矩阵论 (Matrix theory)

数理与统计学院

SUES

第二章

内积空间

矩阵论 MATRIX THEORY

线性空间中向量的运算仅是线性运算。一般而言，我们知道，现实世界是3维欧氏空间。对于n维线性空间，定义了内积以后，向量不仅有了长度（模），还有了两向量之间的夹角等几何性质。特别是有了正交概念后，我们可以得到标准正交基、勾股定理、正交投影等许多优美的结果。

当向量元素在复数域内取值时，欧氏空间就被推广到了酉空间。许多欧氏空间中的定义和性质几乎可以"平滑地"推广到酉空间。欧氏空间和酉空间统称为内积空间。

§1、内积与欧氏空间

向量空间中向量的长度与夹角是用内积定义的，因此要在线性空间中引入相关概念，自然要对内积的概念进行推广。

由于向量的内积与向量的线性运算无关，所以欧氏空间实际上是特殊的线性空间，即定义了内积的线性空间。

矩阵论 MATRIX THEORY

一、内积空间的概念

在线性代数中，我们将 R^3 中的内积推广到 R^n：

$$(x,y) \equiv x_1 y_1 + \cdots + x_n y_n = x^T y = y^T x, \ \forall x, y \in R^n$$

并在此基础上定义了 R^n 中的向量长度、夹角等概念。

我们当然可以将这种定义推广到任意线性空间，但注意到这种定义与向量空间的基有关，我们目前不打算这样做。因此我们需要找出线性空间的内积的定义。

注意到 R^n 中的内积显然具有如下性质：

(1) 对称性：$(x, y) = (y, x)$；

(2) 双线性性：$(x + y, z) = (x, z) + (y, z)$;

$$(x,\ y+z) = (x, y) + (x, z);$$

$$(kx, y) = k(x, y)\ , \quad k \in R;$$

$$(x, ky) = k(x, y)\ , \quad k \in R;$$

(3) 正性：$(x, x) \geq 0$;

(4) 定性：$(x, x)=0 \Leftrightarrow x = O$.

据此，我们可给出线性空间中内积的公理化定义。

定义2. 1. 1　V 是实数域 R 上的线性空间。如果对V 中任意两个向量　α、$\beta\in V$都存在一个实数(记为 (α,β))与之相对应，并满足下面四个条件。则实数　(α,β)　称为向量α　与 β 的内积，

(1) $(\alpha,\beta)=(\beta,\alpha)$；　$\forall\alpha$、β、$\gamma\in V$

(2) $(\lambda\alpha,\beta)=\lambda(\alpha,\beta)$；　$(\forall\lambda\in R)$

(3) $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$；

(4) $(\alpha,\alpha)\geq 0$ ，当且仅当 $\alpha=O$ 时，等号成立。

定义了内积的线性空间V 为实内积空间，简称欧氏空间。

例1 在 R^n 中，对任意两个向量

$$\alpha=(a_1,a_2,\cdots,a_n)^T\in R^n \text{ 及 } \beta=(b_1,b_2,\cdots,b_n)^T\in R^n$$

定义了标准内积

$$(\alpha,\beta)\equiv\beta^T\alpha=\alpha^T\beta$$

$$=a_1b_1+a_2b_2+\cdots+a_nb_n.$$

的 R^n 成为一欧式空间。

例 2 在矩阵空间 $R^{m\times n}$ 中，对任意 A、$B\in R^{m\times n}$

定义

$$(A,B)\equiv tr(B^T A)=tr(A^T B)=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}b_{ij}$$

则 $R^{m\times n}$ 是定义了内积 (A,B) 的内积空间。

例3 在线性空间 $R[a,b]$中

定义内积

$$(f,g) \equiv \int_a^b f(x)g(x)dx\,, \quad f、g \in R[a,b]$$

则 $R[a,b]$ 关于以上内积也构成一欧式空间。

矩阵论 MATRIX THEORY

根据前面的分析，欧氏空间中内积还具有下列性质。

(1) $(\alpha,\lambda\beta)=\lambda(\alpha,\beta)$， $\lambda\in R$;

(2) $(\alpha,O)=(O,\beta)=0$.

(3) $(\alpha,\ \beta+\gamma)=(\alpha,\beta)+(\alpha,\gamma)$;

(4) $(\sum_{i=1}^{n}a_i\alpha_i,\sum_{j=1}^{n}b_j\beta_j)=\sum_{i,j=1}^{n}a_ib_j(\alpha_i,\beta_j)$.

二、欧氏空间的度量

注意到3维空间中，

$$|x| \equiv \sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{(x,x)}, \quad x = [x_1, x_2, x_3]^T$$

定义2.1.2 欧氏空间 V 中的向量 α 的**长度（范数）**为

$$\|\alpha\| \equiv \sqrt{(\alpha,\alpha)}.$$

特别地，称 $\|\alpha\| = 1$ 的向量 α 称为**单位向量**。

任意非零向量 α ，经过**规范化**或**单位化**后可得到单位向量

$$\varepsilon = \frac{\alpha}{\|\alpha\|}.$$

定理2.1.1（柯西-施瓦茨不等式） 若 V 是数域 R 上的欧氏空间，则对 V 中的任意向量 α、$\beta \in V$，有

$$\boxed{|(\alpha,\beta)| \le \|\alpha\|\|\beta\|.}$$

证明： 对任意 $\lambda \in R$

显然
$$0 \le (\alpha+\lambda\beta, \alpha+\lambda\beta)$$
$$=(\alpha,\alpha)+2\lambda(\alpha,\beta)+\lambda^2(\beta,\beta)$$

取 $\lambda = -\dfrac{(\alpha,\beta)}{(\beta,\beta)}$ 时，代入上式 即得不等式成立。

当 $\beta \ne 0$ 时，取 $\alpha = -\lambda\beta$ 即得等式

类似于高等数学，根据柯西-施瓦茨不等式，我们称

$$\theta \equiv \arccos\frac{(\alpha,\beta)}{\|\alpha\|\cdot\|\beta\|},\quad \theta\in[0,\pi],\quad \|\alpha\|、\|\beta\|\neq 0$$

为欧氏空间 V 中向量α与β 的**夹角**。

特别地，当 $(\alpha,\beta)=0$ 时，称 α 与 β **正交**或**垂直**，记为 $\alpha\perp\beta$ 。

另外欧氏空间中的范数显然具有下列性质。

性质1 如果 V 是数域 R 上的欧氏空间，则对 V 中的任意向量 α、$\beta \in V$，具有下列三条性质（非负性、正齐性和三角不等式）：

(1) $\|\alpha\| \geq 0$，当且仅当 $\alpha = O$ 时，等号成立。

(2) $\|\lambda\alpha\| = |\lambda|\,\|\alpha\|$； $(\forall \lambda \in R)$

（3） $\|\alpha+\beta\| \leq \|\alpha\| + \|\beta\|$。

范数还具有下列平行四边形法则和勾股定理。

性质2 如果 V 是数域 R 上的欧氏空间，则对 V 中的任意向量 α、$\beta \in V$，有：

(1) $\|\alpha+\beta\|^2+\|\alpha-\beta\|^2=2(\|\alpha\|^2+\|\beta\|^2)$;

（2）特别地，当 α 与 β 正交时，有

$$\|\alpha+\beta\|^2=\|\alpha\|^2+\|\beta\|^2 .$$

最后我们给出欧氏空间 V 的内积的坐标表示形式。

设 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 为 V 的任意一组基，向量 α,β 在此基下的坐标分别为

$$x=(x_1,x_2,\cdots,x_n)^T,\ y=(y_1,y_2,\cdots,y_n)^T.$$

则内积

$$(\alpha,\beta)=\left(\sum_{i=1}^{n}x_i\alpha_i,\sum_{j=1}^{n}y_j\alpha_j\right)$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}x_iy_j(\alpha_i,\alpha_j)$$

$$\equiv y^TGx$$

这里**度量矩阵**（**metric matrix**，也称**Gram矩阵**）

$$G \equiv \begin{bmatrix} (\alpha_1,\alpha_1) & (\alpha_1,\alpha_2) & \cdots & (\alpha_1,\alpha_n) \\ (\alpha_2,\alpha_1) & (\alpha_2,\alpha_2) & \cdots & (\alpha_2,\alpha_n) \\ \vdots & \vdots & & \vdots \\ (\alpha_n,\alpha_1) & (\alpha_n,\alpha_2) & \cdots & (\alpha_n,\alpha_n) \end{bmatrix}$$

可以证明**Gram矩阵** G 是对称正定矩阵。

§2、欧式空间的正交基

正交性的重要性无论怎么强调都不过分，尤其在数值线性代数和微分方程数值解中，许多重要的算法都与正交性有密切联系。而这两门学科是在工程科学中有着最广泛应用的数学学科之一。

在欧氏空间内引入标准正交基后，欧氏空间内向量的内积运算就转化成了我们熟悉的向量空间内向量的内积运算。

一、标准正交基

定义2.2.1 设 V 为 n 维欧氏空间， $\alpha, \beta \in V$，若 $(\alpha,\beta)=0$ ，则称 α 与 β **正交或垂直，记作** $\alpha \perp \beta$

定义2.2.2 欧氏空间 V 的一组基称为 V 的一个**正交基**，如果它们两两正交。若此正交基的每个基向量又都是单位向量，则称此基为 V 的一个**标准正交基**

。例如 欧氏空间 R^n 的一个标准正交基是

$$e_1=(1,0,\cdots,0)^T, e_2=(0,1,\cdots,0)^T,\cdots,e_n=(0,0,\cdots,1)^T,$$

向量组的正交性与线性无关性有什么联系呢？

命题： 向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是欧氏空间 V的非零的正交向量组，则 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 必线性无关。

证明： 设有 $\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_s\alpha_s=0$

等式两边与 $\alpha_j\ (j=1,2,\cdots,s)$ 作内积，

$$(\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_s\alpha_s,\alpha_j)=\sum_{i=1}^{s}(\lambda_i\alpha_i,\alpha_j)=0$$

注意到 (α_i,α_j)等于0 $(i\neq j)$ 以及 $(\alpha_j,\alpha_j)\neq 0$

从而 $\lambda_j=0\ (j=1,2,\cdots,s)$ ，得证。

注： 反之，不成立。

如何求欧氏空间的标准正交基呢？

问：一个线性无关组，在成为基之后，如何成为标准正交基？

根据定义，标准正交基剩下**两个要件**：

（1）是向量个数与维数相等的**正交向量组**；

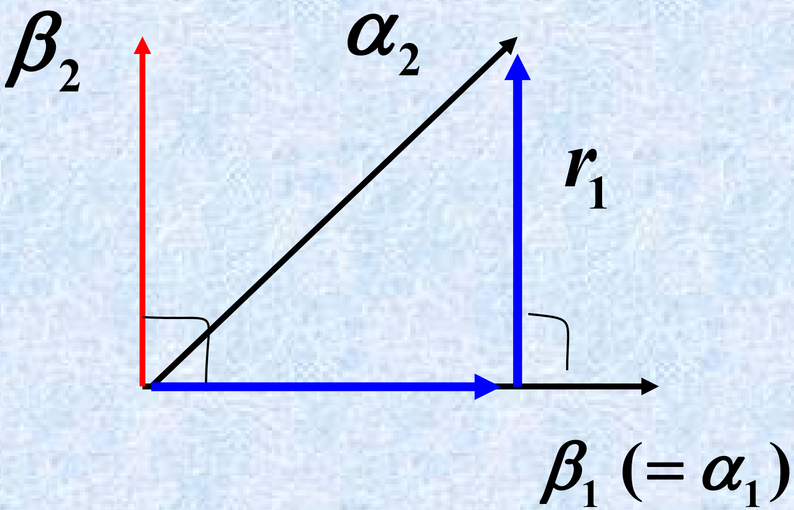
（2）是每个向量都是单位向量的单位向量组。

在标准正交基的两个要件中，正交性显然很不容易达到。下面我们把注意力集中在**如何首先从已知基获得正交基？**

设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是定义了内积的线性空间（即欧氏空间）V 的一个基， $\beta_1,\beta_2,\cdots,\beta_n$ 是我们希望得到的 V 的一个正交基?

显然，我们可令 $\beta_1=\alpha_1$. 如何得到 β_2 呢?

联想到正交分解，我们考察 α_2 在 β_1 即 α_1 上作投影后的**残差向量**，令



$$\boxed{r_1 \equiv \alpha_2-(\alpha_2,\alpha_1)\frac{\alpha_1}{(\alpha_1,\alpha_1)}}$$

$$=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1$$

显然有 $(r_1,\beta_1)=(r_1,\alpha_1) \; =(\alpha_2,\alpha_1)-\dfrac{(\alpha_2,\alpha_1)}{(\alpha_1,\alpha_1)}(\alpha_1,\alpha_1)=0$

故可令 $\beta_2=\alpha_2-\dfrac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1$. 此时 $(\beta_2,\beta_1)=0$

继续考察 α_3 在 β_1,β_2 上作投影后的**残差向量**

$$\boxed{r_2\equiv\alpha_3-(\alpha_3,\beta_1)\frac{\beta_1}{(\beta_1,\beta_1)}-(\alpha_3,\beta_2)\frac{\beta_2}{(\beta_2,\beta_2)}}$$

显然有 $(r_2,\beta_1)=(r_2,\beta_2)=0$ 故可令

$$\beta_3=r_2 \; =\alpha_3-(\alpha_3,\beta_1)\frac{\beta_1}{(\beta_1,\beta_1)}-(\alpha_3,\beta_2)\frac{\beta_2}{(\beta_2,\beta_2)}$$

一般地，可令

$$\beta_1 = \alpha_1$$

$$\beta_2 = \alpha_2 - \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)} \beta_1$$

$$\vdots$$

$$\beta_n = \alpha_n - \frac{(\alpha_n, \beta_1)}{(\beta_1, \beta_1)} \beta_1 - \frac{(\alpha_n, \beta_2)}{(\beta_2, \beta_2)} \beta_2 - \cdots - \frac{(\alpha_n, \beta_{n-1})}{(\beta_{n-1}, \beta_{n-1})} \beta_{n-1}$$

至此，我们就得到了矩阵计算中具有基础性作用的

Gram-Schmidt正交化方法。

定理2.2.2 任一n维欧氏空间 V 都存在标准正交基。

设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 V 的一组基，按施密特正交化过程得正交基

$$\beta_1,\beta_2,\cdots,\beta_n$$

显然，将正交化后得到的正交基再单位化，就得到的了标准正交基

$$\gamma_1=\frac{\beta_1}{\|\beta_1\|},\gamma_2=\frac{\beta_2}{\|\beta_2\|},\cdots,\gamma_n=\frac{\beta_n}{\|\beta_n\|}.$$

例1 试把$\alpha_1=(1,1,0,0),\alpha_2=(1,0,1,0),\alpha_3=(-1,0,0,1)$,
$\alpha_4=(1,-1,-1,1)$正交单位化。

正交化。

$$\beta_1=\alpha_1=(1,1,0,0)$$

$$\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1=(\frac{1}{2},-\frac{1}{2},1,0)$$

$$\beta_3=\alpha_3-\frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2=(-\frac{1}{3},\frac{1}{3},\frac{1}{3},1)$$

$$\beta_4=\alpha_4-\frac{(\alpha_4,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_4,\beta_2)}{(\beta_2,\beta_2)}\beta_2-\frac{(\alpha_4,\beta_3)}{(\beta_3,\beta_3)}\beta_3=(1,-1,-1,1)$$

单位化。

$$\gamma_1=\frac{\beta_1}{\|\beta_1\|}=(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0,0),$$

$$\gamma_2=\frac{\beta_2}{\|\beta_2\|}=(\frac{1}{\sqrt{6}},-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}},0),$$

$$\gamma_3=\frac{\beta_3}{\|\beta_3\|}=(-\frac{1}{\sqrt{12}},\frac{1}{\sqrt{12}},\frac{1}{\sqrt{12}},\frac{3}{\sqrt{12}}),$$

$$\gamma_4=\frac{\beta_4}{\|\beta_4\|}=(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}).$$

例2 在$P_2[t]$定义内积

$$(f(t),g(t))=\int_{-1}^{1}f(t)g(t)\mathrm{d}t,\quad f(t),g(t)\in P_2[t],$$

求其一组标准正交基

解 取$P_2[t]$ 的一组基$1,t,t^2$，将其正交化有

$$g_0(t)=1,\qquad g_1(t)=t-\frac{(t,g_0)}{(g_0,g_0)}g_0=t,$$

$$g_2(t)=t^2-\frac{(t^2,g_0)}{(g_0,g_0)}g_0(t)-\frac{(t^2,g_1)}{(g_1,g_1)}g_1(t)=t^2-\frac{1}{3}.$$

单位化，得$P_2[t]$ 的一组正交基为

$$\varphi_0(t)=\frac{1}{\sqrt{2}},\quad \varphi_1(t)=\frac{\sqrt{6}}{2}t,\quad \varphi_2(t)=\frac{\sqrt{10}}{4}(3t^2-1).$$

为什么总是取标准正交基呢？

二、标准正交基的几个性质

定理 2.2.1 $\gamma_1,\gamma_2,\cdots,\gamma_n$ 是定义了内积的线性空间（即欧氏空间）V 的一组标准正交基，则对任意向量，有$\alpha = x_1\gamma_1 + x_2\gamma_2 + \cdots + x_n\gamma_n \in V$

$$x_i = (\alpha, \gamma_i), \quad i = 1, 2, \cdots, n$$

向量的坐标分量是该向量与相应基向量的内积！

定理 $\gamma_1,\gamma_2,\cdots,\gamma_n$ 是定义了内积的线性空间（即欧氏空间）V 的一组标准正交基，则对任意两个向量

$$\alpha = a_1\gamma_1 + a_2\gamma_2 + \cdots + a_n\gamma_n \in V,$$

$$\beta = b_1\gamma_1 + b_2\gamma_2 + \cdots + b_n\gamma_n \in V,$$

仍然有 $(\alpha, \beta) = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$

向量的内积仍然具有 R^n 中标准内积的简单形式！

三、正交补

定义 设 V_1, V_2 是数域 R 上欧氏空间 V 的两个子空间。向量 $\alpha \in V$ 。如果对任意 $\beta \in V_1$ ，都有 $(\alpha, \beta) = 0$ ，则称 α 与子空间 V_1 正交，记为 $\alpha \perp V_1$ 。如果对任意 $\gamma \in V_2$ ，都有 $\gamma \perp V_1$ ，则称子空间 V_1 与 V_2 正交，记为 $V_1 \perp V_2$

定义 V 中所有与子空间 V_1 正交的向量的集合也构成 V 的子空间，称为 V_1 的正交补，记为 $V_1^{\perp}$ ，即

$$V_1^{\perp} \equiv \{\alpha \in V \mid \alpha \perp V_1\}$$

欧氏空间的正交补是否存在呢？

定理 设 V_1 是数域 R 上欧氏空间 V 的子空间，则存在 V_1 的**唯一**正交补 $V_1^{\perp}$，使得 V 可以正交分解为

$$V = V_1 \oplus V_1^{\perp}.$$

注意：正交分解是特殊的直和分解。

推论 若 V 是有限维欧式空间，W 是 V 的子空间，则 $W = W^{\perp\perp}$.

§3、正交变换

鉴于正交的重要性，所以相应的正交变换显得尤为重要。**Householder变换**（即反射变换）和**Givens变换**（即旋转变换）是两种最重要的正交变换，它们的作用主要是在数值算法中构造正交基。

矩阵论 MATRIX THEORY

二维平面中的图形经过**旋转变换**或**反射变换**后只是位置改变了，形状和大小、长度、角度都保持不变。向量的长度和角度都可以由内积来计算。因此，变换前后的内积保持不变，即**两向量的像的内积与原像的内积相等**。

由于二维平面是特殊的欧氏空间，因此这个想法自然也可以推广到一般的欧氏空间。

定义2.3.1 设 T 是欧式空间V的线性变换，若 T 保持 V中的向量内积不变，即对任何的α、$\beta \in V$ 都有

$$(T(\alpha), T(\beta)) = (\alpha, \beta).$$

则称 T 为V 上的正交变换。

根据定义，正交变换也保持欧氏空间中向量的长度、距离及向量间的夹角等几何属性不变。

正交变换可以从几个不同的方面来加以描述。

定理2.3.1 设 T 是欧氏空间 V 上的一个线性变换，则下列命题是等价的：

（1） T 是正交变换；

（2） T 保持向量的范数不变，即 $\|T(\alpha)\|=\|\alpha\|$；

（3） 若 $e_1, e_2, \cdots, e_n$ 是 V 的一组标准正交基，则 $T(e_1), T(e_2), \cdots, T(e_n)$ 也是 V 的标准正交基；

（4） T 在 V 的任意一组标准正交基下的矩阵表示为正交矩阵。

证明: $(1)\Rightarrow(2)$. 根据正交变换的定义显然成立。

$(2)\Rightarrow(1)$. 若线性变换保持长度不变，即

$$\|T(\alpha)\|^2=(T(\alpha),T(\alpha))=\|\alpha\|^2=(\alpha,\alpha)$$

$$\|T(\beta)\|^2=(T(\beta),T(\beta))=\|\beta\|^2=(\beta,\beta)$$

同样有

$$(T(\alpha+\beta),T(\alpha+\beta))=(\alpha+\beta,\alpha+\beta)$$

展开上式并化简，即得

$$(T\alpha,T\beta)=(\alpha,\beta)$$

证明: $(1)\Rightarrow(3)$.显然成立。

$(3)\Rightarrow(1)$. 对任意 $\alpha,\beta\in V$ ，令

$$\alpha = x_1e_1+\cdots+x_ne_n,\ \beta = y_1e_1+\cdots+y_ne_n,$$

则 $$T(\alpha)=x_1T(e_1)+\cdots+x_nT(e_n),$$

$$T(\beta)=y_1T(e_1)+\cdots+y_nT(e_n),$$

因此

$$(T\alpha,T\beta)=x_1y_1+\cdots+x_ny_n=(\alpha,\beta)$$

证明: **(3) ⇒ (4).** 设 T 在 $e_1,\cdots,e_n$下的矩阵为A，即

$$(T(e_1),\cdots,T(e_n))=(e_1,\cdots,e_n)A,$$

由于 $T(e_1),\cdots,T(e_n)$ 也是标准正交基，所以 A 是两组标准正交基间的过渡矩阵，因此 A 是正交矩阵。

证明： **(4) $\Rightarrow$ (3).** 设 A 是正交矩阵，则

$$(T(e_i),T(e_j))=(a_{1i}e_1+\cdots+a_{ni}e_n,a_{1j}e_1+\cdots+a_{nj}e_n)$$

$$=a_{1i}a_{1j}+\cdots+a_{ni}a_{nj}$$

$$=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}$$

所以 $T(e_1),\cdots,T(e_n)$ 也是标准正交基。

正交变换T在任意一组标准正交基下的矩阵是正交矩阵，但在其他基下的矩阵可能不是正交矩阵

• 正交矩阵的行列式为1 或者 -1.

称行列式为1 的正交变换为第一类正交变换，

称行列式为-1 的正交变换为第二类正交变换.

例1 （**Givens 变换**)将线性空间 R^2 中的所有向量均绕原点顺时针旋转角 θ 的Givens变换

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \to \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \equiv G \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$$

就是一个正交变换。因为此变换的矩阵表示 G 是正交矩阵。

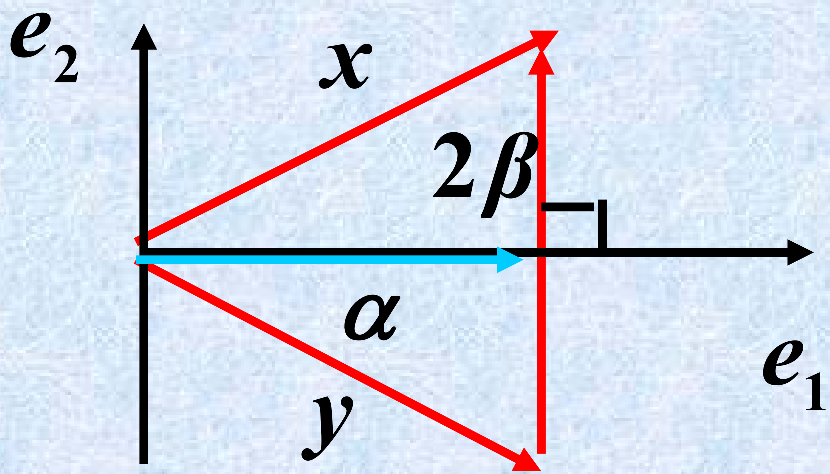
一般的，形如

$$G=\begin{pmatrix} 1 & & & & & & \\ & \cdots & & & & & \\ & & \cos\theta & & \sin\theta & & \\ & & & \cdots & & & \\ & & -\sin\theta & & \cos\theta & & \\ & & & & & \cdots & \\ & & & & & & 1 \end{pmatrix}$$

的矩阵称为Givens矩阵，对应的变换为Givens变换.

例2 HouseHolder变换

如图,



$$x=(x,e_1)e_1+(x,e_2)e_2$$
$$\equiv\alpha+\beta,\quad \alpha\perp\beta$$

因此向量 x 关于“与 e_2 轴正交的直线”对称的镜像向量的表达式为

$$y=x-2\beta=x-2(x,e_2)e_2=x-2e_2(x,e_2)$$
$$=x-2e_2e_2^Tx=(E-2e_2e_2^T)x$$

矩阵论 MATRIX THEORY

类似地，可定义将向量 $x \in R^n$ 变换为关于“与单位向量 $u \in R^n$ 正交的 $n-1$ 维子空间”对称的向量 $y \in R^n$ 的镜像变换。

设 $u \in R^n$ 为单位向量，对任意 $x \in R^n$，定义

$$x \to Hx = (E - 2uu^T)x$$

称 H 为Householder 变换（初等反射变换），则 H 是 R^n 的正交变换。

证：

$$Hx=(E-2uu^T)x=x-2u(u,x)$$

所以，

$$\begin{aligned}&(Hx,Hx)\\&=\big(x-2u(u,x),x-2u(u,x)\big)\\&=(x,x)\end{aligned}$$

从而，H 是 R^n 的一个正交变换。

§4、酉空间(复内积空间)简介

当线性空间中向量的坐标分量的取值由实数域推广为复数域时，欧氏空间中关于向量的内积、标准正交基、向量元素之间的正交变换等概念和结论都可以“平滑地”推广到所谓酉空间。所以欧氏空间与酉空间中的相关结论几乎是“平行的”。

一、酉空间

定义2.4.1 V 是复数域 C 上的线性空间。如果对 V 中任意两个向量 α、$\beta \in V$ 都存在所谓 α 与 β 的内积 $(\alpha,\beta)\in C$，满足下面四个条件。称定义了内积的线性空间 V 为复内积空间，简称酉空间。

$\forall \alpha$、β、$\gamma \in V$

(1) $(\alpha,\beta)=(\beta,\alpha)^*$；

(2) $(\lambda\alpha,\beta)=\lambda(\alpha,\beta)$； $(\forall \lambda \in C)$

(3) $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$；

(4) $(\alpha,\alpha)\geq 0$ ，当且仅当 $\alpha=O$ 时，等号成立。

例1 定义了标准内积的 C^n是一酉空间。这里，对任意两个向量 $\alpha=(a_1,a_2,\cdots,a_n)^T$ 及 $\beta=(b_1,b_2,\cdots,b_n)^T$ 标准内积为

$$(\alpha,\beta)\equiv\beta^H\alpha$$

$$=a_1b_1^*+a_2b_2^*+\cdots+a_nb_n^*,\quad \beta^H\equiv(\beta^*)^T$$

例 2 在矩阵空间 $C^{m\times n}$ 中，对任意 A、$B\in C^{m\times n}$

定义

$$(A,B)\equiv tr(B^H A)=\sum_{j=1}^{n}\sum_{i=1}^{m} b_{ij}^{*}a_{ij}$$

则 $C^{m\times n}$ 是定义了内积 (A,B) 的内积空间。

二、酉空间的一些重要结论

（1） $(\alpha,\lambda\beta)=\lambda^{*}(\alpha,\beta)$;

（2） $(\alpha,\beta+\gamma)=(\alpha,\beta)+(\alpha,\gamma)$;

（3） $(\alpha,O)=(O,\beta)=0$;

（4） $\|\alpha\|=(\alpha,\alpha)$;

（5） $|(\alpha,\beta)|\le\|\alpha\|\|\beta\|$ ，当且仅当 α、β 线性相关时等号成立;

（6） 两个非零向量 α、β 的内积 $(\alpha,\beta)=0$ 时，称 α 与 β 正交；

（7） 任意一组线性无关的向量都可以用Schmidt正交化方法正交化，并扩充成一组标准正交基；

（8）标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的任意两向量

$$\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n$$

$$\beta=b_1\varepsilon_1+b_2\varepsilon_2+\cdots+b_n\varepsilon_n$$

的内积 $(\alpha,\beta)=a_1b_1^*+a_2b_2^*+\cdots+a_nb_n^*$.

（9）任意一个酉空间 V 都可以分解为其子空间 V_1 和 $V_1^{\perp}$ 的直和；（正交分解）

定义 称酉空间 V 中的线性变换 T 为酉变换，如果 T 保持向量的内积不变，即对任意 α、$\beta \in V$，有

$$(T(\alpha),T(\beta)) = (\alpha,\beta).$$

根据定义，显然酉变换也保持酉空间中向量的长度、距离等几何属性不变。不过注意对向量间的夹角的不同定义，未必成立。（酉空间的内积是复数）

定理 设 T 是酉空间 V 上的一个线性变换，则下列命题是等价的：

（1） T 是酉变换；

（2） T 保持向量的范数不变，即 $\|T(\alpha)\|=\|\alpha\|$；

（3） 若 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组标准正交基，则 $T(\varepsilon_1),T(\varepsilon_2),\cdots,T(\varepsilon_n)$ 也是 V 的标准正交基；

（4） T 在 V 的任意一组标准正交基下的矩阵表示 A 为酉矩阵，即

$$A^H A = AA^H = I$$

§5、几种变换与几种矩阵

一、正规变换与正规矩阵

定义 酉空间 V 上的线性变换 T 称为 V 上的一个**正规变换**，如果**存在** V 的标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 及对角矩阵 $D\equiv diag(\lambda_1,\lambda_2,\cdots,\lambda_n)$ 满足

$$(T(\varepsilon_1),T(\varepsilon_2),\cdots,T(\varepsilon_n))=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)D$$

并称 T 在任一标准正交基下的矩阵表示为**正规矩阵**。

定理 方阵 A 是正规的，当且仅当

$$AA^H = A^H A.$$

特别地，

定义 若方阵 A 满足

$$AA^H = A^H A = I.$$

则称 A 为酉矩阵.

例1 判断下列矩阵是不是正规矩阵：

（1）实对称矩阵（ $A^T = A$ ）；

（2）实反对称矩阵（ $A^T = -A$ ）；

（3）正交矩阵 （$A^T = A^{-1}$ ）；

（4）酉矩阵（ $A^H = A^{-1}$ ）；

（5）Hermite 矩阵（ $A^H = A$ ）；

（6）反Hermite 矩阵（ $A^H = -A$ ）；

（7）形如 $a\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, a \in R \text{ or } C$ 的矩阵。

二、Hermite变换与Hermite 矩阵

定义 设 T 是酉空间 上的线性变换，称 T 为 V 上的 Hermite 变换，如果对任意 α、$\beta \in V$ ，都有

$$(T(\alpha),\beta)=(\alpha,T(\beta)).$$

并称 T 在 V 的任意一组标准正交基下的矩阵为 Hermite 矩阵。

酉空间 V 上的线性变换 T 是 Hermite 变换的充要条件是 T 在 V 的任意一组标准正交基下的矩阵 A 满足

$$A^H = A$$

三、反Hermite变换与反Hermite 矩阵

定义 设 T 是酉空间 V 上的线性变换，称 T 为 V 上的反Hermite 变换，如果对任意 α、$\beta \in V$ ，都有

$$(T(\alpha),\beta) = -(\alpha, T(\beta)).$$

并称 T 在 V 的任意一组标准正交基下的矩阵表示为反Hermite 矩阵。

酉空间上的线性变换 T 是 反Hermite 变换的充要条件是 T 在 V 的任意一组标准正交基下的矩阵 A 满足

$$A^H = -A$$