ОБЩАЯ КВАДРАТИЧНАЯ ЗАДАЧА МАТЕМАТИЧЕСКОЙ ДИАГНОСТИКИ*

В. Н. Малоземов — v.malozemov@spbu.ru

Н. А. Соловьева — 4vinyo@gmail.com

11 мая 2022 г.

1°. Простейшая квадратичная задача математической диагностики ставится так [1, 2]: найти ортогональную проекцию произвольной точки на выпуклую оболочку конечного числа известных точек. Сформулируем общую квадратичную задачу математической диагностики.

В пространстве $\mathbb{R}^n$ заданы два конечных множества

$$P_1 = \{p_j\}_{j=1}^{s} \quad \text{и} \quad P_2 = \{p_j\}_{j=s+1}^{m},$$

где $s \in 1 : m-1$. Обозначим через C_1 и C_2 выпуклые оболочки множеств P_1 и P_2 соответственно. Требуется решить экстремальную задачу

$$\tfrac{1}{2}\,\|x-y\|^2 \to \min_{x\in C_1,\, y\in C_2}. \tag{1}$$

Очевидно, что задача (1) имеет решение (x^*, y^*). Если минимальное значение целевой функции равно нулю, то $x^* = y^*$. В этом случае найдена общая точка множеств C_1 и C_2.

2°. По определению выпуклой оболочки точки x и y допускают представления

$$x = \sum_{j=1}^{s} u_j p_j, \quad y = \sum_{j=s+1}^{m} u_j p_j,$$

где

$$\begin{aligned} &\sum_{j=1}^{s} u_j = 1, \quad u_j \geqslant 0 \ \text{ при } j \in 1:s; \\ &\sum_{j=s+1}^{m} u_j = 1, \quad u_j \geqslant 0 \ \text{ при } j \in s+1:m. \end{aligned} \tag{2}$$

*Семинар по оптимизации, машинному обучению и искусственному интеллекту «O&ML» http://www.apmath.spbu.ru/oml/

Введем вектор $\xi = (\xi_1, \dots, \xi_m)$ с компонентами $\xi_j = 1$ при $j \in 1 : s$, $\xi_j = -1$ при $j \in s+1 : m$. Тогда условия (2) можно переписать в эквивалентном виде:

$$\sum_{j=1}^{m} u_j = 2, \quad \sum_{j=1}^{m} \xi_j u_j = 0,$$
$$u_j \geqslant 0, \quad j \in 1 : m.$$

Обозначим через A матрицу со столбцами $\xi_1 p_1, \dots, \xi_m p_m$ и преобразуем целевую функцию задачи (1):

$$\tfrac{1}{2}\,\|x-y\|^2 = \tfrac{1}{2}\Big\|\sum_{j=1}^{m} u_j \xi_j p_j\Big\|^2 = \tfrac{1}{2}\,\|Au\|^2 = \tfrac{1}{2}\langle A^T A\, u, u\rangle.$$

Придем к подробной записи задачи (1),

$$\begin{gathered} \tfrac{1}{2}\langle A^T A\, u, u\rangle \to \min, \\ \sum_{j=1}^{m} u_j = 2, \\ \sum_{j=1}^{m} \xi_j u_j = 0, \\ u_j \geqslant 0, \quad j \in 1 : m. \end{gathered} \tag{3}$$

Очевидно, что задача (3) имеет решение u^*.

Запишем критерий оптимальности для задачи квадратичного программирования (3) (условия Куна–Таккера):

$$\begin{gathered} A^T A\, u = \mu e + \lambda \xi + \sum_{j=1}^{m} t_j e_j, \\ t_j u_j = 0, \quad t_j \geqslant 0 \text{ при } j \in 1 : m, \end{gathered} \tag{4}$$

где $e \in \mathbb{R}^m$ — вектор, все компоненты которого равны единице, и $e_j \in \mathbb{R}^m$ — j-й орт. Вектор u^* удовлетворяет условиям (4) при некоторых μ^*, λ^*, $t_1^*, \dots, t_m^*$.

ЛЕММА 1. *Справедливо равенство*

$$\tfrac{1}{2}\langle A^T A\, u^*, u^*\rangle = \mu^*. \tag{5}$$

Доказательство. На основании (4) и (3) имеем

$$\begin{gathered} \langle A^T A\, u^*, u^*\rangle = \mu^* \langle e, u^*\rangle + \lambda^* \langle \xi, u^*\rangle + \sum_{j=1}^{m} t_j^* u_j^* = \\ = \mu^* \langle e, u^*\rangle + \lambda^* \langle \xi, u^*\rangle = \mu^* \langle e, u^*\rangle = 2\mu^*. \end{gathered}$$

Поделив на 2, придем к (5). □

Перепишем равенство (5) в терминах векторов p_j,

$$\mu^* = \tfrac{1}{2}\,\|Au^*\|^2 = \tfrac{1}{2}\Big\|\sum_{j=1}^{s} u_j^* p_j - \sum_{j=s+1}^{m} u_j^* p_j\Big\|^2.$$

Напомним, что вектор u^* удовлетворяет условиям (2). Становится очевидным следующее утверждение.

ЛЕММА 2. *Справедливо неравенство $\mu^* \geqslant 0$. Равенство $\mu^* = 0$ выполняется тогда и только тогда, когда выпуклые оболочки C_1 и C_2 имеют непустое пересечение.*

3°. Наряду с общей квадратичной задачей математической диагностики (3) рассмотрим еще одну задачу квадратичного программирования:

$$\begin{gathered} \tfrac{1}{2}\langle A^T A\, u, u\rangle - \sum_{j=1}^{m} u_j \to \min, \\ \sum_{j=1}^{m} \xi_j u_j = 0, \\ u_j \geqslant 0, \quad j \in 1:m. \end{gathered} \tag{6}$$

Запишем для нее критерий оптимальности:

$$\begin{gathered} A^T A\, u - e = \lambda \xi + \sum_{j=1}^{m} t_j e_j, \\ t_j u_j = 0, \quad t_j \geqslant 0 \ \text{ при } j \in 1:m. \end{gathered} \tag{7}$$

ТЕОРЕМА 1. *Пусть $C_1 \cap C_2 = \varnothing$ и u^* — решение задачи* (3)*. Тогда вектор*

$$u^0 = \frac{1}{\mu^*}\, u^*$$

будет решением задачи (6)*.*

Доказательство. Очевидно, что u^0 — план задачи (6). Проверим, что это оптимальный план.

По лемме 2 имеем $\mu^* > 0$. Обозначим

$$\lambda^0 = \frac{1}{\mu^*}\,\lambda^*, \quad t_j^0 = \frac{1}{\mu^*}\,t_j^* \ \text{ при } j \in 1:m.$$

При $u = u^0$, $\lambda = \lambda^0$, $t_j = t_j^0$ выполняется критерий оптимальности (7) в силу того, что при $u = u^*$, $\mu = \mu^*$, $\lambda = \lambda^*$, $t_j = t_j^*$ справедливы соотношения (4). Значит, u^0 — оптимальный план задачи (6).

Теорема доказана. □

Отметим, что μ^* зависит от u^*.

4°. Задача (6) появилась не случайно. Она является вспомогательной при линейном отделении множеств P_1 и P_2 SVM-методом [3]. Напомним постановку задачи линейного отделения с максимальной шириной разделяющей полосы:

$$\begin{gathered} \tfrac{1}{2}\,\|w\|^2 \to \min, \\ \xi_j\big(\langle w, p_j\rangle + \beta\big) \geqslant 1, \quad j \in 1:m. \end{gathered} \tag{8}$$

Задача (8) при выполнении условия $C_1 \cap C_2 = \varnothing$ имеет единственное решение (w^0, β^0). В докладе [3] показано, что если u^0 — решение задачи (6), то

$$w^0 = \sum_{j \in J^0} u_j^0 \xi_j p_j,$$

где $J^0 = \{j \in 1:m \mid u_j^0 > 0\}$. Параметр β^0 определяется из равенства

$$\xi_j\big(\langle w^0, p_j\rangle + \beta^0\big) = 1 \quad \text{при } j \in J^0.$$

Таким образом, решение общей квадратичной задачи математической диагностики (3) позволяет найти решение задачи (8) линейного отделения двух конечных множеств P_1 и P_2 с наибольшей шириной разделяющей полосы.

5°. Рассмотрим обратную задачу — по решению задачи (6) восстановить решение задачи (3). Условие $C_1 \cap C_2 = \varnothing$ сохраняется.

По теореме 1 решение задачи (6) существует, но его единственность, вообще говоря, не гарантируется. Возьмем любое решение $\widetilde{u}$ задачи (6). Покажем, что $\widetilde{u} \neq \mathbb{O}$. В противном случае векторное равенство в критерии оптимальности (7) примет вид

$$-e = \widetilde{\lambda}\xi + \sum_{j=1}^{m} \widetilde{t}_j e_j.$$

Выделим равенства первой и последней координат:

$$-1 = \widetilde{\lambda} + \widetilde{t}_1, \quad -1 = -\widetilde{\lambda} + \widetilde{t}_m.$$

Сложив эти равенства, получим $-2 = \widetilde{t}_1 + \widetilde{t}_m$, что невозможно, так как $\widetilde{t}_1 \geqslant 0$ и $\widetilde{t}_m \geqslant 0$. Значит, $\widetilde{u} \neq \mathbb{O}$.

В силу ограничения задачи (6) имеем

$$\sum_{j=1}^{s} \widetilde{u}_j = \sum_{j=s+1}^{m} \widetilde{u}_j =: b,$$

причем $b > 0$.

ТЕОРЕМА 2. *Пусть $C_1 \cap C_2 = \varnothing$ и $\widetilde{u}$ — решение задачи* (6). *Тогда вектор*

$$\widetilde{u}^* = \frac{1}{b}\,\widetilde{u}$$

будет решением задачи (3).

Доказательство. Очевидно, что $\widetilde{u}^*$ — план задачи (3). Проверим, что это оптимальный план. Обозначим

$$\widetilde{\mu}^* = \frac{1}{b}, \quad \widetilde{\lambda}^* = \frac{1}{b}\,\widetilde{\lambda}, \quad \widetilde{t}_j^* = \frac{1}{b}\,\widetilde{t}_j \ \text{ при } j \in 1 : m.$$

При $u = \widetilde{u}^*$, $\mu = \widetilde{\mu}^*$, $\lambda = \widetilde{\lambda}^*$, $t_j = \widetilde{t}_j^*$ выполняется критерий оптимальности (4) в силу того, что при $u = \widetilde{u}$, $\lambda = \widetilde{\lambda}$, $t_j = \widetilde{t}_j$ справедливы соотношения (7). Значит, $\widetilde{u}^*$ — оптимальный план задачи (3).

Теорема доказана. □

6°. В заключение подчеркнем важную роль вспомогательной задачи (3). По ее оптимальному плану восстанавливается как решение задачи (1), так и решение задачи (8).

Задача (3) является задачей квадратичного программирования с неотрицательно определенной, но вырожденной матрицей A^TA, с двумя линейными ограничениями-равенствами и знаковыми ограничениями на все переменные. Такую задачу рекомендуется решать методом Данцига, сводя ее к решению системы Куна–Таккера [4, с. 64–85]. В случае простейшей квадратичной задачи математической диагностики предлагаемый подход подробно описан в докладе [2].

ЛИТЕРАТУРА

1. Певный А. Б. *К нахождению точки многогранника, ближайшей к началу координат* / В сб.: Оптимизация. Вып. 10 (27). Новосибирск, 1973. С. 54–58.

2. Котелина Н. О., Певный А. Б. *Квадратичная задача математического моделирования* // Семинар «O & ML». Избранные доклады. 23 марта 2022 г. (http://apmath.spbu.ru/oml/reps22.shtml#0323)

3. Малоземов В. Н., Плоткин А. В. *SVM-метод строгого линейного отделения двух конечных множеств*// Семинар «O & ML». Избранные доклады. 30 марта 2022 г. (http://apmath.spbu.ru/oml/reps22.shtml#0330)

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