正交性和最小二乘法

#数学 #线性代数

内积，长度，正交性

内积

$$设u, v \in R^n, 则u \cdot v = u^T v称为u和v的内积$$

性质

- 是线性的
- $u \cdot u \geq 0, 且u \cdot u = 0 \Leftrightarrow u = 0$

长度

$$设u = (x_1, x_2 \ldots, x_n) \in R$$

$$则||u|| = \sqrt{u \cdot u} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}称为u的长度$$

性质

$$设u \in R^n, 则||cu|| = |c|||u||$$

- $设u \in R^n, 若||u|| = 1, 则称u为单位向量$
- $设0 \neq u \in R^n, 称\frac{u}{||u||}为u的单位化$
- $||u - v||称为u和v的距离$

正交

$$设u, v \in R^n, 若u \cdot v = 0, 则称u和v正交，记作u \perp v$$

正交补空间

$$如果向量u \in R^n与R^n子空间W中任何向量都正交，称u和W正交$$

$$与W正交的向全体称为W的正交补，记作W^{\perp}$$

性质

$设W = Span\{\alpha_1, \alpha_2, \ldots, \alpha_n\} \subset R^n$

- $x \in W^{\perp} \Leftrightarrow x和\alpha_i正交，\forall i = 1, 2, \ldots$
- $W^{\perp}是R^n的子空间$

正交集

正交集

定义

$$\text{称}R^n\text{中的向量组}\{u_1,u_2,\ldots,u_s\}\text{为正交集，}$$
$$\text{集合中的任意两个向量正交}$$

$$R^n\text{的正交集是}W\text{的一个基且是正交集}$$

定理 设$\{u_1,u_2,\cdots,u_s\}$是$\mathbb{R}^n$的子空间W的一个正交基，

$$\forall\alpha\in W,\alpha=c_1u_1+c_2u_2+\cdots+c_su_s,\text{有}c_j=\frac{\alpha\cdot u_j}{u_j\cdot u_j},\ j=1,2,\cdots,s$$

正交投影

定义 设u是$\mathbb{R}^n$中的非零向量，$\forall\alpha\in\mathbb{R}^n$，寻找

$$\alpha=\hat{\alpha}+v,\ \text{其中}\hat{\alpha}=ku,\ k\in\mathbb{R},v\text{与}u\text{正交}$$

则$\hat{\alpha}$称为α在u上的正交投影，v称为与u正交的分量

事实 $\hat{\alpha}=\dfrac{\alpha\cdot u}{u\cdot u}u \overset{\text{记}}{=\!=} \mathrm{proj}_L\,\alpha$　　令$L=\mathrm{Span}\{u\}$

单位正交集

定义 单位正交集：单位向量构成的正交集

单位(或标准)正交基：基+单位正交集

定理 $U_{m\times n}$具有单位正交列向量$\Leftrightarrow U^TU=I$

性质 设$U_{m\times n}$具有单位正交列向量, $x,y\in\mathbb{R}^n$,

(1) $\|Ux\|=\|x\|$;

(2) $Ux\cdot Uy=x\cdot y$;

(3) $Ux\cdot Uy=0\Leftrightarrow x\cdot y=0$

定义 $A_{n\times n}$称为正交阵,若$A^TA=I$

定理 设A是n阶方阵，则

A为正交阵$\Leftrightarrow$ A的列向量组是$\mathbb{R}^n$的一个单位(或标准)正交基

注意：若2阶正交阵A满足$|A|=1$，则T是线性变换

扩充定理

扩充定理

1、有限维向量空间V中,任意线性无关集可扩充成V的一个基

可以使用数学归纳法证明

设$\dim W=r$, 非零正交集$\alpha_1,\cdots,\alpha_p\in W$, 对$r-p$数归

若$r-p=0$,结论显然成立

假设$r-p=k>0$,结论成立

考虑$r-p=k+1$ 此时，存在β不可由$\alpha_1,\cdots,\alpha_p$线性表示

令$\alpha_{p+1}=\beta-l_1\alpha_1-\cdots-l_p\alpha_p$,

则$\alpha_i\cdot\alpha_{p+1}=\alpha_i\cdot\beta-l_1\alpha_i\cdot\alpha_1-\cdots-l_p\alpha_i\cdot\alpha_p=\alpha_i\cdot\beta-l_i\alpha_i\cdot\alpha_i$, $(i=1,\cdots,p)$

取$l_i=\dfrac{\alpha_i\cdot\beta}{\alpha_i\cdot\alpha_i}$, $(i=1,\cdots,p)$ 则$\alpha_i\cdot\alpha_{p+1}=0$ 故$\alpha_1,\cdots,\alpha_p,\alpha_{p+1}$为$W$中的非零正交集

由假设知$\alpha_1,\cdots,\alpha_p,\alpha_{p+1}$可扩充成$W$的一个正交基

故命题得证

正交分解定理

定理（正交分解定理）

（1）设W是$\mathbb{R}^n$的子空间，则$\forall y\in\mathbb{R}^n$, y可唯一的表示成

$$y=\hat{y}+z，其中\hat{y}\in W, z\in W^{\perp},$$

(2) 设W有一个正交基$\{u_1,u_2,\cdots,u_p\}$, 则

$$\hat{y}=\frac{y\cdot u_1}{u_1\cdot u_1}u_1+\cdots+\frac{y\cdot u_p}{u_p\cdot u_p}u_p$$

且$z=y-\hat{y}$.

$\hat{y}$ 称为y在W上的正交投影

记$\mathrm{proj}_W\, y\triangleq\hat{y}$

分析 只需证$y-\hat{y}\in W^{\perp}$

定理（最佳逼近定理）

设W是$\mathbb{R}^n$的子空间，$\forall y\in\mathbb{R}^n$，$\hat{y}$是y在W上的正交投影，

则$\hat{y}$是W中最接近y的点，

即 $\forall v\in W\setminus\{\hat{y}\}$, $\|y-\hat{y}\|<\|y-v\|$

总结 设W是$\mathbb{R}^n$的子空间，设W有一个正交基$\{u_1,u_2,\cdots,u_s\}$,

$$U\triangleq[u_1,u_2,\cdots,u_s]$$

投影矩阵$P=U(U^TU)^{-1}U^T$　　　投影变换 $y\mapsto Py$

（1）若$W=\mathrm{Span}\{u\}$，　$P=\dfrac{u\,u^T}{u^Tu}$

（2）若W的基是单位正交基，则$P=UU^T$

(3) $P^2=P$,　$P^T=P$

格拉姆-施密特方法

定理

定理 (格拉姆-施密特方法)

$\mathbb{R}^n$中，子空间W有一个基$\{\alpha_1,\cdots,\alpha_p\}$，令

$$
\begin{aligned}
v_1 &= \alpha_1 \\
v_2 &= \alpha_2 - \frac{\alpha_2 \cdot v_1}{v_1 \cdot v_1} v_1 \\
v_3 &= \alpha_3 - \frac{\alpha_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{\alpha_3 \cdot v_2}{v_2 \cdot v_2} v_2 \\
&\cdots \\
v_p &= \alpha_p - \frac{\alpha_p \cdot v_1}{v_1 \cdot v_1} v_1 - \cdots - \frac{\alpha_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}
\end{aligned}
$$

则 (1) $\{v_1,v_2,\cdots,v_p\}$是W的一个正交基，

(2) $\mathrm{Span}\{v_1,\cdots,v_k\} = \mathrm{Span}\{\alpha_1,\cdots,\alpha_k\}$

$k=1,\cdots,p$

$$A_{m\times n} = Q_{m\times n} R_{n\times n}$$

Q的列形成$\mathrm{Col}\,A$ 的一个标准正交基

R是上三角可逆矩阵，且在对角线上的元素为正数

算法

算法 $A=QR$

1.寻找Q：对A的列利用格拉姆-施密特方法正交化、单位化

2.寻找R：$R=Q^TA$

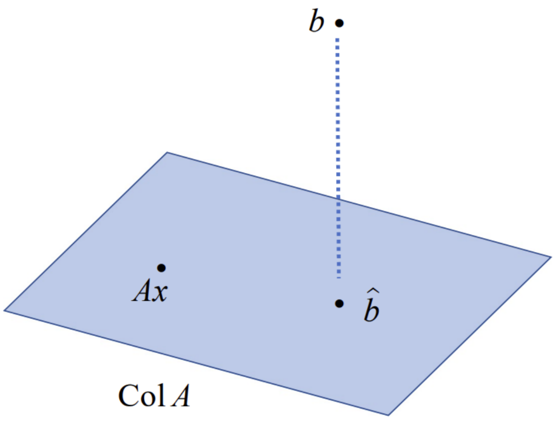
最小二乘问题

最小二乘法

定义

设$m\times n$矩阵A, 向量$b\in R$, 则$Ax=b$的最小二乘解是R^n中的$\vec{x}$

使得$||b - A\vec{x}|| \le ||b - Ax||, \forall x \in R^n$



$\begin{cases} W = \operatorname{Col} A \text{是} \mathbb{R}^m \text{的子空间} \\ b \in \mathbb{R}^m \end{cases}$

$\Downarrow$ **最佳逼近定理**

$\hat{b} = \operatorname{proj}_W b \in W$

$\Downarrow$

$A\hat{x} = \hat{b}$ 相容

$\boxed{Ax = b\text{的最小二乘解是}A\hat{x} = \operatorname{proj}_{\operatorname{Col} A} b\text{的解}}$

定理　设$m \times n$阶矩阵A, $b \in \mathbb{R}^m$,则

方程$Ax = b$的最小二乘解集和方程 $A^T Ax = A^T b$的非空解集一致

定义　方程$A^T Ax = A^T b$称为方程 $Ax = b$的法方程

方程$A^T Ax = A^T b$的解记作$\hat{x}$

定义　b到$A\hat{x}$的距离称为最小二乘误差

唯一最小二乘解

定理 设$m\times n$阶矩阵A的列线性无关，A的QR分解为 $A=QR$，则$\forall b\in\mathbb{R}^m$，方程 $Ax=b$有唯一最小二乘解

$$\hat{x}=R^{-1}Q^Tb$$

证明 $Ax=b$有唯一最小二乘解

$$\begin{aligned}\hat{x}&=(A^TA)^{-1}A^Tb\\&=[(QR)^T(QR)]^{-1}(QR)^Tb\\&=[R^TQ^TQR]^{-1}R^TQ^Tb\\&=R^{-1}\left(R^T\right)^{-1}R^TQ^Tb\\&=R^{-1}Q^Tb\end{aligned}$$

若R是可逆阵则有唯一最小二乘解

内积空间

内积空间

定义 向量空间V上的内积是一个函数，$\forall u,v,w\in V,\ c\in\mathbb{R}$，存在一个实数$\langle u,v\rangle$满足

(1) $\langle u,v\rangle=\langle v,u\rangle$ 对称性

(2) $\langle u+v,w\rangle=\langle u,w\rangle+\langle v,w\rangle$

(3) $\langle cu,v\rangle=c\langle u,v\rangle$

(2)(3) 线性性

(4) $\langle u,u\rangle\geq 0$，且$\langle u,u\rangle=0\Leftrightarrow u=0$ 非负性

一个赋予上面内积的向量空间称为内积空间

例 $\mathbb{R}^{n\times n}:=\{$定义在$\mathbb{R}$上的全体$n\times n$阶矩阵$\}$

$$\langle A,B\rangle\triangleq\mathrm{tr}(AB^T)$$

$U\triangleq\{$全体n阶对称阵$\}$ 则 $U^{\perp}=\{$全体n阶反对称阵$\}$

例 $\mathbb{R}[x]_n \triangleq \{$次数不超过$n$的一元多项式全体$\}$

应用：趋势分析

取定不同的实数$t_0, t_1, \cdots, t_n$

$\forall p, q \in \mathbb{R}[x]_n$， $\langle p, q \rangle \triangleq p(t_0)q(t_0) + p(t_1)q(t_1) + \cdots + p(t_n)q(t_n)$

例 $C[a,b] \triangleq \{$定义在区间$[a,b]$上的全体实连续函数$\}$

$$\langle f(x), g(x) \rangle \triangleq \int_a^b f(x)g(x)\,\mathrm{d}x$$

应用：傅里叶级数

定义 (1) u的长度（或范数）定义为 $\|u\| \triangleq \sqrt{\langle u, u \rangle}$

(2) 若$\|u\| = 1$，则称u为单位向量.

(3) u和v的距离是 $\mathrm{dist}(u, v) \triangleq \|u - v\|$

(4) 若$\langle u, v \rangle = 0$，则称u和v正交，记作$u \perp v$

奇异值分解

定义

定义 设$m \times n$阶矩阵A,

A的奇异值指$A^T A$的特征值的平方根，记作$\sigma_1, \cdots, \sigma_n$，且按递减排序

可假设$\sigma_i = \sqrt{\lambda_i}, i = 1, \cdots, n$

定理（奇异值分解）

设$m\times n$阶矩阵A的秩为r，则$A=U\Sigma V^T$

其中 $\begin{cases} U,V\text{分别为}m\text{阶和}n\text{阶正交阵} \\ \Sigma\text{为}m\times n\text{阶矩阵，形为}\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \end{cases}$ $\quad D=\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$

A的正奇异值：$\sigma_1\geq\cdots\geq\sigma_r>0$

U的列称为A的左奇异向量

V的列称为A的右奇异向量

算法 设$m\times n$阶矩阵A

1 计算A^TA

2 求A^TA的特征值$\lambda_1,\cdots,\lambda_n$，且满足$\lambda_1\geq\cdots\geq\lambda_n\geq 0$

3 求对应的单位正交特征向量$\{v_1,\cdots,v_n\}$，是A的右奇异向量，构成V

4 将$\{Av_1,\cdots,Av_r\}$单位化，扩充成$\mathbb{R}^m$的标正基，是A的左奇异向量，构成U

5 求正特征值$\lambda_1,\cdots,\lambda_r$的平方根，是$A$的正奇异值$\sigma_1\geq\cdots\geq\sigma_r>0$，构成$\Sigma$

应用

应用四　基本子空间的基

设$m\times n$阶矩阵A的奇异值分解为 $A=U\Sigma V^T$

A的左奇异向量：U的列$\{u_1,\cdots,u_m\}$

A的右奇异向量：V的列$\{v_1,\cdots,v_n\}$

A的秩为r,

$\Rightarrow$

- $Col\,A$的一个标正基：$\{u_1,\cdots,u_r\}$
- $Nul\,A^T$的一个标正基：$\{u_{r+1},\cdots,u_m\}$
- $Nul\,A$的一个标正基：$\{v_{r+1},\cdots,v_n\}$
- $Row\,A$的一个标正基：$\{v_1,\cdots,v_r\}$