Лекция 4

Основы матричных вычислений

Рахуба М.В.

08.02.21

Малоранговое приближение матриц 3

① Alternating least squares (ALS)

Рассмотрим $f : \mathbb{R}^{m \times n} \to \mathbb{R}$

Хотим $f(X) \to \min\limits_{X:\ \operatorname{rank}(X) \le r}$

$f(X) = \|A - X\|_F^2$ — задача о наилучш. прибл. A матрицей ранга r

поэлем. (Адамарово) произв.

$f(X) = \|P_\Omega \circ (A - X)\|_F^2$ — задача восст. матрицы (matrix completion)

$$(P_\Omega)_{ij} = \begin{cases} 1, & (i,j) \in \Omega \\ 0, & \text{иначе} \end{cases}$$

	user1	user2	user3
movie1	?	?	5
movie2	3	2	2
movie3	5	?	5

$$P_\Omega = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

$$\min_{X:\,\mathrm{rank}(X)\le r} f(X) = \min_{\underbrace{U,V}_{r \text{ столбцов}}} f(UV^T)$$

$\boxed{\text{Алгоритм 1}}$ (ALS vanilla)

U_0, V_0, например, случайные.

$$U_{k+1} = \operatorname*{argmin}_{U\in\mathbb{R}^{n\times r}} f(U V_k^T)$$

$$V_{k+1} = \operatorname*{argmin}_{V\in\mathbb{R}^{n\times r}} f(U_{k+1} V^T)$$

$\boxed{\text{Алгоритм 2}}$ (ALS с ортогонализацией)

$$\hat{U} = \operatorname*{argmin}_{U} f(U V_k^T)$$

$$\hat{U} = Q_1 R_1 \quad - \quad QR \text{ разлож}$$

Не теряем в общности т.к.

$\{Q_1 R_1 V^T : V\in\mathbb{R}^{n\times r}\} \subseteq \{Q_1 V^T : V \in \mathbb{R}^{n\times r}\}$ $\forall\, R_1 \in \mathbb{R}^{r\times r}$

$$\hat{V} = \operatorname*{argmin}_{V} f(Q_1 V^T)$$

$$\hat{V} = Q_2 R_2$$

$$\left(X \approx Q_1 \hat{V}^T = \underbrace{(Q_1 R_2^T)}_{U_{k+1}} \underbrace{Q_2^T}_{V_{k+1}^T}\right)$$

$$U_{k+1} = Q_1 R_2^T, \qquad V_{k+1} = Q_2$$

(не используется дальше. Нужно только если выходим из процесса)

Как решать $f(UV^T) \to \min_V$, где $U^T U = I$ и $f(UV^T) = \|A - UV^T\|_F^2$?

1. $\langle A, B \rangle_F = \mathrm{Tr}(A^T B)$

Фроб. скаляр. произв.

$$\langle A, A \rangle_F = \mathrm{Tr}(A^T A) = \mathrm{Tr}(A A^T) = \|A\|_F^2$$

2. $f(X + \Delta X) = f(x_{11} + \Delta x_{11}, \ldots, x_{mn} + \Delta x_{mn}) =$

$$= \underbrace{f(x_{11}, \ldots, x_{mn})}_{= f(X)} + \overbrace{\sum_{i,j} \frac{\partial f}{\partial x_{ij}} \Delta x_{ij}}^{df(\Delta x_{ij})} + o(\|\Delta X\|_F)$$

$$\sum_{i,j} \frac{\partial f}{\partial x_{ij}} \Delta x_{ij} = \left\langle \frac{\partial f}{\partial X}, \Delta X \right\rangle_F$$

$$\frac{\partial f}{\partial X} \overset{def}{=} \left\{ \frac{\partial f}{\partial x_{ij}} \right\}_{i,j=1}^{m,n}$$

В точке оптимума

$$\frac{\partial f}{\partial X} = 0$$

Утв.

1) $\frac{\partial}{\partial X} \underbrace{\langle A, X \rangle_F}_{f_1(X)} = A$

2) $\frac{\partial}{\partial X} \underbrace{\langle X, X \rangle_F}_{f_2(X)} = 2X$

$$\square \quad f_1(X+dX) = \langle A, X+dX\rangle_F =$$

$$= \langle A, X\rangle_F + \langle \overset{\partial f_1/\partial X}{A}, dX\rangle_F + 0$$

$$f_2(X+dX) = \langle X+dX, X+dX\rangle_F =$$

$$= \langle X, X\rangle_F + \overbrace{2\langle X, dX\rangle_F}^{\partial f_2/\partial X} + \underbrace{\langle dX, dX\rangle_F}_{}\quad \|dX\|_F^2 = o(\|dX\|_F)$$

$$\blacksquare$$

3. Пусть $V^T V = I$

$$f(UV^T) = \|A - UV^T\|_F^2 =$$

$$= \langle A - UV^T, A - UV^T\rangle_F =$$

$$= \langle A, A\rangle_F - 2\langle A, UV^T\rangle_F + \langle UV^T, UV^T\rangle_F$$

$$\langle A, UV^T\rangle_F = \mathrm{Tr}(A^T U V^T) = \mathrm{Tr}(V^T A^T U) = \langle AV, U\rangle_F$$

$$\langle UV^T, UV^T\rangle_F = \mathrm{Tr}(V U^T U V^T) = \mathrm{Tr}(U^T U) = \langle U, U\rangle_F$$

$$= \langle A, A\rangle_F - 2\langle AV, U\rangle_F + \langle U, U\rangle_F$$

$$\frac{\partial f}{\partial U} = -2AV + 2U = 0$$

$$\Downarrow$$

$U_* = AV$ — оптимальное U при $V^TV = I$

$\left(V_* = A^TU \right.$ для $\min\limits_{V} \|A - UV^T\|_F^2$ при $U^TU = I \left. \right)$

② Кронекерово произведение

Опр.

$$\underbrace{\underset{\in \mathbb{C}^{m\times n}}{A} \otimes \underset{\in \mathbb{C}^{p\times q}}{B}}_{\in \mathbb{C}^{mp\times nq}} = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}$$

Пример

$$I \otimes A = \begin{pmatrix} A & & 0 \\ & A & \\ 0 & & \ddots \\ & & & A \end{pmatrix}$$

Утв.

1) $(A+B)\otimes C = A\otimes C + B\otimes C$

$A\otimes(B+C) = A\otimes B + A\otimes C$

2) $(A\otimes B)^* = A^* \otimes B^*$

3) $A\otimes B = (A\otimes I_p)(I_n\otimes B) = (I_m\otimes B)(A\otimes I_q)$

4) $(AB)\otimes(CD) = (A\otimes C)(B\otimes D)$

5) $(A\otimes B)^{-1} = A^{-1}\otimes B^{-1}$

$\square$

3) $$(A\otimes I_p)(I_n\otimes B) = \begin{pmatrix} a_{11}I_p & \dots & a_{1n}I_p \\ \vdots & & \vdots \\ a_{m1}I_p & \dots & a_{mn}I_p \end{pmatrix}\begin{pmatrix} B & & 0 \\ & \ddots & \\ 0 & & B \end{pmatrix} =$$

$$= A\otimes B$$

4), 5) — семинар

$\blacksquare$

$\otimes$ — ассоц.: $(A\otimes B)\otimes C = A\otimes(B\otimes C)$,

но не коммутатив.: например, $A\otimes I \neq I\otimes A$ вообще говоря

$\boxed{\text{Опр.}}$ $\mathrm{vec}(\cdot): \mathbb{C}^{m\times n} \to \mathbb{C}^{mn}$ — (векторизация)

$$\mathrm{vec}\left(\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}\right) = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \\ a_{12} \\ \vdots \\ a_{n2} \\ \vdots \\ a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Vec в Python:

.) np. reshape (A, [m*n, 1], order='f'), (mn,1)

.) np. flatten (A, order='f'), (mn,)

vec и $\otimes$:

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \otimes \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} u_1 v_1 \\ u_2 v_1 \\ u_1 v_2 \\ u_2 v_2 \end{pmatrix} = \operatorname{vec}\begin{pmatrix} u_1 v_1 & u_1 v_2 \\ u_2 v_1 & u_2 v_2 \end{pmatrix} = u v^T$$

$$\operatorname{vec}(u v^T) = v \otimes u$$

$$u v^T = \begin{pmatrix} u_1 v_1 & \dots & u_1 v_n \\ \vdots & & \vdots \\ u_m v_1 & \dots & u_m v_n \end{pmatrix}, \qquad v \otimes u = \begin{pmatrix} u_1 v_1 \\ \vdots \\ u_m v_1 \\ \vdots \\ u_1 v_n \\ \vdots \\ u_m v_n \end{pmatrix}$$

$\boxed{\text{Утв}}$ $\operatorname{vec}(AXB) = (B^T \otimes A)\operatorname{vec}(X)$

$\square$ Пусть $X = u v^T$

$$\operatorname{vec}(A u v^T B) = \operatorname{vec}\left((Au)(B^T v)^T\right) =$$

$$= (B^T v) \otimes (Au) = (B^T \otimes A)(v \otimes u) =$$

$$= B^T \otimes A \operatorname{vec}(u v^T)$$

любой X — сумма ранг-1 матриц $\blacksquare$

3

$m = n \gg r$

np.linalg.svd $\qquad O(n^3)$

randomized $\qquad O(n^2 r)$

ALS $\qquad O(n^2 r)$

$C \hat{A}^{-1} R \qquad O(n r^2)$

кратко в начале след. лекции