УДК 519.8

П.И. СТЕЦЮК, Е.А. НУРМИНСКИЙ

НЕГЛАДКИЙ ШТРАФ И СУБГРАДИЕНТНЫЕ АЛГОРИТМЫ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ПРОЕКЦИИ НА ПОЛИТОП[1]

Ключевые слова: *задача проекции, негладкие точные штрафные функции, r-алгоритмы.*

Постановка задачи. В работе рассматривается задача проекции начала координат на выпуклый многогранник. Задано множество n-мерных векторов $\mathcal{X}=\{\hat{x}^1, \hat{x}^2, \ldots, \hat{x}^m\}$, требуется найти вектор с минимальной нормой в их выпуклой оболочке (политопе):

$$\min_{x\in \operatorname{co}(\mathcal{X})} \|x\|^2 = \min_{\substack{x=\sum_{i=1}^m \lambda_i \hat{x}^i, \\ \sum_{i=1}^m \lambda_i=1,\ \lambda_i\ge 0,\ i=1,2,\ldots,m}} \|x\|^2 . \tag{1}$$

Данная задача представляет интерес для многих приложений, а также может служить основой для решения аналогичной проблемы для многогранников, заданных иным способом [1].

Для задачи проекции (в постановке (1) и других) создано значительное количество различных алгоритмов [2], однако в связи с многими специфическими приложениями продолжается разработка новых методов. Потребность в них мотивируется чрезвычайно высокой размерностью (1) в практических задачах обработки изображений, планировании радиационной терапии в медицинской информатике [3], аппроксимации геофизических полей [4] и др.

Целью настоящей работы является построение семейства алгоритмов для решения задачи (1) методами минимизации негладких выпуклых функций с проектированием на неотрицательный ортант. С помощью метода негладких штрафных функций задача (1) сведена к негладкой экстремальной задаче (минимизации негладкой выпуклой функции при ограничениях на неотрицательность переменных) и исследован вопрос выбора конечной штрафной константы, обеспечивающей такую эквивалентность. Обсуждается также применение r-алгоритмов [5] для решения негладкой экстремальной задачи.

Проекция как задача квадратичного программирования. Пусть матрица X размера $m\times n$ составлена из вектор-столбцов $\hat{x}^1, \hat{x}^2, \ldots, \hat{x}^m$, а вектор λ размера m состоит из компонент $\lambda_1, \lambda_2, \ldots, \lambda_m$. Тогда задаче (1) соответствует задача квадратичного программирования

$$\min_{\lambda} \|\sum_{i=1}^m \lambda_i \hat{x}^i\|^2 = \min_{\lambda} \|X\lambda\|^2 = \|X\lambda^*\|^2 = \|x^*\|^2 = d_* \tag{2}$$

при ограничениях

$$\sum_{i=1}^m \lambda_i = 1, \tag{3}$$

$$\lambda_i \ge 0, \quad i=1,\ldots,m. \tag{4}$$

[1]Работа выполнена при финансовой поддержке совместного украинско-российского проекта ДФФД–Ф28.1/005 и РФФИ–09–01–90413–Укр_ф_а «Субградиентные методы ускоренной сходимости в задачах выпуклого программирования».

Здесь λ^* — необязательно единственное оптимальное решение (барицентрические координаты) задачи (2)–(4), а x^* — вектор минимальной нормы, который является выпуклой комбинацией векторов $\{\hat{x}^1, \hat{x}^2, \ldots, \hat{x}^m\}$, т.е. $x^* = X\lambda^* = \sum_{i=1}^{m} \lambda_i^* \hat{x}^i$, где $\sum_{i=1}^{m} \lambda_i^* = 1$, $\lambda_i^* \geq 0, i = 1, \ldots, m$. Вектор x^* — единственный в силу строгой выпуклости целевой функции (1) по переменным x.

Прямолинейному решению задачи (2)–(4) как задачи квадратичного программирования с использованием стандартных средств [6] препятствует ее вырожденность как в формулировке (1), так и формулировке (2)–(4), где она неизбежно возникает в наиболее интересном с точки зрения приложений случае $m > n$. Поэтому далее модифицируем (2)–(4) в целях построения в конечном итоге эквивалентной негладкой задачи.

Рассмотрим задачу квадратичного программирования

$$d_* = \min \| X\lambda \|^2 \tag{5}$$

при ограничениях

$$-\sum_{i=1}^{m} \lambda_i + 1 \leq 0, \tag{6}$$

$$\lambda_i \geq 0, \quad i = 1, \ldots, m. \tag{7}$$

Она получена из задачи (2)–(4) заменой ограничения-равенства (3) ограничением-неравенством $\sum_{i=1}^{m} \lambda_i \geq 1$, которое записано в форме (6) для удобства дальнейшего изложения.

Утверждение 1. Пусть $d_* > 0$. Тогда задача (5)–(7) полностью эквивалентна задаче (2)–(4) в том смысле, что множества оптимальных решений обеих задач совпадают.

Доказательство. Пусть у задачи (5)–(7) существует оптимальное решение $\tilde{\lambda}^* = (\tilde{\lambda}_1^*, \tilde{\lambda}_2^*, \ldots, \tilde{\lambda}_m^*)^{\mathrm{T}}$ такое, что $\sum_{i=1}^{m} \tilde{\lambda}_i^* > 1$. Пусть также ему соответствует оптимальное значение целевой функции $\tilde{d}_* = \| \sum_{i=1}^{m} \tilde{\lambda}_i^* \hat{x}^i \|^2$. Тогда вектор $\bar{\lambda}^*$ с компонентами

$$\bar{\lambda}_i^* = \frac{\tilde{\lambda}_i^*}{\sum_{i=1}^{m} \tilde{\lambda}_i^*}, \quad i = 1, \ldots, m,$$

для которых $\sum_{i=1}^{m} \bar{\lambda}_i^* = 1$, удовлетворяет ограничениям (6), (7) и для целевой функции (5) обеспечивает значение

$$\bar{d}_* = \| X\bar{\lambda}^* \|^2 = \frac{\tilde{d}_*}{\sum_{i=1}^{m} \tilde{\lambda}_i^*} < \tilde{d}_*,$$

что противоречит оптимальности точки $\tilde{\lambda}^*$ в задаче (5)–(7).

Если $d_* = 0$, то задача (5)–(7) не эквивалентна задаче (2)–(4), так как множество ее решений расширяется за счет того, что любому оптимальному решению λ^* задачи (2)–(4) в задаче (5)–(7) соответствует множество оптимальных решений $\tilde{\lambda}^* = \mu\lambda^*$, где $\mu \geq 1$. Обе задачи эквивалентны только в том смысле, что им соответствует одно и то же значение целевой функции $d_* = 0$.

Негладкий штраф в задаче (5)–(7) и его свойства. Пусть для учета ограничения (6) используется негладкая штрафная функция в форме функции максимума

$$\Phi_P(\lambda) = ||X\lambda||^2 + P\max\left\{0,\ -\sum_{i=1}^{m}\lambda_i + 1\right\}, \tag{8}$$

где $P > 0$ — штрафной параметр.

Справедлива следующая теорема.

Теорема 1. При $P > 2d_*$ задача (5)–(7) эквивалентна задаче

$$d_* = \min_{\lambda \geq 0} \Phi_P(\lambda). \tag{9}$$

Доказательство. С помощью теоремы Б.Н. Пшеничного [7, теорема 2.14, с. 25] установим конечное значение штрафного параметра P, при котором задача (9) эквивалентна задаче (5)–(7). Рассмотрим семейство параметрических задач

$$V(t) = \inf_{\lambda}\left\{||X\lambda||^2 : -\sum_{i=1}^{m}\lambda_i + 1 \leq t;\ \lambda_i \geq 0,\ i = 1, \ldots, m\right\},$$

зависящее от параметра $t \in R$. Очевидно, что задача

$$V(0) = \inf_{\lambda}\{||X\lambda||^2 :\ -\sum_{i=1}^{m}\lambda_i + 1 \leq 0;\ \lambda_i \geq 0,\ i = 1, \ldots, m\}$$

совпадает с задачей (5)–(7), откуда имеем $V(0) = d_*$.

Теорема Пшеничного утверждает следующее. Пусть $\inf\limits_{t>0}\dfrac{V(t) - V(0)}{t} = -L > -\infty$. Если $P > L$, то точки минимума задач $V(0)$ и $\min\limits_{\lambda \geq 0}\Phi_P(\lambda)$, где $\Phi_P(\lambda)$ определяется по формуле (8), совпадают.

Для установки штрафного параметра достаточно вычислить «нужный» инфинум, т.е. $-L$. Представим его в виде

$$-L = \inf_{t>0}\frac{V(t) - V(0)}{t} = \min\left\{\inf_{0<t<1}\frac{V(t) - V(0)}{t},\ \inf_{t \geq 1}\frac{V(t) - V(0)}{t}\right\}. \tag{10}$$

Вторую часть под минимумом в (10) легко посчитать с учетом того, что при $t \geq 1$ имеем $V(t) = 0$, а $V(0) = d_* > 0$, откуда следует

$$\inf_{t \geq 1}\frac{V(t) - V(0)}{t} = \inf_{t \geq 1}\frac{-V(0)}{t} = -V(0).$$

С учетом последнего равенства из (10) получаем

$$-L = \inf_{t>0}\frac{V(t) - V(0)}{t} = \min\left\{\inf_{0<t<1}\frac{V(t) - V(0)}{t},\ -V(0)\right\}. \tag{11}$$

Семейство параметрических задач $V(t)$ при $0 < t < 1$ запишем следующим образом:

$$V(t) = \inf_{\lambda}\left\{||X\lambda||^2 :\ -\sum_{i=1}^{m}\lambda_i + 1 \leq t;\ \lambda_i \geq 0,\ i = 1, \ldots, m\right\} =$$

$$= \inf_{\lambda}\left\{||X\lambda||^2 :\ -\sum_{i=1}^{m}\lambda_i + (1 - t) \leq 0;\ \lambda_i \geq 0,\ i = 1, \ldots, m\right\} =$$

$$= \inf_{\lambda}\left\{||X\lambda||^2 :\ -\sum_{i=1}^{m}\frac{\lambda_i}{1 - t} + 1 \leq 0;\ \frac{\lambda_i}{1 - t} \geq 0, i = 1, \ldots, m\right\}.$$

Отсюда, сделав замену переменных $\tilde{\lambda} = \frac{\lambda}{1-t}$, получим

$$V(t) = (1-t)^2 \inf_{\tilde{\lambda}} \left\{ ||X\tilde{\lambda}||^2 : -\sum_{i=1}^{m} \tilde{\lambda}_i + 1 \le 0;\ \tilde{\lambda} \ge 0 \right\} = (1-t)^2 V(0).$$

В результате имеем

$$\inf_{0<t<1} \frac{V(t)-V(0)}{t} = \inf_{0<t<1} \frac{(1-t)^2 - 1}{t} V(0) = \inf_{0<t<1} (-2+t) V(0) = -2V(0).$$

Подставив это выражение в (11), получим

$$-L = \min\{-2V(0), -V(0)\} = -2V(0),$$

откуда с учетом того, что $V(0) = d_*$, имеем $L = 2V(0) = 2d_*$.

Связь негладкой задачи с задачей (2)–(4). В случае $d_* > 0$ задача (2)–(4) и задача (5)–(7) полностью эквивалентны. Поэтому оптимальное решение задачи (9) при выборе значения P согласно теореме 1 дает оптимальное решение задачи (2)–(4). Субградиент выпуклой негладкой функции $\Phi_P(\lambda)$ в точке $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m)^{\mathrm{T}}$, где $\bar{\lambda}_i \ge 0$, $i = \overline{1, m}$, определяется по правилу

$$g_{\Phi}(\bar{\lambda}) = \begin{cases} 2X^{\mathrm{T}} X \bar{\lambda}, & \text{если} \quad \sum_{i=1}^{m} \bar{\lambda}_i > 1, \\ 2X^{\mathrm{T}} X \bar{\lambda} - Pe, & \text{если} \quad \sum_{i=1}^{m} \bar{\lambda}_i \le 1, \end{cases}$$

где m-мерный вектор e состоит из всех единиц. Свойства субградиента в точке минимума функции $\Phi_P(\lambda)$ позволяют усилить результат теоремы 1, перенеся его на случай задачи (2)–(4), когда $d_* = 0$.

Справедлива следующая теорема.

Теорема 2. Пусть $\tilde{\lambda}^* = (\tilde{\lambda}_1^*, \ldots, \tilde{\lambda}_m^*)^{\mathrm{T}}$ — оптимальное решение задачи (9) при $P > 2d_*$. Тогда вектор λ^* и вектор x^*, которые определяются по правилу

$$\lambda_i^* = \frac{\tilde{\lambda}_i^*}{\sum_{i=1}^{m} \tilde{\lambda}_i^*}, \quad i = 1, \ldots, m, \quad x^* = X\lambda^*,$$

являются оптимальным решением задачи (2)–(4).

Теорема 2 дает достаточные условия для нахождения проекции с помощью методов минимизации негладких функций. Она включает как случай $d_* > 0$, так и $d_* = 0$. Последний в теореме использует критерий останова для методов негладкой оптимизации, который определяется равенством нулю субградиента минимизируемой функции, т.е. $g_{\Phi}(\tilde{\lambda}^*) \equiv 0$. В этом случае, чтобы получить оптимальное решение для задачи (2)–(4), достаточно оптимальную для задачи (9) точку $\tilde{\lambda}^*$ отмасштабировать так, чтобы она удовлетворяла требованию выпуклой комбинации. В случае $d_* > 0$ такое масштабирование излишне, так как оно автоматически выполняется за счет выбора величины штрафа P согласно теореме 1.

О применении субградиентных алгоритмов в задаче проекции. Для нахождения проекции начала координат на выпуклый многогранник вместо задачи (2)–(4) можем решать негладкую выпуклую задачу (9). Для этого подойдет любой из субградиентных методов минимизации негладких выпуклых функций, позволяющий учитывать неотрицательность переменных. Например, для решения задачи (9) можно использовать r-алгоритмы [5], эффективные при решения задач негладкой оптимизации большой размерности. Эти методы нечувствительны к вырожденности матрицы $X^{\mathrm{T}} X$, обладают устойчивостью по отношению к выбору начального приближения и достаточно быстрой сходимостью.

При использовании r-алгоритмов возможны следующие варианты. Первый вариант определяется стандартным приемом решения таких задач и связан с «четным» продолжением выпуклой функции на все пространство путем замены переменных $\lambda_i = |x_i|$, $i = 1, \ldots, m$. В результате получаем многоэкстремальную функцию $F(x) = \Phi_P(|x_1|, \ldots, |x_m|) = \Phi_P(\lambda)$, в которой все локальные минимумы глобальны. Для нахождения локальных минимумов функции $F(x)$ можно применять любой из методов минимизации негладких функций. Практика использования r-алгоритмов показала, что такой подход дает надежные результаты и, в определенном смысле, лучше метода негладких штрафных функций, так как не требует подбора штрафных параметров. Эта схема достаточно просто реализуется и будет эффективной при решении задач (9), когда количество нулевых компонент в оптимальном решении λ^* сравнительно небольшое по отношению к количеству векторов m.

Второй вариант связан со случаем $m >> n$, который характеризуется большим количеством нулевых компонент в оптимальном векторе λ^*. Здесь целесообразнее использовать модификации r-алгоритмов, разработанные специально для учета простейших границ на переменные. В работе [8] предложена модификация r-алгоритма, в которой по определенным правилам происходит временная фиксация нулевых значений ряда переменных либо их освобождение. Более алгоритмически обоснованные правила для фиксации и освобождения переменных предложены в [9]. Модификации r-алгоритмов, базирующиеся на этих правилах, связаны с процессом минимизации по некоторому подмножеству переменных, и естественно, что на ряде практических задач они показали значительное уменьшение числа итераций в сравнении со схемой «четного» продолжения. Их реализация для задачи проекции позволяет также ускорить вычисление субградиента.

В заключение отметим, что ускоренная сходимость r-алгоритмов, устойчивость по отношению к выбору начального приближения и нечувствительность к вырожденности задачи квадратичного программирования позволяют обеспечить реализацию на их основе эффективных алгоритмов для задачи проекции. Эти особенности алгоритма делают его перспективным для реализации проективных задач большой размерности в системах параллельных или распределенных вычислений.

СПИСОК ЛИТЕРАТУРЫ

1. Нурминский Е.А. Проекция на внешне заданные полиэдры // Журн. вычисл. математики и мат. физики. — 2008. — **48**, вып. 3. — С. 387–396.
2. Bauschke H., Borwein J.M. On projection algorithms for solving convex feasibility problems // SIAM Rev. — 1996. — **38**, N 3. — P. 367—426.
3. Censor Y., Jiang M., Louis A.K. (Eds.) Mathematical methods in biomedical imaging and intensity-modulated radiation therapy (IMRT) // Publications of the Scuola Normale Superiore Subseries: CRM Series, 2008. — **7**. — 526 p.
4. Васин В.В. Итерационные процессы фейеровского типа в некорректных задачах с априорной информацией // Изв. ВУЗов. Математика. — 2009. — Вып. 2. — С. 3–24.
5. Шор Н.З. Методы минимизации недифференцируемых функций и их приложения. — К.: Наук. думка, 1979. — 199 с.
6. Гилл Ф., Мюррей У., Райт М. Практическая оптимизация. — М.: Мир, 1974. — 376 с.
7. Пшеничный Б.Н. Метод линеаризации. — М.: Наука, 1983. — 136 с.
8. Приятель А.М. Решение задачи выпуклого программирования при помощи r-алгоритма с проектированием // Исследования методов решения экстремальных задач: Сб. науч. тр. — Киев: Ин-т кибернетики им. В.М. Глушкова АН УССР, 1986. — С. 24–30.
9. Кунцевич А.В. К вопросу об эффективности применения алгоритмов оптимизации с растяжением пространства // Кибернетика. — 1989. — № 2. — С. 116–117.

Поступила 27.04.2009