Гл 4 Разложения матриц

§ 1 Сингулярные разложение

Пусть $A \in M_{m,n}(\mathbb{C})$, $X \in \mathbb{C}^n$, $y \in \mathbb{C}^m$.

Определим $\mathcal{A}: \mathbb{C}^n \to \mathbb{C}^m$ по правилу $\mathcal{A}(x) = Ax$

Обозн. $A^* = \overline{A}^T$, т.е. $a^*_{ij} = \overline{a_{ji}}$. Определим $\mathcal{A}^*: \mathbb{C}^m \to \mathbb{C}^n$ по правилу $\mathcal{A}^*(y) = A^*y$. Тогда $A^*A: \mathbb{C}^n \to \mathbb{C}^n$, $AA^*: \mathbb{C}^m \to \mathbb{C}^m$.

<u>Утв</u> Матрицы A^*A и AA^* явл. эрмитовыми и неотрицательными.

Д-во: Эрмитово скал. произв на $\mathbb{C}^n$ определяется по формуле $\langle v, w \rangle = \sum\limits_{j=1}^n v_j \overline{w_j}$, где $v = (v_1, \dots, v_n)^T \in \mathbb{C}$, $w = (w_1, \dots, w_n)$

Покажем эрмитово-сть $A^*A = A^*(A^*)^* = A^*A$.

Аналогично для эрм. AA^*. Покажем неотр. A^*A, т.е. $\langle A^*Av, v \rangle \geq 0$. Имеем

$\langle A^*Av, v \rangle = (A^*(Av))^T \overline{v} = (Av)^T (A^*)^T \overline{v} = (Av)^T \overline{Av} = \langle Av, Av \rangle \geq 0$.

Неотр. AA^* док-ся аналог.

Лемма. Собств. векторы, отвечающие разл. с.ч. эрмитовой матрицы ортогональны, и все с.ч. её вещ.

След-но $\exists$ ОНБ $e_1, \ldots, e_n \in \mathbb{C}^n$ такой, что $A^* A e_k = p_k^2 e_k$, $k = 1, \ldots$, и для некот. $p_k \geq 0$.

Кроме того, $|Ae_k| = \sqrt{\langle Ae_k, Ae_k \rangle} = \sqrt{\langle A^* A e_k, e_k \rangle} = \sqrt{p_k^2 \langle e_k, e_k \rangle} = p_k$

ОПР. Величины p_k наз. сингулярными числами м-цы A.

Т. Ненулевые собств. числа м-цы AA^* и A^*A совпадают.

Д-во: Пусть $v = Ae_k \in \mathbb{C}^m \setminus \{0\}$. $AA^* v = A(A^*A)e_k = p_k^2 Ae_k = p_k^2 v$ $\boxtimes$

Пусть таких (ненулевых) собств. чис. будет t штук $p_1^2 \geq \ldots \geq p_t^2 > 0$. Возьм. следующую с-тему единичных векторов

$\varepsilon_1 = \frac{Ae_1}{|Ae_1|} = \frac{Ae_1}{p_1}, \ldots, \varepsilon_t = \frac{Ae_t}{|Ae_t|} = \frac{Ae_t}{p_t}$. Тогда

$$A e_k = \begin{cases} \rho_k \varepsilon_k, & k \le r \\ 0, & k > r \end{cases} \quad (1), \qquad A^* \varepsilon_k = \begin{cases} \rho_k e_k, & k \le r \\ 0, & k > r \end{cases} \quad (2)$$

Набор $\varepsilon_1, \dots, \varepsilon_r$ дополним ОНБ.

$\varepsilon_1, \dots, \varepsilon_r, \varepsilon_{r+1}, \dots, \varepsilon_m$

ОПР ОНБ $e_1, \dots, e_n$ и $\varepsilon_1, \dots, \varepsilon_m$ наз. сингулярными базисами матрицы A

Обозн. $m \times n$-матрицу $\begin{pmatrix} \rho_1 & & 0 \\ & \rho_2 & \\ 0 & & \ddots \end{pmatrix}$ через Λ,

и через V обозн. квадр. м-цу, столбцы которой составлены из столбцов ε_k, а через V^* – квадр м-цы, составл. из столбцов e_k. Тогда соотношения (2) в матр. виде означают $A V^* = V \Lambda$. По построению м-цы V и V^* явл. унитарными. Поэтому $V^* V = E$ и $V = (V^*)^*$ – унитарная. Умножая (3) справа на V, получаем

$$A = V \Lambda V \quad (4)$$

Опр Разложение (4) наз. ~~сит~~ сингулярным разложением матрицы A

1) При $m=n$ $\prod_{i=1}^{n} \rho_i^2 = |\det A|^2$

2) Сумма квадратов модулей всех сингулярных чисел м-цы A равна сумме квадратов модулей всех элементов м-цы A

Д-во:

1) $|\det A|^2 = (\det A)\overline{(\det A)} = \det A \cdot \det A^* = \det(A^*A) = \prod_{i=1}^{n} \rho_i^2$

2) Следует из очевидного замечания, что длина кажд. столбца (строки) не меняется при умн-ии слева (справа) ~~A~~ на унитарную матрицу

Полярное разложение

Т. Любую квадр. м-цу A можно представить в виде $A = HU$ (1), где H - эрмитова матри-

$\Leftarrow$ для неотр. собств. чисел, V - унитарная

$\Rightarrow$ во $A = V \Lambda V^* = (V \Lambda V^*)(V V^*) = HQ$

Q - унитарна, H - эрм. $\square$

ОПР Разложения (1) наз-ся полярным разложением матрицы A

Пример $A = z \in \mathbb{C} = M_{1,1}(\mathbb{C})$ $\quad A^* A = \bar{z} z = |z|^2$

$\Lambda = |z|$. $\quad A = z = \underbrace{|z|}_{H} \cdot \underbrace{e^{i\alpha}}_{Q}$

гл. 5 Итерационные методы

§ 1. Метод градиентного спуска

Пусть $f: \mathbb{R}^n \to \mathbb{R}$ - непр. дифф. Рассм. $\frac{dx}{dt} +$

$+ \nabla f(x(t)) = 0$

$$\frac{df(x(t))}{dt} = \left\langle \nabla f(x(t)), \frac{dx}{dt} \right\rangle \overset{(1)}{=} \langle \nabla f(x), -\nabla f(x) \rangle =$$

$$= -\langle \nabla f(x), \nabla f(x) \rangle \leq 0$$

Значит $\frac{d}{dt} f(x(t)) < 0$ вне за искл. стац. точки ф-ции $f(x)$.

На ЭВМ $t_k = t_0 + k\tau$ $\quad \frac{dx}{dt}\Big|_{t=t_k} \simeq \frac{x(t_{k+1}) - x(t_k)}{2}$

Поэтому $x(t_k) \approx x_k$, где

$$x_{k+1} = x_k - \alpha \nabla f(x_k)$$

$$x_0 \xrightarrow{(2)} x_1 \xrightarrow{(2)} \dots \to x_n \to x_k$$

$$\|\nabla f(x_k)\| < \varepsilon$$

При разных способах выбора ~~тр~~ параметра α получаем алгоритмы градиентного спуска.

§ 2 Метод тяжёлого шарика

Рассм. движение матер. т. поверхности $y = f(x)$ в поле тяжести. ~~Преди~~ Предп., что трение пропорц. скорости и точка не может отрываться от поверхности. Тогда её движение описывается дифф. ур-ем

$$\frac{d^2x}{dt^2} + \gamma \frac{dx}{dt} + \frac{\nabla f(x)}{|f| \, \|\nabla f\|^2} = 0, \quad \gamma > 0 \quad (1)$$

Ясно, что решение этого ур-я с течением времени устанавливается в некоторой т. минимума ф-ии $f(x)$. Вблизи т. экстр. $\|\nabla f(x)\| \ll 1$ и дифф. ур-е (1) близко к уравнению

$$\frac{d^2x}{dt^2} + \gamma \frac{dx}{dt} + \nabla f(x) = 0$$

$$\frac{d}{dt}\left(\frac{1}{2}\left\langle\frac{dx}{dt},\frac{dx}{dt}\right\rangle + f(x(t))\right) = \left\langle\frac{dx}{dt},\frac{d^2x}{dt^2}\right\rangle$$

$$+\left\langle\nabla f,\frac{dx}{dt}\right\rangle = \left\langle\frac{dx}{dt}, -\gamma\frac{dx}{dt}\right\rangle = -\gamma\left\langle\frac{dx}{dt},\frac{dx}{dt}\right\rangle,$$

если $\frac{dx}{dt} \neq 0$. Значит $f(x(t))$ возр.,

что $\left\|\frac{dx}{dt}\right\| \to 0$ при $t \to \infty$

Пусть $a = \lim\limits_{t\to\infty} x(t)$. Тогда из (2)

$\nabla f|_{x=a} = 0$. Значит, мы нашл. т. стац. т. ф-ии $f(x)$

$$\frac{d^2x}{dt}(t_k) = \frac{x(t_{k+1}) - 2x(t_k) + x(t_{k-1})}{t^2}$$

$x(t_k) \approx x_k$, где

$x_{k+1} = x_k - \alpha\nabla f(x_k) + B(x_k - x_{k+1})$, $0 \le \beta < 1$

§ Ускоренный спуск Нестерова

$y_0 = x_0$

$x_1 = y_0 - \alpha\nabla f(y_0)$

$y_1 = x_1 + \beta(x_1 - x_0)$

$x_{k+1} = y_k - \alpha\nabla f(y_m)$

$y_{k+1} = x_{k+1} + \beta(x_{+1}, x_k)$