Сингулярное разложение и решение линейных систем

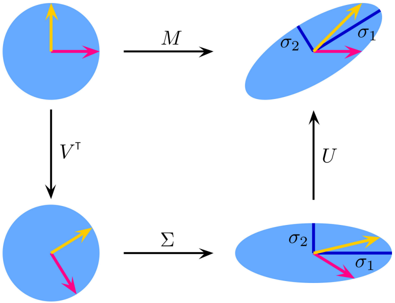
Сергей Кривохатский

СПбГУ, Современное программирование

7 октября 2022 г.

Сингулярное разложение I

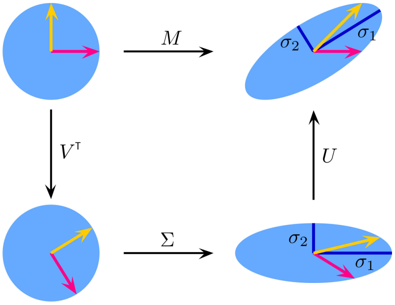
- пусть $\boldsymbol{M}$ — вещественнозначная матрица $m \times n$
- тогда $\boldsymbol{M} = \boldsymbol{U\Sigma V}^{\mathsf{T}}$ — *сингулярное разложение* (*SVD*)
- $\boldsymbol{U}$ — ортогональная матрица $m \times m$
- $\boldsymbol{\Sigma}$ — диагональная матрица $m \times n$ с неотрицательными элементами
- $\boldsymbol{V}$ — ортогональная матрица $n \times n$



$$M = U \cdot \Sigma \cdot V^{\mathsf{T}}$$

Сингулярное разложение II

- диагональные элементы называются *сингулярными числами*
- будем считать, что сингулярные числа расположены в порядке невозрастания



$$M = U \cdot \Sigma \cdot V^{\mathsf{T}}$$

Про матрицы вида $\boldsymbol{A}^\intercal \boldsymbol{A}$ I

Рассмотрим сингулярное разложение

$$\boldsymbol{A} = \boldsymbol{U}\Sigma\boldsymbol{V}^\intercal = \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \\ & & \end{pmatrix} \begin{pmatrix} v_1^\intercal \\ \vdots \\ v_m^\intercal \end{pmatrix}$$

где

- u_i — вектора-столбцы размерности n
- v_i — вектора-столбцы размерности m
- σ_i — сингулярные числа

Про матрицы вида $\boldsymbol{A}^\intercal \boldsymbol{A}$ II

Подставим сингулярное разложение матрицы $\boldsymbol{A}$ в $\boldsymbol{A}^\intercal \boldsymbol{A}$

$$\begin{aligned}
\boldsymbol{A}^\intercal \boldsymbol{A} &= (\boldsymbol{V}\Sigma^\intercal \boldsymbol{U}^\intercal)(\boldsymbol{U}\Sigma\boldsymbol{V}^\intercal) = \\
&= \boldsymbol{V}\Sigma^\intercal\Sigma\boldsymbol{V}^\intercal = \\
&= \boldsymbol{V}\operatorname{diag}(\sigma_1^2, \ldots, \sigma_m^2)\boldsymbol{V}^\intercal \\
&= \boldsymbol{V}\Lambda\boldsymbol{V}^\intercal
\end{aligned}$$

$\boldsymbol{V}$ — ортогональная матрица, значит

- v_i — собственные вектора матрицы $\boldsymbol{A}^\intercal \boldsymbol{A}$,
- $\sigma_i^2 \geqslant 0$ — собственные числа

Линейные наименьшие квадраты I

Задача линейных наименьших квадратов

$$(\boldsymbol{A}x - b)^\intercal(\boldsymbol{A}x - b) \to \min$$

Из лекции про оптический поток мы помним: чтобы найти минимум, нужно решить систему

$$\boldsymbol{A}^\intercal\boldsymbol{A}x = \boldsymbol{A}^\intercal b$$

Если матрица $\boldsymbol{A}^\intercal\boldsymbol{A}$ обратима, то

$$x = (\boldsymbol{A}^\intercal\boldsymbol{A})^{-1}\boldsymbol{A}^\intercal b$$

Линейные наименьшие квадраты II

Подставим в $(\boldsymbol{A}^\intercal\boldsymbol{A})^{-1}\boldsymbol{A}^\intercal b$ сингулярное разложение $\boldsymbol{A}$

$$x = (\boldsymbol{V}\Lambda\boldsymbol{V}^\intercal)^{-1}(\boldsymbol{U}\Sigma\boldsymbol{V}^\intercal)^\intercal b = (\boldsymbol{V}\Lambda^{-1}\boldsymbol{V}^\intercal)(\boldsymbol{V}\Sigma^\intercal\boldsymbol{U}^\intercal)b =$$
$$= \boldsymbol{V}\begin{pmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_m \end{pmatrix}\boldsymbol{U}^\intercal b = \boldsymbol{V}\Sigma^+\boldsymbol{U}^\intercal b$$

Получили псевдообратную к $\boldsymbol{A}$ матрицу $\boldsymbol{A}^+$

$$\boldsymbol{A}^+\boldsymbol{A} = (\boldsymbol{V}\Sigma^+\boldsymbol{U}^\intercal)(\boldsymbol{U}\Sigma\boldsymbol{V}^\intercal) = \mathbf{1}$$

Линейные наименьшие квадраты III

Если у $\boldsymbol{A}$ сингулярные числа $\sigma_i = 0$ для $i > r$, матрица $\boldsymbol{A}^\mathsf{T}\boldsymbol{A}$ вырождена, задача имеет множество решений, а $\boldsymbol{\Sigma}^+$ имеет вид

$$\begin{pmatrix} 1/\sigma_1 & & & \\ & \ddots & & \\ & & 1/\sigma_r & \\ & & & \\ \end{pmatrix}$$

Можно показать, что и в таком случае $x = \boldsymbol{A}^+ b$ будет решением, причем минимальным в смысле $\|x\|$

Линейные наименьшие квадраты IV

Рассмотрим систему $\boldsymbol{S}y = c$

$$\begin{pmatrix} s_1 & & & \\ & \ddots & & \\ & & s_r & \\ & & & \\ \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_r \\ \vdots \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_r \\ \vdots \end{pmatrix}$$

$$\boldsymbol{S}^+ c = \begin{pmatrix} c_1/s_1 & \cdots & c_r/s_r & 0 & \cdots \end{pmatrix}^\intercal$$

будет ее точным решением, если $c_i = 0$ для $i > r$, иначе решением в смысле наимешьших квадратов

$$\begin{aligned} \boldsymbol{A}x &= b \\ \boldsymbol{U\Sigma V}^\intercal x &= b \\ \boldsymbol{\Sigma}(\boldsymbol{V}^\intercal x) &= \boldsymbol{U}^\intercal b \end{aligned}$$

$\implies$ $\boldsymbol{\Sigma}^+\boldsymbol{U}^\intercal b$ — решение для $\boldsymbol{V}^\intercal x$, значит

$$x = \boldsymbol{V\Sigma}^+\boldsymbol{U}^\intercal b = \boldsymbol{A}^+ b$$

будет решением $\boldsymbol{A}x = b$

Решение СЛАУ с помощью SVD

СЛАУ можно решить с помощью сингулярного разложения

- решить переопределенную систему $\boldsymbol{A}x = b$ в смысле $\|\boldsymbol{A}x - b\|^2 \to \min$ можно вычислив $\boldsymbol{A}^+$: $x = \boldsymbol{A}^+ b$
- недоопределенная система решается аналогично
- легко заметить, что так же можно решить и определенную систему (вычислив $\boldsymbol{A}^{-1}$)
- такое решение подходит как метод по умолчанию
- для частных случаев бывают более эффективные методы

Спасибо за внимание!