北京航空航天大学
BEIHANG UNIVERSITY

矩阵理论

王磊

自动化科学与电气工程学院

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
- 直积与投影
- 应用：多项式插值

内 积 空 间

定义1.4.1（内积空间） 设F是实数域或复数域, V是F上的线性空间, 若对V中任意两个向量$\boldsymbol{\alpha}$和$\boldsymbol{\beta}$, 定义了一个数$(\boldsymbol{\alpha},\boldsymbol{\beta})\in F$, 使得对任意向量$\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in V$和$k\in F$满足

（1）**共轭对称性**: $(\boldsymbol{x},\boldsymbol{y})=\overline{(\boldsymbol{y},\boldsymbol{x})}$;

（2）**可加性**: $(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{z})=(\boldsymbol{x},\boldsymbol{z})+(\boldsymbol{y},\boldsymbol{z})$;

（3）**齐次性**: $(k\boldsymbol{x},\boldsymbol{y})=k(\boldsymbol{x},\boldsymbol{y})$;

（4）**正定性**: $(\boldsymbol{x},\boldsymbol{x})\geq 0$,当且仅当$\boldsymbol{x}=\boldsymbol{\theta}$时等号成立.

此时, V称为一个**内积空间**, 数$(\boldsymbol{x},\boldsymbol{y})$称为$\boldsymbol{x}$与$\boldsymbol{y}$的**内积**. 有限维的实内积空间称为**欧几里得空间**. 有限维的复内积空间称为**酉空间**.

注1: 欧氏（酉）空间的维数指它作为线性空间的维数; 它们的线性子空间仍是欧氏（酉）子空间.

注2: 定义1.4.1中性质（3）**齐次性仅对第一个向量成立**, 对第二个向量则是“**共轭齐次性**”.

例1.4.1 设V是F上的线性空间, $\forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in V$和$k\in F$,

$$(\boldsymbol{x},\boldsymbol{y}+\boldsymbol{z})=$$

$$(\boldsymbol{x},k\boldsymbol{y})=$$

注1: 欧氏（酉）空间的维数指它作为线性空间的维数; 它们的线性子空间仍是欧氏（酉）子空间.

注2: 定义1.4.1中性质（3）**齐次性仅对第一个向量成立**, 对第二个向量则是“**共轭齐次性**”.

例1.4.1 设V是F上的线性空间, $\forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in V$和$k\in F$,

$$(\boldsymbol{x},\boldsymbol{y}+\boldsymbol{z})=(\boldsymbol{x},\boldsymbol{y})+(\boldsymbol{x},\boldsymbol{z})$$

$$(\boldsymbol{x},k\boldsymbol{y})=\bar{k}\overline{(\boldsymbol{y},\boldsymbol{x})}=\bar{k}(\boldsymbol{x},\boldsymbol{y})$$

思考: $\forall \boldsymbol{y}\in V,(\boldsymbol{\theta},\boldsymbol{y})=?$

例1.4.2 $\mathbb{R}^n$中, 取$\boldsymbol{x}=[x_1,\cdots,x_n]^T$, $\boldsymbol{y}=[y_1,\cdots,y_n]^T$, 令

$$(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{y}^T\boldsymbol{x}=x_1y_1+\cdots+x_ny_n$$

满足内积四条性质, 故$\mathbb{R}^n$是欧氏空间, 仍记为$\mathbb{R}^n$.

例1.4.3 $\mathbb{C}^n$中$\boldsymbol{x}=[x_1,\cdots,x_n]^T$, $\boldsymbol{y}=[y_1,\cdots,y_n]^T$,令

$$(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{y}^H\boldsymbol{x}=x_1\bar{y}_1+\cdots+x_n\bar{y}_n$$

式中, $\boldsymbol{y}^H=[\bar{y}_1,\cdots,\bar{y}_n]$表示为向量$\boldsymbol{y}$的**共轭转置**.

满足内积四条性质, 故$\mathbb{C}^n$为酉空间, 仍记为$\mathbb{C}^n$.

定义1.4.2（Hermite矩阵） 设矩阵$A=\left(a_{ij}\right)\in\mathbb{C}^{n\times n}$, 若$A^H=A$, 则称矩阵$A$是**Hermite矩阵**; 若$A^H=-A$, 则称$A$是**反Hermite矩阵**.

定义1.4.3（正定矩阵） 设$\boldsymbol{x}=[x_1,\cdots,x_n]^T\in\mathbb{C}^n$是未定元向量, $A=\left(a_{ij}\right)\in\mathbb{C}^{n\times n}$是Hermite矩阵, 定义**复二次型**

$$f(\boldsymbol{x})=\boldsymbol{x}^H A\boldsymbol{x}$$

称A为$f(\boldsymbol{x})$的矩阵. 若$\forall\boldsymbol{x}\in\mathbb{C}^n$, $f(\boldsymbol{x})\geq 0$, 且$f(\boldsymbol{x})=0$当且仅当$\boldsymbol{x}=0$, 则称$f(\boldsymbol{x})$是**正定二次型**, A是**正定矩阵**.

例1.4.4 设复二次型$f(\boldsymbol{x})=\boldsymbol{x}^H A\boldsymbol{x}$, 其中, $\boldsymbol{x}=[x_1,x_2]^T\in\mathbb{C}^2$, 矩阵$A$为

$$A=\begin{bmatrix}1 & \mathrm{i}\\ -\mathrm{i} & 2\end{bmatrix}$$

例1.4.4 设复二次型$f(\boldsymbol{x}) = \boldsymbol{x}^H A\boldsymbol{x}$, 其中, $\boldsymbol{x} = [x_1, x_2]^T \in \mathbb{C}^2$, 矩阵$A$为

$$A = \begin{bmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 2 \end{bmatrix}$$

则$f(\boldsymbol{x}) = |x_1 + \mathrm{i}x_2|^2 + |x_2|^2$. 故$f(\boldsymbol{x})$是正定二次型, A是正定矩阵.

例1.4.4 设复二次型$f(\boldsymbol{x}) = \boldsymbol{x}^H A\boldsymbol{x}$, 其中, $\boldsymbol{x} = [x_1, x_2]^T \in \mathbb{C}^2$, 矩阵$A$为

$$A = \begin{bmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 1 \end{bmatrix}$$

则$f(\boldsymbol{x}) = |x_1 + \mathrm{i}x_2|^2$.

例1.4.4 设复二次型$f(\boldsymbol{x}) = \boldsymbol{x}^H A\boldsymbol{x}$, 其中, $\boldsymbol{x} = [x_1, x_2]^T \in \mathbb{C}^2$, 矩阵$A$为

$$A = \begin{bmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 1 \end{bmatrix}$$

则$f(\boldsymbol{x}) = |x_1 + \mathrm{i}x_2|^2$. 显然, $f(\boldsymbol{x})$是半正定二次型, A是半正定矩阵.

例1.4.5 $\mathbb{C}^n$中, 取$\boldsymbol{x}=[x_1,\cdots,x_n]^T$, $\boldsymbol{y}=[y_1,\cdots,y_n]^T$, 并设Hermite矩阵$A\in\mathbb{C}^{n\times n}$是正定的. 令

$$(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{y}^H A\boldsymbol{x}$$

满足内积的四条性质, 故$\mathbb{C}^n$为酉空间.

例1.4.5 $\mathbb{C}^n$中, 取$\boldsymbol{x}=[x_1,\cdots,x_n]^T$, $\boldsymbol{y}=[y_1,\cdots,y_n]^T$, 并设Hermite矩阵$A\in\mathbb{C}^{n\times n}$是正定的. 令

$$(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{y}^H A\boldsymbol{x}$$

满足内积的四条性质, 故$\mathbb{C}^n$为酉空间.

注4: 同一线性空间可定义不同的内积.

思考: 为什么在酉空间中采用共轭转置而不用普通的向量转置?

定义1.4.4（度量矩阵） 设$\epsilon_1,\cdots,\epsilon_n$是内积空间$V$中的一组基, 称$n$阶矩阵

$$A=\left((\epsilon_i,\epsilon_j)\right)_{n\times n}=\begin{bmatrix}(\epsilon_1,\epsilon_1) & (\epsilon_1,\epsilon_2) & \cdots & (\epsilon_1,\epsilon_n)\\ (\epsilon_2,\epsilon_1) & (\epsilon_2,\epsilon_2) & \cdots & (\epsilon_2,\epsilon_n)\\ \vdots & \vdots & \vdots & \vdots\\ (\epsilon_n,\epsilon_1) & (\epsilon_n,\epsilon_2) & \cdots & (\epsilon_n,\epsilon_n)\end{bmatrix}$$

为V关于基$\epsilon_1,\cdots,\epsilon_n$的**度量矩阵**（或**Gram矩阵**），常记为$G(\epsilon_1,\cdots,\epsilon_n)$.

注5: 内积空间中内积与度量矩阵是一一对应的.

命题1.4.1（度量矩阵的性质）设$G(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)$和$G(\boldsymbol{\epsilon}_1,\cdots,\boldsymbol{\epsilon}_n)$均为内积空间$V$的度量矩阵, 则有

（1）$G(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)$和$G(\boldsymbol{\epsilon}_1,\cdots,\boldsymbol{\epsilon}_n)$是正定Hermite矩阵;

（2）$G(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)$和$G(\boldsymbol{\epsilon}_1,\cdots,\boldsymbol{\epsilon}_n)$**合同**（或**相合**），即存在非奇异矩阵$P$使得

$$P^H G(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)P = G(\boldsymbol{\epsilon}_1,\cdots,\boldsymbol{\epsilon}_n)$$

式中, P是由基$\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n$到基$\boldsymbol{\epsilon}_1,\cdots,\boldsymbol{\epsilon}_n$的过渡矩阵.

定义1.4.5（长度） 设V是内积空间, $\forall \boldsymbol{x} \in V$,实数$\sqrt{(\boldsymbol{x}, \boldsymbol{x})}$称为$\boldsymbol{x}$的**长度（模）**, 记为$\|\boldsymbol{x}\|$. 长度为1的向量称为**单位向量**.

命题1.4.2（长度性质） $\forall \boldsymbol{x}, \boldsymbol{y} \in V$和$k \in F$,有

（1）**正定性**: $\|\boldsymbol{x}\| \geq 0$当且仅当$\boldsymbol{x} = \boldsymbol{0}$时有$\|\boldsymbol{x}\| = 0$;

（2）**齐次性**: $\|k\boldsymbol{x}\| = |k|\|\boldsymbol{x}\|$;

（3）**三角不等式**: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$;

（4）**平行四边形法则**:

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 + \|\boldsymbol{x} - \boldsymbol{y}\|^2 = 2(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2).$$

定理1.4.1（Cauchy不等式）设V是数域F上的内积空间, 对任意$\boldsymbol{x}, \boldsymbol{y} \in V$,

$$|(\boldsymbol{x}, \boldsymbol{y})| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|$$

其中, $|(\boldsymbol{x}, \boldsymbol{y})| = \|\boldsymbol{x}\|\|\boldsymbol{y}\|$当且仅当$\boldsymbol{x}, \boldsymbol{y}$线性相关成立.

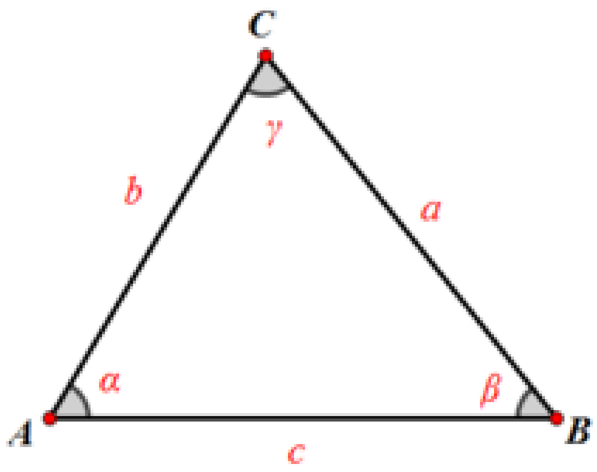
注6: 定义不同内积可得到不同的Cauchy不等式.

（1）对$\mathbb{R}^n$中任两向量$\boldsymbol{x}=[x_1,\cdots,x_n]^T$和$\boldsymbol{y}=[y_1,\cdots,y_n]^T$, 有

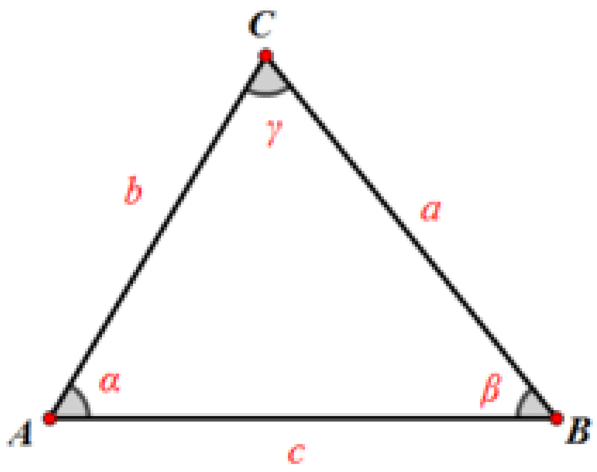
$$\left|\sum_{i=1}^{n} x_i y_i\right| \le \sqrt{\sum_{i=1}^{n} {x_i}^2}\sqrt{\sum_{i=1}^{n} {y_i}^2}$$

（2）对$C_{[a,b]}$中任两函数$f(x)$和$g(x)$, 有

$$\left|\int_a^b f(x)g(x)dx\right| \le \sqrt{\int_a^b f^2(x)dx}\sqrt{\int_a^b g^2(x)dx}$$



$$\cos\alpha = \frac{b^2 + c^2 - a^2}{2bc}$$



$$\cos\alpha = \frac{b^2 + c^2 - a^2}{2bc}$$

若定义$\overrightarrow{AC}$和$\overrightarrow{AB}$为$\mathbb{R}^2$空间的两个向量, 则角α即为向量的夹角.

$$\cos\alpha = \frac{(\overrightarrow{AC}, \overrightarrow{AB})}{\left\|\overrightarrow{AC}\right\| \left\|\overrightarrow{AB}\right\|}$$

定义1.4.6（向量夹角） 设V是欧氏空间, 对V中任意向量$\boldsymbol{x}$和$\boldsymbol{y}$, 定义向量$\boldsymbol{x}$和$\boldsymbol{y}$的**夹角**为

$$\alpha =< \boldsymbol{x},\boldsymbol{y} >= \arccos\frac{(\boldsymbol{x},\boldsymbol{y})}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} \in [0,\pi]$$

例1.4.7 $\mathbb{R}^2$关于基$\boldsymbol{x}_1,\boldsymbol{x}_2$的度量矩阵为

$$G(\boldsymbol{x}_1,\boldsymbol{x}_2) = \begin{bmatrix} (\boldsymbol{x}_1,\boldsymbol{x}_1) & (\boldsymbol{x}_2,\boldsymbol{x}_1) \\ (\boldsymbol{x}_1,\boldsymbol{x}_2) & (\boldsymbol{x}_2,\boldsymbol{x}_2) \end{bmatrix}$$

行列式$|G(\boldsymbol{x}_1,\boldsymbol{x}_2)|$恰好是以向量$\boldsymbol{x}_1$与$\boldsymbol{x}_2$为邻边的平行四边形的面积的平方.

定义1.4.7（正交向量和正交向量组） 设V是内积空间, 对V中任意向量$\boldsymbol{x}$和$\boldsymbol{y}$, $(\boldsymbol{x}, \boldsymbol{y}) = 0$, 则称向量$\boldsymbol{x}$与$\boldsymbol{y}$**正交**（或**垂直**）, 记为$\boldsymbol{x} \perp \boldsymbol{y}$. 若一组非零向量两两正交, 则称为**正交向量组**. 单位向量构成的正交向量组称为**标准正交向量组**.

注7: 零向量$\boldsymbol{\theta}$与任何向量均正交. **正交向量组要求向量均为非零向量**.

注8: 向量$\boldsymbol{x}$与$\boldsymbol{y}$正交当且仅当$\|\boldsymbol{x}+\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$（**勾股定理**）. 该结果可推广至一般情形.

例1.4.8 求$\mathbb{R}^4$中的单位向量$\boldsymbol{x}$使它与$\boldsymbol{\alpha}_1 = [1,1,-1,1]^T$, $\boldsymbol{\alpha}_2 = [1,-1,-1,1]^T$, $\boldsymbol{\alpha}_3 = [2,1,1,3]^T$均正交.

定理1.4.2 正交向量组线性无关.

推论1.4.1 在n维内积空间中, 正交向量组中的向量不会超过n个.

定义1.4.8（标准正交基）在n维内积空间中，由n个向量组成的正交向量组称为**正交基**. 由单位向量组成的正交基称为**标准正交基**.

例1.4.9 考查$[-\pi,\pi]$上的三角函数组

$$1, \cos t, \sin t, \cdots, \cos kt, \sin kt, \cdots$$

定义1.4.8（标准正交基）在n维内积空间中，由n个向量组成的正交向量组称为**正交基**. 由单位向量组成的正交基称为**标准正交基**.

例1.4.9 考查$[-\pi,\pi]$上的三角函数组

$$1, \cos t, \sin t, \cdots, \cos kt, \sin kt, \cdots$$

解: 在区间$[-\pi,\pi]$上的三角函数组

$1, \cos t, \sin t, \cdots, \cos kt, \sin kt, \cdots$,是正交向量组.

光滑函数$f(x)$展开为（**傅里叶三角级数**）

$$f(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx)$$

例1.4.10 设$\boldsymbol{\epsilon}_1,\cdots,\boldsymbol{\epsilon}_n$是线性空间$V$的标准正交基，并定义$\boldsymbol{\alpha}=[\alpha_1,\alpha_2,\cdots,\alpha_n]^T$和$\boldsymbol{\beta}=[\beta_1,\beta_2,\cdots,\beta_n]^T$分别是$V$中向量$\boldsymbol{x}$和$\boldsymbol{y}$在基$\boldsymbol{\epsilon}_1,\cdots,\boldsymbol{\epsilon}_n$下的坐标, 则

$$(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{\beta}^H\boldsymbol{\alpha}$$

注9: 线性空间V的向量组$\boldsymbol{\epsilon}_1,\cdots,\boldsymbol{\epsilon}_n$是标准正交基当且仅当$V$关于$\boldsymbol{\epsilon}_1,\cdots,\boldsymbol{\epsilon}_n$的度量矩阵为单位矩阵. 在标准正交基下, 内积的运算变得更加简单.

定义1.4.9（向量正交于集合） 设V是内积空间, W是V的子集, 若对任意向量$\boldsymbol{y} \in W$和$\boldsymbol{x} \in V$有$\boldsymbol{x} \perp \boldsymbol{y}$, 则称$\boldsymbol{x}$正交（或垂直）于集合$W$, 记为$\boldsymbol{x} \perp W$.

定义1.4.10（两集合正交） 设W_1与W_2是内积空间V的两个子集, 若对任意向量$\boldsymbol{x} \in W_1$和任意向量$\boldsymbol{y} \in W_2$有$\boldsymbol{x} \perp \boldsymbol{y}$, 则称$W_1$与$W_2$是正交的, 记为$W_1 \perp W_2$.

例1.4.11 在直角坐标系$O-xyz$中, 定义过坐标原点的直线$\boldsymbol{x}$和过坐标原点的平面W_1和W_2. 直线$\boldsymbol{x}$在几何上垂直平面W_1, 平面W_1与W_2在几何上相互垂直.

定义1.4.11（子空间正交补） 设W是V的线性子空间, 则集合$W^{\perp}=\{\boldsymbol{x}\in V|\boldsymbol{x}\perp W\}$称为$W$的**正交补**.

定义1.4.11（子空间正交补）设W是V的线性子空间，则集合$W^{\perp}=\{\boldsymbol{x}\in V|\boldsymbol{x}\perp W\}$称为$W$的**正交补**.

定理1.4.3 设W是线性空间V的子空间，则集合$W^{\perp}$是V的线性子空间.

思考: 与W正交的子空间是W的正交补空间吗？

思考: 与W正交的子空间是W的正交补空间吗？

定理1.4.4 设W是线性空间V的线性子空间, 则$V = W + W^{\perp}$.

注10: 我们在说明线性空间V的两个子空间W_1和W_2相等常采用以下两种方法:

（1）W_1和W_2均是由同一向量组所张成的子空间，或更确定的说明它们有相同的一组基;

（2）W_1是W_2的子集（或W_2是W_1的子集）且W_1和W_2的维数相等.

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.

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）$\dim(W_1+W_2)=\dim W_1+\dim W_2$.

Need to prove

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$V=W_1\dot{+}W_2$.

思考: 在线性空间V中, W_1, W_2的维数分别是多少?

定理1.5.2 设W_1与W_2是线性空间V的两个子空间，则$W_1+W_2=W_1\dot{+}W_2$.

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

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

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

（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

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

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$A=(a_{ij})\in\mathbb{R}^{n\times n}$, 则

$$N(A)\oplus R\left(A^T\right)=\mathbb{R}^n\text{且}R\left(A^T\right)=(N(A))^{\perp}$$

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$\boldsymbol{y}=[1,2,3]^T$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$W=\mathrm{span}(\boldsymbol{x}_1,\boldsymbol{x}_2)$, 其中, $\boldsymbol{x}_1=[2,5,-1]^T$, $\boldsymbol{x}_2=[-2,1,1]^T$. 求向量$\boldsymbol{y}=[1,2,3]^T$在$W$和$W^\perp$的正交投影.

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

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

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

再由分解的唯一性知,

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

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$\boldsymbol{y}=[1,2,3]^T$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

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

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

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

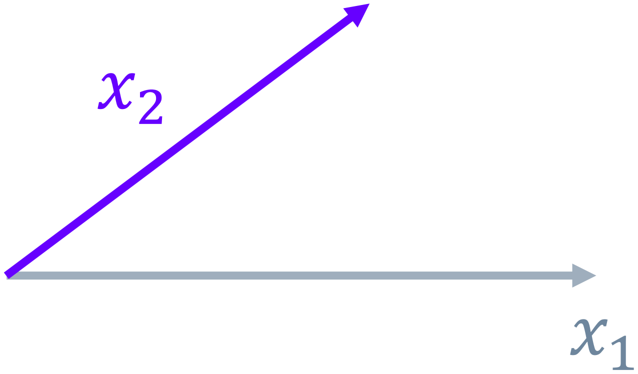
其中$\mathrm{Proj}_W\boldsymbol{y}$和$\mathrm{Proj}_{W^\perp}\boldsymbol{y}$分别为向量$\boldsymbol{y}$在空间$W$和补空间$W^\perp$上的正交投影, 且

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

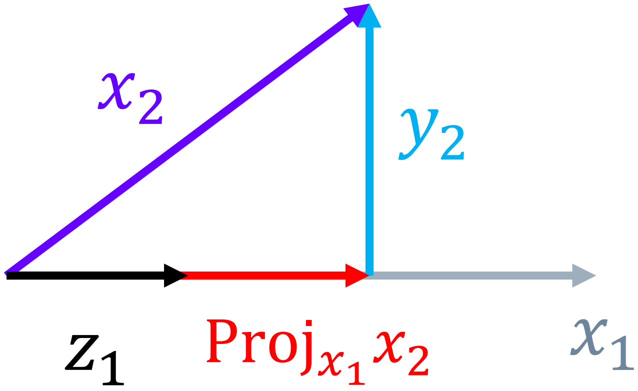
特别地,若$\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$是$W$的一组标准正交基,则

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

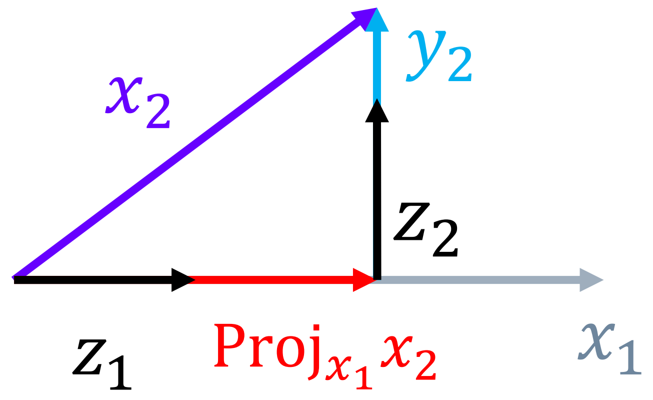
定理1.5.4 内积空间必存在标准正交基.



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.

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

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

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

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

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$\boldsymbol{\alpha}$在W上都有唯一的最佳逼近, 且$\boldsymbol{\alpha}$在W上的最佳逼近是$\boldsymbol{\alpha}$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：

$$y=\lambda_1 x_1+\cdots+\lambda_n x_n$$

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

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

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**

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

则D可逆, 方程有唯一解$\boldsymbol{a} = \left(D^T\right)^{-1}\boldsymbol{y}$, 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:

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

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**.