回顾上次课内容：

1. 实对称矩阵的谱定理：特征值都是实数；可由正交矩阵对角化$S = Q\Lambda Q^T$

2. 对角化实对称矩阵的方法: 属于不同特征值的特征向量互相正交；特征子空间内部 Gram-Schmidt正交化

3. 实对称矩阵正定的判定定理:

(1) 能量判定：对所有非零向量$\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x}^T S \boldsymbol{x} > 0$

(2) 特征值判定：S的所有特征值$> 0.$

(3) 存在可逆矩阵A使得$S = A^T A.$

(4) 主元判定：$S = LDL^T$，且D的对角元均为正数。

L为对角线元素均为1的下三角矩阵，D为对角矩阵对角线元素为S的主元。

(5) 顺序主子式判定：S的n个顺序主子式均>0.

4. 对称性、正定性在合同变换下不变, $S \mapsto X^T S X$

半正定对称矩阵的判定定理:

定义：A一个$\boldsymbol{k}$**阶主子式**是指去掉$i_1, \dots, i_{n-k}$行与$i_1, \dots, i_{n-k}$列剩下方阵的行列式。

对于对称实矩阵，下列叙述等价：

(1)$\forall \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x}^T S \boldsymbol{x} \geq 0$;

(2)S的特征值均≥ 0;

(3)$S = A^T A, A$为n阶方阵;

(4)S所有主元非负;

(5)S的所有主子式均≥ 0;

注意：在半正定下需要考虑所有主子式，而不仅仅是顺序主子式。

例如：$\begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$半负定，顺序主子式$\geq 0$

(c) Rayleigh商与特征值:

给定实矩阵A与非零向量$x \in \mathbb{R}^n$，实数$\frac{\boldsymbol{x}^T A\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}$称为$\boldsymbol{x}$对于$A$的**Rayleigh商**。

实对称矩阵的Rayleigh商与特征值的关系:

设实对称矩阵S的特征值为$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$，相应的特征向量为$\boldsymbol{q}_1, \dots, \boldsymbol{q}_n$，则

$$\lambda_1 = \max_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\boldsymbol{x}^T S\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}, \quad \lambda_i = \max_{\boldsymbol{x}\neq\boldsymbol{0}, \boldsymbol{x}\in Span(\boldsymbol{q}_1,\dots,\boldsymbol{q}_{i-1})^{\perp}} \frac{\boldsymbol{x}^T S\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}, \ 2 \le i \le n.$$

$$\lambda_n = \min_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\boldsymbol{x}^T S\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}, \quad \lambda_i = \min_{\boldsymbol{x}\neq\boldsymbol{0}, \boldsymbol{x}\in Span(\boldsymbol{q}_{i+1},\dots,\boldsymbol{q}_{n})^{\perp}} \frac{\boldsymbol{x}^T S\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}, \ 1 \le i \le n-1.$$

证明:

$$S=Q\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}Q^T, Q=[\boldsymbol{q}_1,\dots,\boldsymbol{q}_n].$$

$$\max_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{\boldsymbol{x}^T S\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}=\max_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{\boldsymbol{x}^T Q\Lambda Q^T\boldsymbol{x}}{\boldsymbol{x}^T QQ^T\boldsymbol{x}}=\max_{\boldsymbol{y}=Q^T\boldsymbol{x}\neq\boldsymbol{0}}\frac{\boldsymbol{y}^T\Lambda\boldsymbol{y}}{\boldsymbol{y}^T\boldsymbol{y}}=\max_{\boldsymbol{y}\neq\boldsymbol{0}}\frac{\lambda_1 y_1^2+\cdots+\lambda_n y_n^2}{y_1^2+\cdots+y_n^2}\leq\frac{\lambda_1(y_1^2+\cdots+y_n^2)}{y_1^2+\cdots+y_n^2}=\lambda_1.$$

令$\boldsymbol{y}=[1,0,\dots,0]^T$可取得等号。

$$\max_{\boldsymbol{x}\neq\boldsymbol{0},\boldsymbol{x}\in Span(\boldsymbol{q}_1,\dots,\boldsymbol{q}_{i-1})^\perp}\frac{\boldsymbol{x}^T S\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}=\max_{\boldsymbol{x}\neq\boldsymbol{0},\boldsymbol{x}\in Span(\boldsymbol{q}_1,\dots,\boldsymbol{q}_{i-1})^\perp}\frac{\boldsymbol{x}^T Q\Lambda Q^T\boldsymbol{x}}{\boldsymbol{x}^T QQ^T\boldsymbol{x}}=$$
$$\max_{\boldsymbol{y}=Q^T\boldsymbol{x}\neq\boldsymbol{0},y_1=\cdots=y_{i-1}=0}\frac{\boldsymbol{y}^T\Lambda\boldsymbol{y}}{\boldsymbol{y}^T\boldsymbol{y}}=\max_{\boldsymbol{y}\neq\boldsymbol{0},y_1=\cdots=y_{i-1}=0}\frac{\lambda_i y_i^2+\cdots+\lambda_n y_n^2}{y_i^2+\cdots+y_i^2}\leq\frac{\lambda_i(y_i^2+\cdots+y_n^2)}{y_i^2+\cdots+y_n^2}=\lambda_i.$$

取$\boldsymbol{y}=\boldsymbol{e}_i$得到等号。

(d) 合同标准形与惯性定理

回顾：合同关系是一个等价关系，$B = X^T AX$，X可逆

实对称矩阵的合同标准形: 对实对称矩阵S，存在可逆矩阵X，使得

$$X^T SX = \begin{bmatrix} I_p & & \\ & -I_{r-p} & \\ & & o \end{bmatrix} =: J$$ 称为S的合同标准形。

证明:

设S 的特征值$\lambda_1, \dots, \lambda_p > 0, \lambda_{p+1}, \dots, \lambda_r < 0, \lambda_{r+1}, \dots, \lambda_n = 0$

$Q^T SQ = \Lambda = Diag(\lambda_1, \dots, \lambda_p, \lambda_{p+1}, \dots, \lambda_r, 0 \dots, 0)$

令$Y = diag(\frac{1}{\sqrt{|\lambda_1|}}, \dots, \frac{1}{\sqrt{|\lambda_r|}}, 1, \dots, 1)$

$Y^T \Lambda Y = J.$

令$X = QY$,则$X^T SX = (QY)^T SQY = J.$

Sylvester惯性定理:

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。

$$X^TSX=\begin{bmatrix} I_p & & \\ & -I_{r-p} & \\ & & O \end{bmatrix}$$

$p=$ 正惯性指数

$r-p=$ 负惯性指数

证明见教材定理6.2.9

实对称矩阵总结:

1. 实对称矩阵的谱定理
2. 对角化实对称矩阵的方法
3. 对称正定矩阵的判定定理
4. 对称性、正定性在合同变换下不变, $S \mapsto X^T S X$
5. 实对称矩阵Rayleigh商与特征值的关系
6. 合同标准形与惯性定理

6.3奇异值分解Singular Value Decomposition

奇异值分解在现实生活中有许多应用，例如图像处理等

奇异值考虑的矩阵都是$m \times n$阶实矩阵

主要内容：

(a) 奇异值分解定理（SVD）

(b) 奇异值分解的几何解释

(c) 广义逆及其应用

(d) 矩阵的谱范数与低秩逼近

(e) 线性代数基本定理（v.2.0）

(a) 奇异值分解定理(SVD):

给定$A \in M_{m\times n}(\mathbb{R})$秩为$r$，存在$m$阶正交矩阵$U$和$n$阶正交矩阵$V$使得

$A = U\Sigma V^T$,

其中，$\Sigma = \begin{bmatrix} \Sigma_r & O_{r\times(n-r)} \\ O_{(m-r)\times r} & O_{(m-r)\times(n-r)} \end{bmatrix}$, $\Sigma_r = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \in M_r(\mathbb{R})$

且$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.　按重要性顺序对奇异值排序

$\sigma_1, \dots, \sigma_r$称为$A$的奇异值,

注意：分解中正交矩阵U, V不唯一，但是奇异值σ_i由A唯一决定。

$\Sigma^T\Sigma = \begin{bmatrix} \Sigma_r^T\Sigma_r & O_{r\times(n-r)} \\ O_{(n-r)\times r} & O_{(n-r)\times(n-r)} \end{bmatrix}$为对角阵

证明思路:

假设如果有定理中的分解$A=U\Sigma V^T$，那么$A^TA=V\Sigma^T\Sigma V^T$,

$$\Sigma^T\Sigma=\begin{bmatrix}\Sigma_r^T\Sigma_r & O_{r\times(n-r)}\\ O_{(n-r)\times r} & O_{(n-r)\times(n-r)}\end{bmatrix},\quad \Sigma_r^T\Sigma_r=\begin{bmatrix}\sigma_1^2 & & \\ & \ddots & \\ & & \sigma_r^2\end{bmatrix}\text{对角线元素}>0.$$

$A^TAV=V\Sigma^T\Sigma$得到$V=[\boldsymbol{v}_1,\dots,\boldsymbol{v}_n]$满足$A^TA\boldsymbol{v}_i=\sigma_i^2\boldsymbol{v}_i$. ← 找$\sigma_i$和$V$

因此，$\boldsymbol{v}_i$为半正定对称矩阵A^TA的特征值为$\boldsymbol{\sigma}_i^2$的特征向量

由$AV=U\Sigma$表明：$1\le i\le r$时,$A\boldsymbol{v}_i=\sigma_i\boldsymbol{u}_i$, 即$\boldsymbol{u}_i=\frac{A\boldsymbol{v}_i}{\sigma_i}$. ← 找$U$

因此，我们证明的思路即是从半正定矩阵A^TA的对角化开始。

奇异值分解定理的证明:

记$r = rank\left(A^T A\right) = rank(A)$. 考查对称半正定矩阵$A^T A$（特征值$\geq 0$）

令$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$为$A^T A$的所有非零特征值

由对称矩阵的谱定理，有正交矩阵$V = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_n]$使得

$$A^T A V = V \begin{bmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & O \end{bmatrix}$$

于是$A^T A \boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i, 1 \leq i \leq r, A^T A \boldsymbol{v}_i = \mathbf{0}, r+1 \leq i \leq n$.

由于$\mathcal{N}\left(A^T A\right) = \mathcal{N}(A), A\boldsymbol{v}_i = \mathbf{0}, r+1 \leq i \leq n$.

令$\sigma_i = \sqrt{\lambda_i}, 1 \leq i \leq r$. 定义$\boldsymbol{u}_i = \frac{A\boldsymbol{v}_i}{\sigma_i}, 1 \leq i \leq r$.

验证$\boldsymbol{u}_1, \dots, \boldsymbol{u}_r \in \mathbb{R}^m$为正交单位向量组

验证$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r \in \mathbb{R}^m$为正交单位向量组，即

(1) $\|\boldsymbol{u}_i\| = 1$. 这是因为$\boldsymbol{u}_i^T \boldsymbol{u}_i = \frac{\boldsymbol{v}_i^T A^T A \boldsymbol{v}_i}{\sigma_i^2} = \frac{\boldsymbol{v}_i^T \sigma_i^2 \boldsymbol{v}_i}{\sigma_i^2} = 1.$

(2) $\boldsymbol{u}_i^T \boldsymbol{u}_j = 0, i \neq j$. 这是因为$\boldsymbol{u}_i^T \boldsymbol{u}_j = \frac{\boldsymbol{v}_i^T A^T A \boldsymbol{v}_j}{\sigma_i \sigma_j} = \frac{\lambda_j \boldsymbol{v}_i^T \boldsymbol{v}_j}{\sigma_i \sigma_j} = 0.$

将$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$扩充为$\mathbb{R}^m$一组标准正交基$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$.

令$U = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m]$.

我们有$A[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n] = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m] \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & O \end{bmatrix}$

即$A = U \Sigma V^T$.

简化奇异值分解:

$A \in M_{m\times n}(\mathbb{R})$秩为$r$. 由奇异值分解

$$A = U\Sigma V^T = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_m]\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix}\begin{bmatrix} \boldsymbol{v}_1^T \\ \vdots \\ \boldsymbol{v}_n^T \end{bmatrix}$$

$$= \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^T + \sigma_2 \boldsymbol{u}_2 \boldsymbol{v}_2^T + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^T.$$

A为r个秩1矩阵的和。

令$U_r = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_r], \Sigma_r = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}, V_r = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_r]$

$A = U_r \Sigma_r V_r^T$ 称为A的**简化奇异值分解**

例: 求$A=\begin{bmatrix}0&1&0&0\\0&0&2&0\\0&0&0&3\end{bmatrix}$的奇异值分解。

$$A^TA=\begin{bmatrix}0&0&0&0\\0&1&0&0\\0&0&4&0\\0&0&0&9\end{bmatrix},\sigma_1=3>\sigma_2=2>\sigma_3=1.$$

求解$A^TA\boldsymbol{v}=\sigma^2\boldsymbol{v}$得到$\boldsymbol{v}_1=\begin{bmatrix}0\\0\\0\\1\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}0\\0\\1\\0\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}0\\1\\0\\0\end{bmatrix}$.

求解$A^TA\boldsymbol{v}=\boldsymbol{0}$得到$\boldsymbol{v}_4=\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$.

令$\boldsymbol{u}=\frac{A\boldsymbol{v}}{\sigma}$得到$\boldsymbol{u}_1=\frac{A\boldsymbol{v}_1}{\sigma_1}=\begin{bmatrix}0\\0\\1\end{bmatrix}, \boldsymbol{u}_2=\frac{A\boldsymbol{v}_2}{\sigma_2}=\begin{bmatrix}0\\1\\0\end{bmatrix}, \boldsymbol{u}_3=\frac{A\boldsymbol{v}_3}{\sigma_3}=\begin{bmatrix}1\\0\\0\end{bmatrix}$

$$A=\begin{bmatrix}0 & 1 & 0 & 0\\0 & 0 & 2 & 0\\0 & 0 & 0 & 3\end{bmatrix}=U\Sigma V^T$$

$$=\begin{bmatrix}0 & 0 & 1\\0 & 1 & 0\\1 & 0 & 0\end{bmatrix}\begin{bmatrix}3 & 0 & 0 & 0\\0 & 2 & 0 & 0\\0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}0 & 0 & 0 & 1\\0 & 0 & 1 & 0\\0 & 1 & 0 & 0\\1 & 0 & 0 & 0\end{bmatrix}$$

奇异值与特征值比较：

1. 奇异值对任意实矩阵定义, 特征值对任意复方阵定义
2. 奇异值均为正数, 特征值可正、负、零
3. 奇异值对矩阵的微扰比特征值稳定

小结: SVD的求法

1. 对半正定对称矩阵$A^T A$做谱分解, 特征值λ, 特征向量$\boldsymbol{v}$
2. 令$\sigma = \sqrt{\lambda}$, 得到奇异值
3. 令$\boldsymbol{u} = \frac{A\boldsymbol{v}}{\sigma}$
4. 将$\boldsymbol{u}$扩充为$\mathbb{R}^m$的一组标准正交基

(b) 奇异值分解的几何的解释

1. 映射复合的角度
2. 空间分解的角度

1. 映射复合的角度看待SVD:

由奇异值分解，矩阵A写成（正交）×（对角）×（正交）

以$A \in M_2(\mathbb{R})$为例，**为简单起见假定两个正交矩阵都为旋转矩阵**

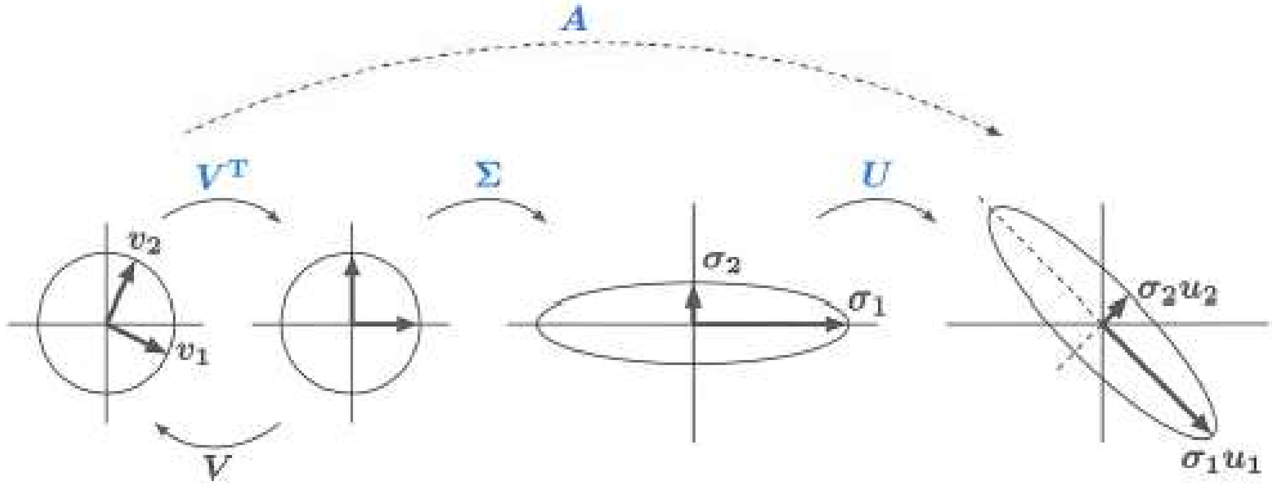
$U = [\boldsymbol{u}_1, \boldsymbol{u}_2], V = [\boldsymbol{v}_1, \boldsymbol{v}_2]$分别为逆时针旋转$\theta, \phi$角。

$$U = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, V = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

$$A = U\Sigma V^T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix} \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}$$

线性映射A由它在标准正交基$\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$上的作用决定

$A[\boldsymbol{v}_1, \boldsymbol{v}_2] = U\Sigma V^T[\boldsymbol{v}_1, \boldsymbol{v}_2]$.



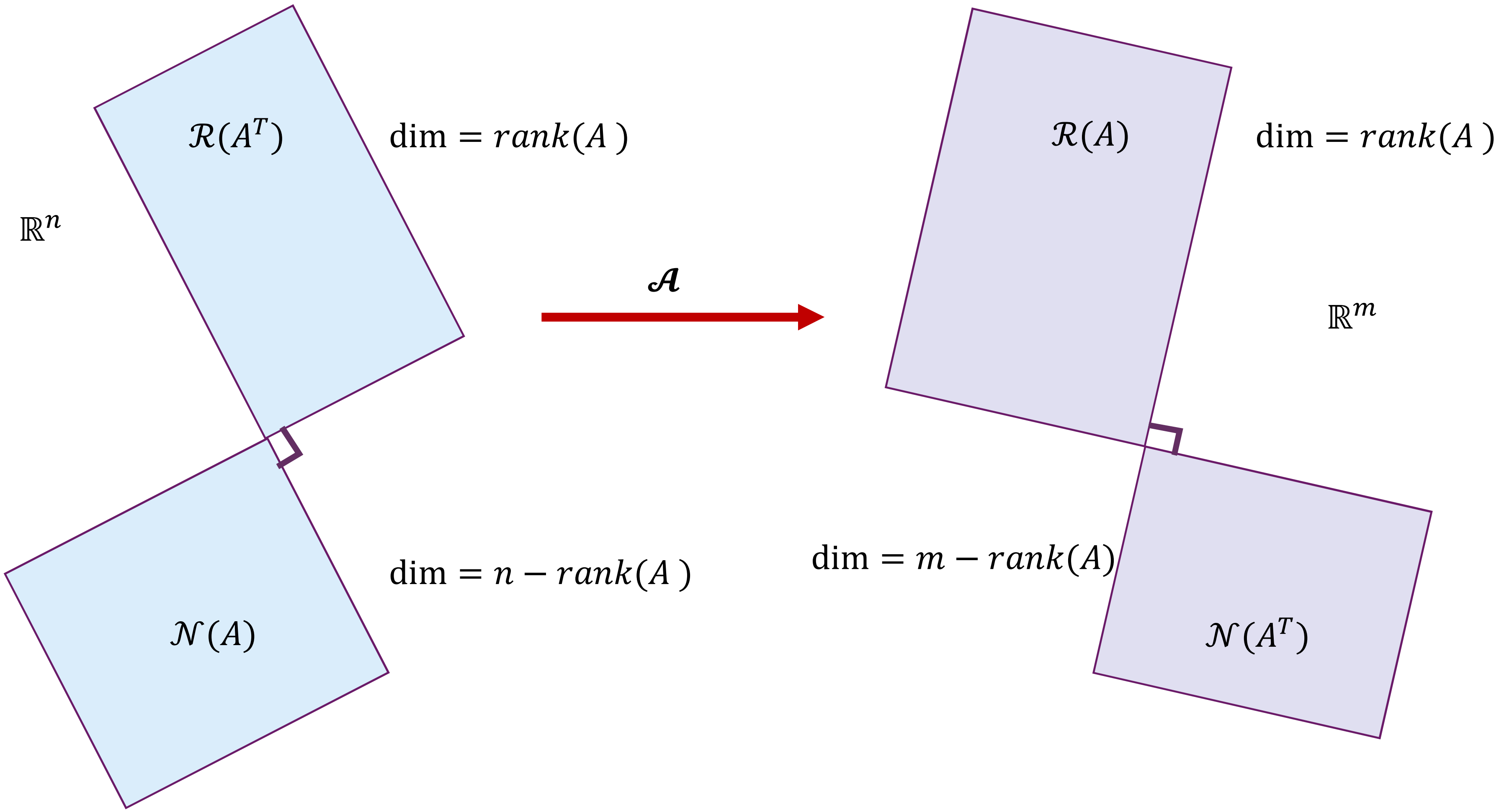
A的作用分解为3步：

1. V^T将$\boldsymbol{v}_1, \boldsymbol{v}_2$顺时针旋转$\phi$角，分别映为$\boldsymbol{e}_1, \boldsymbol{e}_2$；

2. Σ将$\boldsymbol{e}_1, \boldsymbol{e}_2$分别拉伸$\sigma_1, \sigma_2$倍，分别映为$\sigma_1\boldsymbol{e}_1, \sigma_2\boldsymbol{e}_2$；

3. U将$\sigma_1\boldsymbol{e}_1, \sigma_2\boldsymbol{e}_2$逆时针旋转$\theta$角，分别映为$\sigma_1\boldsymbol{u}_1, \sigma_2\boldsymbol{u}_2$.

2. 空间分解的角度看待SVD:



A的奇异值分解给出四个基本子空间的标准正交基:

$A \in M_{m\times n}(\mathbb{R}), r = rank(A). A = U\Sigma V^T \Rightarrow AV = U\Sigma$

令$V = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_n], U = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_m]$.

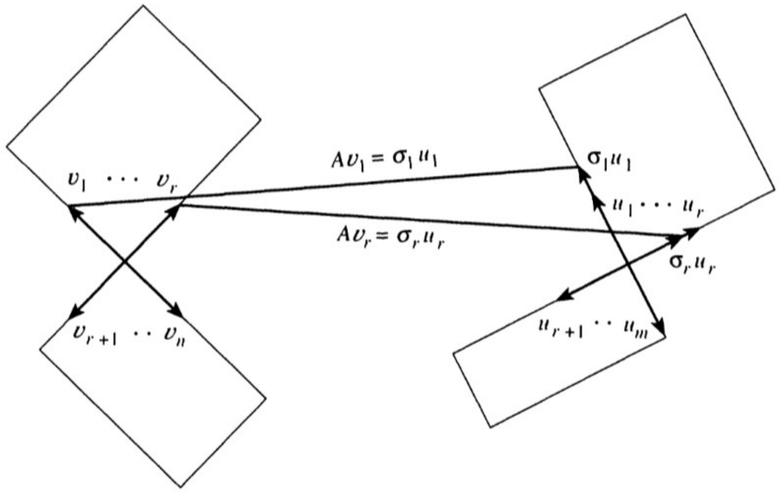
$\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$构成$\mathbb{R}^n$一组标准正交基，$\boldsymbol{u}_1, \dots, \boldsymbol{u}_m$构成$\mathbb{R}^m$一组标准正交基

(1) $\boldsymbol{v}_1, \dots, \boldsymbol{v}_r$构成$\mathcal{R}(A^T)$一组标准正交基

(2) $\boldsymbol{v}_{r+1}, \dots, \boldsymbol{v}_n$构成$\mathcal{N}(A)$一组标准正交基

(3) $\boldsymbol{u}_1, \dots, \boldsymbol{u}_r$构成$\mathcal{R}(A)$一组标准正交基

(4) $\boldsymbol{u}_{r+1}, \dots, \boldsymbol{u}_m$构成$\mathcal{N}(A^T)$一组标准正交基



证明:

$A\boldsymbol{v}_{r+1}=\mathbf{0},\dots,A\boldsymbol{v}_n=\mathbf{0}\Rightarrow \boldsymbol{v}_{r+1},\dots,\boldsymbol{v}_n\in\mathcal{N}(A)$

由$\dim\mathcal{N}(A)=n-r$得到

$\mathcal{N}(A)=Span(\boldsymbol{v}_{r+1},\dots,\boldsymbol{v}_n)$.

$\mathcal{R}(A^T)=\mathcal{N}(A)^{\perp}=Span(\boldsymbol{v}_1,\dots,\boldsymbol{v}_r)$.

当$1\le i\le r$时,$\boldsymbol{u}_i=\frac{A\boldsymbol{v}_i}{\sigma_i}\in\mathcal{R}(A)$

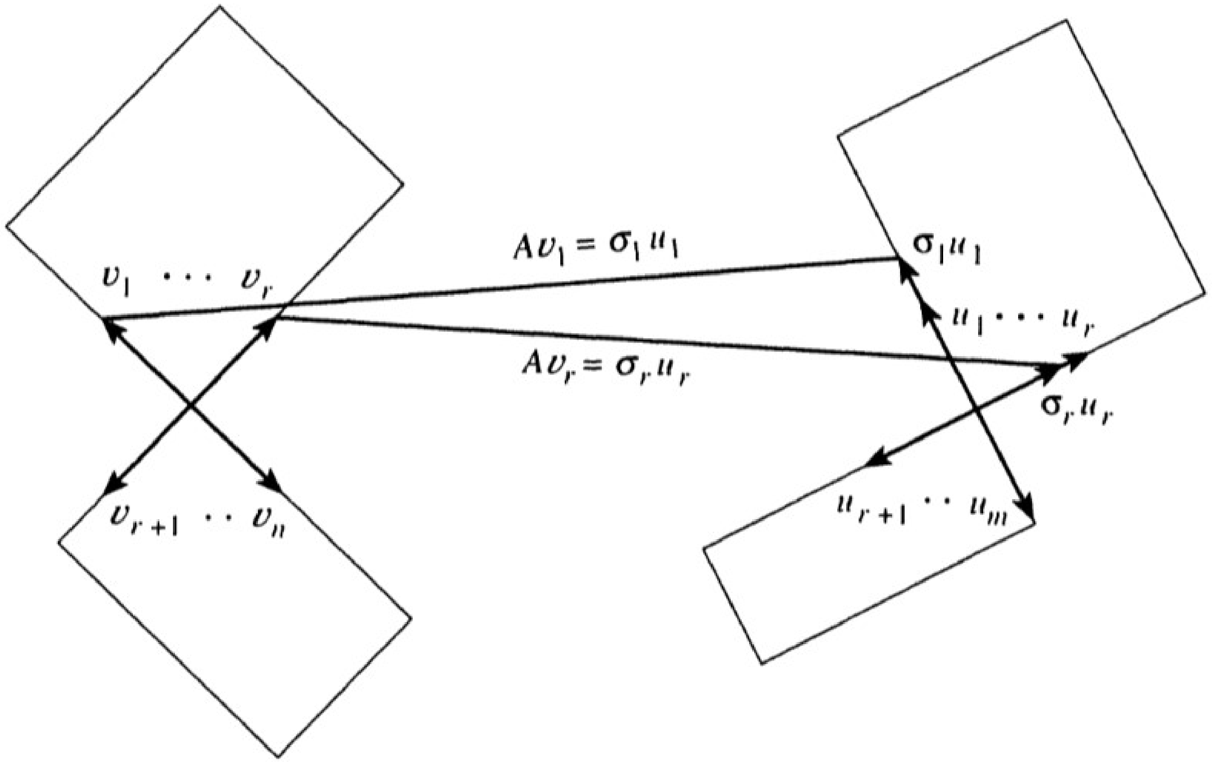
由于$\dim Span(\boldsymbol{u}_1,\dots,\boldsymbol{u}_r)=r=\dim\mathcal{R}(A)$, $\boldsymbol{u}_1,\dots,\boldsymbol{u}_r$构成$\mathcal{R}(A)$一组标准正交基

考虑正交补，$\boldsymbol{u}_{r+1},\dots,\boldsymbol{u}_m$构成$\mathcal{N}(A^T)$一组标准正交基

$\mathcal{R}(A^T), \mathcal{N}(A), \mathcal{R}(A), \mathcal{N}(A^T)$分别存在标准正交基

$\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_r\}$,$\{\boldsymbol{v}_{r+1}, \dots, \boldsymbol{v}_n\}$,$\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_r\}$,$\{\boldsymbol{u}_{r+1}, \dots, \boldsymbol{u}_m\}$使得有$\sigma_i > 0, 1 \le i \le r$满足

$$A\boldsymbol{v}_1 = \sigma_1\boldsymbol{u}_1,\ A\boldsymbol{v}_2 = \sigma_2\boldsymbol{u}_2, \dots,\ A\boldsymbol{v}_r = \sigma_r\boldsymbol{u}_r,\ A\boldsymbol{v}_{r+1} = \boldsymbol{0}, \dots,\ A\boldsymbol{v}_n = \boldsymbol{0}$$



(c) 广义逆与最小二乘法

矩阵的逆只能对可逆方阵谈论

广义逆可以对任意$m \times n$阶矩阵谈论

广义逆是矩阵逆的推广

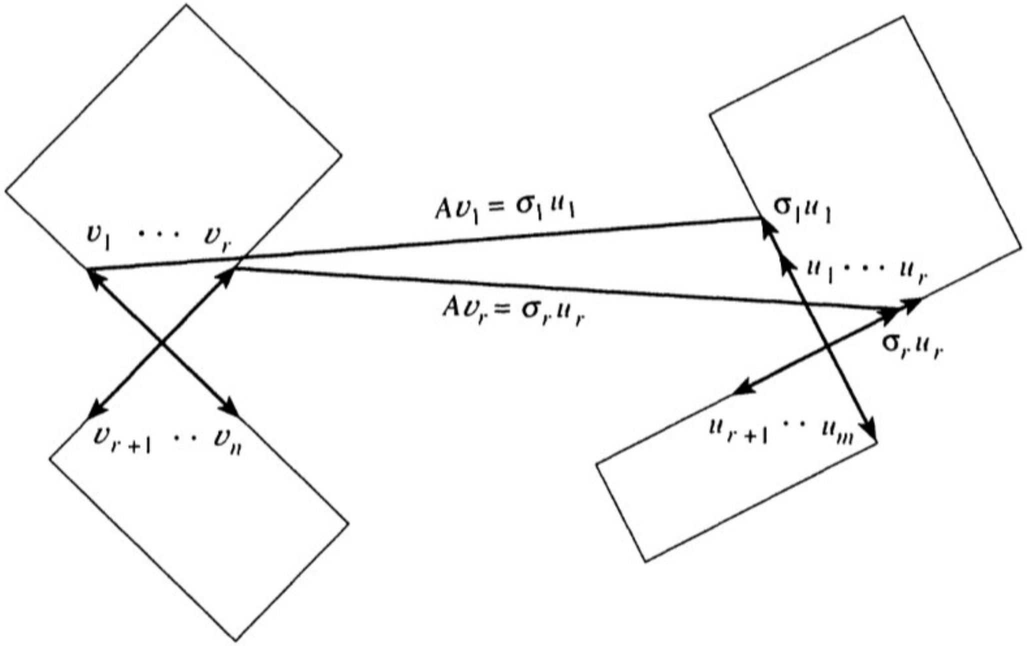
内容：

1. 广义逆的定义
2. 广义逆的应用：正交投影矩阵，最优最小二乘解

$A \in M_{m\times n}(\mathbb{R}) \rightsquigarrow A^+ \in M_{n\times m}(\mathbb{R})$

线性映射角度 $A\colon \mathbb{R}^n \to \mathbb{R}^m \rightsquigarrow A^+\colon \mathbb{R}^m \to \mathbb{R}^n$

广义逆的定义:



$A^+: \mathbb{R}^m \to \mathbb{R}^n$应该满足

$$A^+(\sigma_1 \boldsymbol{u}_1) = \boldsymbol{v}_1, \dots, A^+(\sigma_r \boldsymbol{u}_r) = \boldsymbol{v}_r; A^+(\boldsymbol{u}_{r+1}) = \cdots = A^+(\boldsymbol{u}_m) = \mathbf{0}$$

于是$A^+(\boldsymbol{u}_1) = \frac{\boldsymbol{v}_1}{\sigma_1}, \dots, A^+(\boldsymbol{u}_r) = \frac{\boldsymbol{v}_r}{\sigma_r};\ \ A^+(\boldsymbol{u}_{r+1}) = \cdots = A^+(\boldsymbol{u}_m) = \mathbf{0}.$

写成分块乘法的形式:

$$A^+[\boldsymbol{u}_1, \dots, \boldsymbol{u}_m] = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_n]\begin{bmatrix} \sigma_1^{-1} & & & \\ & \ddots & & \\ & & \sigma_r^{-1} & \\ & & & \\ & & & \end{bmatrix}$$

这提示我们定义$A^+ = V\begin{bmatrix} \sigma_1^{-1} & & & \\ & \ddots & & \\ & & \sigma_r^{-1} & \\ & & & \\ & & & \end{bmatrix}U^T$

广义逆的定义:

设$\Sigma = \begin{bmatrix} \Sigma_r & O_{r\times(n-r)} \\ O_{(m-r)\times r} & O_{(m-r)\times(n-r)} \end{bmatrix} \in M_{m\times n}(\mathbb{R})$, 其中$\Sigma_r = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \in M_r(\mathbb{R})$

令$\Sigma^+ := \begin{bmatrix} \Sigma_r^+ & O_{r\times(m-r)} \\ O_{(n-r)\times r} & O_{(n-r)\times(m-r)} \end{bmatrix} \in M_{n\times m}(\mathbb{R})$, 其中$\Sigma_r^+ = \begin{bmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_r^{-1} \end{bmatrix} \in M_r(\mathbb{R})$

$\Sigma\Sigma^+ = \begin{bmatrix} I_r & O_{r\times(m-r)} \\ O_{(m-r)\times r} & O_{(m-r)\times(m-r)} \end{bmatrix} \in M_m(\mathbb{R})$, $\Sigma^+\Sigma = \begin{bmatrix} I_r & O_{r\times(n-r)} \\ O_{(n-r)\times r} & O_{(n-r)\times(n-r)} \end{bmatrix} \in M_n(\mathbb{R})$,

给定$A \in M_{m\times n}(\mathbb{R})$，有奇异值分解$A = U\Sigma V^T$。定义$A$的广义逆:

$A^+ = V\Sigma^+U^T \in M_{n\times m}(\mathbb{R})$.

广义逆唯一，与A的奇异值分解选取无关。

广义逆与逆:

当A为n阶可逆方阵时，$r = n$，$\Sigma\Sigma^+ = \Sigma^+\Sigma = I_n$

$$AA^+ = \left(U\Sigma V^T\right)\left(V\Sigma^+ U^T\right) = I_n$$

$$A^+A = \left(V\Sigma^+ U^T\right)\left(U\Sigma V^T\right) = I_n$$

于是对于可逆方阵，广义逆就是矩阵的逆。

例: 矩阵的奇异值分解与广义逆

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} = U\Sigma V^T = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

$$A^+ = V\Sigma^+ U^T = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}$$