线性代数 第25讲

12月6日

奇异值分解

上一讲要点回顾

奇异值分解

广义逆

矩阵的谱范数

对称正定矩阵

定义6.2.1 (正定矩阵) 给定 n 阶实矩阵 A, 如果对任意非零向量 $x \in R^n$, 都有$x^T Ax > 0$, 则称 A 正定.

命题 6.2.2 对实对称矩阵 A，以下叙述等价：

1. A 正定；
2. A 的特征值都是正数；
3. 存在可逆矩阵 T，使得 $A = TT^{\mathrm{T}}$；
4. A 有 $\mathrm{LDL}^{\mathrm{T}}$ 分解，且 D 的对角元素都是正数；
5. A 的顺序主子式都是正数；
6. A 的顺序主子阵都正定.

定义 6.2.3 给定 n 阶实矩阵 A，如果对任意非零向量 $\boldsymbol{x} \in \mathbb{R}^n$，都有

1. $\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} > 0$，则称矩阵 A **正定**，如前定义；
2. $\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} \geqslant 0$，则称矩阵 A **半正定**；
3. $\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} < 0$，则称矩阵 A **负定**；
4. $\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} \leqslant 0$，则称矩阵 A **半负定**；

如果 A 不满足以上任何一种条件，则称 A **不定**.

命题 6.2.4 对实对称矩阵 A，以下叙述等价：

1. A 半正定；
2. A 的特征值都是非负数；
3. 存在矩阵 T，使得 $A = TT^{\mathrm{T}}$；
4. A 存在 $\mathrm{LDL}^{\mathrm{T}}$ 分解，且 D 的对角元素都是非负数.

练习6.2.11 举例说明，
实对称矩阵 A 的所有顺序主子式都非负，但 A 并不半正定

练习6.2.13 证明，
矩阵 A 为实对称矩阵半正定，当且仅当它的所有主子式都非负

合同变换

定义 6.2.6 (合同) 对方阵 A，如果存在可逆矩阵 X 使得 $X^{\mathrm{T}}AX = B$，则称 A 和 B **合同**，或 A 合同于 B.

命题 6.2.7 方阵的合同关系是等价关系.

命题 6.2.8 对实对称矩阵 A，存在可逆矩阵 X，使得 $X^{\mathrm{T}}AX = J = \begin{bmatrix} I_p & & \\ & -I_{r-p} & \\ & & 0 \end{bmatrix}$，

其中 $r = \mathrm{rank}(A), 0 \leqslant p \leqslant r$.

命题 6.2.8 中的 J 称为实对称矩阵 A 的**合同标准形**.

定理6.2.9 (Sylvester 惯性定律) 实对称矩阵的合同标准形唯一，且它的合同标准形中正、负、零对角元的个数分别和它的正、负、零特征值的个数相等.

正惯性指数、负惯性指数，三元组 $(p, r-p, n-r)$ 称为 A 的惯性指数或惯量.

容易看出，正惯性指数、负惯性指数分别等于正特征值、负特征值的个数.

n 阶实对称矩阵在合同变换下的等价类数目有限，共有 $\dfrac{(n+1)(n+2)}{2}$ 类 。

$$f(x_1,x_2,x_3)=3x_1^2+x_2^2-x_3^2+2x_1x_2-x_1x_3+4x_2x_3$$

$$=(x_1,x_2,x_3)\begin{bmatrix}3 & 1 & -\frac{1}{2}\\ 1 & 1 & 2\\ -\frac{1}{2} & 2 & -1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$

$$X^{\mathrm{T}}AX=J=\begin{bmatrix}I_p & & \\ & -I_{r-p} & \\ & & 0\end{bmatrix}$$

例 6.2.10 (配平方) 给定 $\mathbb{R}$ 上齐次二次函数 $f(x_1,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$，证明 f 可以写成齐次线性函数的平方的和差形式. ☺

证. 令 $\boldsymbol{x}=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$, $A=\left[\frac{a_{ij}+a_{ji}}{2}\right]$，则 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$. 根据命题 6.2.8 ，存在可逆矩阵 $T=[t_{ij}]$，使得 $A=T^{\mathrm{T}}JT$. 因此 $f(\boldsymbol{x})=(T\boldsymbol{x})^{\mathrm{T}}JT\boldsymbol{x}$. 令 $\boldsymbol{y}=T\boldsymbol{x}=\begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}$，那么 $y_i=\sum_{j=1}^{n}t_{ij}x_j$ 是齐次线性函数，而 $f(\boldsymbol{x})=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2$，其中 p 是 A 的正惯性指数，r 是 A 的秩. □

注6.2.11 齐次二次函数 $f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$，称为自变量 $x_1,x_2,\cdots,x_n$ 的二次型.

练习6.2.16 证明或举出反例：

1. 如果 A 对称正定，则 A^{-1} 正定.
2. 如果 A, B 对称正定，则 $A+B$ 正定.
3. 如果 A, B 对称半正定，则 $A+B$ 半正定.
4. 如果 A, B 对称不定，则 $A+B$ 不定.
5. 如果 A 列满秩，B 对称正定，则 $A^{\mathrm{T}}BA$ 正定.
6. 如果 $S=A^{\mathrm{T}}A$ 且 A 有简化 QR 分解 $A=QR$，则 $S=R^{\mathrm{T}}R$ 是 S 的 Cholesky 分解.
7. 如果 A, B 正定，则 AB 的特征值都是正数.

关于长方阵谱分解问题的讨论

目前对特征值的讨论都局限于方阵.

对 $m \times n$ 矩阵 A，我们能够考虑特征值吗?

特征值的关键是方程 $\boldsymbol{Ax} = \lambda\boldsymbol{x}$. 显然，当 $m \neq n$ 时，$\boldsymbol{Ax} = \lambda\boldsymbol{x}$ 不能成立.

A 定义了一个 R^n 到 R^m 的线性映射 $\boldsymbol{A}$，而 A^{T} 定义了一个 R^m 到 R^n 的线性映射 $\boldsymbol{A}*$,

那么 $\boldsymbol{Ax} = \boldsymbol{y}$，$A^{\mathrm{T}}\boldsymbol{y} = \mu\,\boldsymbol{x}$ 可能成立吗?

这表示 $\boldsymbol{x}$ 被映射 $\boldsymbol{A}$ 变成了 R^m 的向量 $\boldsymbol{y}$，而这个向量被映射 $\boldsymbol{A}*$ 变成了和 $\boldsymbol{x}$ 共线的向量.

事实上，注意到两个线性映射的复合 $\boldsymbol{A}^*\boldsymbol{A}$ 的表示矩阵是 $A^{\mathrm{T}}A$,

而需要求解的是方程 $A^{\mathrm{T}}\boldsymbol{Ax} = \mu\,\boldsymbol{x}$.

可见 μ 恰好是对称半正定矩阵 $A^{\mathrm{T}}A$ 的特征值, 一定有 $\mu \geqslant 0$.

考虑表达的对称性，$Ax = \lambda y$，$A^T y = \lambda x$，其中$\lambda^2 = \mu$

奇异值和左、右奇异向量

定义 6.3.1 (奇异值) 给定 $m\times n$ 矩阵 A，如果存在非零向量 $\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{y}\in\mathbb{R}^m,\sigma\geqslant 0$，使得 $A\boldsymbol{x}=\sigma\boldsymbol{y},A^{\mathrm{T}}\boldsymbol{y}=\sigma\boldsymbol{x}$，则称 σ 为 A 的一个**奇异值**，$\boldsymbol{x}$ 为 A 的属于 σ 的一个**右奇异向量**，$\boldsymbol{y}$ 为 A 的属于 σ 的一个**左奇异向量**.

可以验证，A 的右奇异向量是 $A^{\mathrm{T}}A$ 的特征向量，

类似地，A 的左奇异向量是 AA^{T} 的特征向量，

而 A 的奇异值是 $A^{\mathrm{T}}A$ 或 AA^{T} 的特征值的算术平方根.

定理 6.3.2 (奇异值分解) 给定 $m\times n$ 矩阵 A，存在 m 阶正交矩阵 U 和 n 阶正交矩阵 V，使得 $A=U\Sigma V^{\mathrm{T}}$，其中

$$\Sigma=\begin{bmatrix}\Sigma_r & O\\ O & O\end{bmatrix},\quad \Sigma_r=\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{bmatrix},\quad \sigma_1\geqslant\cdots\geqslant\sigma_r>0.$$

定理 6.3.2 (奇异值分解) 给定 $m \times n$ 矩阵 A，存在 m 阶正交矩阵 U 和 n 阶正交矩阵 V，使得 $A = U\Sigma V^{\mathrm{T}}$，其中

$$\Sigma = \begin{bmatrix} \Sigma_r & O \\ O & O \end{bmatrix}, \quad \Sigma_r = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}, \quad \sigma_1 \geqslant \cdots \geqslant \sigma_r > 0.$$

证. 由于 $A^{\mathrm{T}}A$ 对称半正定，因此有谱分解 $A^{\mathrm{T}}A = V\Lambda V^{\mathrm{T}}$，其中 Λ 的对角元素都是非负数，并由大到小排列. 设 $\Lambda = \begin{bmatrix} \Sigma_r^2 & \\ & O \end{bmatrix}$，其中 Σ_r 是 r 阶方阵. 记 $V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}$，其中 V_1 是 $n \times r$ 矩阵，V_2 是 $n \times (n-r)$ 矩阵. 则 $V_1^{\mathrm{T}}A^{\mathrm{T}}AV_1 = \Sigma_r^2, V_2^{\mathrm{T}}A^{\mathrm{T}}AV_2 = O$，即 $AV_2 = O$. 令 $U_1 = AV_1\Sigma_r^{-1}$，则有 $U_1^{\mathrm{T}}U_1 = \Sigma_r^{-1}V_1^{\mathrm{T}}A^{\mathrm{T}}AV_1\Sigma_r^{-1} = I$，即 U_1 列正交. 把 U_1 补成正交矩阵 $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}$（这相当于把 U_1 的列扩充成一组标准正交基），则有 $U^{\mathrm{T}}AV = \begin{bmatrix} U_1^{\mathrm{T}}AV_1 & U_1^{\mathrm{T}}AV_2 \\ U_2^{\mathrm{T}}AV_1 & U_2^{\mathrm{T}}AV_2 \end{bmatrix} = \begin{bmatrix} \Sigma_r & O \\ O & O \end{bmatrix}$. □

分解 $A = U\Sigma V^{\mathrm{T}}$ 称为 A 的**奇异值分解**，简称 **SVD**.

(Singular Value Decomposition)

由奇异值分解 $\boldsymbol{A} = \boldsymbol{U\Sigma V}^{\mathsf{T}}$，可得$\boldsymbol{AV} = \boldsymbol{U\Sigma}$，$\boldsymbol{A}^{\mathsf{T}}\boldsymbol{U} = \boldsymbol{V\Sigma}^{\mathsf{T}}$，这就说明 $\boldsymbol{\sigma}_i$ 是 $\boldsymbol{A}$ 的奇异值，$\boldsymbol{u}_i$ 是$\boldsymbol{A}$ 的属于 $\boldsymbol{\sigma i}$ 的左奇异向量，$\boldsymbol{vi}$ 是 $\boldsymbol{A}$ 的属于 $\boldsymbol{\sigma i}$ 的右奇异向量.

例 6.3.3 计算 $A=\begin{bmatrix}3&0\\4&5\\0&0\end{bmatrix}$ 的一个奇异值分解.

先计算 $A^{\mathrm{T}}A$ 的谱分解:

$$A^{\mathrm{T}}A=\begin{bmatrix}25&20\\20&25\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}45&\\&5\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{bmatrix}^{\mathrm{T}}=:VDV^{\mathrm{T}}.$$

奇异值是 $\sigma_1=3\sqrt{5},\sigma_2=\sqrt{5}$，记 $\Sigma_r=\begin{bmatrix}3\sqrt{5}&\\&\sqrt{5}\end{bmatrix}$，则

$$U_1=AV\Sigma_r^{-1}=\begin{bmatrix}3&0\\4&5\\0&0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\frac{1}{3\sqrt{5}}&\\&\frac{1}{\sqrt{5}}\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt{10}}&\frac{-3}{\sqrt{10}}\\\frac{3}{\sqrt{10}}&\frac{1}{\sqrt{10}}\\0&0\end{bmatrix}.$$

将 U_1 补成正交矩阵 $U=\begin{bmatrix}\frac{1}{\sqrt{10}}&\frac{-3}{\sqrt{10}}&0\\\frac{3}{\sqrt{10}}&\frac{1}{\sqrt{10}}&0\\0&0&1\end{bmatrix}$，令 $\Sigma=\begin{bmatrix}3\sqrt{5}&0\\0&\sqrt{5}\\0&0\end{bmatrix}$.

矩阵 A 的 SVD 为：

$$A=U\Sigma V^{T}.$$

从例6.3.3 的求解可以看出，

在实际计算奇异值分解时，

只需要计算左右奇异向量中的一组，就可以由定义导出另一组.

对 A 的奇异值分解 $A=U\Sigma V^{\mathrm{T}}$，记 $U=\begin{bmatrix}\boldsymbol{u}_1 & \cdots & \boldsymbol{u}_m\end{bmatrix}, V=\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n\end{bmatrix}$，则

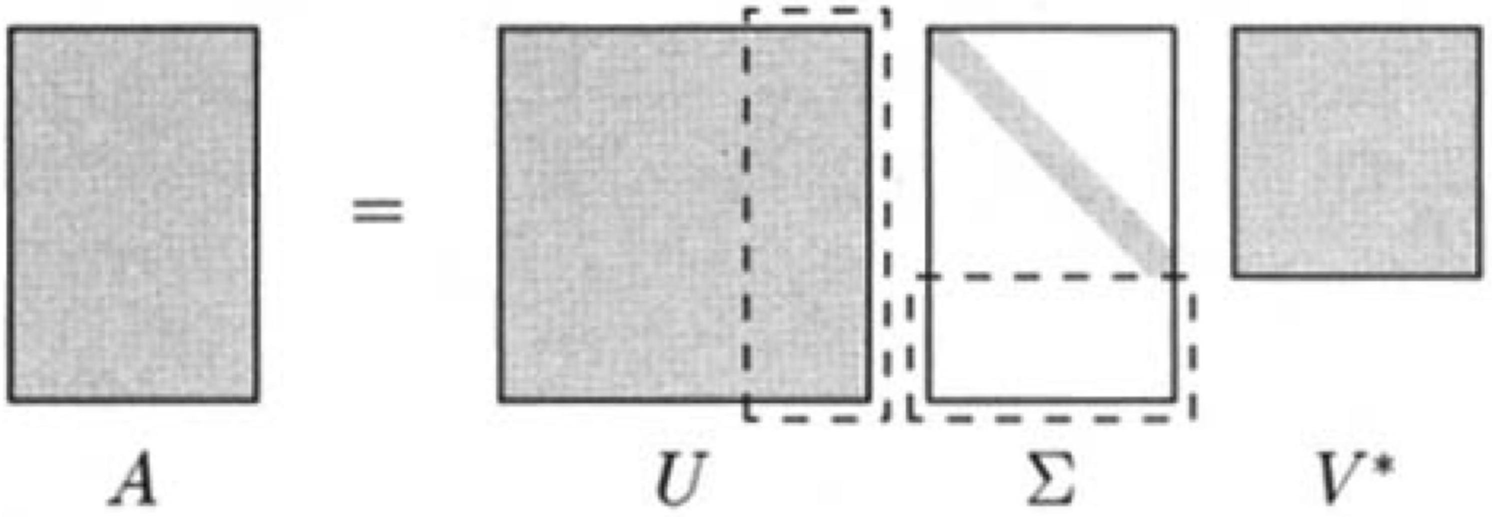
1. $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_r$ 是 $\mathcal{R}(A)$ 的一组标准正交基;
2. $\boldsymbol{u}_{r+1},\cdots,\boldsymbol{u}_m$ 是 $\mathcal{N}(A^{\mathrm{T}})$ 的一组标准正交基;
3. $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r$ 是 $\mathcal{R}(A^{\mathrm{T}})$ 的一组标准正交基;
4. $\boldsymbol{v}_{r+1},\cdots,\boldsymbol{u}_n$ 是 $\mathcal{N}(A)$ 的一组标准正交基.

记 $U_r=\begin{bmatrix}\boldsymbol{u}_1 & \cdots & \boldsymbol{u}_r\end{bmatrix}, V_r=\begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_r\end{bmatrix}$，则

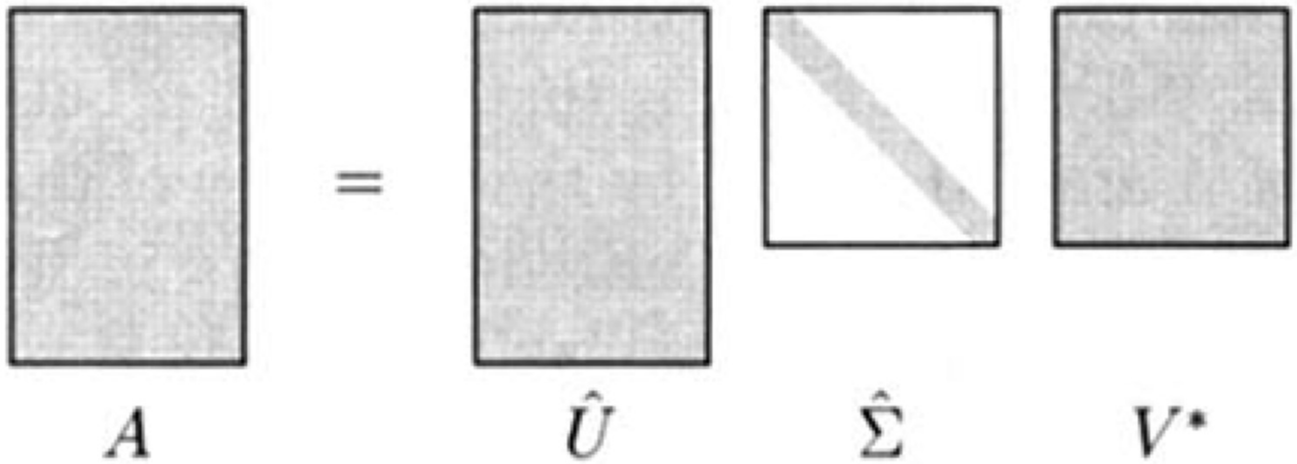
$$A=U_r\Sigma_r V_r^{\mathrm{T}}=\sigma_1\boldsymbol{u}_1\boldsymbol{v}_1^{\mathrm{T}}+\cdots+\sigma_r\boldsymbol{u}_r\boldsymbol{v}_r^{\mathrm{T}},$$

这称为 A 的**简化奇异值分解**.

Full SVD ($m \geq n$)



Reduced SVD ($m \geq n$)



定义3.3.10 给定 $\mathbb{R}^n$ 的子空间 $\mathscr{M}$，对任意 $\boldsymbol{a} \in \mathbb{R}^n$

都有唯一的分解 $\boldsymbol{a} = \boldsymbol{a}_1 + \boldsymbol{a}_2$，其中 $\boldsymbol{a}_1 \in \mathcal{M}, \boldsymbol{a}_2 \in \mathcal{M}^{\perp}$.

线性变换$\boldsymbol{P}_{\mathscr{M}}(\boldsymbol{a}) = \boldsymbol{a}_1$称为子空间 $\mathscr{M}$ 上的**正交投影（变换）**，

而 $\boldsymbol{a}_1 = \boldsymbol{P}_{\mathscr{M}}(\boldsymbol{a})$ 称为向量 $\boldsymbol{a}$ 在 $\mathscr{M}$ 上的**正交投影**.

命题 3.3.12 给定 $\mathbb{R}^n$ 的子空间 $\mathcal{M}$ 和向量 $\boldsymbol{a}$，而 $\boldsymbol{a}_1 = \boldsymbol{P}_{\mathcal{M}}(\boldsymbol{a})$ 为 $\boldsymbol{a}$ 在 $\mathcal{M}$ 上的正交投影，则 $\|\boldsymbol{a} - \boldsymbol{a}_1\| = \min_{\boldsymbol{x} \in \mathcal{M}} \|\boldsymbol{a} - \boldsymbol{x}\|$.

设 $\boldsymbol{q}_1, \cdots, \boldsymbol{q}_r$ 是 $\mathcal{M}$ 的一组标准正交基. 正交投影$\boldsymbol{P}_{\mathcal{M}}$ 的表示矩阵就是 $Q_r Q_r^{\mathrm{T}}$，记为$P_{\mathcal{M}}$.

例 3.3.1 考虑方程组 $A\boldsymbol{x} = \boldsymbol{b}$，其中 $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 3 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 0 \\ 3 \\ 5 \end{bmatrix}$,

$Q = \begin{bmatrix} \frac{1}{\sqrt{6}} & 0 \\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{5}} \end{bmatrix}$. 正交投影矩阵 $P_A = QQ^{\mathrm{T}} = \frac{1}{6}\begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{13}{15} & -\frac{1}{15} \\ \frac{1}{6} & -\frac{1}{15} & \frac{29}{30} \end{bmatrix}$.

向量 $\boldsymbol{b}$ 的正交投影分解为 $\boldsymbol{b} = P_A\boldsymbol{b} + (I_3 - P_A)\boldsymbol{b}$,

如果不计算 $\mathscr{R}(A)$ 的标准正交基，能否求出正交投影矩阵？

对任意向量 $\boldsymbol{b}$，记其在 $\mathcal{R}(A)$ 上的正交投影为 $A\boldsymbol{x}$，

则 $\boldsymbol{b}-A\boldsymbol{x} \perp \mathcal{R}(A)$.

因此 $\boldsymbol{b}-A\boldsymbol{x} \in \mathcal{R}(A)^{\perp} = \mathcal{N}(A^{\mathsf{T}})$，即 $A^{\mathsf{T}}(\boldsymbol{b} - A\boldsymbol{x}) = \mathbf{0}$，于是 $A^{\mathsf{T}}A\boldsymbol{x} = A^{\mathsf{T}}\boldsymbol{b}$.

考虑线性方程组 $A^{\mathsf{T}}A\boldsymbol{x} = A^{\mathsf{T}}\boldsymbol{b}$,

$\mathcal{R}(A^{\mathsf{T}}A) = \mathcal{R}(A^{\mathsf{T}})$，因此 $A^{\mathsf{T}}\boldsymbol{b} \in \mathcal{R}(A^{\mathsf{T}}A)$.

这说明 $A^{\mathsf{T}}A\boldsymbol{x} = A^{\mathsf{T}}\boldsymbol{b}$ 有解.

如果 $A^{\mathsf{T}}A$ 可逆，那么唯一解 $\boldsymbol{x} = (A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}\boldsymbol{b}$,

于是 $\boldsymbol{b}$ 在 $\mathcal{R}(A)$ 上的正交投影是 $A\boldsymbol{x} = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}\boldsymbol{b}$.

这意味着，关于 A 的正交投影矩阵是 $P_A = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$.

最小二乘问题

例 6.3.4 (广义逆) 第 3.3 节中讨论过关于 A 的正交投影矩阵 P_A，若 $\mathcal{R}(A)$ 的一组基排成矩阵 B，则 $P_A = B(B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}$. 而矩阵 U_r 的列是 $\mathcal{R}(A)$ 的一组标准正交基，因此 $P_A = U_rU_r^{\mathrm{T}}$.

若记矩阵 $A^+ = V_r\Sigma_r^{-1}U_r^{\mathrm{T}}$，则 $AA^+ = U_rU_r^{\mathrm{T}}$ 是 $\mathcal{R}(A)$ 上的正交投影的表示矩阵，$A^+A = V_rV_r^{\mathrm{T}}$ 是 $\mathcal{R}(A^{\mathrm{T}})$ 上的正交投影的表示矩阵. 于是 $I - AA^+$ 是 $\mathcal{N}(A^{\mathrm{T}})$ 上的正交投影的表示矩阵，$I - A^+A$ 是 $\mathcal{N}(A)$ 上的正交投影的表示矩阵. 矩阵 $A^+ = V_r\Sigma_r^{-1}U_r^{\mathrm{T}}$ 称为 A 的 **Moore-Penrose 广义逆**，简称**广义逆**. 矩阵的广义逆唯一，证明留给读者.

广义逆在解方程组时很有用.

当 A 可逆时，方程组有唯一解 $A^{-1}\boldsymbol{b}$. 此时，A^+ 就是 A 的逆 A^{-1}.

当 $A\boldsymbol{x} = \boldsymbol{b}$ 有不只一个解时，A 不可逆，但注意到 $\boldsymbol{b} \in \mathcal{R}(A)$，就有 $AA^+\boldsymbol{b} = \boldsymbol{b}$，即 $A^+\boldsymbol{b}$ 是方程组的一个解. 不仅如此，$A^+\boldsymbol{b}$ 还是方程组所有解中长度最小的解. 方程组的解集是 $\{A^+\boldsymbol{b} + \boldsymbol{x}_0 \mid \boldsymbol{x}_0 \in \mathcal{N}(A)\}$，而 $\boldsymbol{x}_0^{\mathrm{T}}A^+\boldsymbol{b} = ((I - A^+A)\boldsymbol{x}_0)^{\mathrm{T}}A^+\boldsymbol{b} = \boldsymbol{x}_0^{\mathrm{T}}(I - A^+A)A^+\boldsymbol{b} = 0$，因此 $\|A^+\boldsymbol{b} + \boldsymbol{x}_0\|^2 = \|A^+\boldsymbol{b}\|^2 + \|\boldsymbol{x}_0\|^2 \geqslant \|A^+\boldsymbol{b}\|^2$，这就说明它是长度最小的解.

再考虑方程组无解的情形，即最小二乘问题. 由于 $AA^+\boldsymbol{b}$ 是 $\boldsymbol{b}$ 在 $\mathcal{R}(A)$ 上的正交投影，因此 $A^+\boldsymbol{b}$ 是一个最小二乘解. 事实上，$A^+\boldsymbol{b}$ 也是 $A\boldsymbol{x} = \boldsymbol{b}$ 的最小二乘解中长度最小的解（为什么?）. ☺

奇异值是良态的，即如果矩阵变化不大，那么奇异值变化也不大.

例 6.3.5 考察 Jordan 块 $J_n = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$，只有一个 n 重特征值 0；有两个奇异值，是 $n-1$ 重的 1 和 1 重的 0.

给 $J_n(0)$ 添加一点微小变化 $\widetilde{J}_n = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 10^{-n} & & & 0 \end{bmatrix}$，有 n 个单特征值 $10^{-1}\mathrm{e}^{2\pi \mathrm{i}\frac{j}{n}}$，$j = 1, \cdots, n$；有两个奇异值，是 $n-1$ 重的 1 和 1 重的 10^{-n}.

可以看到，矩阵添加了一点微小变化，特征值的变化比矩阵的变化要大的多，而奇异值的变化则和矩阵的变化相差不大. ☺

矩阵的谱范数

定义 6.3.6 (矩阵的谱范数) 对任意矩阵 A，非负数 $\max\limits_{\boldsymbol{x}\neq\boldsymbol{0}} \dfrac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|}$ 称为矩阵 A 的**谱范数**，记为 $\|A\|$.

命题 6.3.7 矩阵的谱范数满足:

1. $\|A\| \geqslant 0$，且 $\|A\| = 0$ 当且仅当 $A = O$;
2. $\|kA\| = |k|\|A\|$;
3. $\|A + B\| \leqslant \|A\| + \|B\|$;
4. $\|AB\| \leqslant \|A\|\|B\|$;
5. 如果 U, V 正交，则 $\|UAV^{\mathrm{T}}\| = \|A\|$.

命题 6.3.8 对任意矩阵 A，矩阵的谱范数 $\|A\|$ 等于 A 的最大奇异值.

命题 6.3.9 设实矩阵 A 的奇异值为 $\sigma_1 \geqslant \cdots \geqslant \sigma_n$，相应的右奇异向量为 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$，则

$$\sigma_1 = \max_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|}, \qquad \sigma_i = \max_{\substack{\boldsymbol{x}\neq\boldsymbol{0} \\ \boldsymbol{x}\perp\operatorname{span}(\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{i-1})}} \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|}, i = 2, \cdots, n.$$

作 业（12月6日）

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

练习6.3

1(1,2,3,6)，2，3，4，5，6

12月13日提交

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~