Лекция 3. Основы матричного анализа 3

Максим Рахуба

ФКН ВШЭ

Основы матричных вычислений 24/25

Февраль 4, 2024

Факты из лекции 2

План

QR разложение

Скелетное разложение

Проекторы

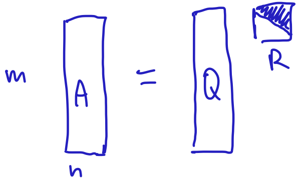
Приближение матрицами меньшего ранга

QR разложение

Теорема (QR разложение)

Для $A \in \mathbb{F}^{m\times n}$, $m \geq n$ найдется $Q \in \mathbb{F}^{m\times n}$ с ортонормированными столбцами ($Q^*Q = I$) и верхнетреугольная $R \in \mathbb{F}^{n\times n}$:

$$A = QR$$



$$Q^*Q = I$$

image

$\{Qx,\ x \in \mathbb{R}^n\}$

$$QQ^* \neq I \text{ при } m > n$$

$$\mathrm{Im}(Q) \supseteq \mathrm{Im}(A)$$

QR разложение

Теорема (QR разложение)

Для $A \in \mathbb{F}^{m\times n}$, $m \geq n$ *найдется* $Q \in \mathbb{F}^{m\times n}$ *с ортонормированными столбцами* ($Q^*Q = I$) *и верхнетреугольная* $R \in \mathbb{F}^{n\times n}$*:*

$$A = QR$$

$$\begin{bmatrix} a_1 \dots a_n \end{bmatrix} = \begin{bmatrix} q_1 \dots q_n \end{bmatrix} \begin{bmatrix} * & * & \dots & * \\ 0 & * & & \vdots \\ \vdots & 0 & & \vdots \\ 0 & 0 & \dots & * \end{bmatrix}$$

Доказательство.

В полноранговом случае Грам-Шмидт (ГШ):

$$\begin{aligned} \widetilde{q}_1 &= a_1, & q_1 &= \widetilde{q}_1/\|\widetilde{q}_1\|_2, \\ \widetilde{q}_2 &= a_2 - q_1(q_1^* a_2), & q_2 &= \widetilde{q}_2/\|\widetilde{q}_2\|_2, \\ &\vdots & &\vdots \\ \widetilde{q}_n &= a_n - q_1(q_1^* a_n) - \ldots - q_{n-1}(q_{n-1}^* a_n), & q_n &= \widetilde{q}_n/\|\widetilde{q}_n\|_2. \end{aligned}$$

В случае неполного ранга: мусто $m = n$

$$A = U \Sigma V^*$$

rank(A)

$$A_k = U \begin{pmatrix} \sigma_1 \dots \sigma_r & \frac{1}{k} & \frac{1}{k} \dots \end{pmatrix} V^* \to A, \quad k \to \infty$$

QR разложение

$$A_k = Q_k R_k$$

$\{Q : Q^* Q = I\}$ — компактное (т.к. конечномерн. то нужно огр. и замк.)

$\|Q\|_2 = 1 \Rightarrow$ огр.

$f(Q) = Q^* Q - I = 0$, f — непрерыв. ф-я

замкн. множ.

$f^{-1}(\{0\})$ — тоже замкнут.

$Q_{k_j} \to Q$, $j \to \infty$

сходящ. подп.

$A_{k_j} = Q_{k_j} R_{k_j} \Rightarrow R_{k_j} = Q_{k_j}^* A_{k_j}$ тоже сх-ся

QR разложение

$[A\ \ldots] = Q\,R$

для прямоугольных как подматрица

1. Процесс Грама-Шмидта неустойчив! (семинар)
2. Мы познакомимся с устойчивыми алгоритмами позже, а пока пользуемся

$$\texttt{Q, R = np.linalg.qr(A)}.$$

3. Количество операций с плавающей точкой:

$$\#\mathtt{FLOPs} = \mathcal{O}(mn^2).$$

План

QR разложение

Скелетное разложение

Проекторы

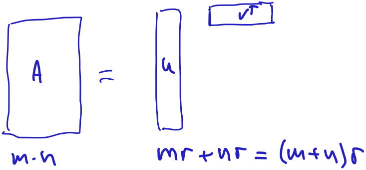
Приближение матрицами меньшего ранга

Скелетное разложение

Теорема

Для любой матрицы $A \in \mathbb{F}^{m \times n}$, $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ ранга r найдутся $U \in \mathbb{F}^{m \times r}$ и $V \in \mathbb{F}^{n \times r}$:

$$A = UV^\top.$$



□ $\exists\ u_1, \ldots, u_r$ — базис пр-ва столбцов A

$\Downarrow$

$$a_i = [u_1 \ldots u_r] \begin{bmatrix} c_{1i} \\ \vdots \\ c_{ri} \end{bmatrix} = U c_i$$

$$[a_1 \ldots a_n] = [Uc_1 \ldots Uc_n] = U \underbrace{[c_1 \ldots c_n]}_{V^T} = UV^T$$

■

Виды записи скелетного разложения

Индексная запись (разделенные переменные i и j):

$$a_{ij} = \sum_{\alpha=1}^{r} u_{i\alpha} v_{j\alpha}.$$

Сумма ранг-1 слагаемых:

$$A = u_1 v_1^\top + \cdots + u_r v_r^\top = \begin{bmatrix} u_1 \cdots u_r \end{bmatrix} \begin{bmatrix} v_1^\top \\ \cdots \\ v_r^\top \end{bmatrix}$$

Неединственность и приведение к виду SVD

$$A = U V^T = \overbrace{(U S)}^{\tilde U}\overbrace{(S^{-1} V^T)}^{\tilde V^T} = \tilde U \tilde V^T$$

Скетное $\to$ SVD

$$A = U V^T = Q_u R_u (Q_v R_v)^T = Q_u \overbrace{(R_u R_v^T)}^{\substack{SVD \\ \hat U \hat \Sigma \hat V^T}} Q_v^T =$$

$$= \underbrace{(Q_u \hat U)}_{\hat U} \hat \Sigma \underbrace{(\hat V^T Q_v^T)}_{\hat V^T} = \hat U \hat \Sigma \hat V^T$$

Вид "компактного" SVD (без 0 стлб. стр.)

План

QR разложение

Скелетное разложение

Проекторы

Приближение матрицами меньшего ранга

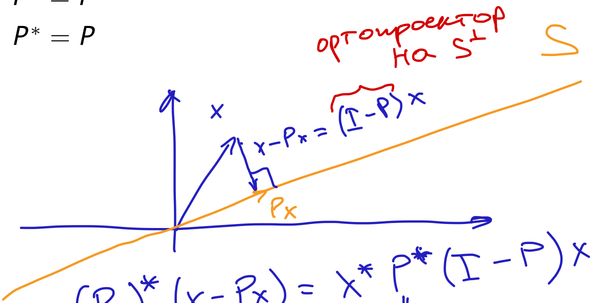
Ортопроектор

Определение

$P \in \mathbb{C}^{n\times n}$ — ортопроектор на $S \subseteq \mathbb{C}^n$, если

1. $\mathrm{Im}(P) = S$
2. $P^2 = P$
3. $P^* = P$

(пункты 1–2:) проектор на S



$$(Px)^*(x - Px) = x^* \underset{P}{\underset{\parallel}{P^*}} (I - P)x = x^*(P - \overset{P}{\overset{\parallel}{P^2}})x = 0$$

Ортопроектор

$A(A^T A)^{-1} A^T$

Теорема

Пусть $A \in \mathbb{C}^{m \times n}$ полного ранга и $m \geq n$. Тогда ортопроектор на $\mathrm{Im}(A)$ *имеет вид:*

$$P = A(A^* A)^{-1} A^*.$$

$\square$ 1) $P^2 = A(A^*A)^{-1} \cancel{A^* A (A^* A)^{-1}} A^* = A(A^*A)A^{-1}$

$P^* = P$

2) $Px = A\left((A^*A)^{-1} A^* x\right) \Rightarrow \mathrm{Im}(P) \subseteq \mathrm{Im}(A)$

$PA = A$

$Ax = PAx \Rightarrow \mathrm{Im}(P) \supseteq \mathrm{Im}(A)$

$\Downarrow$

$\mathrm{Im}(P) = \mathrm{Im}(A)$ $\blacksquare$

Ортопроектор

B – верхнетреуг. для QR
B > 0 для полярного
$B = \Sigma V^*$ для SVD

Но QR самое эффективное ("быстрее" всех)

Теорема

Пусть $A \in \mathbb{C}^{m \times n}$ полного ранга и $m \geq n$. Тогда ортопроектор на $\mathrm{Im}(A)$ *имеет вид:*

$$P = A(A^*A)^{-1}A^*.$$

Если $A = Q\,B$,

То $\mathrm{Im}(A) = \mathrm{Im}(Q)$

Следствие

Для ортогонального базиса Q имеем

$$P = QQ^*.$$

На практике ортогонализуем столбцы A через QR.

$$P = Q(Q^*Q)^{-1}Q^* = QQ^*$$

Проекторы на 4 основных подпространства A

полное SVD

но везде $\mathbb{C}^{m \times n}$

$*$ – эрмит. сопряж.

Утверждение

Пусть SVD матрицы A ранга r имеет вид:

$$A = \begin{bmatrix} U_r & \widetilde{U}_r \end{bmatrix} \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_r & \widetilde{V}_r \end{bmatrix}^\intercal.$$

Тогда

$$\begin{aligned}
U_r U_r^\intercal &= \text{ортопроектор на } \operatorname{Im}(A) \\
\widetilde{V}_r \widetilde{V}_r^\intercal &= \text{ортопроектор на } \ker(A) \\
\widetilde{U}_r \widetilde{U}_r^\intercal &= \text{ортопроектор на } \operatorname{Im}(A)^\perp \equiv \ker(A^\intercal) \\
V_r V_r^\intercal &= \text{ортопроектор на } \ker(A)^\perp \equiv \operatorname{Im}(A^\intercal)
\end{aligned}$$

$$A = U_r U_r^* A = A V_r V_r^* = U_r U_r^* A V_r V_r^*$$

План

QR разложение

Скелетное разложение

Проекторы

Приближение матрицами меньшего ранга

Унитарно-инвариантные нормы

Определение

Норма $\|\cdot\|$ — унитарно-инвариантная, если $\forall\ A \in \mathbb{C}^{m\times n}$ и $\forall$ унитарных $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$:

$$\|UAV\| = \|A\|.$$

вектор сингл. чисел

$$\|A\|_2 = \sigma_1(A) = \|\sigma\|_\infty$$

$$\|A\|_F = \sqrt{\sigma_1^2(A) + \ldots + \sigma_r^2(A)} = \|\sigma\|_2$$

$$\|A\|_* \overset{def}{=} \|\sigma\|_1 = \sigma_1(A) + \ldots + \sigma_r(A)$$ — ядерная норма

$$\|A\|_{p,\,\text{shatten}} = \|\sigma\|_p$$ — Шаттеновские нормы

Унитарно-инвариантные нормы

Определение

Норма $\|\cdot\|$ — унитарно-инвариантная, если $\forall\ A \in \mathbb{C}^{m\times n}$ и $\forall$ унитарных $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$:

$$\|UAV\| = \|A\|.$$

Теорема (Эккарта-Янга-Мирского)

Пусть $k < \operatorname{rank}(A)$ и

$$= U \begin{pmatrix} \sigma_1 \dots \sigma_k & 0 \dots & 0 \end{pmatrix} V^* =$$

$$A_k = U_k \Sigma_k V_k^* \quad \textit{(усеченное SVD)}. \quad = U_k \begin{pmatrix} \sigma_1 \dots \sigma_k \end{pmatrix} V_k^*$$

тогда

$$\min_{B:\ \operatorname{rank}(B)\leq k} \|A - B\| = \|A - A_k\| = \begin{cases} \sigma_{k+1} & \text{для } \|\cdot\|_2 \\ \sqrt{\sigma_{k+1}^2 + \dots} & \text{для } \|\cdot\|_F \\ \sigma_{k+1} + \dots & \text{для } \|\cdot\|_* \end{cases}$$

для любой унитарно-инвариантной $\|\cdot\|$.

Литература

1. Е. Е. Тыртышников. *Методы численного анализа*. Академия, 2007.
2. Е. Е. Тыртышников. *Матричный анализ и линейная алгебра*. Физматлит, 2007.
3. G.H. Golub and C.F. Van Loan. *Matrix Computations*, 4th Edition. The Johns Hopkins University Press, Baltimore, 2013.

Доска