高等代数（上海市精品课程，制作人：朱胜林，mazhusl@fudan.edu.cn ）

朱胜林

第九章
内积空间

我的信息

朱胜林

办公室:	光华东楼 1708
办公电话:	55664896
办公时间:	周二下午 1：30 — 3：30
电子邮件:	mazhusl@fudan.edu.cn
个人主页:	homepage.fudan.edu.cn/~zhusl
本课程主页:	jpkc.fudan.edu.cn/s/100/

内积空间的概念

定义（实内积空间）

设 V 是实数域 $\mathbb{R}$ 上一个线性空间 $(-,-)$ 是 V 上的一个二元实值函数，如果 $(-,-)$ 满足以下规则：

1. 对称性：$(\alpha,\beta)=(\beta,\alpha)$，$\forall \alpha,\beta\in V$；
2. 左侧线性：$(k\alpha+l\beta,\gamma)=k(\alpha,\gamma)+l(\beta,\gamma)$，$\forall \alpha,\beta,\gamma\in V$，$\forall k,l\in\mathbb{R}$；
3. 正定性：$(\alpha,\alpha)\geq 0$，$\forall \alpha\in V$。$(\alpha,\alpha)=0$ 当且仅当 $\alpha=0$。

则称 $(-,-)$ 是 V 上的一个内积，$(V,(-,-))$ 称为一个内积空间。
当 $\dim V<\infty$ 时，内积空间 V 又称为Euclid 空间（欧氏空间）。

例

在 $\mathbb{R}_n$ 上中，对 $\alpha=(a_1,a_2,\cdots,a_n)',\beta=(b_1,b_2,\cdots,b_n)'$, 定义

$$(\alpha,\beta)=a_1b_1+a_2b_2+\cdots+a_nb_n=\beta'\alpha,$$

则 $(\mathbb{R}_n,(-,-))$ 是一个 n 维 Euclid 空间，称为标准内积空间。
如果定义 $\underline{(\alpha,\beta)}=a_1b_1+2a_2b_2+\cdots+na_nb_n$，也会得到一个内积空间。

例

取区间 $[a,b]$ 上的连续函数全体构成的线性空间 $C[a,b]$，定义

$$(f(x), g(x)) = \int_a^b f(x) g(x) \, dx,$$

则 $(C[a,b], (-,-))$ 构成一个无限维内积空间。

定义

设 V 是实数域 $\mathbb{C}$ 上一个线性空间 $(-,-)$ 是 V 上的一个二元复值函数，如果 $(-,-)$ 满足以下规则：

1. 共轭对称性：$(\alpha,\beta) = \overline{(\beta,\alpha)}$，$\forall \alpha,\beta \in V$；
2. 左侧线性：$(k\alpha + l\beta, \gamma) = k(\alpha,\gamma) + l(\beta,\gamma)$，$\forall \alpha,\beta,\gamma \in V$，$\forall k,l \in \mathbb{C}$；
3. 正定性：$(\alpha,\alpha) \geq 0$，$\forall \alpha \in V$。$(\alpha,\alpha) = 0$ 当且仅当 $\alpha = 0$。

则称 $(-,-)$ 是 V 上的一个内积，$(V, (-,-))$ 称为一个内积空间。当 $\dim V < \infty$ 时，内积空间 V 又称为U-空间（酉空间）。

例

在 $\mathbb{C}^n$ 上中，对 $\alpha=(a_1,a_2,\cdots,a_n),\beta=(b_1,b_2,\cdots,b_n)$, 定义

$$(\alpha,\beta)=a_1\overline{b_1}+a_2\overline{b_2}+\cdots+a_n\overline{b_n}=\beta^H\alpha,$$

则 $(\mathbb{C}^n,(-,-))$ 是一个 n 维 U-空间。

由定义，实内积空间上的内积右侧也是线性的，
即 $(\alpha,-):\beta\mapsto(\alpha,\beta)$ 为 $V\to\mathbb{R}$ 的线性变换（函数）；复内积空间上的内积右侧是共轭线性的，
即 $(\alpha,k\beta+l\gamma)=\overline{k}\,(\alpha,\beta)+\overline{l}\,(\alpha,\gamma)$，$\forall\alpha,\beta,\gamma\in V$，$\forall k,l\in\mathbb{C}$。

定义

设 $(V,(-,-))$ 为一个内积空间，且 $\alpha\in V$，非负实数 $\sqrt{(\alpha,\alpha)}$ 称为 α 的长度（范数），记为

$$\|\alpha\|=\sqrt{(\alpha,\alpha)}。$$

若 $\beta\in V$，则定义 α、β 之间的距离

$$d(\alpha,\beta)=\|\beta-\alpha\|;$$

当 V 为实内积空间时，定义 α、β 之间的夹角 $\widehat{(\alpha,\beta)}$ 由下式得出

$$\cos\widehat{(\alpha,\beta)}=\frac{(\alpha,\beta)}{\|\alpha\|\cdot\|\beta\|};$$

当 $(\alpha,\beta)=0$ 时，称 α、β 正交（垂直）。

这里有个约定，当 α 或 β 中有一个为零时，$\widehat{(\alpha,\beta)}$ 可以是任意值。

向量的长度有以下性质：

定理

设 $(V,(-,-))$ 为内积空间，$\alpha,\beta\in V$，k 是定义数域中的数，则

1. 齐次性：$\|k\alpha\|=|k|\ \|\alpha\|$；
2. Cauchy-Schwartz 不等式：$\|(\alpha,\beta)\|\le\|\alpha\|\ \|\beta\|$；
3. 三角不等式：$\|\alpha+\beta\|\le\|\alpha\|+\|\beta\|$。

当 α、β 正交时，$\|\alpha+\beta\|^2=(\alpha,\alpha)+(\beta,\beta)+(\alpha,\beta)+(\beta,\alpha)=\|\alpha\|^2+\|\beta\|^2$，这个结论称为勾股定理。

例

Cauchy-Schwartz 不等式 $\|(\alpha,\beta)\| \le \|\alpha\|\ \|\beta\|$ 在不同空间中的体现：

1. 对于 $\mathbb{R}^n$，标准内积下

$$|a_1b_1+a_2b_2+\cdots+a_nb_n| \le \sqrt{a_1^2+a_2^2+\cdots+a_n^2}\sqrt{b_1^2+b_2^2+\cdots+b_n^2};$$

2. $C[a,b]$ 下

$$\left|\int_a^b f(x)g(x)dx\right| \le \left(\int_a^b f^2(x)dx\right)^{\frac{1}{2}}\left(\int_a^b g^2(x)dx\right)^{\frac{1}{2}}。$$

设 V 是一个 n 维内积空间，在 V 中取一组基 $v_1, v_2, \ldots, v_n$，对于 V 中任意两个向量

$$\alpha = a_1v_1 + a_2v_2 + \cdots + a_nv_n, \beta = b_1v_1 + b_2v_2 + \cdots + b_nv_n,$$

由内积的性质得

$$\begin{aligned}
(\alpha, \beta) &= \langle a_1v_1 + a_2v_2 + \cdots + a_nv_n, b_1v_1 + b_2v_2 + \cdots + b_nv_n\rangle \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}(v_i, v_j)a_i\overline{b_j} \\
&= (a_1, a_2, \ldots, a_n)\begin{pmatrix} (v_1,v_1) & (v_1,v_2) & \cdots & (v_1,v_n) \\ (v_2,v_1) & (v_2,v_2) & \cdots & (v_2,v_n) \\ \vdots & \vdots & \ddots & \vdots \\ (v_n,v_1) & (v_n,v_2) & \cdots & (v_n,v_n) \end{pmatrix}\overline{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}}
\end{aligned}$$

称 $G(v_1, v_2, \ldots, v_n) = \begin{pmatrix} (v_1,v_1) & (v_1,v_2) & \cdots & (v_1,v_n) \\ (v_2,v_1) & (v_2,v_2) & \cdots & (v_2,v_n) \\ \vdots & \vdots & \ddots & \vdots \\ (v_n,v_1) & (v_n,v_2) & \cdots & (v_n,v_n) \end{pmatrix}$ 为内积空间 $(V, (-,-))$ 的基向量的Gram 矩阵。

当 V 是实内积空间时，对任意的非零 $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$，如果取 $\alpha = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n$，则由内积的正定性、对称性知道，

$$(x_1, x_2, \ldots, x_n)\, G\,(v_1, v_2, \ldots, v_n)\,(x_1, x_2, \ldots, x_n)' > 0$$

且 $G\,(v_1, v_2, \ldots, v_n)$ 为一个实对称阵。这样的矩阵称为正定矩阵。这是二次型一章中重点讨论的矩阵。
如果 $u_1, u_2, \cdots, u_n$ 是空间 V 的另一组基，且
从 $v_1, v_2, \cdots, v_n$ 到 $u_1, u_2, \cdots, u_n$ 的过渡矩阵为 P，形式上

$$(u_1, u_2, \cdots, u_n) = (v_1, v_2, \cdots, v_n)P$$

则由内积的左侧线性（右侧共轭线性），可知：

$$G\,(u_1, u_2, \ldots, u_n) = P' G\,(v_1, v_2, \ldots, v_n)\,\overline{P},$$

这样的复矩阵的关系，在后面的章节中称为合同关系。对实方阵而言，两个同阶对称阵 A、B，如果存在一个可逆同阶方阵 P，使得 $P'AP = B$，则称 A 与 B合同。关于合同关系，我们将以后面的章节中详细讨论。

内积空间 V 中 Gram 矩阵是单位阵的基称为标准正交基。

定义

设 $(V,(-,-))$ 是一个内积空间，若 $v_1,v_2,\ldots,v_n$ 是 V 中的相互正交的非零向量组，则称 $v_1,v_2,\ldots,v_n$ 是一个正交向量组；正交向量组若构成 V 的基，则称为正交基；一组正交基中若每个向量都是单位长度，则称为一组标准正交基。

容易看到，一组基是校准正交基的充分必要条件是它的 Gram 矩阵是单位矩阵。

例

若 $v_1,v_2,\ldots,v_n$ 是 V 的一组标准正交基，$x=x_1v_1+\cdots+x_nv_n$，$y=y_1v_1+\cdots+y_nv_n$，则

$$(x,y)=\overline{y_1}x_1+\cdots+\overline{y_n}x_n=([x]_{v's},[y]_{v's})\text{。}$$

正交组在内积空间中有很优点，在计算一个向量的表示系数时，不用再去解一个线性方程组。这在下面的命题中可以看到。

命题

设 V 是一个内积空间，$v_1, v_2, \ldots, v_n$ 是 V 中的正交向量组，$v \in V$。则

1. $v_1, v_2, \ldots, v_n$ 必线性无关，$n \le \dim V$；
2. 若 v 可由 $v_1, v_2, \ldots, v_n$ 线性表出，则必有 $v = \Pr_{\mathcal{L}(v_1,v_2,\ldots,v_n)} v$，其中
$$\Pr_{\mathcal{L}(v_1,v_2,\ldots,v_n)} v = \frac{(v, v_1)}{(v_1, v_1)} v_1 + \frac{(v, v_2)}{(v_2, v_2)} v_2 + \cdots + \frac{(v, v_n)}{(v_n, v_n)} v_n$$
称 $\Pr_{\mathcal{L}(v_1,v_2,\ldots,v_n)} v$ 为 v 在 $v_1, v_2, \ldots, v_n$ 生成的子空间上的投影；
3. $$\left(v - \Pr_{\mathcal{L}(v_1,v_2,\ldots,v_n)} v\right) \perp \mathcal{L}(v_1, v_2, \ldots, v_n),$$
4. $\|v\|^2 = \left\|v - \Pr_{\mathcal{L}(v_1,v_2,\ldots,v_n)} v\right\|^2 + \left\|\Pr_{\mathcal{L}(v_1,v_2,\ldots,v_n)} v\right\|^2 \ge \left\|\Pr_{\mathcal{L}(v_1,v_2,\ldots,v_n)} v\right\|^2$。
最后一个结论称为 Bessel 不等式，该不等式在Fourier 级数展开中有具体应用，也是最小二乘法的理论依据。

正交组是不是存在，下面给出一个方法，它将一个非正交的向量组化成一个正交的向量组。

命题 (Gram-Schmidt 正交化方法)

设 V 是一个内积空间，$v_1, v_2, \ldots, v_n$ 是 V 中的向量组。定义

$$
\begin{aligned}
u_1 &= v_1, \\
u_2 &= v_2 - \frac{(v_2, u_1)}{(u_1, u_1)} u_1 = v_2 - \mathrm{Pr}_{\mathcal{L}(u_1)} v_2, \\
u_3 &= v_3 - \frac{(v_3, u_1)}{(u_1, u_1)} u_1 - \frac{(v_3, u_2)}{(u_2, u_2)} u_2 = v_3 - \mathrm{Pr}_{\mathcal{L}(u_1, u_2)} v_3, \\
&\cdots \\
u_n &= v_n - \sum_{i=1}^{n-1} \frac{(v_n, u_i)}{(u_i, u_i)} u_i = v_n - \mathrm{Pr}_{\mathcal{L}(u_1, u_2, \ldots, u_n)} v_n,
\end{aligned}
$$

则 $u_1, u_2, \ldots, u_n$ 是一个正交向量组且
对 $\forall 1 \le k \le n$，$\mathcal{L}(v_1, \ldots, v_k) = \mathcal{L}(u_1, \ldots, u_k)$。

推论

1. 任一内积空间上必存在标准正交基。
2. 若 A 一个 n 阶可逆实方阵，则存在矩阵 Q 和 R，使得 $A = QR$，且 Q 满足 $Q'Q = I_n$，R 为一个对角元均为正数的上三角矩阵。这个分解式是唯一的，称为矩阵 A 的QR-分解。

如果 U 是内积空间 V 的一个子空间，则 $(-,-)$ 限制在 U 上也是一个内积。记 $U^{\perp} = \{x \in V \mid (x,u) = 0, \forall u \in U\}$，则由内积的左侧线性知道，$U^{\perp}$ 也是 V 的一个子空间。

定理 (子空间垂直分解和正交基扩充定理)

设 U 是内积空间 V 的一个子空间。则

1. $V = U \bigoplus U^{\perp}$；
2. U 的一组标准正交基必可扩充为 V 的一组标准正交基。

当 $V_1,, \ldots, V_k$ 均是 V 的子空间且 $(V_i, V_j) = 0$ $(\forall 1 \le i,j \le k)$ 时，称 $V_1,, \ldots, V_k$ 是正交子空间。根据正交向量组必是线性无关的这一结论，有正交子空间的和必是直和。

内积空间上线性变换的伴随变换

引理

设 $(V,(-,-))$ 是一个内积空间，$\varphi\in\mathcal{L}(V)$，$\beta\in V$，由存在唯一一个 $\beta^*\in V$，使得

$$(\varphi(\alpha),\beta)=(\alpha,\beta^*),\ \forall\alpha\in V。$$

证明：取 V 的一组标准正交基 $v_1,\dots,v_n$。如果 β^* 满足要求，则 $(\varphi(v_i),\beta)=(v_i,\beta^*)$，故有 $(\beta^*,v_i)=(\beta,\varphi(v_i))$，于是由前面正交向量组中向量的表示式，得

$$\beta^*=\sum_{i=1}^n(\beta^*,v_i)v_i=\sum_{i=1}^n(\beta,\varphi(v_i))v_i;$$

反过来，如果 $\beta^*=\sum_{i=1}^n(\beta,\varphi(v_i))v_i$，由 $v_1,\dots,v_n$ 是标准正交基，易知 β^* 满足要求。 □

定义 $\varphi^*:\beta\mapsto\beta^*=\sum_{i=1}^n(\beta,\varphi(v_i))$，由 $(-,-)$ 的左侧线性知，$\varphi^*\in\mathcal{L}(V)$，称为 φ 的伴随变换。

注

注意，虽然 φ^* 称为 φ 的伴随，但不管在什么基下，所对应的矩阵与伴随矩阵一点关系也没有。值得指出的是，在标准正交基下，变换与其伴随的矩阵恰为共轭转置关系。

定理

设 $(V,(-,-))$ 是内积空间，$v_1,\ldots,v_n$ 是 V 的一组标准正交基，$\varphi\in\mathcal{L}(V)$。如果 φ 在 $v_1,\ldots,v_n$ 下的矩阵为 A，则 φ^* 在 $v_1,\ldots,v_n$ 下的矩阵为 A^H。如果 V 是实内积空间，则 φ 和 φ^* 在 $v_1,\ldots,v_n$ 下的矩阵互为转置关系。

因为矩阵互为（共轭）转置关系，所以有如下类似于矩阵转置的关系式成立：

定理

设 $(V,(-,-))$ 是内积空间，$\varphi,\psi\in\mathcal{L}(V)$，$c_1,c_2$ 是数，则

1. $(c_1\varphi+c_2\psi)^*=\overline{c_1}\varphi^*+\overline{c_2}\psi^*$；
2. $(\varphi\psi)^*=\varphi^*\psi^*$；
3. $(\varphi^*)^*=\varphi$。
4. 若 U 是 φ 的不变子空间，则 $U^\perp$ 是 φ^* 的不变子空间；
5. 若 φ 的特征值为 $\lambda_1,\ldots,\lambda_n$，则 φ^* 有特征值 $\overline{\lambda_1},\ldots,\overline{\lambda_n}$。

内积空间的同构，正交变换和酉变换

当两个实（或复）内积空间线性同构，且同构能够保持内积时，我们说这个同构是两个内积空间之间的同构。

定义

设 $(U,(-,-))$ 和 $(V,(-,-))$ 同是实（或复）内积空间，若存在 $\varphi \in \mathcal{L}(V \to U)$，使得

$$(\varphi(x), \varphi(y)) = (x, y), \ \forall x, y \in V,$$

则称 φ 为保持内积的线性映射，如果 φ 还是线性同构，则称 φ 是内积空间 V 到 U 的同构。

命题

$\varphi \in \mathcal{L}(V \to U)$ 保持内积当且仅当 φ 保持范数，即 $\|\varphi(x)\| = \|x\|$，$\forall x \in V$。

内积空间的同构，正交变换和酉变换

当两个内积空间的维数相同时，同构还有其它等价说法：

定理

设 $(U,(-,-))$ 和 $(V,(-,-))$ 都是相同维数的实（或复）内积空间，$\varphi \in \mathcal{L}(V \to U)$。则下列说法等价：

1. φ 保持内积；
2. φ 是同构；
3. φ 把 V 的任一组标准正交基映成 U 的标准正交基；
4. φ 把 V 的某一组标准正交基映成 U 的标准正交基。

于是，两个内积空间同构的充分必要条件是其维数相同。

内积空间的同构，正交变换和酉变换

对于同一个内积空间，其上的保积同构称为正交变换（当内积空间为欧氏空间时）或酉变换（当内积空间为酉空间时）。

定理

设 φ 是内积空间 $(V,(-,-))$ 上的一个线性变换，则 φ 为保积同构当且仅当 $\varphi^* = \varphi^{-1}$。

前面说过，$\mathcal{L}(V)$ 中的 φ 和它的伴随 φ^* 在一组校准正交基下对应的矩阵互为共轭转置，所以我们有如下的定义：

定义

如果 A 是一个 n 阶实方阵且 $A' = A^{-1}$，则称 A 为一个正交矩阵。如果果 A 是一个 n 阶复方阵且 $A^H = A^{-1}$，则称 A 为一个酉矩阵。

引理

设 V 是一个内积空间，$v_1,\ldots,v_n$ 为其上一组标准正交基，$u_1,\ldots,u_n$ 为其上的一组基。则 $u_1,\ldots,u_n$ 也为一组标准正交基的充分必要条件是从 $v_1,\ldots,v_n$ 到 $u_1,\ldots,u_n$ 的过渡矩阵是一个正交矩阵（对实空间而言）或酉矩阵（对复空间而言）。

内积空间的同构，正交变换和酉变换

定理

设 φ 是内积空间 $(V,(-,-))$ 上的一个线性变换，则下列等价：

1. φ 为保积同构；
2. $\varphi^* = \varphi^{-1}$；
3. φ 在任一组标准正交基下的矩阵是正交矩阵（对实内积空间而言）或酉矩阵（对复空间而言）；
4. φ 在某一组标准正交基下的矩阵是正交矩阵（对实内积空间而言）或酉矩阵（对复空间而言）；

正交（酉）矩阵也相一些等价的说法：

定理

一个 n 阶实（复）方阵为正交（酉）矩阵当且仅当它的列向量是 $\mathbb{R}_n$ （ $\mathbb{C}_n$ ）中的一组标准正交基；当且仅当它的行向量是 $\mathbb{R}^n$ （ $\mathbb{C}^n$ ）中的一组标准正交基。

由于一个线性变换的特征值与其在任一组基下的特征值相同，所以有下面的结论：

定理（保积变换的特征值模长为 1）

1. 设 A 为一个 n 阶正交矩阵（或酉矩阵），则 A 的任一特征值的模必为 1，且 $\det A = \pm 1$ （$|\det A| = 1$）；
2. 设 φ 为内积空间 V 上的一个保积变换，则 φ 的特征值的模长为 1。

例

旋转变换 $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ 矩阵是正交矩阵；一个上（或下）三角矩阵是酉矩阵当且仅当它为一个对角阵且对角元的模长都是 1。

例

设 $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$，$B = (kE + A)^2$，$k \in \mathbb{R}$。求对角阵 Λ，使 B 与 Λ 相似，并求 k 为何值时，B 为正交阵。

解 实对称阵 A 的特征多项式为 $\lambda(\lambda-2)^2$，所以 A 相似

例

设 $A=(a_{ij})_{3\times 3}$ 为正交阵，且 $a_{11}=1$，又 $b=(1,0,0)^T$，求 $Ax=b$ 的解。

解： A 的列向量和行向量都构成标准正交基，由于 $a_{11}=0$，所以第一列和第一行中除 $(1,1)$-项外均为零；于是有 $Ab=b$，由解的唯一性得 $x=b$。 $\square$

定理 (Schur 定理)

对复内积空间上的任一线性变换 φ，必存在一组标准正交基，在此基下 φ 的矩阵为上三角矩阵。矩阵形式：对复数域上的任一方阵 A，必存在一个酉矩阵 U，使得 $U^{-1}AU$ 为上三角矩阵。

本定理在实数域上，如果加上 φ 的特征值均为实数，也成立。

自伴随变换（（共轭）对称矩阵）

定义 (自伴随算子)

如果 V 为内积空间且 $\varphi \in \mathcal{L}(V)$ 满足 $\varphi^* = \varphi$，则称 φ 为一个自伴随变换（自伴随算子）。

前面已经证明，一个算子和其伴随算子在一组标准正交基下的矩阵互为（共轭）转置关系，所以自伴随算子在标准正交基下对应的矩阵恰为实对称矩阵（对实空间而言）或Hermite 矩阵（对复空间而言）。
自伴随算子（相应地，Hermite 矩阵）是可以对角化的一类重要例子，且在解析几何中有着重要的应用。就其本身而言，它也有非常突出的性质。

定理

设 $\varphi \in \mathcal{L}(V, (-,-))$ 是一个自共轭算子，则

1. φ 的特征值均为实数；
2. φ 的不同特征值的特征向量均相互正交；
3. φ 有 $\dim V$ 个相互正交的特征向量；
4. φ 在 V 的一组标准正交基下的矩阵为实对角阵。

自伴随变换（（共轭）对称矩阵）

实对称矩阵可以诱导出标准欧氏空间 $(\mathbb{R}_n, (-,-))$ 上的自伴随算子，Hermite 矩阵可以诱导出标准内积空间 $(\mathbb{C}_n, (-,-))$ 上自伴随算子，所以可以把上面的结论应用到这样的矩阵上。

定理

设 A 是一个 n 阶实对称矩阵（ Hermite 矩阵）。则

1. A 的特征值均为实数；
2. A 的不向特征值特征向量均相互正交；
3. A 有 n 个相互正交的特征向量；
4. 存在正交矩阵 Q （酉矩阵 U ），使得 $Q'AQ$ （ U^HAU ）为实对角阵。

注意：在上面的定理中，条件(4) 也是 A 是一个实对称矩阵（ Hermite 矩阵）的等价条件。

例

1. 设 A 是 3 阶实对称矩阵，A 的特征值为 $0,3,3$，已知属于特征值 0 的特征向量为 $v_1=(1,1,1)'$，求矩阵 A 并将 A 正交相似对角化。
2. 设 A 为三阶实对称矩阵，$r(A)=2$，$A\begin{pmatrix}1&1\\0&0\\-1&1\end{pmatrix}=\begin{pmatrix}-1&1\\0&0\\1&1\end{pmatrix}$，
   1. 求 A 的特征值和特征向量；
   2. 求 A。

解： [第 2 题] 易知，A 的两个特征向量为 $\begin{pmatrix}1\\0\\-1\end{pmatrix}\longleftrightarrow -1$，$\begin{pmatrix}1\\0\\1\end{pmatrix}\longleftrightarrow 1$。由于 $r(A)=2$，所以 0 也是一个特征值，与前面两个向量都正交，所以可取 $\begin{pmatrix}0\\1\\0\end{pmatrix}\longleftrightarrow 0$；所以 $A=\begin{pmatrix}1&1&0\\0&0&1\\-1&1&0\end{pmatrix}\begin{pmatrix}-1&0&0\\0&1&0\\0&0&0\end{pmatrix}\begin{pmatrix}1&1&0\\0&0&1\\-1&1&0\end{pmatrix}^{-1}=\begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix}$。 □

复正规算子

前面讨论了自伴随算子，它是在一组标准正交基下矩阵为实对角阵的算子。如果放松要求，把实对称阵改为复对角阵，结果会如何？
显然，如果 $\varphi \in \mathcal{L}(V)$ 在 V 的一组标准正交基 $v_1, \ldots, v_n$ 下的矩阵为对角阵 $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$，则 φ^* 在这组基下的矩阵为

$$\mathrm{diag}(\lambda_1, \ldots, \lambda_n)^H = \mathrm{diag}\left(\overline{\lambda_1}, \ldots, \overline{\lambda_n}\right),$$

于是有 $\varphi(v_i) = \lambda_i v_i$ 及 $\varphi^*(v_i) = \overline{\lambda_i} v_i$，所以 $\varphi^* \varphi(v_i) = \varphi^*(\lambda_i v_i) = \overline{\lambda_i} \lambda_i v_i = |\lambda_i|^2 v_i = \varphi \varphi^*(v_i)$，即得到 $\varphi^* \varphi = \varphi \varphi^*$。
反过来，这样的算子（ $\varphi^* \varphi = \varphi \varphi^*$ ）也正是在一组基下矩阵是对角矩阵的算子。

复正规算子

定义

设 V 为一有限维内积空间，若 $\varphi \in \mathcal{L}(V)$ 满足 $\varphi^*\varphi = \varphi\varphi^*$，则称$\varphi$ 为一个正规变换（算子）。一个方阵 A 若满足 $A^HA = AA^H$，则称 A 为一个正规矩阵。

定理

设 V 为酉空间，$\varphi \in \mathcal{L}(V)$ 是正规算子，则

1. φ 的属于特征值 λ 的特征向量必是 φ^* 的属于特征值 $\overline{\lambda}$ 的特征向量；
2. φ 的属于不同特征值的特征向量相互正交；
3. 存在 V 的一组标准正交基，使得 φ 在这组基下的矩阵为对角阵。

所以，酉空间 V 上的一个算子为正规矩阵的充分必要条件是它在 V 的某组标准正交基下的矩阵为对角阵。

推论

在 $\mathbb{C}$ 上，Hermite 矩阵、酉矩阵、及实反对称阵酉相似于对角阵。

实正规矩阵

这一节要讨论的是实数域上的正规矩阵。

引理

设 A 是一个 n 阶实正规矩阵，且 $\lambda = a + bi$ （其中 $a, 0 \neq b \in \mathbb{R}$ ）是 A 的特征值，$x = u + iv$ （ $u, v \in \mathbb{R}_n$ ）为属于 λ 的一个特征向量，则必有 $(u, v) = 0$ 且 $\|u\| = \|v\| \neq 0$。

证明： $Ax = \lambda x$，两边同取共轭，利用 A 是实矩阵，得到 $A\overline{x} = \overline{\lambda}\overline{x}$，所以 $\overline{x}$ 是 A 的关于特征值 $\overline{\lambda} \neq \lambda$ 的特征向量，于是相互正交；而

$$(x, \overline{x}) = (u + iv, u - iv) = (u, u) - (v, v) + i(u, v) + i(v, u) = 0$$

所以，$(u, v) = 0$ 且 $(u, u) = (v, v)$。 □

定理

设 A 是一个 n 阶实正规矩阵，由存在一个正交矩阵 Q，使得

$$Q'AQ = \operatorname{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \ldots, \begin{pmatrix} a_r & b_r \\ -b_r & a_r \end{pmatrix}, c_1, \ldots, c_s\right),$$

这里 $b_1, , \ldots, b_r$ 为非零数。

实正规矩阵

推论

1. 一个实正规矩阵，如果其所有特征值均为实数，则它必是对称矩阵；
2. 正交矩阵必正交相似于形如

$$\operatorname{diag}\left(\begin{pmatrix}\cos\theta_1 & -\sin\theta_1\\ \sin\theta_1 & \cos\theta_1\end{pmatrix},\dots,\begin{pmatrix}\cos\theta_r & -\sin\theta_r\\ \sin\theta_r & \cos\theta_r\end{pmatrix},\pm1,\dots,\pm1\right)$$

的矩阵；
3. 实反对称矩阵必正交相似于形如

$$\operatorname{diag}\left(\begin{pmatrix}0 & b_1\\ -b_1 & 0\end{pmatrix},\dots,\begin{pmatrix}0 & b_r\\ -b_r & 0\end{pmatrix},0,\dots,0\right)$$

的矩阵，所以其特征值必为零或纯虚数并且其秩必为偶数。

正规算子的谱

设 $(V,(-,-))$ 是一个内积空间，$\varphi \in \mathcal{L}(V)$ 是正规算子。则 φ 在一组基下的矩阵为对角矩阵，设其全部互异特征值 $\lambda_1,\ldots,\lambda_k$。则

$$V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k},$$

这里，由于不同特征值的特征向量正交，所以有特征子空间 $E_{\lambda_i} \perp E_{\lambda_j}$ $(1 \leq i \neq j \leq k)$。
定义 $P_i \in \mathcal{L}(V)$，$P_i(x) = x_i$，这里 $x = x_1 + \cdots + x_k$ 对应于直和分解 $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$。则有

$$\operatorname{Im} P_i = E_{\lambda_i},\ P_iP_j = \delta_{ij}P_i,\ \sum_{i=1}^{k} P_i = I_V \text{为恒等变换。}$$

我们称 P_i 是 V 到子空间 E_{λ_i} 上的正投影算子。实际上，$P_i(x) = Proj_{E_{\lambda_i}} x$。

定理（正规算子的谱分解）

设 $\varphi \in \mathcal{L}(V)$ 是正规算子，$V_1,\ldots,V_k$ 为其特征子空间，相应的特征值为 $\lambda_1,\ldots,\lambda_k$。如果记 P_i 为 V 到 V_i 上的正投影算子，则必有

$$\varphi = \sum_{i=1}^{k} \lambda_i P_i。$$

正规算子的谱

矩阵形式，设 A 是一个正规矩阵，$\lambda_1, \dots, \lambda_k$ 为其互异特征值，相应的特征子空间的维数分别为 $n_1, \dots, n_k$。设

$$v_1, \dots, v_{n_1}; v_{n_1+1}, \dots, v_{n_1+n_2}; \dots; v_{n_1+\dots+n_{k-1}+1}, \dots, v_n;$$

是 A 的构成标准正交基的特征向量，记由这些特征向量作为列构成的矩阵为 U。通过交换特征向量的顺序或不妨设

$$U^H A U = \operatorname{diag}\left(\lambda_1 I_{n_1}, \lambda_2 I_{n_2}, \dots, \lambda_k I_{n_k}\right)。$$

定义 Hermite 矩阵 $P_1 = v_1 v_1^H + \dots + v_{n_1} v_{n_1}^H$，$P_2 = v_{n_1+1} v_{n_1+1}^H + \dots + v_{n_2} v_{n_2}^H$ 等。则有

$$P_i P_j = \delta_{ij} P_i, \ \sum_{i=1}^k P_i = I_n \text{且 } A = \sum_{i=1}^k \lambda_i P_i。$$

表达式 $A = \sum_{i=1}^k \lambda_i P_i$ 称为 A 的谱分解。谱分解是唯一的。

正规算子的谱

当一个正规算子（或正规矩阵）的特征值都为非负实数时，则该算子（矩阵）有同种形式的方根，且方根唯一。

定理

设 $\varphi \in \mathcal{L}(V)$ 为特征值都为非负实数的正规算子。则存在唯一的一个特征值都为非负实数的正规算子 ψ，使得

$$\psi^2 = \varphi.$$

矩阵形式：如果 A 是特征值都为非负实数的正规（实对称）矩阵，则存在唯一一个特征值都为非负实数的正规（实对称）矩阵 B，使得

$$B^2 = A。$$

上面的 ψ （ B ）称作 φ （ A ）的方根。

正规算子的谱

如同矩阵的 QR 分解，即一个可逆实阵可以写成一个正交矩阵和一个对角元均大于零的上三角矩阵之积，一个算子可以分解为一个酉变换与一个特征值均为非负实数的自伴随算子的乘积。

定理

复实空间上任一线性算子 φ 都可写成一个保积算子和一个特征值都为非负实数的自伴随算子的乘积。当 φ 可逆时，这种分解是唯一的。
矩阵形式：任一复方阵都可以写成一个酉阵与特征值都为非负实数的 Hermite 矩阵的乘积；任一实方阵都可以写成一个正交阵与特征值都为非负实数的实对称矩阵的乘积。

定理中的分解称作线性变换或矩阵的极分解。

最小二乘法

对一个实线性方程组 $Ax=b$ 而言，如果将 A 列分块 $A_{m\times n}=\begin{pmatrix}A_1 & A_2 & \cdots & A_n\end{pmatrix}$，则方程组的解是这样一个列向量 $x=\begin{pmatrix}x_1 & x_2 & \cdots & x_n\end{pmatrix}'$，其组合 $x_1A_1+x_2A_2+\cdots+x_nA_n$ 等于 b，或是说，$d\left(x_1A_1+x_2A_2+\cdots+x_nA_n,b\right)=0$。

但在实际生活或实践活动中，方程的导出或是方程中系常数的取得存在一定的误差，使得得到的方程没有解，这样的方程称为矛盾方程组。考虑到人力、物力的原因，需要在矛盾的方程组中找到较好的解。于是就产生了最小二乘法。

其原理是，既然找不到 $d\left(x_1A_1+x_2A_2+\cdots+x_nA_n,b\right)=0$ 的解，就找 $d\left(x_1A_1+x_2A_2+\cdots+x_nA_n,b\right)$ 是所有选择中最小的 x，这样的 x 称为最小二乘解。

最小二乘法

如果取 $U=\{x_1A_1+x_2A_2+\cdots+x_nA_n \mid x_i\in\mathbb{R}\}\subseteq\mathbb{R}_n$，
则 U 是 $\mathbb{R}_n$ 的一个子空间。取 $\mathbb{R}_n$ 的标准内积，则 U 中有一组标准正交基，于是可定义 b 在 U 上的投影 $\mathrm{Pr}_U\, b$。由前面内容已知，$(b-\mathrm{Pr}_U\, b)\perp U$，所以对 U 内的任一向量 v，有

$$\begin{aligned} d(v,b)^2 &= \|b-v\|^2=\|(b-\mathrm{Pr}_U b)+(\mathrm{Pr}_U b-v)\|^2 \\ &= \|b-\mathrm{Pr}_U b\|^2+\|\mathrm{Pr}_U b-v\|^2\geq\|b-\mathrm{Pr}_U b\|^2; \end{aligned}$$

所以 b 到 $\mathrm{Pr}_U b\in U$ 的矩阵最短。如果最小二乘解为 $x^*=\begin{pmatrix}x_1^* & x_2^* & \cdots & x_n^*\end{pmatrix}'$，则应该有 $x_1^*A_1+x_2^*A_2+\cdots+x_n^*A_n=\mathrm{Pr}_U b$，于是得到

$$b-(x_1^*A_1+x_2^*A_2+\cdots+x_n^*A_n)\perp U。$$

但每个 $A_i\in U$，所以得到 $A_i'(b-Ax^*)=0$，由 i 的任意性，得 $A'Ax^*=A'b$。

最小二乘法

定理

设 $Ax = b$ 为实线性方程组。若该方程组为矛盾方程组时，$A'Ax = A'b$ 一定有解 x^*，称为该方程组的最小二乘解。当 A 为列满秩时，最小二乘解还是唯一的。