ПРОЕКТИРОВАНИЕ ТОЧКИ НА ПОДПРОСТРАНСТВО И НА СТАНДАРТНЫЙ СИМПЛЕКС*

В. Н. Малозёмов

malv@math.spbu.ru

28 февраля 2013 г.

В докладе на двух примерах показывается, чем различаются классические и неклассические задачи оптимизации.

1°. Под классической задачей оптимизации понимается задача минимизации гладкой функции при наличии ограничений-равенств. Простейшим примером такой задачи является задача ортогонального проектирования точки на подпространство. Она формализуется следующим образом [1, с. 52–55]:

$$\begin{aligned} &Q(x) := \tfrac{1}{2}\|x-c\|^2 \to \min, \\ &A[M,N] \times x[N] = \mathbb{O}[M]. \end{aligned} \tag{1}$$

Здесь $c \in \mathbb{R}^N$ — фиксированная точка и $A = A[M,N]$ — матрица с линейно независимыми строками. Ограничения задачи (1) определяют подпространство L пространства $\mathbb{R}^N$. Речь идёт о нахождении точки $x_* \in L$, ближайшей к точке c, или, другими словами, об ортогональном проектировании точки c на подпространство L.

Целевая функция задачи (1) представляет собой выпуклую квадратичную функцию, что следует из представления

$$Q(x) = \tfrac{1}{2}\langle x-c, x-c\rangle = \tfrac{1}{2}\langle Ex, x\rangle - \langle c, x\rangle + \tfrac{1}{2}\|c\|^2.$$

Отметим также, что $Q'(x) = x - c$. Запишем критерий оптимальности для задачи (1) [2, с. 90–91]: для того чтобы план $x_* \in L$ был оптимальным, необходимо и достаточно, чтобы нашёлся вектор $u_* = u_*[M]$ со свойством

$$x_* - c = A^T u_*.$$

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

Таким образом, решение задачи (1) сводится к решению системы линейных уравнений

$$A^T u = x - c, \tag{2}$$

$$Ax = \mathbb{O}. \tag{3}$$

Эта система (относительно x и u) имеет единственное решение. Чтобы найти его, умножим уравнение (2) слева на матрицу A. Учитывая (3), получаем

$$(AA^T)u = -Ac. \tag{4}$$

Покажем, что квадратная матрица AA^T обратима. Для этого достаточно проверить, что однородная система $(AA^T)u = \mathbb{O}$ имеет только нулевое решение. Возьмём произвольное решение u_0 этой системы. Запишем

$$\mathbb{O} = \langle AA^T u_0, u_0 \rangle = \langle A^T u_0, A^T u_0 \rangle = \|A^T u_0\|^2.$$

Значит, $A^T u_0 = \mathbb{O}$. По условию строки матрицы A линейно независимы. Это гарантирует линейную независимость столбцов матрицы A^T. Из линейной независимости и условия $A^T u_0 = \mathbb{O}$ следует, что $u_0 = \mathbb{O}$. Тем самым, установлено, что матрица AA^T имеет обратную.

Умножим уравнение (4) слева на матрицу $(AA^T)^{-1}$. Получим

$$u_* = -(AA^T)^{-1} Ac.$$

Подставив это в (2), придём к решению задачи (1):

$$x_* = c + A^T u_* = c - A^T (AA^T)^{-1} Ac. \tag{5}$$

Обозначим

$$P = E - A^T (AA^T)^{-1} A. \tag{6}$$

Тогда формулу (5) можно переписать так:

$$x_* = Pc. \tag{7}$$

Матрица P называется *матрицей ортогонального проектирования на подпространство L.*

Подведём итог.

ТЕОРЕМА 1. *Единственное решение задачи ортогонального проектирования произвольной точки $c \in \mathbb{R}^N$ на подпространство L определяется формулой* (7)*, в которой P — матрица ортогонального проектирования вида* (6)*.*

В частном случае, когда подпространство L задаётся одним уравнением $\langle a, x \rangle = 0$, где a — единичная вектор-строка, $\|a\| = 1$, матрица P принимает вид

$$P = E - a^T a.$$

2°. Остановимся на основных свойствах матрицы ортогонального проектирования.

ТЕОРЕМА 2. *Матрица P вида* (6) *обладает следующими свойствами:*

1) $PA^T = 0$, $PP = P$;

2) *матрица P симметрична и неотрицательно определена;*

3) *ранг матрицы P равен $|N| - |M|$.*

Доказательство. Согласно (6) имеем

$$PA^T = A^T - A^T(AA^T)^{-1}AA^T = 0.$$

Учитывая это равенство, получаем также

$$PP = P(E - A^T(AA^T)^{-1}A) = P - (PA^T)(AA^T)^{-1}A = P.$$

Симметричность матрицы P следует из её определения. Неотрицательная определённость проверяется так:

$$\langle Px, x\rangle = \langle PPx, x\rangle = \langle Px, Px\rangle = \|Px\|^2 \geqslant 0.$$

Чтобы найти ранг матрицы P, введём подпространство

$$\mathcal{P} = \{x \in \mathbb{R}^N \mid Px = \mathbb{O}\}.$$

Как известно (см., например, [1, с. 22]), для размерности этого подпространства справедлива формула.

$$\dim \mathcal{P} = |N| - \operatorname{rank} P. \tag{8}$$

Согласно свойству 1) столбцы матрицы A^T принадлежат $\mathcal{P}$ и по условию задачи (1) они линейно независимы. Покажем, что любой вектор из $\mathcal{P}$ можно представить в виде линейной комбинации столбцов матрицы A^T.

Пусть $x_0 \in \mathcal{P}$. Согласно (6) имеем

$$\mathbb{O} = Px_0 = x_0 - A^T(AA^T)^{-1}Ax_0.$$

Обозначим $\lambda_0 = (AA^T)^{-1}Ax_0$. Тогда $x_0 = A^T\lambda_0$. Это и означает, что x_0 есть линейная комбинация столбцов матрицы A^T.

Установлено, что $|M|$ линейно независимых столбцов матрицы A^T являются базисом подпространства $\mathcal{P}$. По определению размерности линейного множества получаем $\dim \mathcal{P} = |M|$.

Теперь свойство 3) следует из формулы (8). □

3°. К неклассическим относятся экстремальные задачи, у которых в ограничениях присутствуют неравенства. В качестве примера рассмотрим задачу проектирования точки на стандартный симплекс. Задача формализуется так (см. [3]):

$$\begin{gathered} Q(x) := \tfrac{1}{2}\|x-c\|^2 \to \min, \\ \sum_{j=1}^{n} x_j = 1; \\ x_j \geqslant 0,\ j \in 1:n. \end{gathered} \tag{9}$$

В данном случае $N = 1:n$. Матрица ограничений задачи (9) имеет вид

$$A = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Ограничению-равенству сопоставим двойственную переменную λ, а ограничению-неравенству $x_j \geqslant 0$ — двойственную переменную u_j, $j \in 1:n$. Запишем критерий оптимальности [2, с. 91]: для того чтобы план x^* задачи (9) был оптимальным, необходимо и достаточно, чтобы нашлись числа λ^*, $u_1^*, \dots, u_n^*$, такие, что при всех $j \in 1:n$ выполняются соотношения (условия Куна-Таккера):

$$\begin{gathered} x_j^* - c_j = \lambda^* + u_j^*; \\ u_j^* x_j^* = 0,\ u_j^* \geqslant 0. \end{gathered}$$

Таким образом, решение задачи (9) сводится к решению следующей системы равенств и неравенств

$$\lambda + c_j = x_j - u_j,\ j \in 1:n; \tag{10}$$

$$u_j x_j = 0,\ u_j \geqslant 0,\ x_j \geqslant 0,\ j \in 1:n; \tag{11}$$

$$\sum_{j=1}^{n} x_j = 1. \tag{12}$$

Согласно (10) и (11) имеем

$$|\lambda + c_j| = |x_j - u_j| = x_j + u_j, \quad j \in 1:n. \tag{13}$$

Последнее равенство очевидно при $u_j = 0$. В силу условия дополнительности $u_j x_j = 0$ оно выполняется и при $u_j > 0$.

Сложим равенства (10) и (13). Получим

$$x_j = \tfrac{1}{2}\big(\lambda + c_j + |\lambda + c_j|\big). \tag{14}$$

С помощью функции $t_+ = \frac{1}{2}\big(t + |t|\big)$ формулу (14) можно переписать в виде

$$x_j = (\lambda + c_j)_+, \quad j \in 1:n. \tag{15}$$

По определению

$$t_+ = \begin{cases} 0 & \text{при } t \leqslant 0, \\ t & \text{при } t \geqslant 0. \end{cases} \tag{16}$$

Значит,

$$x_j = \begin{cases} 0, & \text{если } \lambda + c_j < 0; \\ \lambda + c_j, & \text{если } \lambda + c_j \geqslant 0. \end{cases}$$

Ясно также, что

$$u_j = \tfrac{1}{2}\big(|\lambda + c_j| - (\lambda + c_j)\big), \quad j \in 1:n.$$

Найденные x_j и u_j удовлетворяют условиям (10) и (11) при всех λ. Остаётся уравнение (12), которое в силу (15) принимает вид

$$\varphi(\lambda) := \sum_{j=1}^{n} (\lambda + c_j)_+ = 1. \tag{17}$$

Покажем, что это уравнение имеет единственное решение.

Поменяем знаки у компонент c_j проектируемой точки c и числа $\{-c_j\}$ упорядочим по неубыванию. Получим последовательность $a_1 \leqslant a_2 \leqslant \ldots \leqslant a_n$. Очевидно, что

$$\varphi(\lambda) = \sum_{j=1}^{n} \big(\lambda - (-c_j)\big)_+ = \sum_{j=1}^{n} (\lambda - a_j)_+. \tag{18}$$

Согласно (16) при всех $j \in 1:n$ имеем

$$(\lambda - a_j)_+ = 0, \quad \text{если } \lambda \leqslant a_1,$$
$$(\lambda - a_j)_+ = \lambda - a_j, \quad \text{если } \lambda \geqslant a_n,$$

поэтому

$$\begin{gathered} \varphi(\lambda) = 0 \quad \text{при } \lambda \leqslant a_1, \\ \varphi(\lambda) = \sum_{j=1}^{n} (\lambda - a_j) = n\lambda - \sum_{j=1}^{n} a_j \quad \text{при } \lambda \geqslant a_n. \end{gathered} \tag{19}$$

По тем же соображениям при $\lambda \in [a_k, a_{k+1}]$, $k \in 1 : n-1$,

$$\varphi(\lambda) = \sum_{j=1}^{k} (\lambda - a_j) = k\lambda - \sum_{j=1}^{k} a_j. \tag{20}$$

Отметим, что

$$\varphi(\lambda) = k(\lambda - a_k) + \varphi(a_k) \tag{21}$$

при $\lambda \in [a_k, a_{k+1}]$, $k \in 1 : n-1$, и при $\lambda \geqslant a_n$, $k = n$. В частности,

$$\varphi(a_{k+1}) = k(a_{k+1} - a_k) + \varphi(a_k), \quad k \in 1 : n-1. \tag{22}$$

На основании (18)–(20) заключаем, что функция $\varphi(\lambda)$ при $\lambda \geqslant a_1$ является непрерывной ломаной, строго возрастающей от 0 до $+\infty$ (см. рис.).

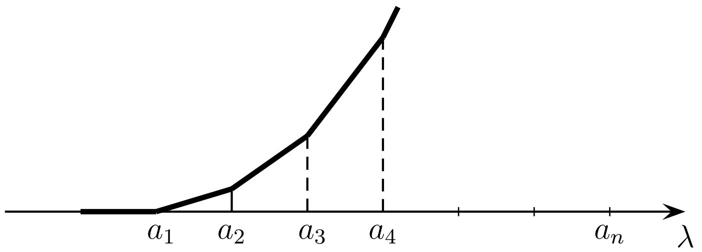

Рис. График ломаной $\varphi(\lambda)$

Это гарантирует существование и единственность точки λ^*, в которой $\varphi(\lambda^*) = 1$.

Установлено, что уравнение (17) имеет единственное решение λ^*. Отсюда следует, что система соотношений (10)–(12) имеет единственное решение и, наконец, что задача (9) имеет единственное решение — вектор x^* с компонентами

$$x_j^* = (\lambda^* + c_j)_+, \quad j \in 1 : n. \tag{23}$$

4°. Опишем простой алгоритм вычисления λ^*.

Ломаная $\varphi(\lambda)$ на отрезке $[a_1, a_n]$ полностью определяется своими значениями $\varphi_k = \varphi(a_k)$ в узлах $\lambda = a_k$, $k \in 1 : n$. Согласно (22) последовательность $\{\varphi_k\}$ можно вычислить рекуррентно:

$$\begin{aligned} \varphi_1 &= 0; \\ \varphi_{k+1} &= \varphi_k + k(a_{k+1} - a_k),\ k = 1, \ldots, n-1. \end{aligned} \tag{24}$$

Будем последовательно вычислять значения $\varphi_1, \varphi_2, \ldots$ по формуле (24), пока не встретим индекс k_0, на котором

$$\varphi_{k_0} < 1 \leqslant \varphi_{k_0+1}.$$

Если и $\varphi_n < 1$, то полагаем $k_0 = n$. На отрезке $[a_{k_0}, a_{k_0+1}]$ в случае $k_0 < n$ и полуоси $[a_{k_0}, +\infty)$ при $k_0 = n$ функция $\varphi(\lambda)$ имеет вид (21) (с заменой k на k_0). Она линейна. Решение уравнения $\varphi(\lambda) = 1$ находим элементарно:

$$\lambda^* = a_{k_0} + \tfrac{1}{k_0}(1 - \varphi_{k_0}). \tag{25}$$

5°. Подведём итог.

АЛГОРИТМ РЕШЕНИЯ ЗАДАЧИ (9)**.**

1) Меняем знаки у компонент c_j проектируемой точки c и числа $\{-c_j\}$ упорядочиваем по неубыванию. Получаем последовательность $a_1 \leqslant a_2 \leqslant \leqslant \ldots \leqslant a_n$.

2) Проводим последовательные вычисления по рекуррентной формуле (24), пока не встретим индекс k_0, на котором $\varphi_{k_0} < 1 \leqslant \varphi_{k_0+1}$. Если и $\varphi_n < 1$, то полагаем $k_0 = n$.

3) Вычисляем λ^* по формуле (25) и компоненты x_j^* проекции точки c на стандартный симплекс по формуле (23).

ПРИМЕР. Найдём проекцию точки

$$c = (1,\ -1,\ 0,\ 1,\ 0,\ \tfrac{2}{3})$$

на стандартный симплекс.

Упорядочим компоненты точки $-c = (-1,\ 1,\ 0,\ -1,\ 0,\ -\frac{2}{3})$ по неубыванию. Получим последовательность $a = (-1,\ -1,\ -\frac{2}{3},\ 0,\ 0,\ 1)$. Составим таблицу

k	1	2	3	4	5	6
a_k	-1	-1	$-\frac{2}{3}$	0	0	1

Проведём вычисления по формуле (24):

$$\begin{aligned}
&\varphi_1 = 0; \quad \varphi_2 = 0 + 1\big(-1 - (-1)\big) = 0;\\
&\varphi_3 = 0 + 2\big(-\tfrac{2}{3} - (-1)\big) = \tfrac{2}{3};\\
&\varphi_4 = \tfrac{2}{3} + 3\big(0 - (-\tfrac{2}{3})\big) = \tfrac{8}{3} > 1.
\end{aligned}$$

Получаем $\varphi_3 < 1 \leqslant \varphi_4$, так что $k_0 = 3$.

Согласно (25)

$$\lambda^* = -\tfrac{2}{3} + \tfrac{1}{3}(1 - \tfrac{2}{3}) = -\tfrac{5}{9}.$$

По формуле (23) находим компоненты проекции x^* точки c на стандартный симплекс:

$$x_1^* = (-\tfrac{5}{9} + 1)_+ = \tfrac{4}{9}; \quad x_2^* = (-\tfrac{5}{9} - 1)_+ = 0; \quad x_3^* = (-\tfrac{5}{9} + 0)_+ = 0;$$
$$x_4^* = (-\tfrac{5}{9} + 1)_+ = \tfrac{4}{9}; \quad x_5^* = (-\tfrac{5}{9} + 0)_+ = 0; \quad x_6^* = (-\tfrac{5}{9} + \tfrac{2}{3})_+ = \tfrac{1}{9}.$$

Таким образом,

$$x^* = (\tfrac{4}{9},\ 0,\ 0,\ \tfrac{4}{9},\ 0,\ \tfrac{1}{9}).$$

6°. В заключение отметим, что решение классической задачи минимизации (1) находится с помощью *формулы* (7), а решение неклассической задачи минимизации (9) — с помощью *алгоритма.*

ЛИТЕРАТУРА

1. Малозёмов В. Н. *Линейная алгебра без определителей. Квадратичная функция.* СПб.: Изд-во СПбГУ, 1997. 80 с.

2. Гавурин М. К., Малозёмов В. Н. *Экстремальные задачи с линейными ограничениями.* Л.: Изд-во ЛГУ, 1984. 176 с.

3. Малозёмов В. Н., Певный А. Б. *Быстрый алгоритм проектирования точки на симплекс* // Вестник СПбГУ. Сер. 1. 1992. Вып. 1 (№ 1). С. 112–113.