О МЕТОДЕ ПЕРЕБОРА ГРАНЕЙ В КВАДРАТИЧНОМ ПРОГРАММИРОВАНИИ*

В. Н. Малозёмов
malv@math.spbu.ru

Е. К. Чернэуцану
katerinache@yandex.ru

10 сентября 2011 г.

1°. Рассмотрим задачу квадратичного программирования

$$Q(x) := \tfrac{1}{2}\langle Dx, x\rangle + \langle c, x\rangle \to \inf_{x\in\Omega}, \tag{1}$$

где $D = D[N,N]$ — симметричная неотрицательно определённая матрица и Ω — выпуклое многогранное множество, определяемое системой линейных неравенств

$$A[M,N] \times x[N] \geqslant b[M].$$

В докладе [1] приведено подробное доказательство следующего утверждения: *для того чтобы задача* (1) *имела решение, необходимо и достаточно, чтобы множество планов* Ω *было непусто и чтобы целевая функция* $Q(x)$ *была ограничена снизу на* Ω.

Упомянутое доказательство конструктивно. В нём предлагается при различных $I \subset M$ (не исключая $I = \varnothing$) исследовать системы линейных уравнений

$$\begin{aligned} D[N,N] \times x[N] - A^T[N,I] \times u[I] &= -c[N], \\ -A[I,N] \times x[N] \qquad\qquad\qquad &= -b[I]. \end{aligned} \tag{2}$$

Назовём множество $I \subset M$ *подходящим*, если система (2) имеет решение (x', u') и $x' \in \Omega$. Соответствующее решение (x', u') назовём *подходящей парой*. Среди подходящих множеств I следует выбрать то, которому соответствует подходящая пара (x', u') с наименьшим значением $Q(x')$. Найденное x' и будет оптимальным планом задачи (1).

*Семинар по дискретному гармоническому анализу и геометрическому моделированию «DHA & CAGD»: `http://www.dha.spb.ru/`

Остаётся организовать перебор конечного числа подходящих множеств I так, чтобы соответствующие значения $Q(x')$ строго убывали. Получим конечный метод решения задачи (1), который и называется *методом перебора граней.*

2°. Описание метода перебора граней будем проводить в предположении, что выполнены *условия невырожденности.* Они включают два пункта:

(a) матрица D положительно определена на $\mathbb{R}^N$;

(b) при каждом $x \in \Omega$ строки матрицы $A\big[I(x), N\big]$, где

$$I(x) = \big\{i \in M \mid A[i,N] \times x[N] = b[i]\big\},$$

линейно независимы.

Выясним сначала, что дают условия невырожденности. Условие (a) гарантирует ограниченность снизу функции $Q(x)$ на $\mathbb{R}^N$. Точнее, выполняется неравенство

$$Q(x) \geqslant -\tfrac{1}{2}\langle D^{-1}c, c\rangle \quad \forall\, x \in \mathbb{R}^N. \tag{3}$$

Действительно, в силу положительной определённости матрицы D существует обратная матрица D^{-1}, которая также является положительно определённой. Точка минимума x_* строго выпуклой на $\mathbb{R}^N$ функции $Q(x)$ определяется из уравнения $Q'(x) = \mathbb{O}$ или $Dx + c = \mathbb{O}$. Отсюда следует, что $x_* = -D^{-1}c$ и

$$Q(x_*) = \tfrac{1}{2}\langle c, D^{-1}c\rangle - \langle c, D^{-1}c\rangle = -\tfrac{1}{2}\langle D^{-1}c, c\rangle.$$

Неравенство $Q(x) \geqslant Q(x_*)$ равносильно (3).

Таким образом, при выполнении условия (a) задача (1) имеет решение, только если множество её планов Ω непусто. Более того, в силу строгой выпуклости функции $Q(x)$ на $\mathbb{R}^N$ решение задачи (1) единственно.

Теперь выясним роль условия (b). Возьмём точку $x_0 \in \Omega$ с множеством индексов активных ограничений $I = I(x_0)$. По условию (b) строки матрицы $A[I,N]$ линейно независимы. Нетрудно проверить, что матрица

$$F[I,I] = A[I,N] \times D[N,N] \times A^T[N,I]$$

симметрична и положительно определена. Симметричность очевидна. Положительная определённость проверяется так: при $u \neq \mathbb{O}$ имеем

$$\langle Fu, u\rangle = \langle ADA^Tu, u\rangle = \langle DA^Tu, A^Tu\rangle > 0,$$

поскольку $A^Tu \neq \mathbb{O}$ в силу линейной независимости строк матрицы $A[I,N]$. Как симметричная и положительно определённая, матрица F имеет обратную матрицу F^{-1}, которая также является симметричной и положительно определённой.

Рассмотрим симметричную матрицу G системы (2):

$$G = \begin{pmatrix} D[N,N] & -A^T[N,I] \\ -A[I,N] & \mathbb{O}[I,I] \end{pmatrix}.$$

При выполнении условий невырожденности матрица G имеет обратную матрицу G^{-1}, при этом [2]

$$G^{-1} = \begin{pmatrix} E[N,N] & D^{-1}[N,N] \times A^T[N,I] \\ \mathbb{O}[I,N] & E[I,I] \end{pmatrix} \times \begin{pmatrix} D^{-1}[N,N] & \mathbb{O}[N,I] \\ \mathbb{O}[I,N] & -F^{-1}[I,I] \end{pmatrix} \times$$
$$\times \begin{pmatrix} E[N,N] & \mathbb{O}[I,N] \\ A[I,N] \times D^{-1}[N,N] & E[I,I] \end{pmatrix}.$$

То, что $G^{-1}G = E$, проверяется непосредственно.

Подчеркнём, что матрица G связана с планом x_0.

3°. Зададимся таким вопросом: когда подходящее индексное множество $I \subset M$ порождает решение задачи (1)?

ПРЕДЛОЖЕНИЕ 1. *Пусть $I \subset M$ — подходящее индексное множество и (x', u') — соответствующая подходящая пара. Тогда x' — оптимальный план задачи* (1)*, если $u'[I] \geqslant \mathbb{O}$.*

Доказательство. Пара (x', u') удовлетворяет соотношениям

$$\begin{aligned} D[N,N] \times x'[N] - A^T[N,I] \times u'[I] &= -c[N], && (4)\\ -A[I,N] \times x'[N] \qquad\qquad\qquad &= -b[I], && (5)\\ A[M \setminus I, N] \times x'[N] \qquad\qquad\qquad &\geqslant b[M \setminus I], && (6)\\ u'[I] &\geqslant \mathbb{O}. && (7) \end{aligned}$$

Расширим множество I, включив в него индексы из $M \setminus I$, на которых неравенство (6) выполняется как равенство (если такие индексы найдутся). На вновь введённых индексах i положим $u'[i] = 0$. Пара (x', u') с преобразованным u' удовлетворяет соотношениям (4), (5), (7) и (6), в котором знак “$\geqslant$” следует заменить на знак строгого неравенства “$>$”. Это значит, что $x' \in \Omega$, $I = I(x')$ и

$$Q'(x') = A^T\big[N, I(x')\big] \times u'\big[I(x')\big],$$
$$u'\big[I(x')\big] \geqslant \mathbb{O}.$$

По теореме Куна-Таккера [3] вектор x' является решением задачи (1). (Для этого достаточно, чтобы матрица D была неотрицательно определённой.)

Предложение доказано. □

4°. Нам потребуется одно вспомогательное утверждение.

ПРЕДЛОЖЕНИЕ 2. *Пусть (x', u') — решение системы (2) и x — произвольный вектор, удовлетворяющий уравнению*

$$A[I, N] \times x[N] = b[I].$$

Тогда при всех вещественных t справедливо равенство

$$Q(x + ts) = Q(x) - t(1 - \tfrac{t}{2})\langle Ds, s\rangle, \tag{8}$$

где $s = x' - x$.

Доказательство. В (4) и (5) подставим $x' = x + s$. Получим

$$\begin{aligned} D[N, N] \times s[N] - A^T[N, I] \times u'[I] &= -Q'(x), &(9)\\ A[I, N] \times s[N] \qquad\qquad\qquad\quad &= \mathbb{O}. &(10)\end{aligned}$$

Умножим (9) слева на $s[N]$. С учётом (10) придём к равенству

$$\langle Ds, s\rangle = -\langle Q'(x), s\rangle. \tag{11}$$

Теперь запишем разложение квадратичной функции

$$Q(x + ts) = Q(x) + t\langle Q'(x), s\rangle + \tfrac{1}{2}t^2\langle Ds, s\rangle. \tag{12}$$

Формула (8) очевидным образом следует из (12) и (11). □

При $t = 1$ равенство (8) принимает вид

$$Q(x') = Q(x) - \tfrac{1}{2}\langle Ds, s\rangle.$$

5°. Построим начальное подходящее множество $I_0 \subset M$. С этой целью возьмём произвольный план $x \in \Omega$ с множеством индексов активных ограничений $I = I(x)$ и решим систему линейных уравнений (2). В силу условия невырожденности система (2) имеет решение (x', u'). Если $x' \in \Omega$, то по определению I — подходящее множество. Положим $I_0 = I$. Соответствующей подходящей парой будет $(x_0, u_0) = (x', u')$.

Пусть $x' \notin \Omega$, то есть при некотором $i \in M \setminus I$ выполняется неравенство

$$A[i, N] \times x'[N] < b[i]. \tag{13}$$

Множество всех таких индексов обозначим J'. Докажем следующее утверждение (оно соответствует лемме об относительной границе [1]): *на интервале (x, x') найдётся точка $\hat{x} \in \Omega$, такая, что $\big|I(\hat{x})\big| > \big|I(x)\big|$ и*

$$Q(\hat{x}) < Q(x). \tag{14}$$

Для доказательства введём обозначения

$$\Delta(x) = Ax - b, \quad s = x' - x.$$

Ясно, что $s \neq \mathbb{O}$. Рассмотрим интервал (x, x'). Его параметрическое представление имеет вид

$$x(t) = tx' + (1-t)x = x + ts, \quad t \in (0,1).$$

При этом

$$\Delta\bigl(x(t)\bigr) = tAx' + (1-t)Ax - b = t\Delta(x') + (1-t)\Delta(x).$$

В силу выбора x и x' при всех $t \in (0,1)$ справедливо равенство

$$\Bigl(\Delta\bigl(x(t)\bigr)\Bigr)[i] = 0, \quad i \in I.$$

Далее, имеем

$$\bigl(\Delta(x)\bigr)[i] > 0,\ i \in M \setminus I; \quad \bigl(\Delta(x')\bigr)[i] \geqslant 0,\ i \in (M \setminus I) \setminus J',$$

поэтому при всех $t \in (0,1)$

$$\Bigl(\Delta\bigl(x(t)\bigr)\Bigr)[i] > 0, \quad i \in (M \setminus I) \setminus J'.$$

При $i \in J'$ неравенство

$$\Bigl(\Delta\bigl(x(t)\bigr)\Bigr)[i] := \bigl(\Delta(x)\bigr)[i] - t\Bigl(\bigl(\Delta(x)\bigr)[i] - \bigl(\Delta(x')\bigr)[i]\Bigr) \geqslant 0$$

равносильно следующему

$$t \leqslant \frac{\bigl(\Delta(x)\bigr)[i]}{\bigl(\Delta(x)\bigr)[i] - \bigl(\Delta(x')\bigr)[i]}.$$

У дроби из правой части этого неравенства числитель и знаменатель положительны (числитель — на основании того, что $J' \subset M \setminus I$, знаменатель — в силу (13)). Значение самой дроби лежит в интервале $(0,1)$. Положим

$$\hat{t} = \min_{i \in J'} \Bigl\{ \frac{\bigl(\Delta(x)\bigr)[i]}{\bigl(\Delta(x)\bigr)[i] - \bigl(\Delta(x')\bigr)[i]} \Bigr\}. \tag{15}$$

Ясно, что $\hat{t} \in (0,1)$. По построению точка $\hat{x} = x(\hat{t})$ принадлежит Ω. Соответствующее множество индексов активных ограничений $\hat{I} = I(\hat{x})$ содержит I и те индексы из J', на которых достигается минимум в правой части (15). По крайней мере, $|\hat{I}| > |I|$.

Справедливость строгого неравенства (14) следует из (8) при $t = \hat{t}$.

Таким образом, от плана x при $x' \notin \Omega$ мы перешли к новому плану $\hat{x}$, такому, что $|I(\hat{x})| > |I(x)|$ и $Q(\hat{x}) < Q(x)$.

Теперь в качестве x возьмём $\hat{x}$ и повторим рассуждения. Тогда либо $x' \in \Omega$, что позволяет в качестве I_0 взять $I(x)$, либо $x' \notin \Omega$. Во втором случае от x мы перейдём к новому плану $\hat{x}$, у которого $|I(\hat{x})| > |I(x)|$ и $Q(\hat{x}) < Q(x)$. Далее процесс повторяется. В силу постоянного расширения множеств $I(\hat{x})$ и строгого убывания целевой функции $Q(x)$ описанная процедура конечна. За конечное число шагов мы придём к начальному подходящему множеству I_0 и соответствующей подходящей паре (x_0, u_0) с $x_0 \in \Omega$.

6°. Если $u_0[I_0] \geqslant \mathbb{O}$, то согласно предложению 1 вектор x_0 является решением задачи (1). Процесс закончен.

Предположим, что

$$u_0[i'] < 0 \quad \text{при некотором} \quad i' \in I_0. \tag{16}$$

Расширим при необходимости множество I_0, включив в него те индексы i из $M \setminus I_0$, на которых

$$A[i, N] \times x_0[N] = b[i].$$

Это действие обеспечивает равенство $I_0 = I(x_0)$. Положив $u_0[i] = 0$ на вновь введённых индексах i, сохраним соотношения

$$\begin{aligned} D[N,N] \times x_0[N] - A^T[N, I_0] \times u_0[I_0] &= -c[N], \\ -A[I_0, N] \times x_0[N] \qquad\qquad\qquad\qquad &= -b[I_0]. \end{aligned} \tag{17}$$

Введём индексное множество $I_0' = I_0 \setminus \{i'\}$. В силу условия невырожденности строки матрицы $A[I_0', N]$ линейно независимы. Решим задачу (2) при $I = I_0'$. Решение обозначим (x_0', u_0'). Таким образом,

$$\begin{aligned} D[N,N] \times x_0'[N] - A^T[N, I_0'] \times u_0'[I_0'] &= -c[N], \\ -A[I_0', N] \times x_0'[N] \qquad\qquad\qquad\qquad &= -b[I_0']. \end{aligned} \tag{18}$$

ПРЕДЛОЖЕНИЕ 3. *Вектор x_0' отличен от x_0. При этом $Q(x_0') < Q(x_0)$ и*

$$A[i', N] \times x_0'[N] > b[i']. \tag{19}$$

Доказательство. Допустив, что $x_0' = x_0$, на основании (17), (18) получим

$$A^T[N, I_0] \times u_0[I_0] = A^T[N, I_0'] \times u_0'[I_0']$$

или
$$A^T[N,I_0'] \times \big(u_0[I_0'] - u_0'[I_0']\big) + A^T[N,i'] \times u_0[i'] = \mathbb{O}.$$
В силу (16) это противоречит линейной независимости строк матрицы $A[I_0,N]$. Значит $x_0' \neq x_0$.

Обозначим $s_0 = x_0' - x_0$. Формулы (17), (18) приводят к равенствам
$$D[N,N] \times s_0[N] - A^T[N,I_0'] \times \big(u_0[I_0'] - u_0'[I_0']\big) + A^T[N,i'] \times u_0[i'] = \mathbb{O}, \quad (20)$$
$$A[I_0',N] \times s_0[N] = \mathbb{O}. \quad (21)$$

Умножим (20) слева на s_0. На основании (21) получим
$$\langle Ds_0, s_0\rangle = -u_0[i'] \times \big(A[i',N] \times s_0[N]\big).$$
Отсюда, с учётом (16) и условия $s_0 \neq \mathbb{O}$, следует неравенство
$$A[i',N] \times s_0[N] > 0,$$
равносильное (19).

Как отмечалось после доказательства предложения 2,
$$Q(x_0') = Q(x_0) - \tfrac{1}{2}\langle Ds_0, s_0\rangle.$$
В частности, $Q(x_0') < Q(x_0)$.

Справедливость всех утверждений предложения 3 установлена. □

7°. Вернёмся к множеству I_0' и соответствующему решению (x_0', u_0') системы (2) при $I_0 = I_0'$. Если $x_0' \in \Omega$, то I_0' — очередное подходящее множество. Обозначим $I_1 = I_0'$, $x_1 = x_0'$, $u_1 = u_0'$. Как доказано в предложении 3, $Q(x_1) < Q(x_0)$.

Предположим, что $x_0' \notin \Omega$. В этом случае существует индекс $i \in M \setminus I_0'$, на котором
$$A[i,N] \times x_0'[N] < b[i].$$
Множество всех таких индексов обозначим J_0'. Отметим, что $i' \in M \setminus I_0'$. Более того, согласно (19)
$$i' \in (M \setminus I_0') \setminus J_0'.$$

На интервале (x_0, x_0') найдём точку $\hat{x}_0 \in \Omega$, такую, что $Q(\hat{x}_0) < Q(x_0)$. Для этого так же, как в п. 5°, запишем параметрическое представление интервала (x_0, x_0')
$$x_0(t) = tx_0' + (1-t)x_0, \quad t \in (0,1),$$
и воспользуемся формулой
$$\Delta\big(x_0(t)\big) = t\Delta(x_0') + (1-t)\Delta(x_0).$$

При всех $t \in (0,1)$ имеем

$$\Big(\Delta\big(x_0(t)\big)\Big)[i] = 0, \quad i \in I_0'. \tag{22}$$

Далее, возьмём $i \in (M \setminus I_0') \setminus J_0'$. Если $i = i'$, то в силу (19) при всех $t \in (0,1)$ получим

$$\Big(\Delta\big(x_0(t)\big)\Big)[i'] = t\big(\Delta(x_0')\big)[i'] > 0.$$

Пусть $i \neq i'$. Тогда $i \in M \setminus I_0$ и $i \notin J_0'$. Эти условия обеспечивают неравенства

$$\big(\Delta(x_0)\big)[i] > 0, \quad \big(\Delta(x_0')\big)[i] \geqslant 0.$$

Отсюда следует, что $\Big(\Delta\big(x_0(t)\big)\Big)[i] > 0$ при всех $t \in (0,1)$. Таким образом, при всех $t \in (0,1)$

$$\Big(\Delta\big(x_0(t)\big)\Big)[i] > 0, \quad i \in (M \setminus I_0') \setminus J_0'. \tag{23}$$

Остаётся разобраться с индексами из J_0'. Для них неравенство

$$\Big(\Delta\big(x_0(t)\big)\Big)[i] := \big(\Delta(x_0)\big)[i] - t\Big(\big(\Delta(x_0)\big)[i] - \big(\Delta(x_0')\big)[i]\Big) \geqslant 0$$

равносильно следующему

$$t \leqslant \frac{\big(\Delta(x_0)\big)[i]}{\big(\Delta(x_0)\big)[i] - \big(\Delta(x_0')\big)[i]}. \tag{24}$$

Индекс i из J_0' принадлежит $M \setminus I_0'$, но отличен от i', так что $i \in M \setminus I_0$. Значит, числитель дроби из правой части (24) положителен. По определению J_0' положителен и знаменатель. При этом значение самой дроби лежит в интервале $(0,1)$.

Положим

$$\hat{t}_0 = \min_{i \in J_0'} \Big\{ \frac{\big(\Delta(x_0)\big)[i]}{\big(\Delta(x_0)\big)[i] - \big(\Delta(x_0')\big)[i]} \Big\}. \tag{25}$$

Очевидно, что $\hat{t}_0 \in (0,1)$. Возьмём на интервале (x_0, x_0') точку $\hat{x}_0 = x_0(\hat{t}_0)$. По построению она принадлежит Ω. Согласно (22) и (23) множество индексов активных ограничений $\hat{I}_0 = I(\hat{x}_0)$ состоит из I_0' и тех индексов из J_0', на которых достигается минимум в правой части (25). По крайней мере, $|\hat{I}_0| > |I_0'|$. На основании предложения 2 (при $x = x_0$ и $I = I_0'$) выполняется строгое неравенство $Q(\hat{x}_0) < Q(x_0)$.

Теперь положим $x = \hat{x}_0$ и запустим процедуру построения подходящего множества из п. 5°. За конечное число шагов получим подходящее множество I_1 с подходящей парой (x_1, u_1), причем гарантируется, что $Q(x_1) < Q(x_0)$.

Аналогичным образом описывается переход от подходящего множества I_k к подходящему множеству I_{k+1} со строгим уменьшением целевой функции. В силу конечности системы подходящих множеств и строгого убывания целевой функции такой процесс конечен. На выходе получим подходящее множество I_* с подходящей парой (x_*, u_*), такой, что $u_*[I_*] \geqslant 0$. Согласно предложению 1, x_* — решение задачи (1).

8°. Решение системы (2) можно заменить решением экстремальной задачи

$$\begin{gathered} Q(x) \to \min, \\ A[I, N] \times x[N] = b[I]. \end{gathered}$$

Это позволяет использовать эффективные методы оптимизации (например, метод сопряжённых градиентов [4, с. 173–191]).

9°. Метод перебора граней без условия невырожденности анализируется в книге [5, гл. 3].

ЛИТЕРАТУРА

1. Малозёмов В. Н. *Теорема существования решения для задачи квадратичного программирования* // Семинар «DHA & CAGD». Избранные доклады. 22 января 2011 г. (`http://dha.spb.ru/reps11.shtml#0122`)

2. Малозёмов В. Н., Монако М. Ф., Петров А. В. *Формулы Фробениуса, Шермана-Моррисона и близкие вопросы.* Журн. вычисл. мат. и матем. физ. 2002. Т. 42. № 10. С. 1459–1465.

3. Малозёмов В. Н. *Теорема Куна-Таккера в дифференциальной форме* // Семинар «DHA & CAGD». Избранные доклады. 27 февраля 2010 г. (`http://dha.spb.ru/reps10.shtml#0227`)

4. Пшеничный Б. Н., Данилин Ю. М. *Численные методы в экстремальных задачах.* М.: Наука, 1975. 320 с.

5. Даугавет В. А. *Численные методы квадратичного программирования.* СПб.: Изд-во СПбГУ, 2004. 128 с.