УДК 519.8

Э.И. НЕНАХОВ

О МАТРИЧНЫХ ЗАДАЧАХ ОПТИМИЗАЦИИ

Рассматриваются задачи выпуклого полуопределенного программирования. Приводятся свойства матричних конусов. Предложен метод решения параметрической матричной задачи с ограничениями на положительную определенность.

Под задачей минимизации матричной функции (или матричной оптимизации) понимается такая задача, которую удобно формулировать в терминах матричного исчисления. Многие важные характеристики матриц (собственные числа симметричных квадратичных матриц, их суммы) представляют собой негладкие функции от элементов матриц. Поэтому матричные задачи оптимизации зачастую оказываются негладкими, что затрудняет их решение классическими методами.

Одна из первых работ, в которой изучались свойства субградиентов негладких функций, возникающих в связи с ограничениями неотрицательной определенности, когда выбираемые параметры входят лишь в диагональные элементы матриц, была работа Флетчера [1]. В работе [2] описаны свойства некоторых матричных функций и приведены алгоритмы решения матричных задач математического программирования. Один из наиболее практически эффективных алгоритмов решения такого рода задач – r-алгоритм [3].

Данная работа посвящена исследованию свойств некоторых матричных конусов и решению задач матричной оптимизации.

Рассмотрим пространство M^n всех матриц A размерностью $n \times n$ с вещественными элементами a_{ij}, где i – индекс строки, j – индекс столбца, $i = 1, \dots, n,\ j = 1, \dots, n$. В M^n введем скалярное произведение матриц

$$A * B = \sum_{i,\ j} a_{i\,j} \quad b_{i\,j}$$

Пусть $S^n \subset M^n$ – пространство симметричных матриц, т. е. $A = A^T$, $S_-^n \subset M^n$ – пространство кососимметричных матриц, т. е. $A = -A^T$. Пространство M^n разлагается в прямую сумму пространств S^n и S_-^n.

Обозначим S_+^n конус положительно определенных симметричных матриц

$$S_+^n = \left\{ A \in S^n : (Ap, p) \geq 0, \, p \in R^n \right\}.$$

Матрица $A \in S_+^n$ строго положительно определена, если здесь имеет место строгое неравенство. Конус строго положительно определенных матриц обозначим S_{++}^n.

Для любого вектора-столбца $p \in R^n$ матрица $pp^T = \{p_i \, p_j\} \in S_+^n$:

$$\left(x, pp^T x\right) = x^T pp^T x = \left(p^T x\right)^2 \geq 0, \; \forall \, x \in R^n.$$

Определим конус, сопряженный к конусу S_+^n:

$$S_+^{n*} = \left\{ Z \in M^n : Z * A \geq 0, \; A \in S_+^n \right\}.$$

Конус S_+^{n*} состоит из может быть несимметричных положительно определенных матриц. Действительно, для произвольных матрицы $Z \in S_+^{n*}$ и вектора $p \in R^n$ выполняется

$$Z * \left(pp^T\right) = \sum_{i\,j} z_{i\,j} \, p_i \, p_j = (Zp, p) \geq 0.$$

Если Z – положительно определенная матрица, а A – симметричная положительно определенная матрица, то $Z * A \geq 0$. Действительно, имеет место представление

$$A = \sum_{i=1}^{n} \lambda_i \, x_i \, x_i^T, \tag{1}$$

где $\lambda_i \geq 0$ – собственные значения матрицы A, а x_i – соответствующие им собственные векторы, причем x_i образуют ортонормированную систему векторов [4]. Так как Z – положительно определенная матрица, то

$$Z * A = \sum_{i=1}^{n} \lambda_i \left(Zx_i, x_i\right) \geq 0.$$

Конус $S_+^n \subset M^n = S^n + S_-^n$ не имеет внутренних точек, так как если матрица симметричная, то малое возмущение делает ее не симметричной. Однако нетрудно показать, что конус S_{++}^n является внутренностью конуса S_+^n

относительно симметричных матриц, т. е. $ri\,S_+^n$ – множество строго положительно определенных симметричных матриц. Кроме того, несущее пространство конуса S_+^n

$$Lin\left(S_+^n\right) = S_+^n - S_+^n .$$

Пусть матрица $A_0 \in S_+^n$. Рассмотрим конус возможных направлений к конусу S_+^n в A_0: $K\left(A_0\right) = con\left(S_+^n - A_0\right)$, а также сопряженный к нему конус $K\left(A_0\right)^*$.

Теорема 1. Сопряженный конус к конусу $K\left(A_0\right)$ имеет вид

$$K\left(A_0\right)^* = \left\{Z \in S_+^{n*} : Z * A_0 = 0\right\}.$$

Пусть теперь $A_0 \in ri\,S_+^n$, т. е. A_0 – строго положительно определенная симметричная матрица. Тогда по любой симметричной матрице B и достаточно малому $\varepsilon > 0$ можно построить такую матрицу $A = A_0 + \varepsilon B$, что $A \in S_+^n$. Следовательно, для $Z \in K\left(A_0\right)^*$ выполняется $Z * \left(A - A_0\right) = Z * A \geq 0$ и, значит,

$$Z * \left(A - A_0\right) = \varepsilon\left(Z * B\right) \geq 0 .$$

Но такие рассуждения верны и для симметричной матрицы $-B$, так что $Z * B = 0$ для произвольной матрицы $B \in S^n$. В частности, положив $B = pp^T$, $p \in R^n$, получаем

$$Z * \left(pp^T\right) = (Zp, p) = \left(p, Z^T p\right) = 0 .$$

Следовательно, для положительно определенной матрицы Z

$$\left(\left(Z + Z^T\right)p, p\right) = 0, \; p \in R^n .$$

Так как $\left(Z + Z^T\right)$ – симметричная матрица, то $Z + Z^T = 0$, т. е. матрица $Z = -Z^T$ – кососимметричная. Итак, получен следующий результат.

Теорема 2. Если матрица $A_0 \in ri\,S_+^n$, то конус $K\left(A_0\right)^*$ состоит из кососимметричных матриц.

Заметим, что любая кососимметричная матрица положительно определена.

В дальнейшем будем вести рассуждения в пространстве симметричных матриц. Здесь имеет место равенство $S_+^{n*} = S_+^n$, где сопряженный конус к исходному есть

$$S_+^{n*} = \left\{B \in S^n : A * B \geq 0, A \in S_+^n\right\}.$$

Пусть $A \in S_+^n$, то аналогично теореме 1 выполняется

Теорема 3. Сопряженный конус к конусу $K(A)$ имеет вид

$$K(A)^* = \{B \in S_+^n : A * B = 0\}.$$

Выведем еще одно свойство матриц, принадлежащих конусу $K(A)^*$. Для матрицы $A \in S_+^n$ в представлении (1) $\lambda_i \geq 0,\ i = 1,\dots,n$, и полагаем $\lambda_i = 0,\ i \leq k$. Тогда для любой матрицы $B \in K(A)^*$ выполняется

$$B * A = B * \sum_{i=1}^{n} \lambda_i\, x_i\, x_i^T = \sum_{i>k} \lambda_i (B\, x_i, x_i) = 0\,.$$

Следовательно, $(B\, x_i, x_i) = 0,\ i > k$. Но $B \in S_+^n$, поэтому $B\, x_i = 0,\ i > k$. Заметим, что в R^2 элементарная матрица поворота на угол $\frac{\pi}{2}$ – кососимметричная. Итак, собственный вектор матрицы A, соответствующий ее положительному собственному значению, является собственным вектором матрицы B, соответствующим ее нулевому собственному значению.

Далее, пусть $y_i,\ i = 1,\dots,n$ – ортонормированная система собственных векторов B, а $\gamma_i \geq 0$ – соответствующие им собственные значения. Как вышепоказано, $\gamma_i = 0,\ y_i = x_i,\ i > k$. Если $i \leq k$, то $\gamma_i \geq 0$ и $(y_i, x_j) = (y_i, y_j) = 0$ для $j > i$. Таким образом, $y_i,\ i \leq k$, лежат в подпространстве, ортогональном подпространству, натянутому на векторы $x_i,\ i > k$. Поэтому $y_i,\ i \leq k$ – собственные вектора матрицы A, соответствующие нулевому собственному значению этой матрицы, а матрица B может быть представлена так:

$$B = \sum_{i \leq k} \gamma_i\, y_i\, y_i^T,\ \gamma_i \geq 0,\ i \leq k\,. \tag{2}$$

Итак, для любой симметричной положительно определенной матрицы A матрица $B \in K(A)^*$ тогда и только тогда, когда B представима в виде (2), где собственные значения матрицы B неотрицательны, y_i – ортонормированная система собственных векторов матрицы A, соответствующие ее нулевому собственному значению.

Далее, рассмотрим матричную задачу математического программирования:

$$\min \{f_0(A) \mid f_i(A) \leq 0,\ i = 1,\dots,k,\ A \in S_+^n\}, \tag{3}$$

где все функции $f_i(A)$ – выпуклые; A_0 – решение этой задачи.

Для построения необходимых и достаточных условий минимума задачи (3) воспользуемся леммой 4.2 [5]. В качестве выпуклого конуса K_M, требуемого для выполнения условий этой леммы, достаточно взять конус $K(A_0)$, а в качестве функции $h_i(P)$, $i=1, \ldots, k$, взять производную по направлению P функции f_i в $A_0: h_i(P)=f_i'(A_0, P)$.

Строим функцию Лагранжа задачи (3) $L(A,u)=\sum_i u_i f_i(A)$ и матрицу $L'(A,u)=\sum_i u_i f_i'(A)$, где $f_i'(A)$ – субградиент функции $f_i(A)$. Тогда согласно лемме 4.2 [5] найдутся такие числа u_i, не все равные нулю, что

$$L'(A_0, u)\in K(A_0)^*, \tag{4}$$

$$u_i \geq 0,\ i=0,1, \ldots ,k;\ u_i f_i(A_0)=0,\ i=1, \ldots ,k. \tag{5}$$

Непосредственное применение теоремы 3 к включению (4) дает положительную определенность матрицы $L'(A_0,u)$ и выполнение условия $L'(A_0,u)*A_0=0$.

Теорема 4. Если A_0 – решение задачи (3), то найдутся не все равные нулю числа u_i такие, что выполняются условия (5), $L'(A_0,u)\in S_+^n$ и $L'(A_0,u)*A_0=0$. Если выполняется условие Слейтера для задачи (3), то эти условия являются и достаточными.

Заметим, что симметричная матрица $L'(A,u)$ не знакоопределена, однако она положительно определена в решении задачи.

Для решения многих задач математического программирования применим метод линеаризации, разработанный для решения гладких задач [6]. На основе этого метода и метода отсечения в [7] построен алгоритм решения негладкой выпуклой задачи, сходящийся при очень общих предположениях.

Рассмотрим задачу

$$\min\left\{ f(x) \mid f_i(x)\leq 0,\ i=1, \ldots ,k,\ x\in M\right\}, \tag{6}$$

где $x\in R^n$, f, f_i – выпуклые непрерывные функции, M – выпуклое множество. Если выпуклая функция $\varphi(x)=\max_i f_i(x)$, то задача (6) эквивалентна задаче

$$\min\left\{ f(x) \middle| \varphi(x)\leq 0,\ x\in M \right\}.$$

Пусть в процессе работы алгоритма построены точки $X_k=\{x_i,\ i=1, \ldots ,k\}$ и число $N_k>0$. Пусть $\bar{x}_k$ – точка минимума $f(x_i)+N_k \max\{0;\varphi(x_i)\}$, $x_i\in X_k$. Следующий массив X_{k+1} и число N_{k+1} строится по такому правилу. Решаем вспомогательную задачу

$$\min \frac{1}{2}\| x-\bar{x}_k \|^2 + \xi_1 + N_k \xi_2$$
$$f(x_i) + (f'(x_i), x - x_i) \le \xi_1,\ i = 1, \dots, k,$$
$$\varphi(x_i) + (\varphi'(x_i), x - x_i) \le \xi_2,\ i = 1, \dots, k,$$
$$\xi_2 \ge 0,\ x \in M. \tag{7}$$

Полагаем $X_{k+1} = X_k \cup \{x_{k+1}\}$, $N_{k+1} = \begin{cases} N_k, & \text{если } \xi_2^k = 0, \\ 2N_k, & \text{если } \xi_2^k > 0, \end{cases}$

где x_{k+1}, ξ_1^k, ξ_2^k – решение задачи (7), $f' \in \partial f$, $\varphi' \in \partial \varphi$.

Если для задачи (6) выполняется условие Слейтера, то существуют множители Куна – Таккера, сумма которых ограничена сверху. Это позволяет утверждать, что начиная с некоторого достаточно большого k N_k будет константой, $\xi_2^k = 0$ и алгоритм индуцирует последовательность точек x_k такую, что $\lim\limits_{k\to\infty} \| x_{k+1} - \bar{x}_k \| = 0$. Критерием останова алгоритма будет малость вектора $p_k = x_{k+1} - \bar{x}_k$.

Если функция f или φ имеет линейный кусок (например, кусочно-линейная функция), то точки x_i, лежащие на этой плоскости, определяют одинаковые ограничения в задаче (7), что упрощает решение вспомогательной задачи на каждой итерации.

К задаче (6) сводятся такие матричные экстремальные задачи на графах: нахождения оценки максимально взвешенного внутренне устойчивого множества вершин графа $G(V, E)$ и построение оценки максимального разреза графа. Например, в первой задаче целевая функция $f(x) = \lambda_{\max}(A(x))$, $\varphi(x) = -\lambda_{\min}(B(x))$, где $\lambda_{\max}$, $\lambda_{\min}$ – максимальные и минимальные собственные значения матрицы, $A(x) = W + U(x)$, $B(x) = 2I + U(x)$, W – весовая матрица, $U(x)$ – матрица, в которой элементы, соответствующие E, переменные, остальные – нулевые (вершины графа соединены последовательно). Здесь для проверки выполнения условия положительной определенности матрицы B используется критерий $\lambda_{\min} \ge 0$. Так как $\lambda_{\min}(B)$ – вогнутая функция от элементов матрицы, то $\varphi(x)$ – выпуклая функция от параметра x. Целевая функция задачи также выпуклая. Поэтому для ее решения можно применять вышеописанный алгоритм.

Вышеуказанные задачи на графах относятся к задачам минимизации матричных параметрических функций с ограничениями на положительную определенность. Большинство таких задач, как было указано ранее, являются негладкими.

В работе [8] построен алгоритм решения общей матричной задачи оптимизации (3), использующий штрафные функции. Если все функции задачи являются непрерывно дифференцируемыми, то алгоритм индуцирует последовательность, сходящуюся к стационарной точке исходной задачи. То есть находится точка, в которой выполняются условия Каруша – Куна – Таккера. Эти условия близки к условиям теоремы 4.

Е.І. Ненахов

ПРО МАТРИЧНІ ЗАДАЧІ ОПТИМІЗАЦІЇ

Розглядаються задачі опуклого напіввизначеного програмування. Наведено властивості матричних конусів. Запропоновано метод розв'язування параметричної матричної задачі з обмеженнями на додатну визначеність.

E.I. Nenakhov

ON MATRIX OPTIMIZATION PROBLEMS

We consider convex semidefinite problems. The properties of matrix cones are derived. Method of solution of parametric matrix problems with positively definite constraints is proposed.

1. *Fletcher R.* Semidefinite matrix constrains in optimization // SIAM J. Control Optim. – 1985. – **23.** – P. 493–513.
2. *Шор Н.З.* Задачи минимизации матричных функций и недифференцируемая оптимизация // Обозрение прикладной и промышленной математики. – 1995. – **2**, вып. 1. – С. 113–138.
3. *Шор Н.З.* Методы минимизации недифференцируемых функций и их приложение. – Киев: Наук. думка, 1979. – 199 с.
4. *Гантмахер Ф.Р.* Теория матриц. – М.: ГИТТЛ, 1953. – 491 с.
5. *Пшеничный Б.Н.* Необходимые условия экстремума. – М.: Наука, 1969. – 151 с.
6. *Пшеничный Б.Н.* Метод линеаризации. – М.: Наука, 1983. – 136 с.
7. *Пшеничный Б.Н., Ненахов Э.И., Кузьменко В.Н.* Комбинированный метод решения общей задачи выпуклого программирования // Кибернетика и системный анализ. – 1998. – № 4. – С. 121–133.
8. *Kanzow C., Nagel C., Kato H., Fukushima M.* Successive linearization methods for nonlinear semidefinite programs // Computational Optim. And Appl. – 2005. – **31.** – P. 251–273.

Получено 23.03.2010