ПРОЕКТИРОВАНИЕ СИММЕТРИЧНОЙ МАТРИЦЫ НА КОНУС НЕОТРИЦАТЕЛЬНО ОПРЕДЕЛЁННЫХ МАТРИЦ И БЛИЗКИЕ ВОПРОСЫ*

В. Н. Малозёмов С. Е. Михеев

v.malozemov@spbu.ru him2@mail.ru

18 декабря 2014 г.

1°. В линейном пространстве $\mathbb{R}^{n\times n}$ квадратных вещественных матриц порядка n введём скалярное произведение

$$\langle A, B\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} A[i,j]\times B[i,j]$$

и норму $\|A\| = \sqrt{\langle A, A\rangle}$. Напомним, что матрица $D \in \mathbb{R}^{n\times n}$ называется *симметричной*, если $D[i,j] = D[j,i]$ при всех $i, j \in 1:n$, и *неотрицательно определённой*, если $\langle Dx, x\rangle \geqslant 0$ при всех $x \in \mathbb{R}^n$. Множество симметричных неотрицательно определённых матриц порядка n обозначим $\mathcal{K}^n$. Очевидно, что $\mathcal{K}^n$ — выпуклый конус.

Возьмём симметричную матрицу $A \in \mathbb{R}^{n\times n}$ и рассмотрим задачу ортогонального проектирования матрицы A на конус $\mathcal{K}^n$ в следующей постановке:

$$F(X) := \|A - X\|^2 \to \inf_{X\in\mathcal{K}^n}. \tag{1}$$

В докладе приводится решение задачи (1), а также ближайших её обобщений.

2°. Нам потребуются некоторые свойства скалярного произведения и нормы матриц.

Из определений непосредственно следует, что

$$\|A + B\|^2 = \|A\|^2 + \|B\|^2 + 2\langle A, B\rangle. \tag{2}$$

*Семинар по конструктивному негладкому анализу и недифференцируемой оптимизации «CNSA & NDO»: `http://www.apmath.spbu.ru/cnsa/`

Напомним, что *следом* квадратной матрицы $A \in \mathbb{R}^{n\times n}$ называется величина

$$\mathrm{Sp}(A) = \sum_{i=1}^{n} A[i,i].$$

ЛЕММА 1. *Справедлива формула*

$$\langle A, B\rangle = \mathrm{Sp}(AB^T). \tag{3}$$

Доказательство. При действиях с матрицами мы будем использовать индексную технику, описанную в [1].

Обозначим $N = 1:n$. Имеем

$$\mathrm{Sp}(AB^T) = \sum_{i=1}^{n} A[i,N] \times B^T[N,i] =$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n} A[i,j] \times B[i,j] = \langle A, B\rangle. \qquad \square$$

Очевидно, что $\mathrm{Sp}(C^T) = \mathrm{Sp}(C)$. Часто используется ещё одно свойство.

ЛЕММА 2. *Справедливо равенство*

$$\mathrm{Sp}(AB) = \mathrm{Sp}(BA). \tag{4}$$

Доказательство. Имеем

$$\mathrm{Sp}(AB) = \sum_{i=1}^{n} A[i,N] \times B[N,i] = \sum_{i=1}^{n}\sum_{j=1}^{n} A[i,j] \times B[j,i];$$
$$\mathrm{Sp}(BA) = \sum_{j=1}^{n} B[j,N] \times A[N,j] = \sum_{j=1}^{n}\sum_{i=1}^{n} B[j,i] \times A[i,j].$$

Отсюда очевидным образом следует (4). $\square$

Матрица $P \in \mathbb{R}^{n\times n}$ называется *ортогональной*, если

$$P^TP = PP^T = E.$$

ЛЕММА 3. *Если P и Q — ортогональные матрицы, то*

$$\|PCQ\| = \|C\|. \tag{5}$$

Доказательство. Согласно (3) и (4) имеем

$$\|PCQ\|^2 = \langle PCQ, PCQ\rangle = \mathrm{Sp}\big(PC(QQ^T)C^TP^T\big) =$$
$$= \mathrm{Sp}\big(P(CC^TP^T)\big) = \mathrm{Sp}\big((CC^TP^T)P\big) = \mathrm{Sp}(CC^T) = \|C\|^2.$$

Остаётся извлечь квадратный корень. $\square$

3°. Переходим к решению задачи (1). Как известно, для симметричной матрицы $A \in \mathbb{R}^{n\times n}$ существует ортогональная матрица P, такая, что

$$AP = P\Lambda, \tag{6}$$

где Λ — диагональная матрица, $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, на диагонали которой стоят вещественные собственные числа матрицы A. Введём диагональную матрицу

$$\Lambda^+ = \operatorname{diag}(\lambda_1^+, \ldots, \lambda_n^+),$$

где $\lambda_i^+ = \max\{0, \lambda_i\}$.

ТЕОРЕМА 1. *Единственным решением задачи* (1) *является матрица*

$$X_* = P\Lambda^+ P^T. \tag{7}$$

При этом

$$F(X_*) = \sum_{\{i|\lambda_i<0\}} \lambda_i^2. \tag{8}$$

Для доказательства нам потребуется следующее вспомогательное утверждение.

ЛЕММА 4. *Пусть D и X — симметричные неотрицательно определённые матрицы порядка n. Тогда*

$$\langle D, X\rangle \geqslant 0.$$

Доказательство. Воспользуемся разложением $D = QVQ^T$, где V — диагональная матрица с неотрицательными диагональными элементами и Q — — ортогональная матрица. Согласно леммам 1 и 2 имеем

$$\langle D, X\rangle = \operatorname{Sp}\bigl(Q(VQ^TX)\bigr) = \operatorname{Sp}\bigl(V(Q^TXQ)\bigr) =$$
$$= \sum_{i=1}^{n} V[i,N] \times (Q^TXQ)[N,i] = \sum_{i=1}^{n} V[i,i] \times (Q^TXQ)[i,i].$$

У матрицы $Y = Q^TXQ$ диагональные элементы неотрицательны. Действительно,

$$Y[i,i] = \langle Q^TXQe_i, e_i\rangle = \langle XQe_i, Qe_i\rangle \geqslant 0.$$

Значит,

$$\langle D, X\rangle = \sum_{i=1}^{n} V[i,i] \times Y[i,i] \geqslant 0. \qquad \square$$

Доказательство теоремы 1. Возьмём произвольную матрицу $X \in \mathcal{K}^n$. В силу (2) имеем

$$F(X) = \big\|(A - X_*) + (X_* - X)\big\|^2 = F(X_*) + \|X_* - X\|^2 + \\ +2\langle A - X_*, X_* - X\rangle.$$

Согласно (6), $A = P\Lambda P^T$. Учитывая формулу (7), получаем

$$\langle A - X_*, X_*\rangle = \mathrm{Sp}\big(P(\Lambda - \Lambda^+)(P^T P)\Lambda^+ P^T\big) = \\ = \mathrm{Sp}\big((\Lambda - \Lambda^+)\Lambda^+\big) = \sum_{i=1}^{n}(\lambda_i - \lambda_i^+)\lambda_i^+ = 0.$$

Значит,

$$F(X) = F(X_*) + \|X_* - X\|^2 + 2\langle X_* - A, X\rangle.$$

Матрица $X_* - A = P(\Lambda^+ - \Lambda)P^T$ симметрична и неотрицательно определена. По лемме 4, $\langle X_* - A, X\rangle \geqslant 0$. Приходим к неравенству

$$F(X) \geqslant F(X_*) + \|X_* - X\|^2. \tag{9}$$

Отсюда следует как оптимальность матрицы X_*, так и её единственность.

Формула (8) проверяется непосредственно:

$$F(X_*) = \|A - X_*\|^2 = \big\|P(\Lambda - \Lambda^+)P^T\big\|^2 = \\ = \|\Lambda - \Lambda^+\|^2 = \sum_{i=1}^{n}(\lambda_i - \lambda_i^+)^2 = \sum_{\{i|\lambda_i<0\}} \lambda_i^2.$$

Теорема доказана. □

Замечание. Неравенство (9) характеризует решение X_* задачи (1) как *сильно единственное.*

4°. В задаче (1) добавим ограничение $\mathrm{Sp}(X) \leqslant T$, где $T > 0$. Получим новую экстремальную задачу:

$$\begin{gathered} F(X) := \|A - X\|^2 \to \inf, \\ \mathrm{Sp}(X) \leqslant T, \quad X \in \mathcal{K}^n. \end{gathered} \tag{10}$$

Найдём её решение.

Обозначим через $\mathcal{G}^n$ множество планов задачи (10). Возьмём $X \in \mathcal{G}^n$. На основании формулы (6) и леммы 3 запишем

$$F(X) = \big\|P^T(A - X)P\big\|^2 = \|\Lambda - P^T X P\|^2.$$

Матрица $Y = P^T X P$ принадлежит конусу $\mathcal{K}^n$. Кроме того,

$$\mathrm{Sp}(Y) = \mathrm{Sp}\big(P^T(XP)\big) = \mathrm{Sp}\big(X(PP^T)\big) = \mathrm{Sp}(X) \leqslant T.$$

Обозначим $y_i = Y[i,i]$. Тогда

$$F(X) = \|\Lambda - Y\|^2 \geqslant \sum_{i=1}^{n} (\lambda_i - y_i)^2, \tag{11}$$

причём

$$\sum_{i=1}^{n} y_i \leqslant T; \quad y_i \geqslant 0,\, i \in 1:n. \tag{12}$$

Рассмотрим вспомогательную задачу: минимизировать функцию

$$f(y) = \sum_{i=1}^{n} (\lambda_i - y_i)^2$$

при ограничениях (12). Решение этой задачи существует и единственно. Обозначим его $\hat{y} = (\hat{y}_1, \ldots, \hat{y}_n)$. На основании (11) получаем

$$F(X) \geqslant f(\hat{y}) \quad \forall\, X \in \mathcal{G}^n. \tag{13}$$

Введём матрицу

$$\hat{X} = P\hat{Y}P^T, \tag{14}$$

где $\hat{Y} = \mathrm{diag}(\hat{y}_1, \ldots, \hat{y}_n)$.

ТЕОРЕМА 2. *Матрица* $\hat{X}$ *вида* (14) *является единственным решением задачи* (10).

Доказательство. Очевидно, что матрица $\hat{X}$ принадлежит множеству $\mathcal{G}^n$. При этом

$$F(\hat{X}) = \big\|P^T(A - \hat{X})P\big\|^2 = \|\Lambda - \hat{Y}\|^2 = \sum_{i=1}^{n} (\lambda_i - \hat{y}_i)^2 = f(\hat{y}).$$

Согласно (13)

$$F(X) \geqslant F(\hat{X}) \quad \forall\, X \in \mathcal{G}^n.$$

Оптимальность матрицы $\hat{X}$ установлена.

Равенство $F(X) = F(\hat{X})$ выполняется только тогда, когда $P^T X P = \hat{Y}$, то есть, когда $X = P\hat{Y}P^T = \hat{X}$. Это доказывает единственность решения задачи (10). □

Замечание. Для решения вспомогательной задачи минимизации функции $f(y)$ при ограничениях (12) разработан быстрый алгоритм [2].

5°. Обозначим через $\mathcal{M}^n_{\alpha,\beta}$ множество симметричных матриц порядка n, все собственные числа которых принадлежат отрезку $[\alpha, \beta]$. Рассмотрим ещё один вариант задачи (1):

$$F(X) := \|A - X\|^2 \to \inf_{X \in \mathcal{M}^n_{\alpha,\beta}}. \tag{15}$$

Укажем явное решение и этой задачи.

Возьмём спектральное разложение симметричной матрицы $A : A = P\Lambda P^T$, где $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ и P — ортогональная матрица. Введём диагональную матрицу

$$\check{\Lambda} = \mathrm{diag}(\check{\lambda}_1, \ldots, \check{\lambda}_n)$$

с диагональными элементами

$$\check{\lambda}_i = \begin{cases} \lambda_i, & \text{если } \lambda_i \in [\alpha, \beta]; \\ \alpha, & \text{если } \lambda_i < \alpha; \\ \beta, & \text{если } \lambda_i > \beta. \end{cases}$$

ТЕОРЕМА 3. *Матрица*

$$\check{X} = P\check{\Lambda}P^T$$

является единственным решением задачи (15).

Доказательство теоремы основано на следующем вспомогательном утверждении.

ЛЕММА 5. *Для диагональных элементов матрицы $D \in \mathcal{M}^n_{\alpha,\beta}$ выполняются неравенства*

$$\alpha \leqslant D[i,i] \leqslant \beta, \quad i \in 1:n. \tag{16}$$

Доказательство. Воспользуемся спектральным разложением симметричной матрицы D: $D = QVQ^T$, где Q — ортогональная матрица и $V = \mathrm{diag}(v_1, \ldots, v_n)$, причём по условию все v_i принадлежат отрезку $[\alpha, \beta]$. Запишем

$$D[i,i] = (QV)[i,N] \times Q^T[N,i] = \\ = \sum_{j=1}^n \Big(\sum_{k=1}^n Q[i,k] \times V[k,j]\Big) \times Q[i,j] = \sum_{j=1}^n v_j \big(Q[i,j]\big)^2. \tag{17}$$

Вместе с тем, из условия $QQ^T = E$ следует, что

$$1 = (QQ^T)[i,i] = Q[i,N] \times Q^T[N,i] = \sum_{j=1}^n \big(Q[i,j]\big)^2,$$

то есть

$$\sum_{j=1}^{n}\bigl(Q[i,j]\bigr)^2 = 1 \quad \text{при всех} \quad i \in 1:n. \tag{18}$$

На основании (17), (18) и неравенств $\alpha \leqslant v_i \leqslant \beta$, $i \in 1:n$, приходим к (16). $\square$

Доказательство теоремы 3. Возьмём план X задачи (15). Как и раньше, имеем

$$F(X) = \|\Lambda - Y\|^2, \tag{19}$$

где $Y = P^T X P$. Покажем, что матрицы Y и X имеют одни и те же собственные числа. Действительно, пусть $XQ = QV$. Тогда

$$P^T X Q = P^T Q V. \tag{20}$$

Обозначим $U = P^T Q$. Очевидно, что U — ортогональная матрица. При этом $Q = PU$. Перепишем равенство (20) в виде

$$(P^T X P)U = UV.$$

Это и означает, что спектры матриц X и $Y = P^T X P$ совпадают. По лемме 5 для величин $y_i = Y[i,i]$ выполняются неравенства

$$\alpha \leqslant y_i \leqslant \beta, \quad i \in 1:n.$$

В силу (19)

$$F(X) \geqslant \sum_{i=1}^{n}(\lambda_i - y_i)^2. \tag{21}$$

Рассмотрим вспомогательную задачу

$$f(y) := \sum_{i=1}^{n}(\lambda_i - y_i)^2 \to \inf$$
$$\alpha \leqslant y_i \leqslant \beta, \quad i \in 1:n.$$

Эта задача имеет единственное решение $\check{y} = (\check{y}_1, \dots, \check{y}_n)$, где $\check{y}_i = \check{\lambda}_i$. Учитывая (21), получаем

$$F(X) \geqslant f(\check{y}) \quad \forall\, X \in \mathcal{M}^n_{\alpha,\beta}. \tag{22}$$

Матрица $\check{X} = P\check{\Lambda}P^T$, указанная в формулировке теоремы 3, принадлежит множеству $\mathcal{M}^n_{\alpha,\beta}$. При этом

$$F(\check{X}) = \bigl\|P^T(A - \check{X})P\bigr\|^2 = \|\Lambda - \check{\Lambda}\| = \sum_{i=1}^{n}(\lambda_i - \check{\lambda}_i)^2 = f(\check{y}).$$

Согласно (22),

$$F(X) \geqslant F(\check{X}) \quad \forall\, X \in \mathcal{M}^n_{\alpha,\beta}.$$

Оптимальность матрицы $\check{X}$ установлена.

Равенство $F(X) = F(\check{X})$ выполняется только тогда, когда $P^T XP = \check{\Lambda}$, то есть, когда $X = P\check{\Lambda}P^T = \check{X}$. Это доказывает единственность решения задачи (15). □

6°. Близкие экстремальные задачи на множестве матриц рассматривались в работе [3].

ЛИТЕРАТУРА

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

2. Малозёмов В. Н., Тамасян Г. Ш. *Проектирование точки на телесный симплекс* // Семинар «DHA & CAGD». Избранные доклады. 11 октября 2013 г. (`http://dha.spb.ru/reps13.shtml#1011`)

3. Coope I. D., Renaud P. F. *Trace inequalities with applications to orthogonal regression and matrix nearness problems* // J. Inequalities in Pure and Applied Math., 2009. Vol. 10, No. 4 (Article 92. 7 pp.). (`http://jipam.vu.edu.au`)