}\right\|. \tag{16}$$

Так как отображение φ является непрерывным по y, то из (16) следует, что существует окрестность $\bar{V}$ точки $\bar{y}$ такая, что для любой точки $\bar{y}' \in \bar{V}$ имеем $\left\|\varphi_{\bar{y}'}(\tilde{x}) - \tilde{z}\right\| < \left\|\tilde{x} - \tilde{z}\right\|$. Выберем окрестность V таким образом, чтобы выполнялось вложение $V \subset \bar{V}$. Тогда в силу (12) имеем $\tilde{y} \in \bar{V}$, откуда следует $\left\|\varphi_{\tilde{y}}(\tilde{x}) - \tilde{z}\right\| < \left\|\tilde{x} - \tilde{z}\right\|$. Получили противоречие с (15). Таким образом, мы доказали, что условие 2) выполняется. Т е о р е м а д о к а з а н а .

В заключение осталось заметить, что фейеровское отображение (5), в соответствии с доказанной теоремой, является устойчиво фейеровским, что обеспечивает устойчивость предложенного алгоритма.

Литература

1. *Еремин И.И.* Фейеровские методы сильной отделимости выпуклых полиэдральных множеств // Известия вузов. Сер. математика, 2006. № 12. С. 33-43.
2. *Boser B., Guyon I., Vapnik V.* A training algorithm for optimal margin classifiers // Proc. of the 5th Annual ACM Workshop on Computational Learning Theory. Pittsburgh: ACM Press, 1992. 144–152.
3. *Еремин И.И., Мазуров В.Д.* Нестационарные процессы математического программирования. М.: Наука, 1979.
4. *Ершова А.В., Соколинская И.М.* Параллельный алгоритм решения задачи сильной отделимости на основе фейеровских отображений // Вычислительные методы и программирование. 2011. Т. 12, № 2. С. 178-189.
5. *Еремин И.И.* Теория линейной оптимизации. Екатеринбург: Изд-во “Екатеринбург”, 1999.
6. *Ершова А.В.* Алгоритм разделения двух выпуклых непересекающихся многогранников с использованием фейеровских отображений // Системы управления и информационные технологии, 2009. №1(35). С. 53-56.
7. *Черников С.Н.* Линейные неравенства. М.: Изд-во "Наука", 1968. 488 с.