

Квадратичная функция на единичной сфере
Лукашевич Александр
МФТИ 476

1. Постановка задачи

Будем рассматривать квадратичную функцию на единичной сфере без линейной части:

$$\begin{aligned} \min_{x \in S} f(x) &= \frac{1}{2} x^T A x \\ S &= \{x \in \mathbb{R}^n \mid \|x\|^2 = 1\} \end{aligned} \quad (1.1)$$

Сразу же отметим, что в конечномерном евклидовом пространстве \mathbb{R}^n единичная сфера является компактным множеством, что говорит о том, что решение задачи существует, поскольку целевая функция непрерывна. Прежде чем приступить к вопросам об условиях экстремума запишем функцию Лагранжа:

$$L(x, \lambda) = \frac{1}{2} x^T A x + \lambda (\|x\|^2 - 1). \quad (1.2)$$

2. Условия экстремума, стационарные точки и точки локального минимума

2.1. Необходимые условия

- Необходимым условием (первого порядка) того, что точка x_0 является точкой экстремума задачи (1.1), является существование вектора множителей Лагранжа λ_0 , такого, что (x_0, λ_0) - стационарная точка функции Лагранжа (2.4):

$$\nabla L(x_0, \lambda_0) = \vec{0}. \quad (2.1)$$

В случае квадратичной функции (2.1) выражается следующий образом:

$$\nabla L(x_0, \lambda_0) = \begin{pmatrix} Ax_0 + 2\lambda_0 x_0 \\ \|x_0\|^2 - 1 \end{pmatrix} = \begin{pmatrix} \vec{0} \\ 0 \end{pmatrix}$$

Из этого условия явно видно, что выполняется условие допустимости.

- Необходимым условием (второго порядка) того, что точка x_0 является точкой экстремума задачи (1.1) является следующее неравенство:

$$\begin{aligned} (L_{xx}(x_0, \lambda_0)s, s) &= ((A + 2\lambda_0 E)s, s) \geq 0 \\ s &\in \mathbb{S} = \{s \mid (2x, s) = 0\}, \end{aligned} \quad (2.2)$$

где $2x$ суть градиент ограничения, отражающего принадлежность точки к сфере.

2.2. Достаточные условия

- Достаточными условиями второго порядка для того, чтобы точка x_0 была точкой локального минимума в задаче (1.1) является существование вектора множителей Лагранжа λ_0 такого, что выполнены следующие условия:

$$\begin{aligned} \nabla L(x_0, \lambda_0) &= \vec{0}, \\ (L_{xx}(x_0, \lambda_0)s, s) &= ((A + 2\lambda_0 E)s, s) > 0, \end{aligned} \quad (2.3)$$

где $s \in \mathbb{S}$, определенном в (2.2)

2.3. Стационарные точки и точки локального минимума

- Эти виды точек в данной задаче определяются условиями (2.1) и (2.3) соответственно. Рассмотрим этот вопрос подробнее.

– Запишем функцию Лагранжа:

$$L = \frac{1}{2} x^T A x + \lambda (\|x\|^2 - 1). \quad (2.4)$$

– Необходимое условие первого порядка:

$$\nabla L = \begin{pmatrix} Ax + 2\lambda x \\ \|x\|^2 - 1 \end{pmatrix} = \vec{0}. \quad (2.5)$$

– Вторая производная L по x :

$$L_{xx} = A + 2\lambda E. \quad (2.6)$$

– Касательное подпространство к \mathbb{S} :

$$E_t = \{s \in \mathbb{R}^n : (s, 2x) = 0\}. \quad (2.7)$$

– Условие минимума второго порядка:

Если x^* такова, что

$$\nabla L(x^*, \lambda^*) = \begin{pmatrix} Ax^* + 2\lambda^* x^* \\ \|x^*\|^2 - 1 \end{pmatrix} = \vec{0}. \quad (2.8)$$

$$\forall s \in E_t \hookrightarrow (L_{xx}(x^*, \lambda^*)s, s) > 0.$$

Тогда x^* - точка локального минимума в задаче (1.1).

Рассмотрим (2.8) подробнее. Учитывая (2.6),

$$(L_{xx}(x^*, \lambda^*)s, s) = ((A + 2\lambda^* E)s, s) = s^T A s + 2\lambda^* s^T s. \quad (2.9)$$

Рассмотрим A как матрицу некоторого оператора. Так как матрица симметрична, то все собственные значения вещественны, помимо этого, собственные векторы попарно ортогональны.

Из (4.20) следует, что

$$Ax^* = -2\lambda^* x^*, \quad (2.10)$$

То есть стационарные точки есть собственные векторы оператора A - $\{h_i\}_{i=1}^n$, соответствующие собственным значениям $\{-2\lambda_i\}_{i=1}^n$.

Поскольку при проверке (2.8) нас будут интересовать лишь $s \in E_t$, рассмотрим это множество подробнее: для фиксированной стационарной точки, E_t представляет собой ортогональное дополнение собственного вектора, соответствующего этой стационарной точке. Поскольку собственные векторы оператора A попарно ортогональны, то s можно разложить по остальным собственным векторам. Пусть стационарной точке соответствует собственный вектор с номером k . Тогда

$$s = \sum_{i=1}^{k-1} \alpha_i h_i + \sum_{i=k+1}^n \alpha_i h_i.$$

Чтобы не загромождать запись, будем считать, что $k = n$, что не умаляет общности.

Используя вышесказанное, перепишем (2.9):

$$s^T A s + 2\lambda s^T s = \left(\sum_{i=1}^{n-1} \alpha_i h_i \right)^T A \left(\sum_{i=1}^{n-1} \alpha_i h_i \right) + 2\lambda^* \left(\sum_{i=1}^{n-1} \alpha_i h_i \right)^T \left(\sum_{i=1}^{n-1} \alpha_i h_i \right). \quad (2.11)$$

Поскольку каждый h_i - собственный вектор оператора A , то

$$A h_i = -2\lambda_i h_i.$$

Итак, (2.11) принимает вид

$$\begin{aligned} & \left(\sum_{i=1}^{n-1} \alpha_i h_i \right)^T \left(\sum_{i=1}^{n-1} (-2\lambda_i) \alpha_i h_i \right) + 2\lambda^* \left(\sum_{i=1}^{n-1} \alpha_i h_i \right)^T \left(\sum_{i=1}^{n-1} \alpha_i h_i \right) = \\ & = \left(\sum_{i=1}^{n-1} \alpha_i h_i, \sum_{i=1}^{n-1} (-2\lambda_i) \alpha_i h_i \right) + 2\lambda^* \left(\sum_{i=1}^{n-1} \alpha_i h_i, \sum_{i=1}^{n-1} \alpha_i h_i \right). \end{aligned} \quad (2.12)$$

Поскольку $\{h_i\}_{i=1}^n$ попарно ортогональны, то (2.12) примет вид

$$\sum_{i=1}^{n-1} 2(-\lambda_i + \lambda^*) \alpha_i^2 (h_i, h_i). \quad (2.13)$$

В силу произвольности вектора $s \in E_t$, для того, чтобы (2.13) было положительным, каждое $(-\lambda_i + \lambda^*)$ должно быть положительно. Таким образом, $\lambda^* > \lambda_i \forall i = 1, \dots, n-1$, $-2\lambda^* < -2\lambda_i$.

Итак, точкой локального минимума будет та точка, множитель Лагранжа которой наибольший, то есть собственный вектор матрица A , с наименьшим собственным значением.

3. Выпуклая функция

Критерий выпуклости второго порядка выражается неравенством :

$$\nabla^2 f(x) = A \succ 0 \iff \forall s \in \mathbb{R}^n \rightarrow s^T A s \geq 0.$$

Факт положительной определенности можно установить с помощью критерий Сильвестра.

4. Методы

4.1. Градиентный метод

Рассмотрим градиентный метод для безусловной оптимизации. Общая схема метода выглядит так:

$$\begin{aligned} x^{k+1} &= x^k - \gamma_k p^k \\ \gamma_k &= const \\ p^k &= \nabla f(x^k) \end{aligned} \tag{4.1}$$

Данный метод сходится при следующих условиях:

- Градиент Липшицев с константой L
- $0 < \gamma < \frac{2}{L}$
- $f(x)$ ограничена снизу

В таком случае градиент стремится к нулю, а $f(x)$ монотонно убывает на последовательности x^k .

Градиент функции $f(x) = \frac{1}{2}x^T A x + b^T x + c$ Липшицев. Действительно:

$$\forall x, y \quad \|\nabla f(x) - \nabla f(y)\| = \|Ax + b - (Ay + b)\| = \|A(x - y)\| \leq \|A\| \cdot \|x - y\|$$

То есть константа Липшица есть норма оператора A . Это число, так как в конечномерных пространствах все операторы непрерывны, что эквивалентно ограниченности.

$f(x)$ ограничена снизу на S , поскольку она непрерывна. В нашем случае стоит применить более общий метод - метод проекции градиента:

$$x^{k+1} = P_S(x^k - \gamma_k \nabla f(x^k)) = P_S((E - \gamma A)x^k - \gamma b), \tag{4.2}$$

где $P_S(x) = \underset{s \in S}{\operatorname{argmin}} \|x - s\| = \frac{x}{\|x\|}$ - проекция x на S .

Поскольку оператор проектирования $P_S(x)$ обладает свойством Липшица с $L = 1$, то есть $\|P_S(x) - P_S(y)\| \leq \|x - y\|$, то условия сходимости те же, что и для безусловного метода.

4.2. Скорейший спуск

4.2.1. Немного о методе

Этот метод отличается от предыдущего тем, что шаг γ не постоянный, а выбирается в соответствии с определенным правилом, например, в соответствии с правилом одномерной минимизации, которое будет описано ниже.

Сразу запишем метод для нашей задачи - с ограничением на принадлежность сфере:

$$\begin{aligned} x^{k+1} &= P_S(x^k - \gamma_k \nabla f(x^k)) \\ \gamma_k &= \underset{\gamma \geq 0}{\operatorname{argmin}} f(P_S(x^k - \gamma \nabla f(x^k))) \end{aligned} \tag{4.3}$$

4.2.2. Выбор шага

$$\begin{aligned} x^{k+1} &= P_S(x^k - \gamma_k \nabla f(x^k)) = \frac{x^k - \gamma_k \nabla f(x^k)}{\|x^k - \gamma_k \nabla f(x^k)\|} \\ \gamma_k &= \underset{\gamma \geq 0}{\operatorname{argmin}} f(P_S(x^k - \gamma \nabla f(x^k))) = \underset{\gamma \geq 0}{\operatorname{argmin}} f\left(\frac{x^k - \gamma \nabla f(x^k)}{\|x^k - \gamma \nabla f(x^k)\|}\right). \end{aligned} \tag{4.4}$$

Учтем вид функции: $f(x) = \frac{1}{2}x^T Ax$ при расчете шага:

$$\gamma_k = \underset{\gamma \geq 0}{\operatorname{argmin}} \frac{1}{2} \left(\frac{x^k - \gamma Ax^k}{\|x^k - \gamma Ax^k\|} \right)^T A \frac{x^k - \gamma Ax^k}{\|x^k - \gamma Ax^k\|} = \quad (4.5)$$

$$= \underset{\gamma \geq 0}{\operatorname{argmin}} \frac{1}{2} \frac{(x^k)^T Ax^k - 2\gamma(x^k)^T A^2 x^k + \gamma^2(x^k)^T A^3 x^k}{\langle x^k - \gamma Ax^k, x^k - \gamma Ax^k \rangle} = \quad (4.6)$$

$$= \underset{\gamma \geq 0}{\operatorname{argmin}} \frac{1}{2} \frac{(x^k)^T Ax^k - 2\gamma(x^k)^T A^2 x^k + \gamma^2(x^k)^T A^3 x^k}{(x^k - \gamma Ax^k)^T E(x^k - \gamma Ax^k)} = \quad (4.7)$$

$$= \underset{\gamma \geq 0}{\operatorname{argmin}} \frac{1}{2} \frac{(x^k)^T Ax^k - 2\gamma(x^k)^T A^2 x^k + \gamma^2(x^k)^T A^3 x^k}{(x^k)^T x^k - 2\gamma(x^k)^T Ax^k + \gamma^2(x^k)^T A^2 x^k} = \quad (4.8)$$

$$= \underset{\gamma \geq 0}{\operatorname{argmin}} \frac{1}{2} F(x^k, \gamma). \quad (4.9)$$

Поскольку в знаменателе стоит квадрат нормы вектора $x^k - \gamma Ax^k$, то, вводя обозначения $a_i(x^k) = (x^k)^T A^i x^k$, $i = 0, \dots, 3$, знаменатель примет следующий вид:

$$a_2(x^k)\gamma^2 - 2\gamma a_1(x^k) + a_0(x^k) > 0. \quad (4.10)$$

Этот трехчлен должен быть строго больше нуля, чтобы сохранить смысл выражения.

Выясним, что нужно для того, чтобы (4.10) выполнялось. Для этого действительных корней не должно быть, а коэффициент при γ^2 должен быть больше нуля:

$$\begin{aligned} d &= (2a_1(x^k))^2 - a_2(x^k)a_0(x^k) < 0, \\ a_2(x^k) &> 0. \end{aligned} \quad (4.11)$$

Заметим, что последнее неравенство означает, что A^2 положительно определена.

Согласно необходимому условию экстремума, γ_k должно быть решением уравнения $\nabla_\gamma F(x^k, \gamma) = 0$:

$$\left(\frac{\gamma^2 a_3 - 2\gamma a_2 + a_1}{a_2 \gamma^2 - 2\gamma a_1 + a_0} \right)' = 2 \frac{\gamma^2(a_2^2 - a_3 a_1) + \gamma(a_3 a_0 - a_2 a_1) + \overset{<0, (4.11)}{(a_1^2 - a_0 a_2)}}{(a_2 \gamma^2 - 2\gamma a_1 + a_0)^2} = 0 \quad (4.12)$$

$$\gamma_k = \frac{(a_2 a_1 - a_3 a_0) \pm \sqrt{(a_2 a_1 - a_3 a_0)^2 - 4(a_2^2 - a_3 a_1)(a_1^2 - a_0 a_2)}}{2(a_2^2 - a_3 a_1)}. \quad (4.13)$$

В случае, когда $(a_2^2 - a_3 a_1) = 0$, получим одно решение

$$\gamma = \frac{a_1^2 - a_0 a_2}{a_3 a_0 - a_2 a_1}$$

Заметим, что первый множитель в (4.13) во втором слагаемом под корнем есть произведение детерминантов квадратных уравнений в (4.9). Итак, получены две стационарные точки. Теперь нужно выяснить какая из них даёт минимум. Для этого нужно определить каков знак второй производной в каждой из стационарных точек. Точка с положительной второй производной будет точкой минимума:

$$F'' = \frac{\gamma^3(-2a_2 \hat{a}_2) + \gamma^2(-3a_2 \hat{a}_1) + \gamma(2\hat{a}_2 a_0 - 4a_2 \hat{a}_0 + 2a_1 \hat{a}_1) + \hat{a}_1 a_0 + 4a_1 \hat{a}_0}{(\gamma^2 a_2 - 2\gamma a_1 + a_0)^3}, \quad (4.14)$$

где $\hat{a}_2 = a_2^2 - a_3 a_1$, $\hat{a}_1 = a_3 a_0 - a_2 a_1$, $\hat{a}_0 = a_1^2 - a_0 a_2$. Можно указать промежутки положительности F'' используя, например, формулу Кардано (напомним, что знаменатель положителен, см. (4.10), поэтому рассматриваем только числитель):

- Сперва приведем числитель к каноническому виду $y^3 + py + q$:

Сделаем замену $\gamma = y - \frac{b}{3a} = y - \frac{3a_2 \hat{a}_1}{6a_2 \hat{a}_2}$.

Тогда

$$\begin{aligned} p &= \frac{6a_2 \hat{a}_2(2\hat{a}_2 a_0 - 4a_2 \hat{a}_0 + 2a_1 \hat{a}_1) - (3a_2 \hat{a}_1)}{(6a_2 \hat{a}_2)} \\ q &= \frac{2(-3a_2 \hat{a}_1)^3 - 9(-2a_2 \hat{a}_2)(-3a_2 \hat{a}_1)(2\hat{a}_2 a_0 - 4a_2 \hat{a}_0 + 2a_1 \hat{a}_1) + 27(-2a_2 \hat{a}_2)^2(\hat{a}_1 a_0 + 4a_1 \hat{a}_0)}{27(-2a_2 \hat{a}_2)^3} \end{aligned}$$

- Вычислим $Q = \left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2$ и определим его знак:

$Q > 0$: один вещественный корень и два сопряженных комплексных корня,

$Q = 0$: один однократный вещественный корень и один двукратный, или, если $p = q = 0$, то один трёхкратный вещественный корень,

$Q < 0$ три вещественных корня.

- Введем обозначения:

$$\alpha = \sqrt[3]{-\frac{q}{2} + \sqrt{Q}}$$

$$\beta = \sqrt[3]{-\frac{q}{2} - \sqrt{Q}}$$

Тогда корни уравнения выражаются следующим образом:

$$\begin{aligned} y_1 &= \alpha + \beta, \\ y_{2,3} &= -\frac{\alpha + \beta}{2} \pm i \frac{\alpha - \beta}{2} \sqrt{3}. \end{aligned} \quad (4.15)$$

Переходя к исходной переменной:

$$\begin{aligned} \gamma_1 &= \alpha + \beta - \frac{3a_2\hat{a}_1}{6a_2\hat{a}_2}, \\ \gamma_{2,3} &= -\frac{\alpha + \beta}{2} \pm i \frac{\alpha - \beta}{2} \sqrt{3} - \frac{3a_2\hat{a}_1}{6a_2\hat{a}_2}. \end{aligned} \quad (4.16)$$

Считая, что имеет место случай $Q < 0$, а так же считая, что корни упорядочены по возрастанию, промежутки выпуклости будут следующими: $(-\inf, \gamma_1) \cup (\gamma_2, \gamma_3)$

Итак, найдя стационарные точки, нужно будет проверить их на принадлежность вышеуказанным интервалам.

Возможно, дабы не погрязть в комплексных корнях, проще будет просто подставить найденные стационарные точки в (4.14) и посмотреть знак выражения.

4.2.3. Сходимость

Покажем, что метод сходится со скоростью геометрической прогрессии со знаменателем

$$q = \sup_{0 \leq \gamma \leq \gamma_3} \|E + \gamma A\|, \quad (4.17)$$

Где γ_3 - максимальное значение γ из (4.16).

Напомним вид метода:

$$x^{k+1} = P_S(x^k - \gamma A x^k) = \frac{x^k - \gamma A x^k}{\|x^k - \gamma A x^k\|}, \quad (4.18)$$

S - единичная сфера,

$$\gamma_k = \underset{\gamma \geq 0}{\operatorname{argmin}} \frac{1}{2} \left(\frac{x^k - \gamma A x^k}{\|x^k - \gamma A x^k\|} \right)^T A \frac{x^k - \gamma A x^k}{\|x^k - \gamma A x^k\|} \quad (4.19)$$

Пусть x^* - стационарная точка. Тогда для нее верно, что

$$(\nabla f(x^*), x - x^*) = (A x^*, x - x^*) \quad \forall x \in S. \quad (4.20)$$

Это условие эквивалентно следующему:

$$x^* = P_S(x^* - \gamma A x^*) \quad \forall \gamma > 0. \quad (4.21)$$

Действительно, исходя из геометрического смысла: (4.20) эквивалентно тому, что

$S \cap Q = \{x \in \mathbb{R}^n : (\nabla f(x^*), x - x^*) < 0\} = \emptyset$. То есть множество направлений локального убывания не пересекается с множеством S . Допустим, что $x^* \in \operatorname{ri} S$ (относительная внутренность). В таком случае $\nabla f(x^*) = 0$ и $x^* = P_S(x^* - \gamma \nabla f(x^*)) = P_S(x^*) = x^*$. Пусть $x^* \in \operatorname{cl} S \setminus \operatorname{ri} S$. Предположим, что утверждение неверно. Тогда $\exists y \in S : y \neq x^*$ и $y = P_S(x^* - \gamma \nabla f(x^*))$. Тогда $(\nabla f(x^*), y - x^*) = (\nabla f(x^*), P_S(x^* - \gamma \nabla f(x^*) - x^*)) \stackrel{(*)}{<} (\nabla f(x^*), x^* - \gamma \nabla f(x^*) - x^*) = -\gamma (\nabla f(x^*), \nabla f(x^*)) < 0$. Неравенство при использовании липшицевости $(*)$ строгое, в силу структуры множества Q . Это означает, что $y \in Q$, то есть $y \notin S$ - противоречие.

Итак, воспользуемся (4.21):

$$\begin{aligned}\|x^{k+1} - x^*\| &= \|P_S(x^k - \gamma_k A x^k) - P_S(x^* - \gamma_k A x^k)\| \leq \\ &\leq \|x^k - \gamma_k A x^k - x^* + \gamma_k A x^k\| = \|(E + \gamma_k A)(x^k - x^*)\| \leq \\ &\leq \|x^k - x^*\| \cdot \|E + \gamma_k A\| \leq \|x^k - x^*\| \cdot q.\end{aligned}$$

Таким образом $\|x^k - x^*\| \leq \|x^0 - x^*\| q^k$, $q = \sup_{0 \leq \gamma \leq \gamma_3} \|E + \gamma A\|$.

5. Метод сопряженных градиентов и метод тяжелого шарика

Немного поясним смысл этих методов. В методах, описанных выше, никак не использовалась информация о том, что проделывал метод ранее. В следующих методах учитывается предыдущий шаг. Можно провести физическую аналогию с учётом информации с предыдущего шага: добавляется инерция, которая улучшает сходимость.

5.1. Метод тяжелого шарика

Рассмотрим метод тяжелого шарика для безусловной минимизации:

$$x^{k+1} = x^k - \alpha \nabla f(x^k) + \beta(x^k - x^{k-1}), \quad \alpha > 0, \quad b \geq 0. \quad (5.1)$$

Слагаемое $\beta(x^k - x^{k-1})$ отражает ту самую "инерцию". Физически она проявляется в следующем: метод при больших β ("инерции") будет проскакивать незаметные, то есть неглубокие локальные (а может и не локальные) минимумы и идти дальше. Перейдем к вопросам о сходимости:

Если x^* - невырожденная точка минимума, то есть в ней выполнено достаточное условие точки минимума второго порядка (2.3), помимо этого выполнены следующие условия:

- $0 \leq \beta < 1$
- $0 < \alpha < 2 \frac{1+\beta}{L}$
- $lE \leq \nabla^2 f(x^*) \leq LE$

тогда $\exists \varepsilon : \forall x^0, x^1 : \|x^0 - x^*\| < \varepsilon, \|x^1 - x^*\| < \varepsilon$ метод сходится к x^* со скоростью геометрической прогрессии. Здесь l и L - наименьшее и наибольшее собственные значения матрицы Гессе соответственно.

В случае, когда мы работаем на сфере, метод стоит переписать следующим образом:

$$x^{k+1} = P_S(x^k - \alpha \nabla f(x^k) + \beta(x^k - x^{k-1})), \quad \alpha > 0, \quad b \geq 0. \quad (5.2)$$

5.2. Метод сопряженных градиентов

Метод запишем сразу для работы на сфере:

$$\begin{aligned}x^{k+1} &= P_S(x^k - \alpha_k p^k) \\ \alpha_k &= \underset{\alpha}{\operatorname{armin}} f(P_S(x^k + \alpha p^k)), \quad x_0 \in S \\ p^k &= P_S(-\nabla f(x^k) - \beta_k p^{k-1}), \quad p^0 = -P_S(\nabla f(x^0)) \\ \beta_k &= \|P_S(\nabla f(x^k))\|^2 / \|P_S(\nabla f(x^{k-1}))\|^2\end{aligned} \quad (5.3)$$

6. Метод Ньютона

$$\begin{aligned}x^{k+1} &= \underset{x \in S}{\operatorname{argmin}} f_k(x) \\ f_k(x) &= f(x^k) + (\nabla f(x^k), x - x^k) + \frac{1}{2}(\nabla^2 f(x^k)(x - x^k), x - x^k)\end{aligned} \quad (6.1)$$

Метод сходится при следующих условиях:

- Функция $f(x)$ достигает минимум на S в точке x^*
- В окрестности x^* функция $f(x)$ - дважды дифференцируема
- Матрица Гессе $\nabla^2 f(x)$ - удовлетворяет условиям Липшица и положительно определена.

Перепишем второе уравнение для квадратичной функции:

$$\begin{aligned}
 f_k(x) &= \frac{1}{2}x^{kT}Ax^k + b^Tx^k + c + (Ax + b, x - x^k) + \frac{1}{2}A(x - x^k), x - x^k = \\
 &= \frac{1}{2}x^{kT}Ax^k + b^Tx^k + c + x^TAx - x^TAx^k + b^Tx - \\
 &\quad - b^Tx^k + \frac{1}{2}x^TAx - \frac{1}{2}x^{kT}Ax - \frac{1}{2}x^TAx^k + \frac{1}{2}x^kAx^k = \\
 &= x^{kT}Ax^k - \frac{3}{2}x^TAx^k - \frac{1}{2}x^{kT}Ax + \frac{3}{2}x^TAx + b^Tx + c.
 \end{aligned}$$

Решить задачу (6.1) можно с помощью метода множителей Лагранжа. Запишем функцию Лагранжа для (6.1):

$$L = x^{kT}Ax^k - \frac{3}{2}x^TAx^k - \frac{1}{2}x^{kT}Ax + \frac{3}{2}x^TAx + b^Tx + c + \lambda(\|x\|^2 - 1) \quad (6.2)$$

Условия оптимальности на функцию Лагранжа будут выглядеть так:

$$\begin{aligned}
 3Ax &= 2Ax^k - b + 2\lambda x, \\
 \|x\|^2 - 1 &= 0.
 \end{aligned}$$

Помимо этого стоит потребовать положительной определенности матрицы, задающей вторую производную по x функции Лагранжа:

$$\forall s \in \mathbb{S} \quad s^TL_{xx}s > 0.$$

Отсюда можно найти требуемый x .