МЕТОД СОПРЯЖЁННЫХ ГРАДИЕНТОВ В КВАДРАТИЧНОМ ПРОГРАММИРОВАНИИ*

В. Н. Малозёмов Е. К. Чернэуцану
malv@math.spbu.ru katerinache@yandex.ru

26 мая 2012 г.

Памяти Б. Н. Пшеничного
(1937–2000)

Данный доклад является продолжением доклада [1]. Здесь анализируется метод сопряжённых градиентов минимизации выпуклой квадратичной функции при наличии линейных ограничений. В идейном плане мы следуем Б. Н. Пшеничному [2, с. 173–191].

1°. Напомним описание метода сопряжённых градиентов решения следующей экстремальной задачи

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

где D — симметричная неотрицательно определённая матрица (см. [1]).

Нулевой шаг. Берём произвольное начальное приближение $x_0 \in \mathbb{R}^n$ и вычисляем градиент $g_0 = Dx_0 + c$. Если $g_0 = \mathbb{O}$, то x_0 — решение задачи (1). Вычисления прекращаются. Иначе полагаем $s_1 = -g_0$.

***k*-й шаг.** Пусть уже имеются x_{k-1}, $g_{k-1} \neq \mathbb{O}$, s_k. Последовательно вычисляем

$$\begin{aligned} t_k &= \frac{\langle g_{k-1}, g_{k-1}\rangle}{\langle Ds_k, s_k\rangle}, \\ x_k &= x_{k-1} + t_k s_k, \\ g_k &= g_{k-1} + t_k Ds_k. \end{aligned}$$

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

Если $g_k = \mathbb{O}$, то x_k — решение задачи (1). Вычисления прекращаются. В противном случае находим

$$b_k = \frac{\langle g_k, g_k\rangle}{\langle g_{k-1}, g_{k-1}\rangle},$$
$$s_{k+1} = -g_k + b_k s_k.$$

Описание метода закончено.

Если квадратичная функция $Q(x)$ ограничена снизу на $\mathbb{R}^n$, то по крайней мере на r-м шаге, где $r = \operatorname{rank} D$, получим решение задачи (1). Если же $Q(x)$ неограничена снизу, то при некотором k выполнится равенство $\langle Ds_k, s_k\rangle = 0$. В этом случае $Q(x)$ на луче $x = x_{k-1}+ts_k$, $t > 0$, стремится к $-\infty$ при $t \to +\infty$.

2°. Усложним задачу (1), введя линейные ограничения-равенства:

$$\begin{gathered} Q(x) := \tfrac{1}{2}\langle Dx, x\rangle + \langle c, x\rangle \to \inf, \\ Ax = b. \end{gathered} \tag{2}$$

Считаем, что D — симметричная неотрицательно определённая матрица, а A — произвольная $(m \times n)$-матрица. Множество планов задачи (2) обозначим ω.

Напомним критерий оптимальности [3, с. 110]: *вектор $x^* \in \omega$ является решением задачи* (2) *тогда и только тогда, когда найдётся вектор $u^* \in \mathbb{R}^m$, такой, что*

$$Q'(x^*) = A^T u^*.$$

Для дальнейшего нам потребуется дополнительное предположение. Будем считать, что строки матрицы A линейно независимы. В этом случае $\operatorname{rank} A = m$.

Введём подпространство $\mathcal{L} = \{x \in \mathbb{R}^n \mid Ax = \mathbb{O}\}$ и обозначим через P матрицу ортогонального проектирования на $\mathcal{L}$. Как известно [4, с. 53],

$$P = E - A^T(AA^T)^{-1}A,$$

где E — единичная матрица порядка n. Матрица P обладает следующими свойствами:

1) $AP = 0, \quad PA^T = 0;$

2) $P^T = P, \quad PP = P;$

3) $\operatorname{rank} P = n - m.$

Сведём задачу (2) к задаче без ограничений. Возьмём произвольный план $x_0 \in \omega$ и сделаем замену переменных $x = x_0 + Py$. Рассмотрим функцию

$$q(y) := Q(x_0 + Py) = Q(x_0) + \langle Dx_0 + c, Py\rangle + \tfrac{1}{2}\langle DPy, Py\rangle = \\ = \tfrac{1}{2}\langle PDPy, y\rangle + \big\langle P(Dx_0 + c), y\big\rangle + Q(x_0).$$

Видим, что $q(y)$ — квадратичная функция с симметричной неотрицательно определённой матрицей PDP. Отметим (см. [4, с. 15]), что

$$\operatorname{rank} PDP \leqslant \min\{\operatorname{rank} P, \operatorname{rank} D\} \leqslant n - m. \tag{3}$$

Кроме того,

$$q'(y) = PDPy + P(Dx_0 + c) = P\big[D(x_0 + Py) + c\big] = P[Dx + c] = PQ'(x). \tag{4}$$

Допустим, что y^* — точка минимума $q(y)$ на $\mathbb{R}^n$. Покажем, что точка $x^* = x_0 + Py^*$ является решением задачи (2). Имеем

$$Ax^* = Ax_0 + (AP)y^* = Ax_0 = b,$$

то есть $x^* \in \omega$. Далее, из условия $q'(y^*) = \mathbb{O}$ в силу (4) следует, что $PQ'(x^*) = \mathbb{O}$ или

$$Q'(x^*) - A^T(AA^T)^{-1}AQ'(x^*) = \mathbb{O}.$$

Положив

$$u^* = (AA^T)^{-1}AQ'(x^*),$$

получим

$$Q'(x^*) = A^T u^*.$$

По критерию оптимальности, x^* — решение задачи (2).

Таким образом, задача (2) сводится к минимизации выпуклой квадратичной функции $q(y)$ на $\mathbb{R}^n$.

3°. Запишем метод сопряжённых градиентов для минимизации $q(y)$ на $\mathbb{R}^n$. Предварительно найдём матрицу ортогонального проектирования P. Напомним, что в определение $q(y)$ входит некоторый план x_0 задачи (2).

Нулевой шаг. В качестве начального приближения возьмём вектор $y_0 = \mathbb{O}$. Вычислим градиент $g_0 = P(Dx_0 + c)$. Если $g_0 = \mathbb{O}$, то y_0 — точка минимума. Вычисления прекращаются. Иначе полагаем $s_1 = -g_0$.

k-й шаг. Пусть уже имеются y_{k-1}, $g_{k-1} \neq \mathbb{O}$, s_k. Вычисляем

$$t_k = \frac{\langle g_{k-1}, g_{k-1}\rangle}{\langle PDPs_k, s_k\rangle}, \tag{5}$$

$$y_k = y_{k-1} + t_k s_k,$$

$$g_k = g_{k-1} + t_k PDPs_k. \tag{6}$$

Если $g_k = \mathbb{O}$, то y_k — точка минимума. Вычисления прекращаются. В противном случае находим

$$b_k = \frac{\langle g_k, g_k\rangle}{\langle g_{k-1}, g_{k-1}\rangle},$$
$$s_{k+1} = -g_k + b_k s_k.$$

Описание метода закончено.

На самом деле, формулы (5) и (6) допускают упрощение. Покажем, что

$$Ps_k = s_k, \quad k = 1, 2, \ldots \tag{7}$$

Имеем

$$Ps_k = s_k - A^T(AA^T)^{-1}(As_k). \tag{8}$$

По построению s_k есть линейная комбинация градиентов $g_0, g_1, \ldots, g_{k-1}$. Вместе с тем, согласно (4)

$$Ag_i = Aq'(y_i) = (AP)\big[D(x_0 + Py_i) + c\big] = 0.$$

Отсюда и из (8) следует (7).

Теперь формулы (5) и (6) принимают более простой вид:

$$t_k = \frac{\langle g_{k-1}, g_{k-1}\rangle}{\langle Ds_k, s_k\rangle},$$
$$g_k = g_{k-1} + t_k PDs_k.$$

4°. Обозначим $x_k = x_0 + Py_k$. Очевидно, что $x_k \in \omega$ и согласно (7)

$$x_k - x_{k-1} = P(y_k - y_{k-1}) = t_k s_k.$$

Перепишем алгоритм предыдущего пункта в терминах x_k.

Предварительно находим матрицу ортогонального проектирования $P = E - A^T(AA^T)^{-1}A$.

Нулевой шаг. Берём произвольное начальное приближение $x_0 \in \omega$ и вычисляем $g_0 = P(Dx_0+c)$. Если $g_0 = \mathbb{O}$, то x_0 — решение задачи (2). Вычисления прекращаются. Иначе полагаем $s_1 = -g_0$.

***k*-й шаг.** Пусть уже имеются x_{k-1}, $g_{k-1} \neq \mathbb{O}$, s_k. Последовательно вычисляем

$$t_k = \frac{\langle g_{k-1}, g_{k-1}\rangle}{\langle Ds_k, s_k\rangle},$$
$$x_k = x_{k-1} + t_k s_k,$$
$$g_k = g_{k-1} + t_k PDs_k.$$

Если $g_k = \mathbb{O}$, то x_k — решение задачи (2). Вычисления прекращаются. В противном случае находим

$$\begin{aligned} b_k &= \frac{\langle g_k, g_k\rangle}{\langle g_{k-1}, g_{k-1}\rangle}, \\ s_{k+1} &= -g_k + b_k s_k. \end{aligned}$$

Описание алгоритма решения задачи (2) закончено.

Если квадратичная функция $Q(x)$ ограничена снизу на ω, то квадратичная функция $q(y)$ ограничена снизу на $\mathbb{R}^n$. В силу (3), по крайней мере на $(n-m)$-м шаге данного алгоритма получим решение задачи (2). Из неограниченности $Q(x)$ снизу на ω следует неограниченность снизу на $\mathbb{R}^n$ функции $q(y)$ (иначе задача (2) имела бы решение). В этом случае при некотором k выполнится равенство $\langle Ds_k, s_k\rangle = 0$, которое гарантирует убывание $Q(x)$ на луче $x = x_{k-1} + ts_k$, $t > 0$, до $-\infty$ при $t \to +\infty$.

5°. В предыдущем пункте мы описали метод сопряжённых градиентов решения задачи (2). Он требует выбора начального приближения $x_0 \in \omega$, то есть решения системы линейных уравнений $Ax = b$. Покажем, что решение такой системы при любой $(m \times n)$-матрице A можно получить с помощью метода сопряжённых градиентов.

Рассмотрим задачу квадратичного программирования

$$\varphi(x) := \tfrac{1}{2}\|Ax - b\|^2 \to \min_{x\in\mathbb{R}^n}. \tag{9}$$

Запишем квадратичную функцию $\varphi(x)$ в стандартной форме:

$$\varphi(x) = \tfrac{1}{2}\langle Ax - b, Ax - b\rangle = \tfrac{1}{2}\langle A^T Ax, x\rangle - \langle A^T b, x\rangle + \tfrac{1}{2}\|b\|^2.$$

Матрица $A^T A$ симметрична и неотрицательно определена. Кроме того, $\varphi(x) \geqslant 0$ при всех $x \in \mathbb{R}^n$. Значит, задача (9) имеет решение. Оно может быть получено методом сопряжённых градиентов. Для этого потребуется не более m шагов, поскольку $\operatorname{rank} A^T A \leqslant m$.

Пусть x_0 — решение задачи (9). Если $\varphi(x_0) = 0$, то $x_0 \in \omega$. В противном случае система $Ax = b$ решений не имеет ($\omega = \varnothing$).

6°. Переходим к общей задаче квадратичного программирования

$$\begin{aligned} Q(x) &:= \tfrac{1}{2}\langle Dx, x\rangle + \langle c, x\rangle \to \inf, \\ \langle a_j, x\rangle &\geqslant b_j, \quad j \in M_1; \\ \langle a_j, x\rangle &= b_j, \quad j \in M_2. \end{aligned} \tag{10}$$

Считаем, что матрица D симметрична и неотрицательно определена. Множество планов задачи (10) обозначим Ω.

При фиксированном $x \in \Omega$ введём индексные множества активных ограничений

$$M_1(x) = \big\{ j \in M_1 \mid \langle a_j, x\rangle = b_j \big\},$$
$$J(x) = M_1(x) \cup M_2.$$

Напомним критерий оптимальности [3, с. 88]: *вектор $x^* \in \Omega$ является решением задачи* (10) *тогда и только тогда, когда найдутся множители Лагранжа u_j^*, $j \in J(x^*)$, со свойствами*

$$Q'(x^*) = \sum_{j \in J(x^*)} u_j^* a_j,$$
$$u_j^* \geqslant 0, \quad j \in M_1(x^*).$$

В дальнейшем будем предполагать, что выполнено условие регулярности ограничений: *при всех $x \in \Omega$ векторы a_j, $j \in J(x)$, линейно независимы.*

7°. Опишем метод решения задачи (10), основанный на переборе активных ограничений.

В качестве начального приближения возьмём произвольный вектор $x_0 \in \Omega$. Обозначим $J_0 = J(x_0)$, так что

$$\langle a_j, x_0\rangle > b_j \quad \text{при} \quad j \notin J_0. \tag{11}$$

Из строк a_j, $j \in J_0$, составим матрицу A_{J_0}. Найдём матрицу ортогонального проектирования

$$P_{J_0} = E - A_{J_0}^T (A_{J_0} A_{J_0}^T)^{-1} A_{J_0}$$

и вектор

$$u^0 = (A_{J_0} A_{J_0}^T)^{-1} A_{J_0} Q'(x_0).$$

В этом случае

$$P_{J_0} Q'(x_0) = Q'(x_0) - A_{J_0}^T u^0.$$

Имеются две возможности.

I) $P_{J_0} Q'(x_0) = \mathbb{O}$. По критерию оптимальности из п. 2° получается, что x_0 доставляет минимум квадратичной функции $Q(x)$ при ограничениях-равенствах

$$\langle a_j, x\rangle = b_j \quad j \in J_0.$$

Если к тому же $u_j^0 \geqslant 0$ при $j \in M_1(x_0)$, то, по критерию оптимальности из п. 6°, x_0 — решение задачи (10). Вычисления прекращаются.

Допустим, что нашёлся индекс $j_0 \in M_1(x_0)$, на котором $u_{j_0}^0 < 0$. Это значит, что ограничение $\langle a_{j_0}, x\rangle = b_{j_0}$ препятствует уменьшению функции $Q(x)$ на Ω.

Уберём его. Введём индексное множество $J_0' = J_0 \setminus \{j_0\}$ и применим метод сопряжённых градиентов для минимизации $Q(x)$ при ограничениях

$$\langle a_j, x\rangle = b_j, \quad j \in J_0'. \tag{12}$$

В качестве начального приближения возьмём известную точку x_0. По алгоритму

$$x_1 = x_0 + t_1 s_1, \tag{13}$$

где $s_1 = -g_0$, $g_0 = P_{J_0'} Q'(x_0)$.

ПРЕДЛОЖЕНИЕ 1. *Вектор g_0 отличен от нулевого и*

$$\langle a_{j_0}, s_1\rangle = -(u_{j_0}^0)^{-1}\|g_0\|^2 > 0, \tag{14}$$

$$\langle a_j, s_1\rangle = 0 \quad \textit{при} \quad j \in J_0'. \tag{15}$$

Доказательство. Имеем

$$g_0 = Q'(x_0) - A_{J_0'}^T v^0, \quad \text{где} \quad v^0 = (A_{J_0'} A_{J_0'}^T)^{-1} A_{J_0'} Q'(x_0).$$

Вместе с тем, по условию I) $Q'(x_0) = A_{J_0}^T u^0$. Получаем

$$g_0 = A_{J_0}^T u^0 - A_{J_0'}^T v^0.$$

В правой части этого равенства стоит линейная комбинация столбцов a_j, $j \in J_0$, причём коэффициент при a_{j_0}, равный $u_{j_0}^0$, отличен от нуля. В силу условия регулярности ограничений векторы a_j, $j \in J_0$, линейно независимы. Значит, $g_0 \neq \mathbb{O}$.

Далее, перепишем условие I) в виде

$$Q'(x_0) - \sum_{j \in J_0'} u_j^0 a_j - u_{j_0}^0 a_{j_0} = \mathbb{O}.$$

Умножим это равенство скалярно на $s_1 = -g_0$. Получим

$$u_{j_0}^0 \langle a_{j_0}, s_1\rangle = \big\langle Q'(x_0), s_1\big\rangle - \sum_{j \in J_0'} u_j^0 \langle a_j, s_1\rangle. \tag{16}$$

Запишем

$$\langle a_j, s_1\rangle = -\big\langle a_j, P_{J_0'} Q'(x_0)\big\rangle = -\big\langle P_{J_0'} a_j, Q'(x_0)\big\rangle.$$

По свойству матрицы ортогонального проектирования $P_{J_0'} A_{J_0'}^T = 0$, поэтому $P_{J_0'} a_j = \mathbb{O}$ при $j \in J_0'$. При тех же j выполнится равенство $\langle a_j, s_1\rangle = 0$, что соответствует (15).

Теперь формула (16) принимает вид

$$u_{j_0}^0\langle a_{j_0}, s_1\rangle = \big\langle Q'(x_0), s_1\big\rangle. \tag{17}$$

Отметим, что

$$\begin{aligned}\big\langle Q'(x_0), s_1\big\rangle &= -\big\langle Q'(x_0), P_{J_0'}Q'(x_0)\big\rangle = -\big\langle Q'(x_0), P_{J_0'}P_{J_0'}Q'(x_0)\big\rangle = \\ &= -\big\langle P_{J_0'}Q'(x_0), P_{J_0'}Q'(x_0)\big\rangle = -\|g_0\|^2. \end{aligned}\tag{18}$$

Отсюда и из (17) следует (14).

Предложение доказано. □

8°. Вернёмся к формуле (13). Условие $g_0 \neq \mathbb{O}$ требуется при описании метода сопряжённых градиентов для перехода к точке x_1. В то же время неравенство (14) гарантирует, что в точках $x(t) = x_0 + ts_1$ при $t > 0$ ограничение с индексом j_0, активным в точке x_0, становится неактивным. Действительно,

$$\big\langle a_{j_0}, x(t)\big\rangle = \langle a_{j_0}, x_0\rangle + t\langle a_{j_0}, s_1\rangle = b_{j_0} + t\langle a_{j_0}, s_1\rangle > b_{j_0}. \tag{19}$$

Отметим, что шаг t_1 может выводить точку x_1 из Ω. Мы же хотим, чтобы все точки минимизирующей последовательности содержались в Ω. В этой связи введём ограничитель шага

$$\hat{t}_1 = \min_{j\in\hat{J}_0} \frac{\langle a_j, x_0\rangle - b_j}{-\langle a_j, s_1\rangle}, \tag{20}$$

где $\hat{J}_0 = \big\{j \in M_1 \cup M_2 \mid j \notin J_0',\ \langle a_j, s_1\rangle < 0\big\}$. (Если $\hat{J}_0 = \varnothing$, то по определению $\hat{t}_1 = +\infty$.)

ПРЕДЛОЖЕНИЕ 2. *Справедливо неравенство*

$$\hat{t}_1 > 0.$$

Д о к а з а т е л ь с т в о. Достаточно проверить, что при всех $j \in \hat{J}_0$ дробь из правой части формулы (20) положительна.

Возьмём $j \in \hat{J}_0$. Тогда $j \notin J_0'$, так что либо $j = j_0$, либо $j \notin J_0$ (см. рис.).

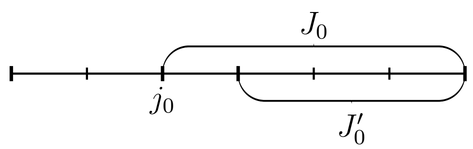

Рис.

Кроме того, $\langle a_j, s_1\rangle < 0$. Равенство $j = j_0$ невозможно в силу (14). Значит, $j \notin J_0$. Остаётся сослаться на неравенство (11). □

ПРЕДЛОЖЕНИЕ 3. *Если $t_1 < \hat{t}_1$, то точка x_1 вида* (13) *принадлежит* Ω *и*

$$J(x_1) = J_0'.$$

Доказательство. При $j \in J_0'$ согласно (15) имеем

$$\langle a_j, x_1\rangle = \langle a_j, x_0\rangle + t_1\langle a_j, s_1\rangle = \langle a_j, x_0\rangle = b_j,$$

то есть ограничения задачи (10) в точке x_1 при $j \in J_0'$ активны. Проверим, что при $j \notin J_0'$ ограничения в точке x_1 неактивны.

Возьмём $j \notin J_0'$ и предположим сначала, что $\langle a_j, s_1\rangle \geqslant 0$. При $j = j_0$ неравенство $\langle a_j, x_1\rangle > b_j$ следует из (19), а при $j \notin J_0$ — из (11).

Пусть теперь $j \notin J_0'$ и $\langle a_j, s_1\rangle < 0$, то есть $j \in \hat{J}_0$. На основании определения $\hat{t}_1$ и условия $t_1 < \hat{t}_1$ заключаем, что

$$\langle a_j, x_0\rangle - b_j \geqslant \hat{t}_1\bigl(-\langle a_j, s_1\rangle\bigr) > t_1\bigl(-\langle a_j, s_1\rangle\bigr).$$

Это неравенство приводится к виду $\langle a_j, x_1\rangle > b_j$.

Установлено, что $x_1 \in \Omega$ и $J(x_1) = J_0'$. Предложение доказано. $\square$

9°. Снова вернёмся к формуле (13). Если $t_1 < \hat{t}_1$, то согласно предложению 3 точка x_1 принадлежит Ω и $J(x_1) = J_0'$. Применение метода сопряжённых градиентов для минимизации квадратичной функции $Q(x)$ при ограничениях (12) можно продолжить. Если же $t_1 \geqslant \hat{t}_1$, то полагаем $x_1 = x_0 + \hat{t}_1 s_1$, и точку x_1 принимаем за новое начальное приближение. В последнем случае множество $J(x_1)$ состоит из J_0' и тех индексов из $\hat{J}_0$, на которых в правой части (20) достигается минимум (множество активных ограничений расширяется). При этом $Q(x_1) < Q(x_0)$, что проверяется следующим образом. Выпуклая квадратичная функция одной переменной

$$\varphi(t) := Q(x_0 + ts_1) = Q(x_0) - t\|g_0\|^2 + \tfrac{1}{2}\langle Ds_1, s_1\rangle$$

достигает глобального минимума при $t = t_1$ (мы считаем, что $\langle Ds_1, s_1\rangle > 0$) и на отрезке $[0,\, t_1]$ строго убывает. В частности, $\varphi(0) > \varphi(\hat{t}_1)$, что равносильно требуемому неравенству.

В результате описанных действий мы либо найдём точку минимума $Q(x)$ при ограничениях (12) с множеством индексов активных ограничений J_0', либо на некоторой итерации метода сопряжённых градиентов выйдем на ограничитель шага

$$\hat{t}_k = \min_{j \in \hat{J}_{k-1}} \frac{\langle a_j, x_{k-1}\rangle - b_j}{-\langle a_j, s_k\rangle}, \tag{21}$$

где $\hat{J}_{k-1} = \bigl\{j \in M_1 \cup M_2 \mid j \notin J_0',\, \langle a_j, s_k\rangle < 0\bigr\}$. При $t_k \geqslant \hat{t}_k$ положим $x_k = x_{k-1} + \hat{t}_k s_k$. В обоих случаях полученную точку принимаем за очередное начальное приближение и работаем с ней как с x_0.

10°. До сих пор предполагалось, что выполнено условие I): $P_{J_0}Q'(x_0) = \mathbb{O}$. Рассмотрим вторую возможность.

II) $P_{J_0}Q'(x_0) \neq \mathbb{O}$. В этом случае методом сопряжённых градиентов решаем задачу минимизации $Q(x)$ при ограничениях

$$\langle a_j, x\rangle = b_j, \quad j \in J_0, \tag{22}$$

начиная с x_0. На каждой итерации вычисляем ограничитель шага $\hat{t}_k$ по формуле (21), заменив в определении $\hat{J}_{k-1}$ индексное множество J_0' на J_0. Если $t_k \geqslant \hat{t}_k$, то полагаем $x_k = x_{k-1} + \hat{t}_k s_k$ и принимаем эту точку за новое начальное приближение. Если же мы ни разу не выйдем на ограничитель шага, то методом сопряжённых градиентов получим точку минимума $Q(x)$ при ограничениях (22), которую примем за новое начальное приближение и будем работать с ней как с x_0.

ЗАМЕЧАНИЕ 1. По ходу процесса может оказаться, что $\langle Ds_k, s_k\rangle = 0$. Тогда функция $Q(x)$ на луче $x = x_{k-1} + ts_k$, $t > 0$, стремится к $-\infty$ при $t_k \to +\infty$. Формально положим $t_k = +\infty$. Если при этом и $\hat{t}_k = +\infty$, то функция $Q(x)$ неограничена снизу на Ω.

11°. Описанным методом будет построена последовательность планов задачи (10), на которой целевая функция $Q(x)$ строго убывает. Покажем, что эта последовательность конечна.

Если минимизация на аффинном множестве идёт до конца, то соответствующее x_k будет точкой минимума $Q(x)$ при ограничениях

$$\langle a_j, x\rangle = b_j, \quad j \in J(x_k). \tag{23}$$

Выходить на ограничитель шага последовательно бесконечное число раз мы не можем, поскольку при этом расширяется множество активных ограничений. Значит, по ходу процесса мы систематически будем получать точки, в которых достигается минимум $Q(x)$ при ограничениях вида (23). Ясно, что таких точек конечное число и они не могут повторяться в силу строгой монотонности процесса.

Последний элемент построенной последовательности будет решением задачи (10). По описанию это единственное условие выхода из процесса (если попутно не выяснится, что задача (10) не имеет решения).

12°. Мы рассмотрели лишь принципиальную схему решения задачи квадратичного программирования (10). Вопроса организации вычислений мы не касались.

13°. Сделаем два заключительных замечания.

ЗАМЕЧАНИЕ 2. В методе сопряжённых градиентов приходится вычислять произведение матрицы ортогонального проектирования P_{J_0} на вектор. Покажем, что вектор вида $P_{J_0}d$ можно найти без вычисления матрицы P_{J_0}. Методом сопряжённых градиентов решим вспомогательную задачу

$$\tfrac{1}{2}\|d - A_{J_0}^T u\|^2 \to \min_{u\in\mathbb{R}^m}. \tag{24}$$

В данном случае критерий оптимальности имеет вид

$$A_{J_0}A_{J_0}^T u = A_{J_0}d,$$

так что решением задачи (24) является вектор

$$u_0 = (A_{J_0}A_{J_0}^T)^{-1}A_{J_0}d.$$

Получаем

$$P_{J_0}d = d - A_{J_0}^T(A_{J_0}A_{J_0}^T)^{-1}A_{J_0}d = d - A_{J_0}^T u_0.$$

ЗАМЕЧАНИЕ 3. Начальное приближение $x_0 \in \Omega$ при решении задачи (10) можно искать, используя идеи из линейного программирования. Например, привести ограничения задачи (10) к каноническому виду и найти начальный базисный план [5].

ЛИТЕРАТУРА

1. Малозёмов В. Н. *О методе сопряжённых градиентов* // Семинар «DHA & CAGD». Избранные доклады. 28 апреля 2012 г. (`http://dha.spb.ru/reps12.shtml#0428`)

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

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

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

5. Малозёмов В. Н. *Модифицированный симплекс-метод* // Семинар «DHA & CAGD». Избранные доклады. 20 ноября 2010 г. (`http://dha.spb.ru/reps10.shtml#1120`)