GSK-АЛГОРИТМ И ЗАДАЧА СИЛЬВЕСТРА*

Н. А. Соловьева
4vinyo@gmail.com

5 октября 2023 г.

В докладе проводится углубленный анализ результатов работы [1], касающихся применения GSK-алгоритма[1] к решению задачи Сильвестра. На основе этого анализа дано конструктивное доказательство сильной сходимости GSK-алгоритма в рассматриваемом случае.

1°. Постановка задачи. В пространстве $\mathbb{R}^n$ с евклидовой нормой заданы m различных точек $a_1, a_2, \ldots, a_m$. Задача Сильвестра ставится так: *найти шар наименьшего объема, содержащий все точки* a_i, $i \in 1:m$. Запишем формализацию этой задачи

$$\max_{i \in 1:m} \left\{ \tfrac{1}{2} \|a_i - x\|^2 \right\} \to \min_{x \in \mathbb{R}^n}. \tag{1}$$

Здесь x — центр шара. Задача (1) имеет решение x^* и оно единственно.

Обозначим через A матрицу со столбцами $a_1, a_2, \ldots, a_m$ и через b — вектор из $\mathbb{R}^m$ с компонентами $b[i] = \frac{1}{2}\|a_i\|^2$. Рассмотрим вспомогательную задачу квадратичного программирования

$$\begin{gathered} Q(u) := \tfrac{1}{2}\langle A^T A u, u\rangle - \langle b, u\rangle \to \min, \\ \sum_{i=1}^m u[i] = 1, \quad u[i] \geqslant 0 \ \text{ при всех } i \in 1:m. \end{gathered} \tag{2}$$

Множество ее планов обозначим через U. Очевидно, что у задачи (2) существует оптимальный план.

В докладе [2] было доказано следующее утверждение.

*Семинар по оптимизации, машинному обучению и искусственному интеллекту «O&ML» `http://oml.cmlaboratory.com/`

[1]GSK — сокращение от Gilbert-Schlesinger-Kozinec.

ТЕОРЕМА 1. *Целевая функция $Q(u)$ задачи (2) на множестве планов U допускает представление*

$$Q(u) = -\tfrac{1}{2} \sum_{i=1}^{m} u[i] \cdot \|a_i - x\|^2 \,, \tag{3}$$

где

$$x = Au = \sum_{i=1}^{m} u[i]\, a_i.$$

СЛЕДСТВИЕ. *Для произвольного орта e_i пространства $\mathbb{R}^m$ справедливо равенство*

$$Q(e_i) = 0.$$

Для остальных планов $u \in U$ задачи (2) выполняется неравенство

$$Q(u) < 0.$$

Доказательство. Определим носитель плана u:

$$I^+(u) = \big\{ i \in 1:m \mid u[i] > 0 \big\}.$$

Рассмотрим вначале случай, когда $u = e_j$ при некотором $j \in 1:m$. Тогда $I^+(u) = \{j\}$. Вычислим

$$x = Ae_j = a_j.$$

С использованием (3) получим

$$Q(e_j) = -\tfrac{1}{2} \sum_{i=1}^{m} e_j[i] \cdot \|a_i - a_j\|^2 = -\tfrac{1}{2}\, \|a_j - a_j\|^2 = 0.$$

Теперь возьмем план $u \in U$, не совпадающий ни с одним из ортов. Так как u — не орт, то в его носителе $I^+(u)$ содержатся как минимум два индекса. Обозначим их j и ℓ. На основании (3) запишем

$$Q(u) = -\tfrac{1}{2}\, u[j] \cdot \|a_j - x\|^2 - \tfrac{1}{2}\, u[\ell] \cdot \|a_\ell - x\|^2 - \tfrac{1}{2} \sum_{i \neq j, i \neq \ell} u[i] \cdot \|a_i - x\|^2 \,. \tag{4}$$

Заметим, что $u[j]$ и $u[\ell]$ положительны, а $\|a_j - x\|^2$ и $\|a_\ell - x\|^2$ не обращаются в ноль одновременно, поскольку все a_i различны. Учитывая неотрицательность суммы

$$\tfrac{1}{2} \sum_{i \neq j, i \neq \ell} u[i] \cdot \|a_i - x\|^2 \,,$$

заключаем, что правая часть равенства (4) отрицательна. $\square$

Решение задачи Сильвестра характеризуют следующие утверждения [2].

ТЕОРЕМА 2. *Если u^* — оптимальный план задачи* (2), *то вектор $x^* = Au^*$ будет решением задачи* (1). *При этом*

$$\max_{i\in 1:m}\left\{\tfrac{1}{2}\left\|a_i - x^*\right\|^2\right\} = -Q(u^*).$$

ТЕОРЕМА 3. *План u задачи* (2) *будет оптимальным тогда и только тогда, когда для вектора $x = Au$ при всех $s \in I^+(u)$ выполняются равенства*

$$\|a_s - x\| = \max_{i\in 1:m}\|a_i - x\|. \tag{5}$$

2°. Оценка плана задачи (2). Возьмем план $u \in U$ задачи (2). Вычислим

$$x = Au = \sum_{i=1}^{m} u[i]\, a_i.$$

Найдем среди a_i, $i \in 1 : m$, точку, наиболее удаленную от x, и обозначим ее индекс через i':

$$R(u) := \|a_{i'} - x\| = \max_{i\in 1:m}\|a_i - x\|.$$

Введем оценку $\Delta(u)$ плана u следующим образом:

$$\Delta(u) = R^2(u) + 2Q(u). \tag{6}$$

Так как (см. формулу (3))

$$-2Q(u) = \sum_{i=1}^{m} u[i]\ \|a_i - x\|^2 \leqslant R^2(u),$$

то $\Delta(u) \geqslant 0$ для любого плана $u \in U$.

ПРЕДЛОЖЕНИЕ 1. *План $u \in U$ является оптимальным планом задачи* (2) *в том и только том случае, когда $\Delta(u) = 0$.*

Д о к а з а т е л ь с т в о. Если u — оптимальный план, то по теореме 2 выполняется соотношение

$$\tfrac{1}{2}R^2(u) = -Q(u),$$

которое равносильно равенству $\Delta(u) = 0$.

Наоборот, возьмем план $u \in U$ с нулевой оценкой. Условие $\Delta(u) = 0$ означает, что

$$R^2(u) - \sum_{i=1}^{m} u[i]\ \|a_i - x\|^2 = 0.$$

Перепишем последнее равенство в эквивалентном виде

$$\sum_{s\in I^+(u)} u[s]\left(R^2(u) - \|a_s - x\|^2\right) = 0.$$

Отсюда с учетом положительности компонент $u[s]$ при всех $s \in I^+(u)$ и неотрицательности разностей $R^2(u) - \|a_s - x\|^2$ следуют равенства

$$\|a_s - x\|^2 = R^2(u), \quad s \in I^+(u).$$

По теореме 3 план u является оптимальным. $\square$

Покажем, как с помощью величины $\Delta(u)$ оценить близость вектора $x = Au$ к решению x^* задачи (1).

ПРЕДЛОЖЕНИЕ 2. *Пусть $u \in U$ — план задачи* (2)*, $x = Au$. Тогда*

$$\|x - x^*\|^2 \leqslant \tfrac{1}{2}\Delta(u). \tag{7}$$

Д о к а з а т е л ь с т в о. Запишем

$$\begin{aligned} \|x - x^*\|^2 &= \|x\|^2 + \|x^*\|^2 - \langle x^*, x\rangle - \langle x, x^*\rangle = \\ &= \|x\|^2 + \|x^*\|^2 - \langle Au^*, x\rangle - \langle Au, x^*\rangle = \\ &= \|x\|^2 + \|x^*\|^2 - \sum_{i=1}^m u^*[i]\,\langle a_i, x\rangle - \sum_{i=1}^m u[i]\,\langle a_i, x^*\rangle. \end{aligned} \tag{8}$$

Напомним, что $b[i] = \frac{1}{2}\|a_i\|^2$ при всех $i \in 1:m$. Поэтому

$$\begin{aligned} \langle a_i, x\rangle - b[i] &= \tfrac{1}{2}\big(\|x\|^2 - \|a_i - x\|^2\big), \\ \langle a_i, x^*\rangle - b[i] &= \tfrac{1}{2}\big(\|x^*\|^2 - \|a_i - x^*\|^2\big). \end{aligned}$$

Отсюда следует, что

$$\begin{aligned} \langle a_i, x\rangle &= \tfrac{1}{2}\big(\|x\|^2 - \|a_i - x\|^2\big) + b[i], \\ \langle a_i, x^*\rangle &= \tfrac{1}{2}\big(\|x^*\|^2 - \|a_i - x^*\|^2\big) + b[i]. \end{aligned}$$

Подставив эти равенства в (8), получим

$$\|x - x^*\|^2 = \|x\|^2 + \|x^*\|^2 -$$
$$- \sum_{i=1}^m u^*[i]\left[\tfrac{1}{2}\big(\|x\|^2 - \|a_i - x\|^2\big) + b[i]\right] - \sum_{i=1}^m u[i]\left[\tfrac{1}{2}\big(\|x^*\|^2 - \|a_i - x^*\|^2\big) + b[i]\right].$$

После раскрытия скобок придем к представлению

$$\|x - x^*\|^2 = \|x\|^2 + \|x^*\|^2 -$$

$$\begin{aligned} &-\tfrac{1}{2}\sum_{i=1}^m u^*[i]\ \|x\|^2 + \tfrac{1}{2}\sum_{i=1}^m u^*[i]\ \|a_i - x\|^2 - \sum_{i=1}^m u^*[i]\,b[i] - \\ &-\tfrac{1}{2}\sum_{i=1}^m u[i]\ \|x^*\|^2 + \tfrac{1}{2}\sum_{i=1}^m u[i]\ \|a_i - x^*\|^2 - \sum_{i=1}^m u[i]\,b[i]. \end{aligned}$$

Скомпонуем восемь слагаемых в последней формуле таким образом: в одну сумму — первое, третье, четвертое и восьмое, в другую сумму — второе, пятое, шестое и седьмое. Запишем

$$\begin{gathered}\|x-x^*\|^2 = \\ = \Big[\|x\|^2 - \tfrac{1}{2}\sum_{i=1}^m u^*[i]\ \|x\|^2 + \tfrac{1}{2}\sum_{i=1}^m u^*[i]\ \|a_i - x\|^2 - \sum_{i=1}^m u[i]\, b[i]\Big] + \\ + \Big[\|x^*\|^2 - \sum_{i=1}^m u^*[i]\, b[i] - \tfrac{1}{2}\sum_{i=1}^m u[i]\ \|x^*\|^2 + \tfrac{1}{2}\sum_{i=1}^m u[i]\ \|a_i - x^*\|^2\Big] =: \\ =: S_1 + S_2.\end{gathered} \tag{9}$$

Разберемся отдельно с S_1 и S_2. Имеем

$$\begin{gathered}S_1 = \|x\|^2 - \tfrac{1}{2}\|x\|^2 - \langle b, u\rangle + \tfrac{1}{2}\sum_{i=1}^m u^*[i]\ \|a_i - x\|^2 \leqslant \\ \leqslant \tfrac{1}{2}\|x\|^2 - \langle b, u\rangle + \tfrac{1}{2}\max_{i\in 1:m}\|a_i - x\|^2 \sum_{i=1}^m u^*[i] = \\ = Q(u) + \tfrac{1}{2}\max_{i\in 1:m}\|a_i - x\|^2 = Q(u) + \tfrac{1}{2}R^2(u).\end{gathered} \tag{10}$$

Аналогично получим

$$\begin{gathered}S_2 = \|x^*\|^2 - \tfrac{1}{2}\|x^*\|^2 - \langle b, u^*\rangle + \tfrac{1}{2}\sum_{i=1}^m u[i]\ \|a_i - x^*\|^2 \leqslant \\ \leqslant \tfrac{1}{2}\|x^*\|^2 - \langle b, u^*\rangle + \tfrac{1}{2}\max_{i\in 1:m}\|a_i - x^*\|^2 \sum_{i=1}^m u[i] = \\ = Q(u^*) + \tfrac{1}{2}\max_{i\in 1:m}\|a_i - x^*\|^2 = Q(u^*) + \tfrac{1}{2}R^2(u^*).\end{gathered} \tag{11}$$

Объединив (9), (10) и (11), придем к неравенству

$$\|x-x^*\|^2 \leqslant \Big[Q(u) + \tfrac{1}{2}R^2(u)\Big] + \Big[Q(u^*) + \tfrac{1}{2}R^2(u^*)\Big].$$

Здесь вторая квадратная скобка по теореме 2 равна нулю. Следовательно,

$$\|x-x^*\|^2 \leqslant Q(u) + \tfrac{1}{2}R^2(u) = \tfrac{1}{2}\Delta(u).$$

Предложение доказано. □

3°. Вариация плана задачи (2). Зафиксируем план $u \in U$ задачи (2), отличный от оптимального. На нем $\Delta(u) > 0$. Вычислим $x = Au$. Возьмем $h \in U$ — произвольный план задачи (2). Рассмотрим вариацию $\widetilde{u}$ плана u вида

$$\widetilde{u} = (1-\lambda)u + \lambda h = u + \lambda\bigl(h-u\bigr), \quad \lambda \in [0,1].$$

При всех значениях λ из отрезка $[0,1]$ вектор $\widetilde{u}$ является планом задачи (2).

Вектор h предлагается искать как решение задачи линейного программирования вида

$$\Phi(h) := \langle A^T Au - b, h\rangle \to \min_{h\in U}.$$

С учетом определения $b[i]$ и равенства

$$\langle a_i, x\rangle - \tfrac{1}{2}\,\|a_i\|^2 = \tfrac{1}{2}\bigl(\|x\|^2 - \|a_i - x\|^2\bigr)$$

запишем

$$\Phi(h) = \langle A^T Au - b, h\rangle = \langle A^T x - b, \sum_{i=1}^m h[i]\, e_i\rangle =$$

$$= \sum_{i=1}^m h[i]\langle A^T x - b, e_i\rangle = \sum_{i=1}^m h[i]\bigl(\langle A^T x, e_i\rangle - \tfrac{1}{2}\,\|a_i\|^2\bigr) =$$

$$= \sum_{i=1}^m h[i]\bigl(\langle a_i, x\rangle - \tfrac{1}{2}\,\|a_i\|^2\bigr) = \tfrac{1}{2}\sum_{i=1}^m h[i]\bigl(\|x\|^2 - \|a_i - x\|^2\bigr) =$$

$$= \tfrac{1}{2}\,\|x\|^2 - \tfrac{1}{2}\sum_{i=1}^m h[i]\,\|a_i - x\|^2.$$

Поменяв знак у целевой функции $\Phi(h)$, придем к задаче вида

$$\sum_{i=1}^m h[i]\,\|a_i - x\|^2 \to \max_{h\in U}. \tag{12}$$

Оценим новую целевую функцию сверху

$$\sum_{i=1}^m h[i]\,\|a_i - x\|^2 \leqslant \max_{i\in 1:m}\|a_i - x\|^2 \cdot \sum_{i=1}^m h[i] = \max_{i\in 1:m}\|a_i - x\|^2.$$

Отсюда следует, что решением задачи (12) является, в частности, вектор $h = e_{i'}$, где i' — индекс максимально удаленной от x точки a_i, $i \in 1:m$:

$$\|a_{i'} - x\| = \max_{i\in 1:m}\|a_i - x\|$$

(индекс i' появлялся ранее в определении величины $R(u)$). Получили, что вариацию $\widetilde{u}$ плана u можно представить в виде

$$\widetilde{u} = (1-\lambda)u + \lambda e_{i'} = u + \lambda\bigl(e_{i'} - u\bigr).$$

Займемся выбором шага $\lambda \in [0,1]$ из точки u в направлении $e_{i'} - u$. Вычислим значение целевой функции задачи (2) на плане $\widetilde{u}$:

$$\Psi(\lambda) = Q(\widetilde{u}) = Q\big((1-\lambda)u + \lambda e_{i'}\big) =$$

$$= \tfrac{1}{2}\left\|A\big((1-\lambda)u + \lambda e_{i'}\big)\right\|^2 - \langle b, (1-\lambda)u + \lambda e_{i'}\rangle =$$

$$= \tfrac{1}{2}(1-\lambda)^2\|Au\|^2 + \lambda(1-\lambda)\langle Au, Ae_{i'}\rangle + \tfrac{1}{2}\lambda^2\|Ae_{i'}\|^2 - (1-\lambda)\langle b,u\rangle - \lambda\langle b, e_{i'}\rangle =$$

$$= \tfrac{1}{2}(1-\lambda)^2\|x\|^2 + \lambda(1-\lambda)\langle x, a_{i'}\rangle + \tfrac{1}{2}\lambda^2\|a_{i'}\|^2 - (1-\lambda)\langle b,u\rangle - \tfrac{1}{2}\lambda\|a_{i'}\|^2 =$$

$$= \tfrac{1}{2}\lambda^2\Big[\|x\|^2 - 2\langle x, a_{i'}\rangle + \|a_{i'}\|^2\Big] + \lambda\Big[-\|x\|^2 + \langle x, a_{i'}\rangle + \langle b,u\rangle - \tfrac{1}{2}\|a_{i'}\|^2\Big] + Q(u) =$$

$$= \tfrac{1}{2}\lambda^2\Big[\|a_{i'} - x\|^2\Big] + \lambda\Big[-\tfrac{1}{2}\|a_{i'} - x\|^2 - Q(u)\Big] + Q(u).$$

Таким образом,

$$\Psi(\lambda) = Q(\widetilde{u}) = \tfrac{1}{2}\lambda^2 R^2(u) - \tfrac{1}{2}\lambda\Delta(u) + Q(u). \tag{13}$$

Заметим, что $\Psi(\lambda)$ является квадратичной по λ функцией. Найдем точку минимума $\Psi(\lambda)$, приравняв к нулю ее производную. Получим

$$\lambda = \frac{\Delta(u)}{2R^2(u)}\,.$$

Ранее предполагалось, что $\Delta(u) > 0$. Значит, $\lambda > 0$. Напомним, что величина $Q(u)$ неположительна. Следовательно,

$$\Delta(u) = R^2(u) + 2Q(u) \leqslant R^2(u)$$

и $0 < \lambda \leqslant \frac{1}{2}$. Приходим к выводу, что вектор

$$\widetilde{u} = u + \frac{\Delta(u)}{2R^2(u)}\big(e_{i'} - u\big)$$

является планом задачи (2).

Посмотрим, насколько уменьшается целевая функция $Q(u)$ задачи (2) при переходе от плана u к плану $\widetilde{u}$. Согласно (13) имеем

$$Q(u) - Q(\widetilde{u}) = -\tfrac{1}{2}\lambda^2 R^2(u) + \tfrac{1}{2}\lambda\Delta(u).$$

Подставив в это равенство найденное значение λ, получим

$$Q(u) - Q(\widetilde{u}) = -\frac{1}{2}\cdot\frac{\Delta^2(u)}{4R^4(u)}\,R^2(u) + \frac{1}{2}\cdot\frac{\Delta(u)}{2R^2(u)}\,\Delta(u) = \frac{\Delta^2(u)}{8R^2(u)}\,. \tag{14}$$

Отметим, что правая часть равенства (14) положительна.

4°. Описание алгоритма. Построим начальное приближение следующим образом [1]. Найдем точку a_α, наиболее удаленную от a_1. Затем найдем точку a_β, наиболее удаленную от a_α. В качестве начального приближения возьмем вектор u_0 с компонентами

$$u_0[i] = \begin{cases} \frac{1}{2} & \text{при } i = \alpha, \\ \frac{1}{2} & \text{при } i = \beta, \\ 0 & \text{при остальных } i \in 1 : m. \end{cases}$$

Очевидно, что u_0 принадлежит множеству планов U задачи (2).

Предположим, что уже имеется k-е приближение $u_k \in U$. Вычислим

$$x_k = Au_k = \sum_{i=1}^{m} u_k[i]\, a_i.$$

Найдем среди a_i, $i \in 1 : m$, точку, наиболее удаленную от x_k. Обозначим ее индекс через i'_k:

$$R(u_k) := \left\| a_{i'_k} - x_k \right\| = \max_{i \in 1:m} \| a_i - x_k \| \,.$$

Вычислим оценку

$$\Delta(u_k) = R^2(u_k) + 2Q(u_k).$$

Если $\Delta(u_k) = 0$, то алгоритм заканчивает работу. Вектор x_k будет решением задачи (1).

Пусть $\Delta(u_k) > 0$. Тогда найдем

$$\lambda_k = \frac{\Delta(u_k)}{2R^2(u_k)}$$

и построим новый план по формуле

$$u_{k+1} = (1 - \lambda_k) u_k + \lambda_k e_{i'_k}.$$

Вычислим

$$x_{k+1} = Au_{k+1} = (1 - \lambda_k) x_k + \lambda_k a_{i'_k}.$$

Описание алгоритма завершено.

С помощью предложенного алгоритма строятся две последовательности

$$u_0, u_1, \ldots, u_k, \ldots,$$
$$x_0, x_1, \ldots, x_k, \ldots$$

Они конечны, если при некотором k выполнится неравенство $\Delta(u_k) = 0$. Они бесконечны, когда

$$\Delta(u_k) > 0 \quad \text{при всех } k = 0, 1, \ldots \tag{15}$$

5°. Сходимость алгоритма. Будем считать выполненным условие (15). Займемся изучением вопроса о сходимости бесконечной последовательности $\{x_k\}$.

ПРЕДЛОЖЕНИЕ 3. *Предположим, что выполняется условие* (15). *Пусть* $\{u_k\}$ *— последовательность, построенная GSK-алгоритмом,* $x_k = Au_k$. *Тогда справедливо предельное соотношение*

$$x_k \to x^* \quad \textit{при } k \to \infty,$$

где x^* *— решение задачи* (1).

Доказательство. Последовательность $\{Q(u_k)\}$ строго убывает и ограничена снизу значением $Q(u^*)$. Значит, она сходится. Как следствие,

$$Q(u_k) - Q(u_{k+1}) \to 0 \quad \text{при } k \to \infty. \tag{16}$$

На основании (14) имеем

$$Q(u_k) - Q(u_{k+1}) = \frac{\Delta^2(u_k)}{8R^2(u_k)}\,.$$

Обозначим через d диаметр выпуклой оболочки точек $\{a_i\}_i^m$. Тогда

$$Q(u_k) - Q(u_{k+1}) \geqslant \frac{\Delta^2(u_k)}{8d^2}.$$

Отсюда и из (16) следует, что

$$\Delta(u_k) \to 0 \quad \text{при } k \to \infty.$$

Остается сослаться на предложение 2. □

Другие алгоритмы решения задачи Сильвестра рассматриваются в обзорной статье [3].

ЛИТЕРАТУРА

1. E. Alper Yildirim. *Two algorithms for the minimum enclosing ball problem* // SIAM J. Optim. Vol. 19. No 3. 2008. Pp. 1368–1391.

2. Малозёмов В. Н., Соловьёва Н. А., Тамасян Г. Ш. *MDM-алгоритм и задача Сильвестра. I* // Семинар «O & ML». Избранные доклады. 14 сентября 2023 г. (`http://oml.cmlaboratory.com/reps23.shtml#0914`)

3. F. Nielsen, R. Nock. *Approximating Smallest Enclosing Balls with Application to Machine Learning* // Int. J. Computational Geometry & Applications. Vol. 19. No. 5. 2009. Pp. 389–414.