Лекция 7

Основы матричных вычислений

Рахуба М.В.

01.03.21

Псевдообратные матрицы и метод наим. квадратов

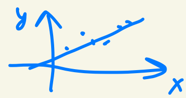
$Ax = b$, $\quad A \in \mathbb{R}^{m \times n}$, $\quad b \in \mathbb{R}^m$

•) $m = n$, $\det(A) \neq 0 \Rightarrow \exists!$ реш., $x = A^{-1} b$

•) $m > n$, вообще говоря $\nexists \Rightarrow \|Ax - b\|_2 \to \min\limits_x$

•) $m < n$, бесконеч. много реш. $\Rightarrow \|x\|_2 \to \min$

$y_i = a x_i + b$, $\begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ — фиттинг прямой $y = ax + b$ по $\{(x_1, y_1), \dots, (x_n, y_n)\}$



1 Полноранговый случай

$m \geq n$, $\quad \operatorname{rank}(A) = n$

$$J(x) = \|Ax - b\|_2^2 \to \min_x$$

$$\frac{\partial J}{\partial x} = 2A^T(Ax - b) = 0, \qquad \frac{\partial^2 J}{\partial x^2} = 2A^T A > 0$$

$$A^T A x = A^T b$$ — система нормал. уравн.

$$x = \underbrace{(A^T A)^{-1} A^T}_{A^+} b$$ — псевдорешение

A^+ — псевдообр. матрица (Мура-Пенроуза) при $m \geq n$, $\operatorname{rank}(A) = n$ — частный случай

$$x = A^+ b$$

Способы решения:

а) $A = QR$ (thin QR), $A^TA = R^T \underbrace{Q^TQ}_{I} R = R^TR$

$$\cancel{R^T} R x = \cancel{R^T} Q^T b$$

$$Rx = Q^T b \qquad x_n = \frac{f_n}{r_{nn}}$$

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}, \qquad x_{n-1} = \frac{f_{n-1} - r_{n-1,n} x_n}{r_{n-1,n-1}}$$

$$x = R^{-1} Q^T b \qquad O(n^2);$$

Итог: $2mn^2 - \frac{2}{3}n^3 + O(mn)$

$$Q^T b = \left(H_n \dots H_1 \begin{bmatrix} I_n \\ 0 \end{bmatrix}\right)^T b$$

б) Через SVD

$$A = U \Sigma V^T \qquad O(mn^2)$$

$$A^TA = V \Sigma^T U^T U \Sigma V^T = V \Sigma^2 V^T$$

$$\cancel{V} \Sigma^{\cancel{2}} V^T x = \cancel{V} \cancel{\Sigma} U^T b$$

$$\Sigma V^T x = U^T b \qquad (V^T)^{-1} = V$$

$$V^T x = \Sigma^{-1} U^T b$$

$$x = \underline{V \Sigma^{-1} U^T} b$$

A^+ через SVD

② Общий случай

Опр. Пусть $A = U \Sigma V^*$, $\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$ — полное SVD ($r = \operatorname{rank}(A)$) и $\Sigma^+ \overset{def}{=} \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}$, тогда

$$A^+ \overset{def}{=} V \Sigma^+ U^*$$ — псевдообратн. Мура – Пенроуза

$$\left(= V_r \Sigma_r^{-1} U_r^* \right)$$

Замеч. A^+ – единств. матрица, удовл. (эквив. способ определить A^+)

- $A A^+ A = A$,
- $A^+ A A^+ = A^+$
- $(A A^+)^* = A A^+$
- $(A^+ A)^* = A^+ A$

$P_1 = A A^+$ – ортопроектор на $\operatorname{Im}(A)$ (ДЗ)

$P_2 = A^+ A$ на $\operatorname{Im}(A^*)$

$\Downarrow$

$P_3 = I - A^+ A$ на $\ker(A)$ $(\operatorname{Null}(A))$

Теор Все решения $\|Ax - b\|_2 \to \min\limits_x$ имеют вид

$$x = A^+ b + (I - A^+ A) y, \ \forall y,$$

а $x_* = A^+ b$ имеет среди реш. $\min \|x\|_2$

□ $A = \underbrace{U}_{m\times m} \Sigma V^T = {}_m\big[\underbrace{U_r}_{r}\ \underbrace{U_r^\perp}_{m-r}\big] \begin{bmatrix}\Sigma_r & 0\\ 0 & 0\end{bmatrix} \big[V_r\ V_r^\perp\big]^*$

$U_r^* U_r^\perp = 0 \qquad \mathrm{Im}(V_r^\perp) = \ker(A)$

$$\|Ax - b\|_2^2 = \|U^*(U\Sigma V^* x - b)\|_2^2 = \|\Sigma \underbrace{V^* x}_{\alpha} - U^* b\|_2^2$$

$$= \left\| \begin{bmatrix}\Sigma_r & 0\\0&0\end{bmatrix}\begin{bmatrix}\alpha_r\\ \alpha_r^\perp\end{bmatrix} - \begin{bmatrix}U_r^* b\\ (U_r^\perp)^* b\end{bmatrix}\right\|_2^2 = \left\|\begin{bmatrix}\Sigma_r\alpha_r - U_r^* b\\ -(U_r^\perp)^* b\end{bmatrix}\right\|_2^2 =$$

$\left\|\begin{bmatrix}y_1\\y_2\end{bmatrix}\right\|_2^2 = \|y_1\|_2^2 + \|y_2\|_2^2$

$$= \|\Sigma_r\alpha_r - U_r^* b\|_2^2 + \|(U_r^\perp)^* b\|_2^2 \to \min_\alpha$$

$$\Sigma_r \alpha_r = U_r^* b$$

$\alpha_r = \Sigma_r^{-1} U_r^* b$, $\quad \alpha_r^\perp$ — любое

$$x = V\alpha = V\begin{bmatrix}\alpha_r\\ \alpha_r^\perp\end{bmatrix} = \overset{[V_r\, V_r^\perp]}{V}\begin{bmatrix}\alpha_r\\0\end{bmatrix} + V\begin{bmatrix}0\\ \alpha_r^\perp\end{bmatrix} =$$

$$= V_r\alpha_r + V_r^\perp \alpha_r^\perp = \underbrace{V_r\Sigma_r^{-1}U_r^* b}_{A^+ b} + V_r^\perp\alpha_r^\perp$$

$$x = V\begin{bmatrix}\alpha_r\\ \alpha_r^\perp\end{bmatrix}$$

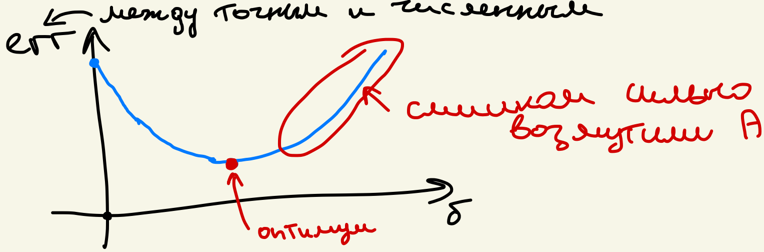
$$\|x\|_2 = \left\|\begin{bmatrix}\alpha_r\\ \alpha_r^\perp\end{bmatrix}\right\|_2 \to \min_\alpha \Rightarrow \alpha_r^\perp = 0 \Rightarrow x = A^+ b$$

$\blacksquare$

$\boxed{\text{Замечание}}$ На практике $A_\delta^+ = V\Sigma_\delta^+ U^*$, где Σ_δ, у котор. отброшены синг. числа $\le \delta$.

np. linalg. pinv(A, δ)

Это регуляризирует задачу



③ Регуляризация

а) Тихонова

парам. регуляриз.

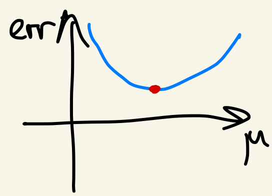
$$J_\mu(x) = \|Ax-b\|_2^2 + \underset{>0}{\mu}\|x\|_2^2 \to \min$$

(a.k.a. ridge regression)

$$\frac{\partial J_\mu}{\partial x} = 2A^T(Ax-b) + 2\mu x = 0$$

$$(A^TA + \mu I)x = A^Tb$$

$$x = \underbrace{(A^TA+\mu I)^{-1}A^T}_{B(\mu)\xrightarrow[\mu\to 0]{} A^+} b$$



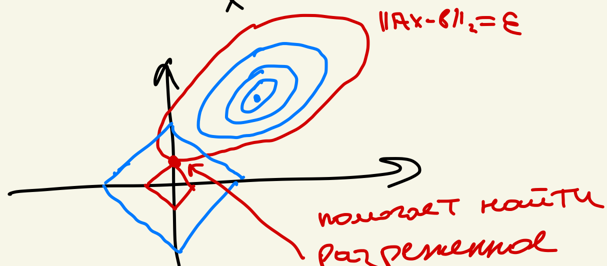
б) Lasso

$$\|Ax-b\|_2^2 + \mu(\varepsilon)\|x\|_1 \to \min_x$$

$$\Updownarrow$$

$$\min \|x\|_1$$

$$\|Ax-b\|_2 \le \varepsilon$$

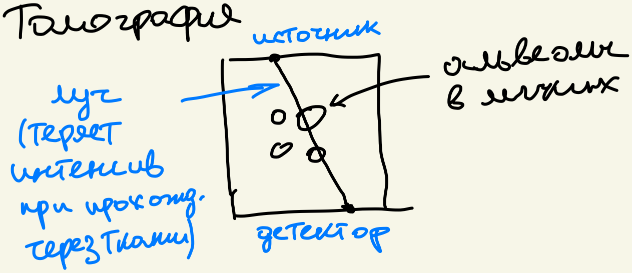


Смотрите

https://www.researchgate.net/publication/234022297_Compressed_Sensing_Theory_and_Applications

Compressed sensing - очень популярная тема в начале 2000-х

Томография



Почти вся картинка из 0, поэтому хорошо подходит lasso (Д3)

любая картинка ≈ разрежена в вейвлет базисе (JPEG2000).

https://sundoc.bibliothek.uni-halle.de/diss-online/02/03H033/t4.pdf

Литература

[1]: Sec. 5.5.1, 5.5.2, 5.5.4, 6.1.4

[3]: Lecture 11

(см Вики страницу)

$$\underset{m\times m}{U}\ \underset{m\times n}{\Sigma}\ \underset{n\times n}{V^T} = U\begin{bmatrix}\Sigma_n\\ 0\end{bmatrix}_{m\times n} V^T =$$

$$= \underset{m\times n}{U_n}\ \underset{n\times n}{\Sigma_n}\ \underset{n\times n}{V^T} = \underset{m\times r}{U_r}\ \underset{r\times r}{\Sigma_r}\ \underset{r\times n}{V_r^T}$$

Thin SVD   Compact SVD

$$\underset{m\times n}{A} = \underset{m\times m}{\begin{bmatrix} Q_1 & Q_2 \end{bmatrix}} \underset{m\times n}{\begin{bmatrix} R_1 \\ 0 \end{bmatrix}} = Q_1 R_1$$

— thin QR