УДК 519.8

Э.И. НЕНАХОВ

ЭКСТРЕМАЛЬНЫЕ ЗАДАЧИ С КОНИЧЕСКИМИ ОГРАНИЧЕНИЯМИ

Рассматриваются задачи нелинейного полуопределенного программирования. Построены условия оптимальности первого порядка. Для выпуклой задачи с коническим ограничением предложен алгоритм ее решения, который является глобально сходящимся. Исследована общая параметрическая задача.

Введение. В настоящей работе изучаются задачи нелинейного программирования, когда переменные есть матрицы. Они являются частным случаем общей задачи оптимизации. Особенности таких задач состоят в том, что матрицы имеют некоторые численные характеристики, которые достаточно сложным образом определяются через их параметры. Это требуется учитывать, в частности, при построении необходимых и достаточных условий экстремума.

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 $A * B = \sum_{i,j} a_{ij} b_{ij}$ и норму $\| A \|^2 = A * A = = \sum_{i,j} a_{ij}^2$.

Пусть теперь $S^n \subset M^n$ – пространство симметричных матриц, $S_+^n \subset S^n$ – конус положительно определенных симметричных матриц. По определению $A \in S_+^n$, если $(Ap, p) \geq 0 \ \forall p \in R^n$, и писать $A \geq 0$. Матрица $A \in S_+^n$ строго положительно определена $(A > 0)$, если здесь имеет место строгое неравенство при $p \neq 0$.

Конус строго положительно определенных матриц обозначим S_{++}^n.

При рассмотрении оптимизационных задач необходимо исследовать конус, сопряженный к конусу S_+^n:

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

Теорема 1 [1]. Имеет место равенство $S_+^n * = S_+^n$.

Для матрицы $A_0 \in S_+^n$ введем конус $K(A_0) = \left\{ \lambda (A - A_0) : A \in S_{++}^n, \lambda > 0 \right\}$.

Теорема 2. Сопряженный конус к конусу $K(A_0)$ есть $K(A_0)* = \left\{ B \geq 0 : A_0 * B = 0 \right\}$.

Доказательство. Пусть матрица $B \in K(A_0)*$, то из определения сопряженного конуса

$$B * \lambda (A - A_0) \geq 0, \forall A \in S_{++}^n, \lambda > 0.$$

Учитывая, что S_{++}^n – внутренность конуса S_+^n и любая матрица A из S_+^n – предельная для матриц из S_{++}^n, получаем

$$A * B \geq A_0 * B, A \in S_+^n. \tag{1}$$

Покажем, что $A * B \geq 0, \forall A \geq 0$. Пусть это не выполняется, т. е. для некоторой матрицы $A_1 \geq 0$ будет $A_1 * B < 0$. Тогда для любого $\lambda > 0$ $(\lambda A_1) * B < 0$, и при $\lambda \to +\infty$ $(\lambda A_1) * B \to -\infty$. Но $\lambda A_1 \geq 0$ для $\lambda > 0$, так, что приходим в противоречие с (1). Таким образом, по теореме 1 $B \geq 0$.

Далее, полагая в (1) $A = 0$ получаем, что $A_0 * B \leq 0$. Так как $A_0 \in S_+^n$, то в соответствии с теоремой 1 $A_0 * B \geq 0$. Значит, $A_0 * B = 0$, что и требовалось доказать.

Рассмотрим функции от матриц. Пусть $A \in S^n$, $\varphi(A)$ – некоторая функция от ее элементов $a_{ij}, i \leq j$. Вначале предположим, что она дифференцируема и обозначим $\varphi'_{ij} = \dfrac{\partial \varphi}{\partial a_{ij}}$, $i \leq j$. Для произвольного приращения $P \in S^n$ выполняется

$$\varphi(A + P) = \varphi(A) + \sum_{i \leq j} \varphi'_{ij}(A) p_{ij} + O(P) = \varphi(A) + \sum_{i=1}^{n} \varphi'_{ii}(A) p_{ii} +$$

$$+ \frac{1}{2} \sum_{i < j} \varphi'_{ij}(A) p_{ij} + \frac{1}{2} \sum_{i > j} \varphi'_{ij}(A) p_{ij} + O(P) = \varphi(A) + \varphi'(A) * P + O(P),$$

где матрица $\varphi'(A)$ имеет элементы φ'_{ij} на диагонали и $\frac{1}{2} \varphi'_{ij}, i \neq j$.

Пусть теперь $\varphi(A)$ – выпуклая функция. Тогда существует ее субградиент $\{\varphi'_{ij}\}_{i\le j}$ [2], для приращения $P\in S^n$ выполняется

$$\varphi(A+P)\ge\varphi(A)+\sum_{i\le j}\varphi'_{ij}(A)p_{ij}=\varphi(A)+\varphi'(A)*P,$$

так, что матрица $\varphi'(A)$ – субградиент функции $\varphi(A)$ в пространстве S^n.

В качестве примера выпуклой функции рассмотрим наибольшее собственное значение матрицы A

$$\lambda(A)=\max_{z\in B}(Az,z),\text{ где } B=\{z:\|z\|=1\}.$$

Производная по направлению P от этой функции в точке A [3] есть

$$\lambda'(A,P)=\max_{z\in B(A)}(Pz,z)=\max_{z\in B(A)}\left(zz^T\right)*P,$$

где $B(A)=\{z\in B:\lambda(A)=(Az,z)\}$.

Следовательно, субдифференциал функции $\lambda(A)$ [2, 3] есть

$$\partial\lambda(A)=co\left\{zz^T:z\in B(A)\right\},$$

так, что субградиент функции $\lambda(A)$ имеет вид

$$\lambda'(A)=\sum_{k=1}^{m}\gamma_k z_k z_k^T,\ \gamma_k\ge 0,\ \sum_{k=1}^{m}\gamma_k=1,\ z_k\in B(A).$$

Заметим, что каждый z_k – собственный вектор матрицы A, соответствующий ее максимальному собственному значению. Кроме того, для множества матриц ранга 1 $R=\left\{zz^T:\|z\|=1\right\}\subset S_+^n$ выполняется

$$\lambda(A)=\max_{z\in B}A*\left(zz^T\right)=\max_{C\in R}A*C=W_R(A),$$

где $W_R(A)$ – опорная функция к множеству R в точке A [3]. Так как R – компакт, то его выпуклая оболочка $Q=coR=\{C\ge 0:S_pC=1\}$ замкнута (S_pC – след матрицы C). Следовательно,

$$\lambda(A)=\max_{C\in Q}A*C=W_Q(A),$$

а экстремальные точки множества Q есть матрицы из множества R.

Рассмотрим следующую задачу:

$$\min\left\{\varphi_0(A)\,|\,\varphi_k(A)\le 0,\ k=1,\dots,l,\ A\in S_+^n\right\},\tag{2}$$

где функции $\varphi_k(A),\ k=0,1,\dots,l,$ – выпуклые, A_0 – ее решение. Пусть выполняется условие Слейтера: существует такая матрица $A\in S_+^n$, что $\varphi_k(A)<0,\ k=1,\dots,l$.

Для построения необходимых и достаточных условий минимума задачи (2) воспользуемся леммой 4.2 [4]. В качестве выпуклого конуса K_M, требуемого для выполнения условий этой леммы, можно взять конус возможных направлений в точке A_0 к множеству S_+^n, а в качестве функций $h_k(P)$ – производные по направлению P в точке A_0 $\varphi'_k(A_0, P)$, $k = 0, 1, \ldots, l$.

Далее, строим функцию Лагранжа $L(A,u) = \sum_{k=0}^{l} u_k \varphi_k(A)$ и матрицу $L'(A,u) = \sum_{k=0}^{l} u_k \varphi'_k(A)$, где $\varphi'_k(A)$ – некоторый субградиент функции $\varphi_k(A)$. Тогда согласно упомянутой лемме найдутся такие числа u_k, не все равные нулю, что

$$L'(A_0, u) \in K(A_0)^*, \tag{3}$$

$$u_k \varphi_k(A_0) = 0,\ k = 1, \ldots, l;\ u_k \geq 0,\ k = 0, 1, \ldots, l. \tag{4}$$

В силу теоремы 2 из включения (3) следует, что

$$L'(A_0, u) \geq 0,\ L'(A_0, u) * A_0 = 0. \tag{5}$$

Поскольку выполняется условие Слейтера, то множитель $u_0 = 1$.

Теорема 3. Для того, чтобы матрица A_0 была решением задачи (2) необходимо и достаточно, чтобы выполнялись условия (4), (5).

Исследуем общую параметрическую задачу. Предположим, что элементы матрицы A зависят от вектора $x \in R^q$ $A(x) = \{a_{ij}(x)\}$, $i = 1, \ldots, n$, $j = 1, \ldots, n$, $a_{ij}(x) = a_{ji}(x)$. Пусть a_{ij} – непрерывно дифференцируемые функции x, a'_{ij} – вектор с компонентами $\frac{\partial a_{ij}}{\partial x_k}$, $k = 1, \ldots, q$. Обозначим $\frac{\partial A}{\partial x_k} = \left\{ \frac{\partial a'_{ij}}{\partial x_k} \right\}$, $i = 1, \ldots, n$, $j = 1, \ldots, n$. Теперь для $p \in R^q$ определим

$$A'(x, p) = \lim_{t \to +0} \frac{A(x + tp) - A(x)}{t} = \{(a'_{ij}(x), p)\}_{i,j}.$$

Тогда для функции $\varphi(x) = \max_{z \in B} (A(x)z, z)$ производная по направлению p в точке x будет

$$\varphi'(x,p)=\lim_{t\to+0}\frac{\varphi(x+t\,p)-\varphi(x)}{t}=\max_{z\in B(A(x))}\lim_{t\to+0}\frac{(A(x+t\,p)z,z)-(A(x)z,z)}{t}=$$

$$=\max_{z\in B(A(x))}(A'(x,p)z,z)=\max_{z\in B(A(x))}\left(z\,z^T\right)*A'(x,p),$$

где $B(A(x))=\{z\in B:\varphi(x)=(A(x)z,z)\}$.

Общая параметрическая задача имеет следующий вид:

$$\min\left\{\varphi_0(x)\,\middle|\,\varphi_k(x)=0,\ k=-m,\dots,-1;\ \varphi_k(x)\le 0,\ k=1,\dots,l;\ A(x)\in S_+^n;\ x\in R^q\right\}. \quad (6)$$

Пусть для $k<0$ $\varphi_k(x)$ – гладкая функция, $h_k(x,p)=\varphi_k'(x,p)$. Для $k\ge 0$ в решении x задачи (6)

$$\lim_{t\to+0}\frac{\varphi_k(x+t\,p+o(t))-\varphi_k(x)}{t}\le h_k(x,p),$$

где $h_k(x,p)$ – непрерывная, выпуклая и положительно однородная функция по p, $o(t)/t\to 0$ при $t\to+0$.

Кроме того, обозначим $\partial\varphi_k(x)$ субдифференциал выпуклой по p функции $h_k(x,p)$ при $p=0\left(\partial\varphi_k(x)=\{\varphi_k'(x)\}\right), k<0$, а вектор $U*\dfrac{\partial A}{\partial x}=\left\{U*\dfrac{\partial A}{\partial x_k}\right\}\in R^q$.

Тогда справедлив следующий результат.

Теорема 4. Для того, чтобы точка x была решением задачи (6), необходимо, чтобы нашлись такие не все равные нулю числа u_k, $k=-m,\dots,l$ и матрица $U\in S_+^n$, что

$$u_k\,\varphi_k(x)=0,\ k>0;\ u_k\ge 0,\ k\ge 0;$$

$$U*A(x)=0;\ \sum_{k=-m}^{l}u_k\,\varphi_k'(x)=U*\frac{\partial A}{\partial x},$$

где $\varphi_k'(x)\in\partial\varphi_k(x)$.

В линейном случае $\varphi_0(x)=(c,x)$, $A(x)=A_0+\sum\limits_{k=1}^{q}x_k\,A_k$ точка x^* удовлетворяет условию оптимальности, если найдется такое $U\in S_+^n$, что

$$c_k=U*A_k,\ k=1,\dots q;\ U*A_0+\sum_{k=1}^{q}x_k^*\left(U*A_k\right)=0.$$

Замечание 1. Пусть в задаче (6) отсутствуют функциональные ограничения, а $A(x)$ – положительное полуопределенное выпуклое $(psd-convex)$ отображение, т.е.

$$tA(x)+(1-t)A(y)-A(tx+(1-t)y)\in S_+^n,\ \forall x,y\in R^q,\ t\in[0,1].$$

Для случая, когда целевая функция и $A(x)$ достаточно гладкие на R^q необходимые условия минимума первого и второго порядка построены в [5]. При этом условия первого порядка вытекают из теоремы 4.

Для решения задачи (2) построим алгоритм, основанный на комбинации метода линеаризации и метода отсекающих гиперплоскостей. Для этого определим выпуклую функцию $\varphi(A)=\max\limits_{1\le k\le l}\varphi_k(A)$. Тогда задача (2) эквивалентна задаче

$$\min\left\{\varphi_0(A)\,\middle|\,\varphi(A)\le 0,\ A\in S_+^n\right\}.$$

Введем теперь штрафную функцию $\Phi_\alpha(A)=\varphi_0(A)+\alpha\max\{0,\varphi(x)\}$ и рассмотрим экстремальную задачу

$$\min\left\{\Phi_\alpha(A)\,\middle|\,A\in S_+^n\right\}=\min\left\{\xi_1+\alpha\xi_2\,\middle|\,\varphi_0(A)\le\xi_1,\ \varphi(A)\le\xi_2,\ \xi_2\ge 0,\ A\in S_+^n\right\}. \quad (7)$$

Пусть, по прежнему, A_0 – решение задачи (2), u_k – ее множитель Лагранжа.

Лемма 1. Если число α достаточно велико $\left(\alpha>\sum\limits_{k=1}^{l}u_k\right)$, то решение задачи (7) есть $A_0,\ \xi_1=\varphi_0(A_0),\ \xi_2=0$.

Рассмотрим вспомогательную задачу. Исходная матрица $A_1\in S_+^n$ и число α_1 заданы. Начальное множество X_1 состоит из этой единственной матрицы. Пусть на j–й итерации алгоритма построено множество $X_j=\{A_s,\ s=1,\dots,j\}$, число α_j, $\bar A_j$ – точка минимума $\Phi_\alpha(A)$ на X_j. Следующее множество X_{j+1} и число α_{j+1} строятся по такому правилу. Решаем вспомогательную задачу

$$\begin{gathered}\min\frac{1}{2}\left\|A-\bar A_j\right\|^2+\xi_1+\alpha_j\xi_2,\\ \varphi_0(A_s)+\varphi_0'(A_s)*(A-A_s)\le\xi_1,\ s=1,\dots,j,\\ \varphi(A_s)+\varphi'(A_s)*(A-A_s)\le\xi_2,\ s=1,\dots,j,\\ \xi_2\ge 0,\ A\in S_+^n.\end{gathered} \quad (8)$$

Пусть $A_{j+1},\ \xi_1^{j+1},\xi_2^{j+1}$ – решение задачи (8), полагаем множество $X_{j+1}=X_j\cup\{A_{j+1}\}$, $\alpha_{j+1}=2\alpha_j$ при $\xi_2^{j+1}>0$, иначе $\alpha_{j+1}=\alpha_j$.

Лемма 2. Начиная с некоторого достаточно большого j индуцируемая алгоритмом α_j будет константой α.

В силу выпуклости функций φ_0 и φ при $A = \bar{A}_j$, $\xi_1 = \varphi_0(\bar{A}_j)$, $\xi_2 = \max\{0, \varphi(\bar{A}_j)\}$ все неравенства в (8) выполняются, поэтому для матрицы $P_{j+1} = A_{j+1} - \bar{A}_j$ имеет место оценка

$$\frac{1}{2}\left\| P_{j+1} \right\|^2 \le \Phi_{\alpha_j}(\bar{A}_j) - (\xi_1^j + \alpha_j \xi_2^j).$$

Отсюда, согласно лемме 2 для достаточно больших j

$$\frac{1}{2}\left\| P_{j+1} \right\|^2 \le \Phi_{\alpha_j}(A_j) - (\xi_1^j + \alpha \xi_2^j).$$

Но правая часть этого неравенства при $j \to \infty$ (как и ξ_2^j) стремится к нулю. Следовательно, $\lim\limits_{j\to\infty} P_j = 0$. Используя данное равенство, аналогично схеме доказательства [6], можно показать, что последовательность $\{A_j\}$ сходится к решению задачи (2).

Замечание 2. Для случая, когда в задаче (2) функции непрерывно дифференцируемые в [7] построен алгоритм, близкий к вышеописанному.

Заключение. В работе исследованы экстремальные задачи, в которых переменными являются симметричные матрицы. Условие положительной определенности существенно усложняет получение необходимых (и достаточных) условий экстремума. Дальнейшее развитие данной работы будет направлено на построение необходимых условий экстремума второго порядка.

Е.І. Ненахов

ЕКСТРЕМАЛЬНІ ЗАДАЧІ З КОНІЧНИМИ ОБМЕЖЕННЯМИ

Розглядаються задачі нелінійного напіввизначеного програмування. Побудовані умови оптимальності першого порядку. Для опуклої задачі з конічним обмеженням запропоновано алгоритм її розв′язку, який є глобально збіжним. Досліджена загальна параметрична задача.

E.I. Nenakhov

EXTREMAL PROBLEMS WITH CONE CONSTRAINTS

We consider nonlinear semidefinite problems. First-order optimality conditions are derived. An algorithm for solving convex programming with cone constraints is proposed. The algorithm globally converges. General parametric problem was investigated.

1. *Fletcher R.* Semidefinite matrix constraints in optimization // SIAM J. Control Optim. – 1985. – **23**. – P. 493–513.

2. *Шор Н.З.* Методы минимизации недифференцируемых функций и их применение. – Киев: Наук. думка, 1979. – 199 с.
3. *Пшеничный Б.Н.* Выпуклый анализ и экстремальные задачи. – М.: Наука, 1980. – 319 с.
4. *Пшеничный Б.Н.* Необходимые условия экстремума. – М.: Наука, 1969. – 151 с.
5. *Shapiro A.* First and second order analysis of nonlinear semidefinite programs // Math. progr. – 1997. – **77**. – P. 301 – 320.
6. *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.
7. *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 31.05.2006