矩阵理论

刘克新

自动化科学与电气工程学院

答疑邮箱：skxliu@163.com

第一章 线性空间引论

- 线性空间
- 线性子空间
- 基与坐标
- 内积空间
- 直和与投影
- 应用：多项式插值

1.5 直和与投影

定义1.5.1（直和与正交直和） 设W_1与W_2是线性空间V的子空间, 若和空间W_1+W_2中任意向量均唯一地表示成W_1中的一个向量和W_2中的一个向量之和, 则称W_1+W_2是W_1与W_2的**直和**, 记为$W_1\dot{+}W_2$. 进一步, 若$W_1\perp W_2$, 则称直和$W_1\dot{+}W_2$是W_1与W_2的**正交直和**, 记为$W_1\oplus W_2$.

例1.5.1 在直角坐标系$O-xyz$中，若$W_1=xoy$平面，$W_2=yoz$平面，试判断W_1+W_2是否为直和.

x,y,0
0,y,z

若$W_1=x$轴，$W_2=yoz$平面，试判断W_1+W_2是否为直和.

x,0,0
0,y,z

若W_1为过原点且不在yoz平面的一条直线，$W_2=yoz$平面，试判断W_1+W_2是否为直和.

定理1.5.1（直和判定定理）设W_1与W_2是线性空间V的两个子空间, 则以下命题等价:

（1）W_1+W_2是直和;

（2）W_1+W_2中零元素表法唯一;

（3）$W_1\cap W_2=\{\boldsymbol{\theta}\}$;

（4）$\dim(W_1+W_2)=\dim W_1+\dim W_2$.

定理1.5.1（直和判定定理） 设W_1与W_2是线性空间V的两个子空间, 则以下命题等价:

（1）W_1+W_2是直和;

（4）$\dim(W_1+W_2)=\dim W_1+\dim W_2$.

$$\dim(W_1 \cap W_2)=0$$

例1.5.1 在直角坐标系$O-xyz$中，若$W_1=xoy$平面，$W_2=yoz$平面，试判断W_1+W_2是否为直和.

$$W_1 \cap W_2$$

若$W_1=x$轴，$W_2=yoz$平面，试判断W_1+W_2是否为直和.

若W_1为过原点且不在yoz平面的一条直线，$W_2=yoz$平面，试判断W_1+W_2是否为直和.

例1.5.2 取$V=\mathbb{R}^{n\times n}$, $W_1=\{A\in V|A^T=A\}$, $W_2=\{A\in V|A^T=-A\}$, 则$V=W_1\dot{+}W_2$.

W_1, W_2的维数分别是多少？

例1.5.3 定义$\mathbb{R}^{2\times 2}$的两个线性子空间分别为

$$W_1=\left\{\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}\in\mathbb{R}^{2\times 2}\middle|\begin{array}{r} 2x_1+3x_2-x_3=0 \\ x_1+2x_2+3x_3-x_4=0 \end{array}\right\}$$

$$W_2=\operatorname{span}\left(\begin{bmatrix} 2 & -1 \\ a+2 & 1 \end{bmatrix},\begin{bmatrix} -1 & 2 \\ 4 & a+8 \end{bmatrix}\right)$$

（1）求W_1的一组基;

（2）当a取何值时, W_1+W_2是直和.

例1.5.4 设$A \in \mathbb{C}^{n\times n}$且满足$A^2 = A$, 试证明

$$\mathbb{C}^n = R(A) \dot{+} R(I - A)$$

证明: (1)$R(A) + R(I - A) = R(A) \dot{+} R(I - A)$

(2) $\mathbb{C}^n = R(A) + R(I - A)$

(a) $\dim(\mathbb{C}^n) = \dim(R(A) + R(I - A))$

或 (b) $\mathbb{C}^n \subset R(A) + R(I - A)$

例1.5.4 设$A \in \mathbb{C}^{n\times n}$且满足$A^2 = A$, 试证明

$$\mathbb{C}^n = R(A)\dot{+}R(I-A)$$

证明: (1)$R(A) + R(I-A) = R(A)\dot{+}R(I-A)$

$$R(A) \cap R(I-A) = \{0\}$$

例1.5.4 设$A \in \mathbb{C}^{n\times n}$且满足$A^2 = A$, 试证明

$$\mathbb{C}^n = R(A) \dotplus R(I-A)$$

证明: (2) $\mathbb{C}^n = R(A) + R(I-A)$

(a) $\dim(\mathbb{C}^n) = \dim(R(A) + R(I-A))$

$$\dim\big(R(I-A)\big) = n - rank(A)$$

或 (b) $\mathbb{C}^n \subset R(A) + R(I-A)$

例1.5.4 设$A \in \mathbb{C}^{n\times n}$且满足$A^2=A$, 试证明

$$\mathbb{C}^n = R(A)\dot{+}R(I-A)$$

证明: (2) $\mathbb{C}^n = R(A) + R(I-A)$

(a) $\dim(\mathbb{C}^n) = \dim(R(A) + R(I-A))$

$$\dim\big(R(I-A)\big) = n - rank(A)$$

例1.5.4 设$A \in \mathbb{C}^{n\times n}$且满足$A^2=A$, 试证明

$$\mathbb{C}^n = R(A)\dot{+}N(A),$$

逆命题是否成立呢？

$$A=\begin{bmatrix}0 & 1\\ 0 & 2\end{bmatrix}, A^2=\begin{bmatrix}0 & 2\\ 0 & 4\end{bmatrix}, \mathbb{C}^n = R(A)\dot{+}N(A).$$

补充：设$A \in \mathbb{C}^{n\times n}$, 则$\mathrm{r}(A^2)=\mathrm{r}(A)$当且仅当 $\mathbb{C}^n = R(A)\dot{+}N(A)$

例1.5.4 设$A \in \mathbb{C}^{n\times n}$且满足$A^2 = A$, 试证明

$$\mathbb{C}^n = R(A)\dot{+}N(A),\ R(I - A) = N(\mathrm{A})$$

$$\mathbb{C}^n = R(A)\dot{+}R(I - A)$$

$$\mathbb{C}^n = R(A)\dot{+}N(A),$$

"消元"，直接推出$R(I - A) = N(\mathrm{A})$? 可以吗

定理1.5.2

若子空间$W_1 \perp W_2$，则$W_1 + W_2 = W_1 \oplus W_2$.

注1：对于子空间W，$V = W \oplus W^{\perp}$.

注2：正交直和分解中，若$V = W_1 \oplus W_2$，W_1给定，则W_2唯一确定，且$W_2 = W_1^{\perp}$（正交补空间的唯一性）. 普通的直和分解不具备这样的性质.

例1.5.5 在直角坐标系$O-xyz$中，假设W_1是位于ox轴上所有向量的集合，W_2是过坐标原点且不包括ox轴的任一平面，则$\mathbb{R}^3 = W_1 \dot{+} W_2$（直和分解不唯一）．假若定义$W_1^{\perp} = yoz$平面．此时,$V = W_1 \oplus W_1^{\perp}$（正交直和分解）．

定理1.5.3 设$A=(a_{ij})\in\mathbb{C}^{n\times n}$, 则

$$N(A)\oplus R(A^H)=\mathbb{C}^n \text{且} R(A^H)=(N(A))^{\perp}$$

$N(A)\oplus R(A^H)=\mathbb{C}^n$

当且仅当

$R(A^H)\perp N(A)$ 且$R(A^H)+N(A)=\mathbb{C}^n$

当且仅当

$R(A^H)=(N(A))^{\perp}$.

证明$V = W_1 \oplus W_2$, 先证明$W_1 \perp W_2$， 然后证明

（1）$V = W_1 + W_2$

或者

（2）$\dim(V) = \dim W_1 + \dim W_2$， 从而

$$W_2 = W_1^{\perp}$$

证明$V = W_1 \oplus W_2$, 先证明$W_1 \perp W_2$， 然后证明

（1）$V = W_1 + W_2$

$$\mathbb{C}^n = N(A) + R(A^H)$$

$$\dim\left(N(A)\right) = \mathrm{r}(A), \dim\left(R(A^H)\right) = \mathrm{n} - \mathrm{r}(\mathrm{A})$$

证明$V = W_1 \oplus W_2$, 先证明$W_1 \perp W_2$，然后证明

（1）$V = W_1 + W_2$

$$\mathbb{C}^n = N(A) + R(A^H)$$

证明 $x \in \mathbb{C}^n, x = y + z, y \in N(A), z \in R(A^H)$

假设$z = A^H \xi$, 证明$y = x - z = x - A^H \xi \in N(A)$, 即

$$Ay = A(x - A^H \xi) = 0,$$

$$AA^H \xi = Ax$$

$$\mathrm{rank}(AA^H) = rank(AA^H \vdots A)$$

该方程一定有解

定义1.5.3（投影与正交投影）设W_1与W_2是线性空间V的两个子空间且$V = W_1 \dot{+} W_2$, 对任意向量$\boldsymbol{x} \in V$均可唯一地分解成$\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}$, 其中$\boldsymbol{y} \in W_1$, $\boldsymbol{z} \in W_2$, 此时称向量$\boldsymbol{y}$为向量$\boldsymbol{x}$在W_1上的**投影**. 特别地, 若$V = W_1 \oplus W_2$, 则称向量$\boldsymbol{y}$为向量$\boldsymbol{x}$在W_1上的**正交投影**.

例1.5.6 定义$\mathbb{R}^3$的线性子空间$W = \mathrm{span}(\boldsymbol{x}_1, \boldsymbol{x}_2)$，其中，$\boldsymbol{x}_1 = [2,5,-1]^T$，$\boldsymbol{x}_2 = [-2,1,1]^T$. 求向量$\boldsymbol{y} = [1,2,3]^{\boldsymbol{T}}$在$W$和$W^{\perp}$的正交投影.

例1.5.6 定义$\mathbb{R}^3$的线性子空间$W=\text{span}(\boldsymbol{x}_1,\boldsymbol{x}_2)$, 其中, $\boldsymbol{x}_1=[2,5,-1]^T$, $\boldsymbol{x}_2=[-2,1,1]^T$. 求向量$\boldsymbol{y}=[1,2,3]^T$在$W$和$W^\perp$的正交投影.

解: 方法1. 先求$W^\perp=\text{span}\{\boldsymbol{x}_3\}$, $\boldsymbol{x}_3=(1,0,2)^T$.

再求$\boldsymbol{y}=(1,2,3)^T$在空间$W^\perp$上的投影:

$$\text{Proj}_{W^\perp}\boldsymbol{y}=\frac{(\boldsymbol{y},\boldsymbol{x}_3)}{\|\boldsymbol{x}_3\|}\frac{\boldsymbol{x}_3}{\|\boldsymbol{x}_3\|}=(\frac{7}{5},0,\frac{14}{5})^T.$$

尽管我们利用的欧式空间中夹角的概念求解投影,这里数域并不局限于实数域

例1.5.6 定义$\mathbb{R}^3$的线性子空间$W = \operatorname{span}(\boldsymbol{x}_1, \boldsymbol{x}_2)$, 其中, $\boldsymbol{x}_1 = [2,5,-1]^T$, $\boldsymbol{x}_2 = [-2,1,1]^T$. 求向量$\boldsymbol{y} = [1,2,3]^T$在$W$和$W^\perp$的正交投影.

解:方法1. 先求$W^\perp = \operatorname{span}\{\boldsymbol{x}_3\}$, $\boldsymbol{x}_3 = (1,0,2)^T$.

再求$\boldsymbol{y} = (1,2,3)^T$在空间$W^\perp$上的投影:

$$\operatorname{Proj}_{W^\perp}\boldsymbol{y} = \frac{(\boldsymbol{y}, \boldsymbol{x}_3)}{\|\boldsymbol{x}_3\|}\frac{\boldsymbol{x}_3}{\|\boldsymbol{x}_3\|} = \left(\frac{7}{5}, 0, \frac{14}{5}\right)^T.$$

再由分解的唯一性知,

$$\operatorname{Proj}_W\boldsymbol{y} = \boldsymbol{y} - \operatorname{Proj}_{W^\perp}\boldsymbol{y} = \left(-\frac{2}{5}, 2, \frac{1}{5}\right)^T.$$

例1.5.6 定义$\mathbb{R}^3$的线性子空间$W=\text{span}(\boldsymbol{x}_1,\boldsymbol{x}_2)$，其中，$\boldsymbol{x}_1=[2,5,-1]^T$，$\boldsymbol{x}_2=[-2,1,1]^T$. 求向量$\boldsymbol{y}=[1,2,3]^T$在$W$和$W^{\perp}$的正交投影.

解：方法2. 先求$\boldsymbol{y}=(1,2,3)^T$在空间W上的投影

$$\text{Proj}_W\boldsymbol{y}=\frac{(\boldsymbol{y},\boldsymbol{x}_1)}{\|\boldsymbol{x}_1\|^2}\boldsymbol{x}_1+\frac{(\boldsymbol{y},\boldsymbol{x}_2)}{\|\boldsymbol{x}_2\|^2}\boldsymbol{x}_2=(-\frac{2}{5},2,\frac{1}{5})^T.$$

余下略.

命题1.5.1 若W是V的子空间,$\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$是$W$的一组正交基. 对于$V$中任一向量$\boldsymbol{y}$均可唯一地表示为

$$\boldsymbol{y}=\mathrm{Proj}_W\boldsymbol{y}+\mathrm{Proj}_{W^\perp}\boldsymbol{y}$$

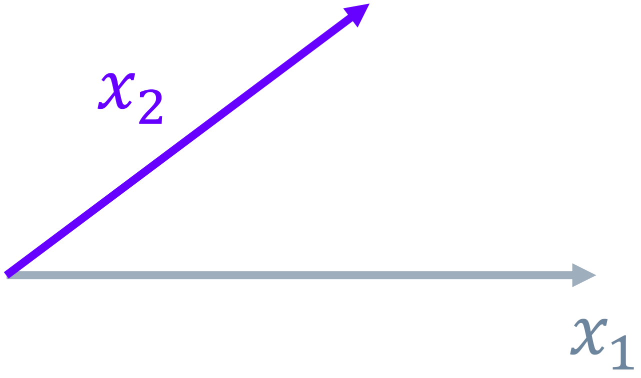
其中$\mathrm{Proj}_W\boldsymbol{y}$和$\mathrm{Proj}_{W^\perp}\boldsymbol{y}$分别为向量$\boldsymbol{y}$在空间$W$和补空间$W^\perp$上的正交投影, 且

$$\mathrm{Proj}_W\boldsymbol{y}=\frac{(\boldsymbol{y},\boldsymbol{x}_1)}{(\boldsymbol{x}_1,\boldsymbol{x}_1)}\boldsymbol{x}_1+\cdots+\frac{(\boldsymbol{y},\boldsymbol{x}_n)}{(\boldsymbol{x}_n,\boldsymbol{x}_n)}\boldsymbol{x}_n$$

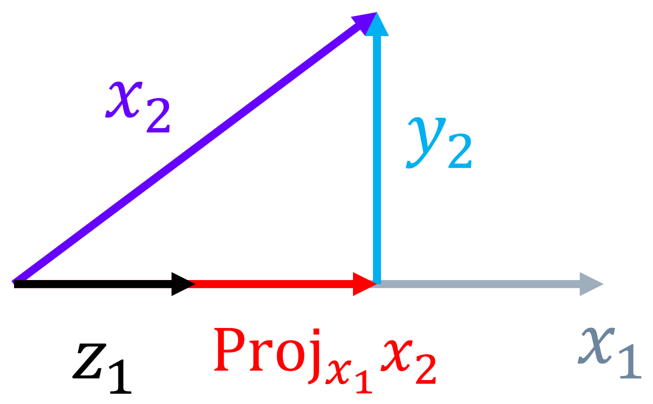
特别地,若$\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$是$W$的一组标准正交基,则

$$\mathrm{Proj}_W\boldsymbol{y}=(\boldsymbol{y},\boldsymbol{x}_1)\boldsymbol{x}_1+\cdots+(\boldsymbol{y},\boldsymbol{x}_n)\boldsymbol{x}_n$$

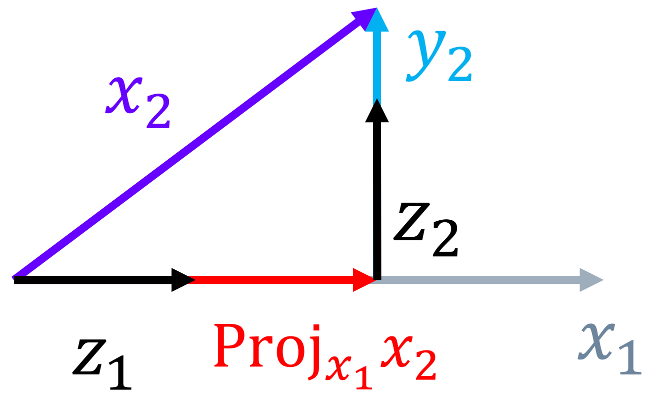
定理1.5.4 内积空间必存在标准正交基.



定理1.5.4 内积空间必存在标准正交基.



定理1.5.4 内积空间必存在标准正交基.



定理1.5.4 内积空间必存在标准正交基.

证明方法是构造性的, 称为Gram-Schmidt正交化方法.

$$\boldsymbol{y}_k = \boldsymbol{x}_k - \sum_{i=1}^{k-1} \frac{(\boldsymbol{x}_k, \boldsymbol{y}_i)}{(\boldsymbol{y}_i, \boldsymbol{y}_i)} \boldsymbol{y}_i$$

$$= \boldsymbol{x}_k - \sum_{i=1}^{k-1} (\boldsymbol{x}_k, \boldsymbol{z}_i)\, \boldsymbol{z}_i$$

$$\boldsymbol{z}_k = \frac{\boldsymbol{y}_k}{\|\boldsymbol{y}_k\|}$$

定理1.5.4 内积空间必存在标准正交基.

构造性证明方法, 称为Gram-Schmidt正交化方法.

$(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n)=(\boldsymbol{z}_1,\cdots,\boldsymbol{z}_n)A$, 其中, 过渡矩阵$A$为正线上三角矩阵, 定义为

$$A=\begin{bmatrix} \|\boldsymbol{y}_1\| & (\boldsymbol{x}_2,\boldsymbol{z}_1) & \cdots & (\boldsymbol{x}_n,\boldsymbol{z}_1) \\ & \|\boldsymbol{y}_2\| & \cdots & (\boldsymbol{x}_n,\boldsymbol{z}_2) \\ & & \ddots & \vdots \\ 0 & & & \|\boldsymbol{y}_n\| \end{bmatrix}.$$

例1.5.7 已知$\mathbb{R}^4$中的一组基$\boldsymbol{x}_1=[1,1,0,0]^T$,$\boldsymbol{x}_2=[1,0,1,0]^T$, $\boldsymbol{x}_3=[-1,0,0,1]^T$, $\boldsymbol{x}_4=[1,-1,-1,1]^T$, 求$\mathbb{R}^4$的一组标准正交基.

定义1.5.3（最佳逼近）设W是线性空间V的非空子集, $\boldsymbol{\alpha} \in V$为给定向量, 若存在$\boldsymbol{x} \in W$满足如下不等式

$$\|\boldsymbol{\alpha} - \boldsymbol{x}\| \le \|\boldsymbol{\alpha} - \boldsymbol{y}\|, \forall \boldsymbol{y} \in W$$

则称$\boldsymbol{x}$是$\boldsymbol{\alpha}$在W的**最佳逼近（最佳近似）向量**.

定理1.5.5（最佳逼近定理）设W是线性空间V的线性子空间, 则V中任一向量$\boldsymbol{x}$在W上都有唯一的最佳逼近, 且$\boldsymbol{x}$在W上的最佳逼近是$\boldsymbol{x}$在W上的正交投影.

例1.5.8（最小二乘问题）在许多实际观测数据的处理中, 若已知量y与量$x_1,\cdots,x_n$间呈线性关系:

$$y=\lambda_1x_1+\cdots+\lambda_nx_n$$

但不知道系数$\lambda_1,\cdots,\lambda_n$. 为确定这些系数, 通常做$m\geq n$次试验, 得到$m$组观测数据

$$\left(x_1^{(k)},x_2^{(k)},\cdots,x_n^{(k)},y^{(k)}\right),k=1,\cdots,m$$

通常按如下意义确定系数: 求$\lambda_1,\cdots,\lambda_n$使得

$$\min_{c_k\in F}\sum_{k=1}^{m}\left|y^{(k)}-\sum_{i=1}^{n}c_ix_i^{(k)}\right|^2$$

设函数$y=f(x)$定义在区间$[a,b]$上,$x_0,x_1,\cdots,x_n$是$[a,b]$上取定的$(n+1)$个互异点, 且仅在这些点处的函数值$y_i=f(x_i)$已知,要构造函数$g(x)$使得

$$g(x_i)=y_i, i=0,1,\cdots,$$

且要求误差$r(x)=f(x)-g(x)$的绝对值在区间$[a,b]$上比较小. 点$x_0,x_1,\cdots,x_n$称为**插值基点**,由插值基点确定的区间为端点的区间）称为**插值区间**, $f(x)$称为**求插函数**, $g(x)$称为**插值函数**, $r(x)$称为**插值余项**.

若选定$P_n(x)$中的一组基为$1, x, x^2, \cdots, x^n$,则

$$g(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

得如下方程组

$$D^T \boldsymbol{a} = \boldsymbol{y}$$

式中,$\boldsymbol{a} = [a_0, a_1, \cdots, a_n]^T$为待定向量,$\boldsymbol{y} = [y_0, y_1, \cdots, y_n]^T$为给定向量, 系数矩阵$D$定义为

$$D = \begin{bmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & \cdots & x_n^n \end{bmatrix}$$

形如D的矩阵称为**Vardermonde矩阵**

$$|D^T| = |D| = \prod_{0\le j<i\le n}(x_i - x_j)$$

则D可逆, 方程有唯一解$\boldsymbol{a} = (D^T)^{-1}\boldsymbol{y}$, 可唯一确定$g(x)$. 为避开矩阵求逆, 重新选取$P_n(x)$的一组基:

$$l_i(x) = \prod_{\substack{j=0,\\ j\neq i}}^{n}\frac{(x - x_j)}{(x_i - x_j)}, i = 1, \cdots, n$$

称为**Lagrange基本多项式**.

因此,我们将

$$p_n(x) = g(x) = \sum_{i=0}^{n} f(x_i) l_i(x)$$

称为**Lagrange插值多项式**.