ПРОЕКТИРОВАНИЕ СИММЕТРИЧНОЙ МАТРИЦЫ НА КОНУС НЕОТРИЦАТЕЛЬНО ОПРЕДЕЛЕННЫХ МАТРИЦ

В.Н.Малоземов
(Доклад на семинаре, 20.10.98)

1^o. Положим $N = 1 : n$. В линейном вещественном пространстве $\mathbb{R}^{NN}$ квадратных матриц $A = A[N, N]$ введем скалярное произведение

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

и норму $\|A\| = \langle A, A \rangle^{\frac{1}{2}}$. Через $\mathrm{tr}\,(A)$ обозначим след матрицы A :

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

ЛЕММА 1. *Справедлива формула*

$$\langle A, B \rangle = \mathrm{tr}\,(AB^T). \tag{1}$$

Доказательство. Имеем

$$\mathrm{tr}\,(AB^T) = \sum_{i=1}^{n} A[i, N] \cdot B^T[N, i] =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} A[i, j] \cdot B[i, j] = \langle A, B \rangle. \quad \blacksquare$$

ЛЕММА 2. *Справедлива формула*

$$\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA). \tag{2}$$

Д о к а з а т е л ь с т в о. Имеем

$$\mathrm{tr}\,(AB) = \sum_{i=1}^{n} A[i,N]\cdot B[N,i] = \sum_{i=1}^{n}\sum_{j=1}^{n} A[i,j]\cdot B[j,i],$$

$$\mathrm{tr}\,(BA) = \sum_{j=1}^{n} B[j,N]\cdot A[N,j] = \sum_{j=1}^{n}\sum_{i=1}^{n} B[j,i]\cdot A[i,j].$$

Отсюда очевидным образом следует (2). $\blacksquare$

ЛЕММА 3. *Справедлива формула*

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

Д о к а з а т е л ь с т в о. Очевидно, что

$$\mathrm{tr}\,(AB^T) = \mathrm{tr}\,\big((AB^T)^T\big) = \mathrm{tr}\,(BA^T).$$

Учитывая этот факт, получаем

$$\|A+B\|^2 = \langle A+B, A+B\rangle = \mathrm{tr}\,\big[(A+B)(A^T+B^T)\big] =$$

$$\mathrm{tr}\,(AA^T) + \mathrm{tr}\,(BB^T) + 2\mathrm{tr}\,(AB^T) = \|A\|^2 + \|B\|^2 + 2\langle A, B\rangle. \blacksquare$$

Матрица $P \in \mathbb{R}^{NN}$ называется ортогональной, если ее столбцы $p_1, \ldots, p_n$ ортонормированы. Другими словами, если $P^T P = E$. Поскольку в этом случае и $PP^T = E$, то наряду с P ортогональной является и матрица P^T.

ЛЕММА 4. *Пусть P и Q ортогональные матрицы. Тогда*

$$\|PAQ\| = \|A\|. \tag{3}$$

Д о к а з а т е л ь с т в о. Согласно (1) и (2) имеем

$$\|PAQ\|^2 = \mathrm{tr}\,\big(PA(QQ^T)A^T P^T\big) = \mathrm{tr}\,\big(AA^T(P^T P)\big) =$$

$$= \mathrm{tr}\,(AA^T) = \|A\|^2. \quad \blacksquare$$

2^o. Обозначим через $\mathcal{M}^n \subset \mathbb{R}^{NN}$ подпространство симметричных матриц, а через $\mathcal{K}^n \subset \mathcal{M}^n$ – конус симметричных неотрицательно определенных матриц. Очевидно, что $\mathcal{K}^n$ – выпуклый конус.

Известно [1, с.47], что для каждой матрицы $A \in \mathcal{M}^n$ существует ортогональная матрица P, такая что

$$P^T A P = \Lambda, \tag{4}$$

где $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ – диагональная матрица, на диагонали которой стоят собственные числа матрицы A. Из (4) следует, что

$$A = P \Lambda P^T = \sum_{i=1}^{n} \lambda_i p_i p_i^T. \tag{5}$$

Учитывая (3) и (5), приходим к равенству

$$\|A\|^2 = \|\Lambda\|^2 = \operatorname{tr}(\Lambda^2) = \sum_{i=1}^{n} \lambda_i^2.$$

ЛЕММА 5. *Пусть A и B матрицы из конуса $\mathcal{K}^n$, причем A ненулевая, а B положительно определенная. Тогда $\langle A, B\rangle > 0$.*

Доказательство. Так как матрица B положительно определена, то найдется неособенная матрица C, такая что

$$C^T A C = V, \quad C^T B C = E, \tag{6}$$

где V – диагональная матрица, на диагонали которой стоят обобщенные собственные числа $v_1, \ldots, v_n$ матрицы A [1, с.45]. Неотрицательная определенность матрицы A гарантирует неотрицательность чисел $v_1, \ldots, v_n$. Среди них есть и положительное, поскольку $A \neq \mathbf{0}$.

Согласно (6) имеем

$$\langle A, B\rangle = \operatorname{tr}\big((C^T)^{-1} V C^{-1} (C^T)^{-1} C^{-1}\big) =$$

$$= \operatorname{tr}\big(V (C^T C)^{-1} C^{-1} (C^T)^{-1}\big) = \operatorname{tr}\big(V (C^T C)^{-2}\big).$$

Матрица $C^T C$ симметричная и положительно определенная, поэтому в ее спектральном разложении

$$C^T C = \sum_{i=1}^{n} \mu_i q_i q_i^T \tag{7}$$

все собственные числа μ_i положительны. Из (7) следует равенство

$$(C^T C)^{-2} = \sum_{i=1}^{n} (\mu_i)^{-2} q_i q_i^T.$$

Учитывая его, получаем

$$\langle A, B\rangle = \mathrm{tr}\left(V(C^T C)^{-2}\right) = \sum_{i=1}^{n} (\mu_i)^{-2} \mathrm{tr}\,(V q_i q_i^T) =$$

$$= \sum_{i=1}^{n} (\mu_i)^{-2} \sum_{j=1}^{n} v_j (q_i[j])^2.$$

Числа $h_i = \sum_{j=1}^{n} v_j (q_i[j])^2$ неотрицательны. Остается показать, что не все они равны нулю.

Матрица Q, составленная из столбцов $q_1, \ldots, q_n$, ортогональна, поэтому $QQ^T = E$. В частности,

$$\sum_{i=1}^{n} (q_i[j])^2 = 1 \text{ при всех } \ j \in 1:n.$$

Теперь имеем

$$\sum_{i=1}^{n} h_i = \sum_{j=1}^{n} v_j \sum_{i=1}^{n} (q_i[j])^2 = \sum_{j=1}^{n} v_j.$$

Как отмечалось, $\sum_{j=1}^{n} v_j > 0$. Значит, среди неотрицательных чисел h_i имеется ненулевое. Лемма доказана. ■

Обозначим через $(\mathcal{K}^n)^+$ конус, сопряженный $\mathcal{K}^n$ в подпространстве $\mathcal{M}^n$, т.е.

$$(\mathcal{K}^n)^+ = \left\{ A \in \mathcal{M}^n \,\middle|\, \langle A, X\rangle \geq 0 \ \ \forall \ \ X \in \mathcal{K}^n \right\}.$$

ТЕОРЕМА 1. *Справедливо равенство*

$$(\mathcal{K}^n)^+ = \mathcal{K}^n.$$

Доказательство. Начнем с включения $\mathcal{K}^n \subset (\mathcal{K}^n)^+$. Зафиксируем матрицу A из $\mathcal{K}^n$. Нужно проверить, что

$$\langle A, X\rangle \geq 0 \ \ \forall \ \ X \in \mathcal{K}^n. \tag{8}$$

Если A – ненулевая матрица, то это очевидно. В дальнейшем считаем, что A – ненулевая матрица.

Матрице $X \in \mathcal{K}^n$ сопоставим матрицу $X_\varepsilon = X + \varepsilon E$, где $\varepsilon > 0$. Матрица X_ε симметрична и положительно определена. Согласно лемме 5, $\langle A, X_\varepsilon\rangle > 0$. Более подробно: $\langle A, X\rangle + \varepsilon\langle A, E\rangle > 0$. Переходя в этом неравенстве к пределу при $\varepsilon \to +0$, получаем (8).

Обратное включение $(\mathcal{K}^n)^+ \subset \mathcal{K}^n$ будем доказывать от противного. Пусть $A \in (\mathcal{K}^n)^+$, но $A \notin \mathcal{K}^n$. Последнее означает, что у A имеется отрицательное собственное число, например, $\lambda_1 < 0$. Запишем спектральное разложение матрицы A

$$A = \sum_{i=1}^{n} \lambda_i p_i p_i^T$$

и введем матрицу $X = p_1 p_1^T$. Имеем $X \in \mathcal{K}^n$ и

$$\langle A, X\rangle = \sum_{i=1}^{n} \lambda_i \langle p_i p_i^T, p_1 p_1^T\rangle = \sum_{i=1}^{n} \lambda_i \cdot \mathrm{tr}\,(p_i p_i^T p_1 p_1^T) =$$

$$= \lambda_1 \cdot \mathrm{tr}\,(p_1 p_1^T) = \lambda_1 < 0.$$

Но это противоречит условию $A \in (\mathcal{K}^n)^+$. Теорема доказана. ■

3^o. Зафиксируем симметричную матрицу A и, следуя С. Е. Михееву, рассмотрим экстремальную задачу

$$F(X) := \|X - A\| \to \inf_{X \in \mathcal{K}^n}. \tag{9}$$

Матрица A имеет спектральное разложение (5). Введем матрицу

$$A_+ = \sum_{i=1}^{n} (\lambda_i)_+ p_i p_i^T,$$

где $(\lambda_i)_+ = \max\{0, \lambda_i\}$.

ТЕОРЕМА 2. *Матрица A_+ является единственным решением задачи* (9). *При этом*

$$F(A_+) = \left[\sum_{\{i \,|\, \lambda_i < 0\}} (\lambda_i)^2 \right]^{\frac{1}{2}}.$$

Мы приведем два доказательства данной теоремы. Первое основано на теореме 1, второе использует подход С.Е.Михеева.

Первое доказательство. Обозначим

$$\Lambda_+ = \operatorname{diag}\big((\lambda_1)_+, \ldots, (\lambda_n)_+\big).$$

Имеем

$$A_+ - A = \sum_{i=1}^{n} [(\lambda_i)_+ - \lambda_i] p_i p_i^T = P(\Lambda_+ - \Lambda)P^T.$$

Очевидно, что собственные числа симметричной матрицы $A_+ - A$ неотрицательны. Это гарантирует ее неотрицательную определенность. Согласно теореме 1

$$\langle X, A_+ - A \rangle \geq 0 \quad \forall \quad X \in \mathcal{K}^n. \tag{10}$$

Отметим также, что

$$\langle A_+, A_+ - A \rangle = \operatorname{tr}\big(P\Lambda_+ P^T P(\Lambda_+ - \Lambda)P^T\big) =$$

$$= \operatorname{tr}\big(\Lambda_+(\Lambda_+ - \Lambda)\big) = \sum_{i=1}^{n} (\lambda_i)_+ \cdot [(\lambda_i)_+ - \lambda_i] = 0. \tag{11}$$

В силу леммы 3

$$\|X - A\|^2 = \|(X - A_+) + (A_+ - A)\|^2 =$$

$$= \|X - A_+\|^2 + \|A_+ - A\|^2 + 2\langle X - A_+, A_+ - A \rangle.$$

Для всех $X \in \mathcal{K}^n$ согласно (10) и (11) выполняется неравенство $\langle X - A_+, A_+ - A \rangle \geq 0$. Поэтому

$$\|X - A\|^2 \geq \|X - A_+\|^2 + \|A_+ - A\|^2 \quad \forall \quad X \in \mathcal{K}^n.$$

Отсюда очевидным образом следует заключение теоремы. ■

Второе доказательство. Согласно лемме 4 и формуле (4)

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

При $X \in \mathcal{K}^n$ матрица $Y = P^TXP$ также принадлежит $\mathcal{K}^n$. Учитывая это, получаем

$$\|X-A\|^2 \geq \inf_{Y\in\mathcal{K}^n} \|Y-\Lambda\|^2 \quad \forall \quad X \in \mathcal{K}^n. \tag{12}$$

Далее

$$\|Y-\Lambda\|^2 \geq \sum_{i=1}^n (Y[i,i]-\lambda_i)^2.$$

При $Y \in \mathcal{K}^n$ имеем $Y[i,i] = \langle Ye_i, e_i\rangle \geq 0$, поэтому

$$\|Y-\Lambda\|^2 \geq \inf_{y_i\geq 0} \sum_{i=1}^n (y_i-\lambda_i)^2 \quad \forall \quad Y \in \mathcal{K}^n. \tag{13}$$

Экстремальная задача

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

$$y_i \geq 0, \ i \in 1:n,$$

распадается на n независимых подзадач

$$(y_i-\lambda_i)^2 \to \inf,$$
$$y_i \geq 0,$$

решение которых очевидно: $y_i^* = (\lambda_i)_+$. При этом

$$\inf_{y_i\geq 0} \sum_{i=1}^n (y_i-\lambda_i)^2 = \sum_{i=1}^n ((\lambda_i)_+-\lambda_i)^2 = \sum_{\{i|\lambda_i<0\}} (\lambda_i)^2. \tag{14}$$

На основании (12) – (14) приходим к неравенству

$$\|X-A\|^2 \geq \sum_{\{i|\lambda_i<0\}} (\lambda_i)^2 \quad \forall \quad X \in \mathcal{K}^n. \tag{15}$$

Положим

$$Y^* = \operatorname{diag}(y_1^*, \dots, y_n^*) = \operatorname{diag}((\lambda_1)_+, \dots, (\lambda_n)_+) = \Lambda_+,$$

$$X^* = PY^*P^T = P\Lambda_+P^T = A_+.$$

Очевидно, что X^* – единственная матрица из $\mathcal{K}^n$, на которой неравенство (15) обращается в равенство. Теорема доказана. ■

4^o. Обозначим через G^n множество симметричных матриц, все собственные числа которых по модулю не превосходят единицы. Зафиксируем симметричную матрицу $A \in \mathcal{M}^n$ и рассмотрим экстремальную задачу

$$\|X - A\| \to \inf_{X \in G^n}. \tag{16}$$

Допустим, что A имеет спектральное разложение

$$A = \sum_{k=1}^{n} \lambda_k p_k p_k^T.$$

Введем функцию

$$s(t) = \begin{cases} t & \text{при} \quad |t| \le 1, \\ \operatorname{sign}(t) & \text{при} \quad |t| > 1. \end{cases}$$

ТЕОРЕМА. *Единственным решением задачи* (16) *является матрица* X_* *вида*

$$X_* = \sum_{k=1}^{n} s(\lambda_k) p_k p_k^T.$$

Доказательство. Отметим, что

$$\|X - A\|^2 = \|P^T(X - A)P\|^2 = \|P^TXP - \Lambda\|^2. \tag{17}$$

Матрица $Y = P^TXP$ имеет те же собственные числа, что и X. Действительно, пусть $XQ = QV$, где V – диагональная матрица, а Q – ортогональная. Тогда $P^TXQ = P^TQV$. Положим $U = P^TQ$. В этом случае $Q = PU$ и

$$(P^TXP)U = (P^TP)UV = UV,$$

что и требовалось проверить.

Согласно (17) для любой матрицы $X \in G^n$ справедливо неравенство
$$\|X - A\|^2 \geq \inf_{Y \in G^n} \|Y - \Lambda\|^2. \tag{18}$$

Далее,
$$\|Y - \Lambda\|^2 \geq \sum_{i=1}^{n} (Y[i,i] - \lambda_i)^2. \tag{19}$$

Положим $y_i = Y[i,i]$ и покажем, что $|y_i| \leq 1$ при всех $i \in 1:n$. Разложим орты e_i по ортонормированному базису, состоящему из собственных векторов матрицы Y :

$$e_i = \sum_{j=1}^{n} \alpha_{ij} u_j, \quad \sum_{j=1}^{n} \alpha_{ij}^2 = 1.$$

Поскольку собственные числа v_j матрицы Y по модулю не превосходят единицы, то

$$|y_i| = |\langle Ye_i, e_i\rangle| = \left|\langle \sum_{j=1}^{n} v_j \alpha_{ij} u_j, \sum_{j=1}^{n} \alpha_{ij} u_j \rangle\right| =$$

$$= \left|\sum_{j=1}^{n} v_j \alpha_{ij}^2\right| \leq \sum_{j=1}^{n} \alpha_{ij}^2 = 1.$$

Теперь из (19) следует, что

$$\|Y - \Lambda\|^2 \geq \inf_{|y_i| \leq 1} \sum_{i=1}^{n} (y_i - \lambda_i)^2. \tag{20}$$

Экстремальная задача

$$\sum_{i=1}^{n} (y_i - \lambda_i)^2 \to \inf,$$

$$|y_i| \leq 1, \quad i \in 1:n,$$

распадается на n независимых подзадач

$$(y_i - \lambda_i)^2 \to \inf_{|y_i| \le 1}.$$

Решение каждой такой задачи очевидно:

$$y_i^* = s(\lambda_i),$$

где s – введенная выше функция. Таким образом,

$$\inf_{|y_i| \le 1} \sum_{i=1}^{n} (y_i - \lambda_i)^2 = \sum_{i=1}^{n} (s(\lambda_i) - \lambda_i)^2. \tag{21}$$

На основании (18)–(21) заключаем, что

$$\|X - A\|^2 \ge \sum_{i=1}^{n} (s(\lambda_i) - \lambda_i)^2 \quad \forall \quad X \in G^n.$$

Равенство в этом неравенстве достигается лишь тогда, когда

$$X_* = PY_*P^T, \quad Y_* = \mathrm{diag}\,(s(\lambda_1), \ldots, s(\lambda_n)).$$

Теорема доказана. ■

ЗАМЕЧАНИЕ. Определим $D^n \subset \mathcal{M}^n$ как множество симметричных матриц, спектр (вектор собственных чисел) которых принадлежит n - мерному множеству Ω. Предположим, что для каждой матрицы $Y \in D^n$ диагональный вектор

$$(Y[1,1], \ldots, Y[n,n])$$

также принадлежит Ω. Тогда экстремальная задача

$$\|X - A\|^2 \to \inf_{X \in D^n}$$

сводится к более простой задаче:

$$\begin{gathered} \textstyle\sum_{i=1}^{n} (y_i - \lambda_i)^2 \to \inf, \\ (y_1, \ldots, y_n) \in \Omega. \end{gathered}$$

ЛИТЕРАТУРА

1. Малоземов В.Н. Линейная алгебра без определителей. Квадратичная функция. СПб, 1997. 80 с.