§4.4 标准正交基和Gram-Schmidt正交化

回顾: $A\in M_{m\times n}(\mathbb{R})$, $A=[\vec{\alpha}_1\ \vec{\alpha}_2\cdots\vec{\alpha}_n]$ 列满秩 ($n\le m$). $\vec\alpha_1,\vec\alpha_2,\cdots,\vec\alpha_n$ 为$C(A)$一组基.

$\forall \vec b\in\mathbb{R}^m$, 则 $\begin{cases} A\vec x=\vec b \text{ 的最小二乘解为: } \hat{\vec x}=(A^TA)^{-1}A^T\vec b, \quad (\text{满足 } (A^TA)\hat{\vec x}=A^T\vec b) \\ \vec b \text{ 在 } C(A) \text{ 上的投影为 } \vec p=A\hat{\vec x}=A(A^TA)^{-1}A^T\vec b. \\ \text{投影矩阵 } P=A(A^TA)^{-1}A^T\in M_{m\times m}(\mathbb{R}) \end{cases}$

当$\{\vec\alpha_i, i=1,2,\cdots,n\}$两两正交时. A^TA为对角阵. 易于计算.

1. 标准正交向量组

定理. 设$\vec\alpha_1,\vec\alpha_2,\cdots,\vec\alpha_n\in\mathbb{R}^m$为两两正交的非零向量, 则$\vec\alpha_1,\vec\alpha_2,\cdots,\vec\alpha_n$线性无关.

Pf. 设 $x_1\vec\alpha_1+x_2\vec\alpha_2+\cdots+x_n\vec\alpha_n=\vec 0$

两边分别点乘$\vec\alpha_i$, 得: $x_1(\vec\alpha_1\cdot\vec\alpha_i)+x_2(\vec\alpha_2\cdot\vec\alpha_i)+\cdots+x_n(\vec\alpha_n\cdot\vec\alpha_i)=\vec 0$

两两正交: $\vec\alpha_i\cdot\vec\alpha_j=0$, 对 $i\ne j$

$\Rightarrow x_i\|\vec\alpha_i\|^2=0$. 又 $\vec\alpha_i\ne\vec0$. $\therefore \|\vec\alpha_i\|^2>0$

$\Rightarrow x_i=0$. $i=1,2,\cdots n$

$\therefore \vec\alpha_1,\vec\alpha_2,\cdots,\vec\alpha_n$线性无关. #

推论: $\mathbb{R}^m$中两两正交的非零向量个数不超过m.

定义. 若$\vec\alpha_1,\vec\alpha_2,\cdots,\vec\alpha_n$两两正交. 即 $\vec\alpha_i\cdot\vec\alpha_j=0$. $\forall i\ne j$. 称向量组$\vec\alpha_1,\vec\alpha_2,\cdots,\vec\alpha_n$为正交(orthogonal)向量组.

进一步地, 若 $\|\vec\alpha_i\|=1$, $i=1,2,\cdots,n$, 称$\vec\alpha_1,\vec\alpha_2,\cdots,\vec\alpha_n$为标准正交(orthonormal)向量组.

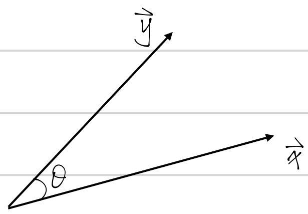
定理 设$Q=[\vec q_1,\vec q_2,\cdots,\vec q_n]\in M_{m\times n}(\mathbb R)$, $n\le m$; 若$\vec q_1,\vec q_2,\cdots,\vec q_n$为$\mathbb R^m$中的标准正交向量组, 则 $Q^TQ=I_n$.

进一步地, 若$m=n$, 则Q为正交阵, 即 $QQ^T=Q^TQ=I_n$, 而 $Q^T=Q^{-1}$

Pf. $(Q^TQ)_{ij}=\vec q_i^T\vec q_j=\vec q_i\cdot\vec q_j=\delta_{ij}\Rightarrow Q^TQ=I_n$. #

定理. $Q\in M_{m,n}(\mathbb R)$定义了一个从$\mathbb R^n$到$\mathbb R^m$的线性映射($\vec x\in\mathbb R^n\to Q\vec x\in\mathbb R^m$). 若$Q^TQ=I_n$, 则该线性映射$Q$保持点积

即 $(Q\vec x)\cdot(Q\vec y)=\vec x\cdot\vec y$. $\forall\vec x,\vec y\in\mathbb R^n$.

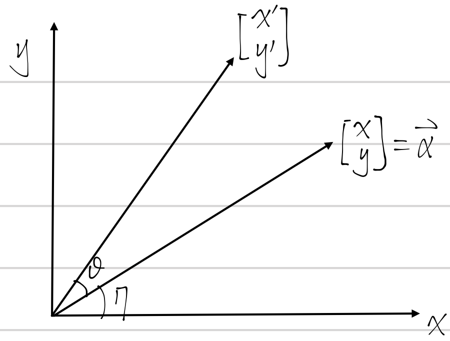


Pf. $(Q\vec x)\cdot(Q\vec y)=(Q\vec x)^TQ\vec y=\vec x^TQ^TQ\vec y=\vec x^T\vec y=\vec x\cdot\vec y$. #

推论. (1). Q保持长度: $\|Q\vec x\|^2=(Q\vec x)\cdot(Q\vec x)=\vec x\cdot\vec x=\|\vec x\|^2$. $\forall \vec x\in\mathbb R^n$.

(2) Q保持角度: $\cos\theta=\dfrac{(Q\vec x)\cdot(Q\vec y)}{\|Q\vec x\|\,\|Q\vec y\|}=\dfrac{\vec x\cdot\vec y}{\|\vec x\|\|\vec y\|}$

例. $\mathbb{R}^2$ 中的旋转变换 $(m=n=2)$



将 $\begin{bmatrix}x\\y\end{bmatrix}\in\mathbb{R}^2$ 绕原点逆时针旋转θ角，得 $\begin{bmatrix}x'\\y'\end{bmatrix}$.

$$\begin{cases} x' = |\vec{\alpha}|\cos(\theta+\eta) = |\vec{\alpha}|\cos\theta\cos\eta - |\vec{\alpha}|\sin\theta\sin\eta = x\cos\theta - y\sin\theta \\ y' = |\vec{\alpha}|\sin(\theta+\eta) = |\vec{\alpha}|\sin\theta\cos\eta + |\vec{\alpha}|\cos\theta\sin\eta = x\sin\theta + y\cos\theta \end{cases}$$

$$\Rightarrow \begin{bmatrix}x'\\y'\end{bmatrix} = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}. \qquad Q = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix} := [\vec{q}_1\ \vec{q}_2]$$

① $\vec{q}_1\cdot\vec{q}_2 = 0 \Rightarrow \vec{q}_1\perp\vec{q}_2$ 列向量正交

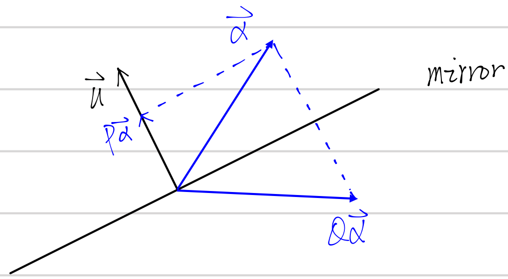
② $\|\vec{q}_1\|^2 = \|\vec{q}_2\|^2 = \sin^2\theta+\cos^2\theta = 1$ 单位向量

$\Rightarrow Q^TQ = I_n$. 又 Q 是方阵 $\Rightarrow$ Q 是正交矩阵. $Q^{-1}=Q^T$.

$\vec{q}_1, \vec{q}_2$ 构成 $\mathbb{R}^2$ 中一组标准正交基, $\vec{q}_1 = Q\vec{e}_1$, $\vec{q}_2 = Q\vec{e}_2$. 分别由 $\vec{e}_1, \vec{e}_2$ 旋转而得. Fig 4.10

$$Q^{-1} = Q^T = \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix} = \begin{bmatrix}\cos(-\theta) & -\sin(-\theta)\\ \sin(-\theta) & \cos(-\theta)\end{bmatrix}$$，即逆时针旋转 $-\theta$ 角，还原为 $\begin{bmatrix}x\\y\end{bmatrix}$.

例. $\mathbb{R}^2$ 中的镜面反射 (reflection)



$\vec{u}$ 为正交于mirror的单位向量. $\|\vec{u}\|=1$

$\forall\vec{\alpha}\in\mathbb{R}^2$, 其关于mirror的镜面反射: $Q\vec{\alpha}\in\mathbb{R}^2$.

$\vec{\alpha}$ 向 $\vec{u}$ 所在直线的投影: $P\vec{\alpha} = \frac{\vec{u}\vec{u}^T}{\vec{u}^T\vec{u}}\vec{\alpha} = (\vec{u}\vec{u}^T)\vec{\alpha}$

$Q\vec{\alpha} = \vec{\alpha} - 2P\vec{\alpha} = (I-2P)\vec{\alpha} = (I-2\vec{u}\vec{u}^T)\vec{\alpha}$

$\Rightarrow Q = I_2 - 2\vec{u}\vec{u}^T = Q^T$ symmetric.

$$Q^TQ = (I_2 - 2\vec{u}\vec{u}^T)(I_2 - 2\vec{u}\vec{u}^T) = I_2 - 4\vec{u}\vec{u}^T + 4\vec{u}\underbrace{\vec{u}^T\vec{u}}_{\|\vec{u}\|^2=1}\vec{u}^T = I_2.$$

$\Rightarrow$ Q 为正交矩阵, $Q^TQ = QQ^T = I_2$. $\Rightarrow Q^T = Q^{-1}$

Moreover, $Q^2 = Q^TQ = I$. reflecting twice through a mirror brings back the original.

2. Projections Using Othonormal Bases

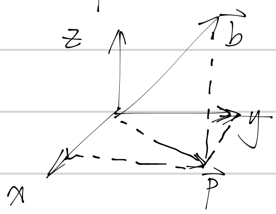
W 为 $\mathbb{R}^m$ 的 n 维 $(n\le m)$ 子空间. 取 W 中一组标准正交基 $\vec{q}_1, \vec{q}_2, \cdots, \vec{q}_n \in\mathbb{R}^m$

则 $W = C(Q)$, where $Q = [\vec{q}_1\ \vec{q}_2 \cdots \vec{q}_n] \in M_{m\times n}(\mathbb{R})$.

$\Rightarrow Q^TQ = I_n$

$\forall\vec{b}\in\mathbb{R}^m$, 则
$$\begin{cases} Q\vec{x} = \vec{b} \text{ 的最小二乘解满足 normal eqn. } \overbrace{Q^TQ}^{I_n}\hat{\vec{x}} = Q^T\vec{b} \Rightarrow \hat{\vec{x}} = Q^T\vec{b} \\ \vec{b} \text{ 在 } C(Q) \text{ 上的投影为 } \vec{p} = Q\hat{\vec{x}} = QQ^T\vec{b} \quad (m\ne n \text{ 时}, QQ^T\ne I_m\,!) \\ \text{投影矩阵 } P = Q(Q^TQ)^{-1}Q^T = QQ^T \end{cases}$$

$\hat{\vec{x}} = Q^T\vec{b} \Rightarrow \begin{bmatrix}\hat{x}_1\\ \hat{x}_2\\ \vdots\\ \hat{x}_n\end{bmatrix} = \begin{bmatrix}\vec{q}_1^T\\ \vec{q}_2^T\\ \vdots\\ \vec{q}_n^T\end{bmatrix}\vec{b} \Rightarrow \hat{x}_1 = \vec{q}_1^T\vec{b},\ \hat{x}_2 = \vec{q}_2^T\vec{b},\ \cdots,\ \hat{x}_n = \vec{q}_n^T\vec{b}$ decoupled.



$\vec{p} = Q\hat{\vec{x}} = \hat{x}_1\vec{q}_1 + \hat{x}_2\vec{q}_2 + \cdots + \hat{x}_n\vec{q}_n = \vec{q}_1\vec{q}_1^T\vec{b} + \vec{q}_2\vec{q}_2^T\vec{b} + \cdots + \vec{q}_n\vec{q}_n^T\vec{b}$ (*)

即将 $\vec{b}$ 分别向一维子空间 $span\{\vec{q}_1\}$, $span\{\vec{q}_2\}$, $\cdots$, $span\{\vec{q}_n\}$ 上投影, 再相加!

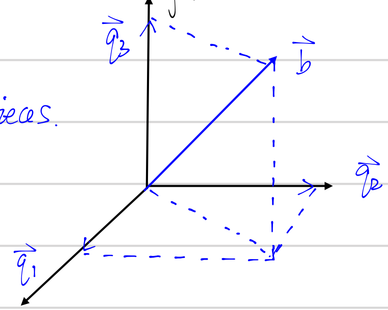
特别地, 当 $m=n$ 时, Q 为正交矩阵. $W = C(Q) = \mathbb{R}^m$. the subspace is the whole space. $Q^T = Q^{-1}$.

① $\hat{\vec{x}} = Q^T\vec{b} = Q^{-1}\vec{b}$. the solution is exact

② $\vec{p} = Q\hat{\vec{x}} = QQ^{-1}\vec{b} = \vec{b}$ the projection of $\vec{b}$ onto the whole space is $\vec{b}$ itself.

③ $P = QQ^T = I_n$ 恒等变换.

(*) $\Rightarrow \vec{b} = (\vec{q}_1^T\vec{b})\vec{q}_1 + (\vec{q}_2^T\vec{b})\vec{q}_2 + \cdots + (\vec{q}_n^T\vec{b})\vec{q}_n$ break $\vec{b}$ into perpendicular pieces.



$\vec{b}$ 展开成 $\mathbb{R}^m$ 中标准正交基的组合, 组合系数 $x_i = \vec{b}\cdot\vec{q}_i$.

也可如下推导:

设 $\vec{b} = x_1\vec{q}_1 + x_2\vec{q}_2 + \cdots + x_n\vec{q}_n$. $\vec{x}$ 为 $\vec{b}$ 的坐标

$\Rightarrow \vec{b}\cdot\vec{q}_1 = x_1\|\vec{q}_1\|^2 \Rightarrow x_1 = \vec{b}\cdot\vec{q}_1$

$\vec{b}\cdot\vec{q}_2 = x_2\|\vec{q}_2\|^2 \Rightarrow x_2 = \vec{b}\cdot\vec{q}_2$

$\vdots$

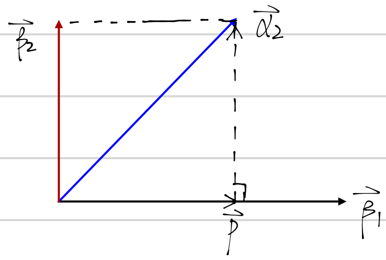
$\vec{b}\cdot\vec{q}_n = x_n\|\vec{q}_n\|^2 \Rightarrow x_n = \vec{b}\cdot\vec{q}_n$

3. Gram-Schmidt 过程

n 个线性无关的向量 $\vec{\alpha}_1, \vec{\alpha}_2, \cdots, \vec{\alpha}_n \in \mathbb{R}^m$ $(n \le m)$ $\longrightarrow$ 标准正交向量组 $\vec{q}_1, \vec{q}_2, \cdots, \vec{q}_n \in \mathbb{R}^m$.

Step 1 正交化.

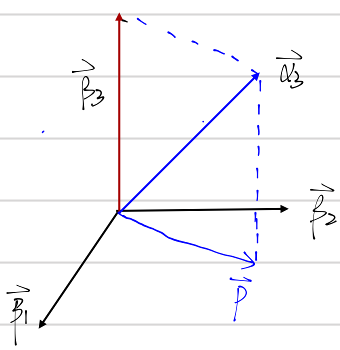
1°. 取 $\vec{\beta}_1 = \vec{\alpha}_1 \neq \vec{0}$



2°. 从 $\vec{\alpha}_2$ 中减去其在 $\vec{\beta}_1$ 上的投影部分, 得 $\vec{\beta}_2$

$$\vec{\beta}_2 = \vec{\alpha}_2 - \frac{\vec{\beta}_1^T\vec{\alpha}_2}{\vec{\beta}_1^T\vec{\beta}_1}\vec{\beta}_1, \quad \Rightarrow \vec{\beta}_2 \perp \vec{\beta}_1$$

这里 $\vec{\beta}_2 \neq \vec{0}$, 否则 $\vec{\alpha}_1$ 与 $\vec{\alpha}_2$ 线性相关.



3°. 从 $\vec{\alpha}_3$ 中减去其在 $span\{\vec{\beta}_1, \vec{\beta}_2\}$ 上的投影, 得 $\vec{\beta}_3$

$\Rightarrow \vec{\beta}_3 \perp span\{\vec{\beta}_1, \vec{\beta}_2\} \Rightarrow \vec{\beta}_3 \perp \vec{\beta}_1$ 且 $\vec{\beta}_3 \perp \vec{\beta}_2$.

又有 $\vec{\beta}_1 \perp \vec{\beta}_2$ $\therefore$ $\vec{\alpha}_3$ 在 $span\{\vec{\beta}_1, \vec{\beta}_2\}$ 上的投影即为 $\vec{\alpha}_3$ 分别在 $\vec{\beta}_1$ 及 $\vec{\beta}_2$ 上的投影之和.

$$\vec{\beta_3} = \vec{\alpha_3} - \frac{\vec{\beta_1}^T \vec{\alpha_3}}{\vec{\beta_1}^T \vec{\beta_1}} \vec{\beta_1} - \frac{\vec{\beta_2}^T \vec{\alpha_3}}{\vec{\beta_2}^T \vec{\beta_2}} \vec{\beta_2} \neq \vec{0}$$，否则 $\vec{\alpha_1}, \vec{\alpha_2}, \vec{\alpha_3}$ 线性相关.

$\vdots$

一般地，$$\vec{\beta_k} = \vec{\alpha_k} - \frac{\vec{\beta_1}^T \vec{\alpha_k}}{\vec{\beta_1}^T \vec{\beta_1}} \vec{\beta_1} - \cdots - \frac{\vec{\beta_{k-1}}^T \vec{\alpha_k}}{\vec{\beta_{k-1}}^T \vec{\beta_{k-1}}} \vec{\beta_{k-1}}$$

即从 $\vec{\alpha_k}$ 中分别减去其在 $\vec{\beta_1}, \vec{\beta_2}, \cdots, \vec{\beta_{k-1}}$ 方向上的投影.

$\vdots$

可得 $\vec{\beta_1}, \vec{\beta_2}, \cdots, \vec{\beta_n}$ 正交

Step 2. 单位化

$\vec{q_i} = \vec{\beta_i} / \|\vec{\beta_i}\|$，$i=1, 2, \cdots, n$.