Лекция 3

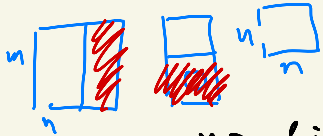
Основы матричных вычислений

Рахуба М.В.

01.02.21

Малоранговое приближение матриц 2

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



thin SVD

np.linalg.svd(A, full_matrices=False)

Сложность: $O(mn \min(m,n)) \overset{m=n}{=} O(n^3)$

Применимо для матриц с $m, n \sim 10^3 - 10^4$.

1 QR разложение

$$A = [a_1, \dots, a_n] \in \mathbb{C}^{m \times n}, \quad m \geq n, \quad \operatorname{rank}(A) = n$$

Хотим получить

$$Q = [q_1, \dots, q_n] \in \mathbb{C}^{m \times n} \qquad (Q^* Q = I)$$

ортог. базис в $\operatorname{Im}(A)$

Грам-Шмидт:

$$\tilde{q}_1 = a_1, \qquad q_1 = \frac{\tilde{q}_1}{\|\tilde{q}_1\|_2}$$

$$\tilde{q}_2 = a_2 - q_1 (q_1^* a_2), \qquad q_2 = \frac{\tilde{q}_2}{\|\tilde{q}_2\|_2}$$

$\vdots$

$$[a_1 a_2 \dots a_n] = [q_1 q_2 \dots q_n] \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & * \end{bmatrix}$$

$A = QR$ — QR-разложение.

$O(mn^2) = O(mn\min(m,n))$

Сложность: np.linalg.qr

② Операции с малоранговыми матрицами

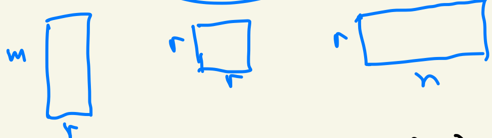
a) Приведение скелетного разл. к компактному SVD

$$A = UV^*, \quad U \in \mathbb{C}^{m\times r}, \quad V \in \mathbb{C}^{n\times r}$$

$$Q_u R_u = U \quad - \quad O(mr^2)$$

$$Q_v R_v = V \quad - \quad O(nr^2)$$

$$A = UV^* = Q_u \underbrace{R_u R_v^*}\, Q_v^*$$



$\hat{U}\Sigma\hat{V}^* =$ np.linalg.svd$(R_u R_v^*)$: $O(r^3)$

$$A = \underbrace{(Q_u \hat{U})}_{\tilde{U}}\,\Sigma\,\underbrace{(Q_v \hat{V})}_{\tilde{V}}{}^* = \tilde{U}\Sigma\tilde{V}^*$$

$$\tilde{U}^*\tilde{U} = \hat{U}^* \underbrace{Q_u^* Q_u}_{I} \hat{U} = I$$

$O((n+m)r^2 + r^3)$

$\ll O(mn\min(m,n))$

при $r \ll \min(m,n)$

(В этих нер-вах воспринимайте $O(\cdot)$ как константу)

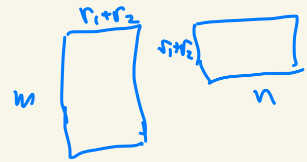
б) Уменьшение ранга

$$A = U V^*$$

- •) Приводим к SVD   а)
- •) Зануляем "маленькие" синг. числа

в) Арифм. операции

$$A_1 + A_2 = U_1 V_1^* + U_2 V_2^* = [U_1 U_2][V_1 V_2]^*$$



и т.д. $(A_1 A_2,\ A_1 \odot A_2, \dots)$

поэлементное (Адамарово произв.)

3 Ортопроектор

Опр. 1) $P \in \mathbb{C}^{n \times n}$ — ортопроектор, если

$P^2 = P$ (косой проектор)

и

$P^* = P$

2) P — ортопроектор на S, если $\mathrm{Im}(P) = S$

вект. подпр-во

Утв. 1 Если $v \in \mathrm{Im}(P)$, то $Pv = v$

проектор

$\square$ $Pv = PPx = P^2 x = Px = v$ $\blacksquare$

Утв. 2 Если P – ортопроектор, то $(I-P)$ – ортопроектор

$\square$ $(I-P)^* = I - P$

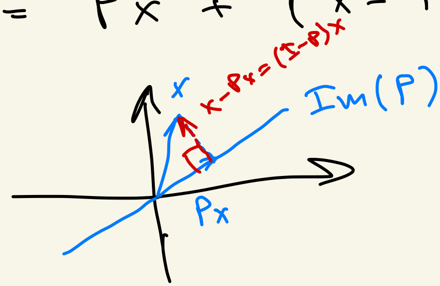
$$(I-P)^2 = I - P - P + P^2 = I - P \quad \blacksquare$$

Утв. 3 $\forall\, x, y \in \mathbb{C}^n \quad Px \perp (I-P)y$

$$\square \quad (Px)^*(I-P)y = x^* \underset{\overset{\shortparallel}{P}}{P^*} (I-P)y =$$

$$= x^*(P-P^2)y = x^*(P-P)y = 0 \quad \forall\, x, y \in \mathbb{C}^n \quad \blacksquare$$

$$x = Px + (x - Px) = Px + (I-P)x$$



Утв. 4 Пусть $A \in \mathbb{C}^{m \times n}$ – полного ранга, $m \geq n$. Тогда ортопроектор на $\mathrm{Im}(A)$:

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

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

$$P^* = P$$

$\mathrm{Im}(P) \subseteq \mathrm{Im}(A)$ т.к. $Px = A[(A^*A)^{-1}A^*x]$.

$$PA = A\cancel{(A^*A)^{-1}}\cancel{A^*A} = A$$

$$\Downarrow$$

$$\mathrm{Im}(A) \subseteq \mathrm{Im}(P) \Rightarrow \mathrm{Im}(P) = \mathrm{Im}(A)$$

$\blacksquare$

На практике не надо считать $(A^*A)^{-1}$!

$$A = QR$$

$$P = QR\big(\underbrace{(QR)^*(QR)}_{R^*Q^*QR}\big)^{-1}(QR)^* =$$

$$= QR(R^*R)^{-1}R^*Q^* =$$

$$= Q\cancel{R}\cancel{R^{-1}}\cancel{(R^*)^{-1}}\cancel{R^*}Q^* = QQ^*$$

4 Простейший рандом. алгоритм SVD

$\boxed{\text{Утв. 5}}$ Пусть $Q \in \mathbb{C}^{m\times r}$, $Q^* Q = I_r$ — задано.

$$\min_{B:\ \mathrm{Im}(B) \subseteq \mathrm{Im}(Q)} \|\underbrace{A}_{\in \mathbb{C}^{m\times n}} - B\|_F = \|\underbrace{A - QQ^*A}_{(I - QQ^*)A}\|_F$$

$$\square \quad \|A - \underbrace{QC}_{=B}\|_F^2 = \|A - [Q\ Q^\perp]\begin{bmatrix} C \\ 0 \end{bmatrix}\|_F^2 =$$

$m \times n$, $m \times r$, $r \times n$

$$[Q\ Q^\perp]^* = \begin{bmatrix} Q^* \\ (Q^\perp)^* \end{bmatrix}$$

$$= \left\| \begin{bmatrix} Q^* \\ (Q^\perp)^* \end{bmatrix} A - \begin{bmatrix} C \\ 0 \end{bmatrix} \right\|_F^2 = \left\| \begin{bmatrix} Q^*A - C \\ (Q^\perp)^* A \end{bmatrix} \right\|_F^2 =$$

$$= \|Q^*A - C\|_F^2 + \underbrace{\|(Q^\perp)^* A\|_F^2}_{\text{const}} \to \min_C$$

$$C = Q^*A \Rightarrow B = QC = QQ^*A \qquad \blacksquare$$

Как выбрать Q?

Идеально выбрать $Q = U_r$ — r первых сингл. левых вект. A.

$$QQ^*A = U_r U_r^* \underbrace{U}_{[\underset{r}{U_r}\ \underset{m-r}{U_r^\perp}]} \Sigma V^* = U_r \Sigma_r V_r^*$$

$$\left(\begin{array}{l} U_r^* U = [\,I \;\; 0\,] \\ U_r^* U \Sigma = [\,I \;\; 0\,] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} = [\,\Sigma_1 \;\; 0\,] \end{array}\right)$$

U_r мы не знаем (как-раз хотим найти).

Будем действовать по-другому:

1) $\Omega = [\omega_1, \ldots, \omega_r] \in \mathbb{C}^{n \times r}$, $r = \underset{\text{желаемый ранг}}{k} + p$

np.random.randn(n,r) (каждый элемент из $N(0,1)$)

2) $\underset{[y_1 \ldots y_n]}{Y} = A\Omega$ $\quad$ ($y_i = A\omega_i \in \mathrm{Im}(A)$) ~ лин. комб. столбцов A со случайными коэффициентами.

3) $QR = Y$ $\quad$ $O(mr^2)$ – QR-разлож Y

$\Downarrow$

умножение Q^*A $O(mnr)$ $\quad$ SVD от Q^*A $O(nr^2)$ $\quad$ QW: $O(mr^2)$

$$B = Q(Q^*A) = Q(W\Sigma V^*) = (QW)\Sigma V^*$$

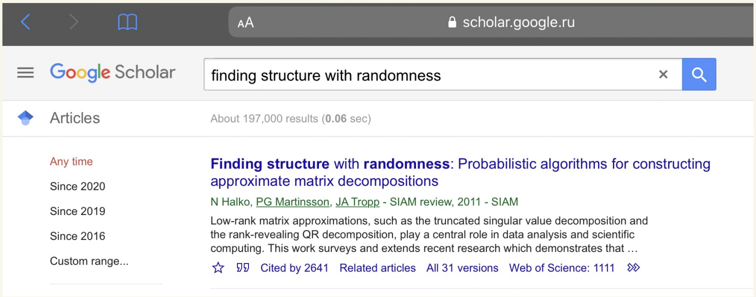
4) B Σ оставить k первых сингл. чисел.

Итог: $O(mnr) \ll O(mn\min(m,n))$

Можно генерировать Ω с доп. структурой и ещё уменьшить сложность.

Ранд. алгоритмы хорошо подходят для матриц, которые не помещаются в быструю память.

Сходимость алгоритма (теор 1.1)
(не входит в программу курса)



https://arxiv.org/pdf/0909.4061.pdf