与内积空间有关的最深刻的结果，涉及我们现在所要讨论的主题——内积空间上的线性映射和算子．我们将会看到，利用伴随的性质，可以证明一些很好的定理

极其重要的谱定理将全面地描述实内积空间上的自伴算子和复内积空间上的正规算子．然后我们将借助谱定理以理解正算子和幺正算子，这将引出幺正矩阵和矩阵分解．谱定理也将引出相当常用的奇异值分解，奇异值分解又将引出极分解

本书剩余部分最重要的结果都只在有限的维数下成立．所以从现在开始我们假设 V 和 W是有限维的

7A 自伴算子和正规算子

伴随

7.1 伴随

$T \in \mathcal{L}(V, W)$ ． T 的伴随是使得对任 $\nu \in V$ 和任 $w \in W$ 都有

$$\langle T\nu, w\rangle = \langle \nu, T^*w\rangle$$

的函数 $T^*: W \to V$

下面看看以上定义为什么是有意义的：设 $T \in \mathcal{L}(V, W)$ ，取定 $w \in W$ ，考虑 V 上的线性泛函

“伴随”这个词在线性代数里还有另一种含义，你要是在别的地方遇到了，要注意它和此处的“伴随”没有联系

$$\nu \mapsto \langle T\nu, w\rangle$$

它将 $\nu \in V$ 映射到 $\langle T\nu, w\rangle$ ．该线性泛函依赖于 T 和 w ．根据里斯表示定理（6.42）， V 中存在唯一的向量，使得该线性泛函由与它的内积给出．我们称这唯一的向量为 T^*w ．换句话说， T^*w 是 V 中使得对任一 $\nu \in V$ 都有

$$\langle T\nu, w\rangle = \langle \nu, T^*w\rangle$$

的唯一向量

上式中，左侧的内积是在 W 上的，右侧的内积是在 V 上的．不过，我们对这两种内积用相同的记号 $\langle \cdot, \cdot\rangle$

7.2 从 $\mathbf{R}^3$ 到 $\mathbf{R}^2$ 的一线性映射之伴随

$T: \mathbf{R}^3 \to \mathbf{R}^2$

$$T(x_1, x_2, x_3) = (x_2 + 3x_3, 2x_1)$$

为了计算 T^* ，设 $(x_1, x_2, x_3) \in \mathbf{R}^3$ 和 $(y_1, y_2) \in \mathbf{R}^2$ ．从而有

$$\begin{aligned}\big\langle T(x_1, x_2, x_3), (y_1, y_2)\big\rangle &= \big\langle (x_2 + 3x_3, 2x_1), (y_1, y_2)\big\rangle \\ &= x_2y_1 + 3x_3y_1 + 2x_1y_2 \\ &= \big\langle (x_1, x_2, x_3), (2y_2, y_1, 3y_1)\big\rangle.\end{aligned}$$

根据上式以及伴随的定义可得

$$T^*(y_1, y_2) = (2y_2, y_1, 3y_1).$$

7.3 值域维数至多为1 的一线性映射之伴随

取定 $u \in V$ 和 $x \in W$ ．定义 $T \in \mathcal{L}(V, W)$ 为：对任一 $\nu \in V$

$$T\nu = \langle \nu, u\rangle x$$

为了计算 T^* ，设 $\nu \in V$ 和 $w \in W$ ．从而有

$$\begin{aligned}\langle T\nu, w\rangle &= \big\langle \langle \nu, u\rangle x, w\big\rangle \\ &= \langle \nu, u\rangle\langle x, w\rangle \\ &= \big\langle \nu, \langle w, x\rangle u\big\rangle\end{aligned}$$

因此

$$T^* w = \langle w, x \rangle u$$

以上两例中最后得到的 T^* 不但是从W 到 V 的函数，还是从 W 到 V 的线性映射
下个结果表明，这在一般情况下都是成立的以上两例以及下面的证明用到了计算 T^* 的常用方法：从 $\langle T\nu, w\rangle$ 的表达式开始，然后处理一番，使得 $\langle\cdot,\cdot\rangle$ 的前一个位置里只有 ν ，那么后一个位置里就是 $T^* w$ 了

> **7.4 线性映射的伴随是线性映射**
>
> $T \in \mathcal{L}(V, W) \Rightarrow T^* \in \mathcal{L}(W, V)$

证

设 $T \in \mathcal{L}(V, W)$ ．如果 $\nu \in V$ 且 $w_1, w_2 \in W$ ，那么

$$\begin{aligned}\langle T\nu, w_1 + w_2\rangle &= \langle T\nu, w_1\rangle + \langle T\nu, w_2\rangle \\ &= \langle \nu, T^* w_1\rangle + \langle \nu, T^* w_2\rangle \\ &= \langle \nu, T^* w_1 + T^* w_2\rangle\end{aligned}$$

上式表明

$$T^*(w_1 + w_2) = T^* w_1 + T^* w_2$$

如果 $\nu \in V$ 、 $\boldsymbol{\lambda} \in \mathbf{F}$ 且 $w \in W$ ，那么

$$\begin{aligned}\langle T\nu, \lambda w\rangle &= \bar{\lambda}\langle T\nu, w\rangle \\ &= \bar{\lambda}\langle \nu, T^* w\rangle \\ &= \langle \nu, \lambda T^* w\rangle\end{aligned}$$

上式表明

$$T^*(\lambda w) = \lambda T^* w$$

因此， T^* 是线性映射，命题得证

> **7.5 伴随的性质**
>
> 设 $T \in \mathcal{L}(V, W)$
> (a) $(S+T)^* = S^* + T^*$ 对所有 $S \in \mathcal{L}(V, W)$ 成立
> (b) $(\lambda T)^* = \bar{\lambda} T^*$ 对所有 $\boldsymbol{\lambda} \in \mathbf{F}$ 成立
> (c) $(T^*)^* = T$
> (d) $(ST)^* = T^* S^*$ 对所有 $S \in \mathcal{L}(W, U)$ 成立（这里 U 是 $\mathbf{F}$ 上的有限维内积空间）
> (e) $I^* = I$ ，其中 I 是 V 上的恒等算子
> (f) 如果 T 可逆，那么 T^* 可逆且 $(T^*)^{-1} = \left(T^{-1}\right)^*$

证

设 $\nu \in V$ 且 $w \in W$ (a) 如果 $S \in \mathcal{L}(V, W)$ ，那么

$$\begin{aligned}&(S+T)\nu, w = \langle S\nu, w\rangle + \langle T\nu, w\rangle \\ &= \langle \nu, S^* w\rangle + \langle \nu, T^* w\rangle \\ &= \langle \nu, S^* w + T^* w\rangle.\end{aligned}$$

因此 $(S+T)^* w = S^* w + T^* w$ ，命题得证
(b) 如果 $\boldsymbol{\lambda} \in \mathbf{F}$ ，那么

$$\Big\langle (\lambda T)\nu, w \Big\rangle = \lambda\langle T\nu, w\Big\rangle = \lambda\langle \nu, T^* w\Big\rangle = \langle \nu, \bar{\lambda} T^* w\rangle.$$

因此 $(\lambda T)^* w = \bar{\lambda} T^* w$ ，命题得证

(c) 因为

$$T^*w,\nu = \overline{\nu,T^*w} = \overline{T\nu,w} = w,T\nu,$$

所以 $(T^*)^*\nu = T\nu$ ，命题得证

(d) 设 $S \in \mathcal{L}(W,U)$ 和 $u \in U$ ，从而有

$$(ST)\nu,u = S(T\nu),u = T\nu,S^*u = \nu,T^*(S^*u).$$

因此 $(ST)^*u = T^*(S^*u)$ ，命题得证

(e) 设 $u \in V$ ．从而有

$$\langle Iu,\nu\rangle = \langle u,\nu\rangle.$$

因此 $I^*\nu = \nu$ ，命题得证

(f) 设 T 可逆．对等式 $T^{-1}T = I$ 两边取伴随，然后用(d) 和(e) 证得 $T^*(T^{-1})^* = I$ ．类似地，由$TT^{-1} = I$ 可得 $(T^{-1})^*T^* = I$ ．因此， $(T^{-1})^*$ 是 T^* 的逆，命题得证

如果 $\mathbf{F} = \mathbf{R}$ ，那么根据以上结果的 (a) 和 (b)， $T \mapsto T^*$ 这一映射是从 $\mathcal{L}(V,W)$ 到 $\mathcal{L}(W,V)$ 的线性映射．然而，如果 $\mathbf{F} = \mathbf{C}$ ，那么该映射不是线性的，这是由于(b) 中出现的复共轭
下面的结果揭示了线性映射同其伴随的零空间和值域之间的关系

7.6 $T*$ 的零空间和值域

$T \in \mathcal{L}(V,W)$
(a) null $T^* = (\mathrm{range}T)^\perp$
(b) range $T^* = (\mathrm{null}T)^\perp$
(c) null $T = (\mathrm{range}T^*)^\perp$
(d) range $T = (\mathrm{null}T^*)^\perp$

证

(a) 令 $w \in W$ ，从而有

$$\begin{aligned} w \in \mathrm{null}T^* &\iff T^*w = 0 \\ &\iff w \in (\mathrm{range}T)^\perp. \end{aligned}$$

因此， $\mathrm{null}T^* = (\mathrm{range}T)^\perp$ ，(a) 得证
如果对(a) 的两侧求正交补，我们就得到了(d)，其中用到了6.52．在(a) 中将 T 替换成 T^* ，就得到了(c)，其中用到了7.5 (c)．最后，在(d) 中将 T 替换成 T^* ，就得到了(b)

我们很快会看到，接下来的定义与线性映射的伴随的矩阵密切相关

7.7 共轭转置

$m \times n$ 矩阵 A 的共轭转置是将其行列互换再对每个元素取复共轭得到的 $n \times m$ 矩阵 A^* 换句话说，如果 $j \in \{1,\ldots,n\}$ 且 $k \in \{1,\ldots,m\}$ ，那么

$$(A^*)_{j,k} = \overline{A_{k,j}}$$

7.8 $\begin{pmatrix} 2 & 3+4\mathrm{i} & 7 \\ 6 & 5 & 8\mathrm{i} \end{pmatrix}$ 的共轭转置

$$\begin{pmatrix} 2 & 6 \\ 3-4\mathbf{i} & 5 \\ 7 & -8\mathbf{i} \end{pmatrix}$$

接下来这条结果展示了如何从 T 的矩阵计算出 T^* 的矩阵．注意：关于非规范正交基，T^* 的矩阵不一定等于 T 的矩阵的共轭转置
线性映射的伴随不依赖于基的选取．因此我们经常强调的是线性映射的伴随，而不是矩阵的转置或共轭转置

7.9 $T*$ 的矩阵等于T 的矩阵的共轭转置

$T \in \mathcal{L}(V,W)$．设 $e_1,\ldots,e_n$ 是 V 的规范正交基， $f_1,\ldots,f_m$ 是 W 的规范正交基．那么 $M(T^*,(f_1,\ldots,f_m),(e_1,\ldots,e_n))$ 是 $\mathcal{M}(T,(e_1,\ldots,e_n),(f_1,\ldots,f_m))$ 的共轭转置．换句话说，

$$M(T^*) = \big(M(T)\big)^*$$

证

在本证明中，我们把 $\mathcal{M}(T,(e_1,\ldots,e_n),(f_1,\ldots,f_m))$ 和 $M(T^*,(f_1,\ldots,f_m),(e_1,\ldots,e_n))$ 这两个较长的表达式，分别写成 $\nabla\mathcal{M}(T)$ 和 $\boldsymbol{\mathcal{M}}(\boldsymbol{T}^*)$
回顾一下， $\nabla\mathcal{M}(T)$ 的第 k 列，是通过把 e 写成各 f_j 的线性组合得到的——该线性组合中所用的标量就成为 $\nabla\mathcal{M}(T)$ 的第 k 列．因为 $f_1,\ldots,f_m$ 是 W 的规范正交基，所以我们知道如何把 Te_k 写成各 f_j 的线性组合【见6.30 (a)】:

$$Te_k = \langle Te_k, f_1\rangle f_1 + \cdots + \langle Te_k, f_m\rangle f_m.$$

因此
在以上表述中，将 T 替换成 T^*，并互换 $e_1,\ldots,e_n$ 和 $f_1,\ldots,f_m$，就证得 $\mathcal{M}(T^*)$ 的 j 行 k 列元素为 $\langle T^*f_k, e_j\rangle$．这等于 $\langle f_k, Te_j\rangle$，进而等于 $\overline{\langle Te_j, f_k\rangle}$，而这就等于 $\nabla\mathcal{M}(T)$ 的 k 行 j 列元素的复共轭．因此 $\boldsymbol{\mathcal{M}}(T^*) = (\boldsymbol{\mathcal{M}}(T))^*$

里斯表示定理（6.58 中所述）指出了 V 和其对偶空间 V'（在 3.110 中定义）的等同关系在这一等同关系下，子集 $U \subseteq V$ 的正交补 $U^\perp$ 对应于 U 的零化子 U^0．如果 U 是 V 的子空间，那么在这一等同关系下， $U^\perp$ 和 U^0 的维数公式就等同起来了——见3.125 和6.51
设 $T: V \to W$ 是一线性映射．在 V 和 V' 的等同关系与W 和 W' 的等同关系之下，伴随映射 $T^*: W \to V$ 对应于对偶映射 $T': W' \to$
比起零化子和对偶映射，正交补和伴随更易处理。所以在内积空间的背景下，就不需处理零化子和对偶映射了
V'（在 3.118 中定义），这也是你在本节习题 32 中所要验证的．在这一等同关系下，null T^* 和range T^* 的公式【7.6 (a) 和 (b)】就跟 null T' 和 range T' 的公式【3.128 (a) 和 3.130 (b)】完全一致．此外，有关 T^* 的矩阵的定理（7.9）也类似于有关 T' 的矩阵的定理（3.132）

自伴算子

现在，我们把注意力转向内积空间上的算子

7.10 自伴

算子 $T \in \mathcal{L}(V)$ 自伴的若 $T = T^*$

如果 $T \in \mathcal{L}(V)$ 且 $e_1,\ldots,e_n$ 是 V 的规范正交基，那么 T 是自伴的当且仅当 $M(T,(e_1,\ldots,e_n)) = M(T,(e_1,\ldots,e_n))^*$．这是由7.9 推得的

7.11 根据 T 的矩阵确定其是否自伴

$c \in \mathbf{F}$， T 是 $\mathbf{F}^2$ 上的算子，其关于标准基的矩阵是

$$M(T) = \begin{pmatrix} 2 & c \\ 3 & 7 \end{pmatrix}$$

T^* 关于标准基的矩阵是

$$M(T^*) = \begin{pmatrix} 2 & 3 \\ \bar{c} & 7 \end{pmatrix}$$

因此 $\mathcal{M}(T)=\mathcal{M}(T^*)$ 当且仅当 $c=3$．于是，算子 T 自伴当且仅当 $c=3$
有个值得一记的好类比：$\mathcal{L}(V)$ 上的伴随扮演了类似于 $\mathbf{c}$ 上的复共轭的角色．复数 z 是实的当且仅当 $z=\bar{z}$；自伴算子（$T=T^*$）就类似于实数
算子 $T\in\mathcal{L}(V)$ 自伴，当且仅当
我们将看到，以上所讨论的类比，会在自伴算子的一些重要性质上体现出来，首先是下一条结果中的特征值

$$\langle T\nu,w\rangle=\langle\nu,Tw\rangle$$

如果 $\mathbf{F}=\mathbf{R}$，那么根据定义每个特征值都是实的．所以接下来这条结果只在 $\mathbf{F}=\mathbf{C}$ 的情形下引人兴趣
对所有 $\nu,w\in V$ 成立

7.12 自伴算子的特征值

自伴算子的每个特征值都是实的

证

设 T 是 V 上的自伴算子．令 λ 是 T 的特征值，再令 ν 是 V 中使得 $T\nu=\lambda\nu$ 的非零向量那么有

$$\lambda\|\nu\|^2=\langle\lambda\nu,\nu\rangle=\langle T\nu,\nu\rangle=\langle\nu,T\nu\rangle=\langle\nu,\lambda\nu\rangle=\bar{\lambda}\|\nu\|^2$$

因此 $\lambda=\bar{\lambda}$，意即 λ 是实的，命题得证

接下来这条结果对于实内积空间而言是假命题．举个例子，考虑"绕原点逆时针旋转 $90^{\circ\prime\prime}$ 这一算子 $T\in\mathcal{L}(\mathbf{R}^2)$，那么 $T(x,y)=(-y,x)$．注意到，$T\nu$ 正交于 ν 对任一 $\boldsymbol{\nu}\in\mathbf{R}^2$ 都成立，尽管 $T\neq 0$

7.13 对所有 v 都有 Tv 正交于 v $T=0$（假设 F = C）

设 V 是复内积空间以及 $T\in\mathcal{L}(V)$，那么

证

如果 $u,w\in V$，那么有

$$\begin{aligned}\langle Tu,w\rangle=&\frac{\langle T(u+w),u+w\rangle-\langle T(u-w),u-w\rangle}{4}\\&+\frac{\langle T(u+\mathrm{i}w),u+\mathrm{i}w\rangle-\langle T(u-\mathrm{i}w),u-\mathrm{i}w\rangle}{4}\mathrm{i}.\end{aligned}$$

这可以通过计算右侧来验证．注意到，右侧的每一项都具有 $\langle T\nu,\nu\rangle$（$\nu\in V$）的形式
现设 $\langle T\nu,\nu\rangle=0$ 对任一 $\nu\in V$ 都成立，那么上式意味着 $\langle Tu,w\rangle=0$ 对所有 $u,w\in V$ 成立．这意味着 $Tu=0$ 对任一 $u\in V$ 都成立（取 $w=Tu$
）．于是 $T=0$，命题得证

接下来这条结果对于实内积空间而言是假命题．考虑实内积空间上任何一个非自伴的算子，都能说明这点

关于自伴算子如何表现得像实数，下面这条结果给出了又一个好例子

7.14 Tv,v 对所有 v 都是实的 T 是自伴的（假设F = C）

设 V 是复内积空间以及 $T\in\mathcal{L}(V)$，那么

证

如果 $\nu \in V$，那么

$$\langle T^*\nu, \nu\rangle = \overline{\langle \nu, T^*\nu\rangle} = \overline{\langle T\nu, \nu\rangle}$$

这样一来，就有
其中第二个等价关系通过对 $T - T^*$ 应用 7.13 得到，第三个等价关系由式(7.15) 得到

实内积空间 V 上非零算子 T 也许会对所有 $\nu \in V$ 都满足 $\langle T\nu, \nu\rangle = 0$
但接下来这条表明这不会发生在自伴算子上

7.16 T 自伴且 Tv, v = 0 对所有 v 成立 T = 0

设 T 是 V 上的自伴算子，那么

证

在 V 是复内积空间的条件下，我们已经证明过这一定理，且无需假设 T 自伴（见7.13）因此我们可以假设 V 是实内积空间．如果 $u, w \in V$，那么

$$\langle Tu, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4}$$

这可以通过计算右侧来证明．计算过程需用到下式

$$Tw, u = w, Tu = Tu, w,$$

其中第一个等号成立是因为 T 自伴，第二个等号成立是因为我们在实内积空间里操作
现设 $T\nu, \nu = 0$ 对任一 $\nu \in V$ 都成立．因为式(7.17) 的右侧每一项都具有 $\langle T\nu, \nu\rangle\,(\nu \in V)$ 的形式，所以这意味着 $\langle Tu, w\rangle = 0$ 对所有 $u, w \in V$ 成立．这又意味着 $Tu = 0$ 对任一 $u \in V$ 都成立（取 $w = Tu$）．于是，$T = 0$，命题得证

正规算子

7.18 正规

内积空间上的算子被称为正规的，如果它与它的伴随可交换

换句话说，$T \in \mathcal{L}(V)$ 是正规的，如果 $TT^* = T^*T$
每个自伴算子都是正规的，这是因为，如果 T 自伴，那么 $T^* = T$，于是 T 和 T^* 可交换

7.19 正规但不自伴的算子

令 T 是 $\mathbf{F}^2$ 上的算子，其关于标准基的矩阵是

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}$$

因此 $T(w, z) = (2w - 3z, 3w + 2z)$
该算子 T 并不是自伴的，因为其矩阵第 2 行第 1 列的元素（等于 3）不等于第 1 行第 2列的元素（等于 $^{-3}$）的复共轭
TT^* 的矩阵等于

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}$$

类似地，T^*T 的矩阵等于

$$\begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}$$

因为 TT^* 和 T^*T 的矩阵相同，我们得知 $TT^* = T^*T$． 因此 T 是正规的

下一节我们将看到正规算子为什么值得特别关注．接下来这条结果给出了对正规算子的一个很有用的刻画

7.20 T 是正规的当且仅当Tv 和$T{*}v$ 的范数相同

$T \in \mathcal{L}(V)$ 则 T 是正规的 $\iff$ $\|T\nu\| = \|T^*\nu\|$ 对任一 $\nu \in V$ 都成立

证

其中，我们用到了 7.16 来建立第二个等价关系（注意到算子 $T^*T - TT^*$ 是自伴的）

下一条结果展示了以上结果的几个推论．将下一条结果的 (e) 和本节习题 3 相比较：那道习题指出每个算子的伴随的特征值（作为集合）都等于该算子特征值的复共轭 1，而没有提及特征向量，因为算子同它的伴随可以有不同的特征向量．然而，下一条结果中的 (e) 表明，正规算子和它的伴随有完全相同的特征向量

7.21 正规算子的值域、零空间和特征向量

$T \in \mathcal{L}(V)$ 是正规的，则
(a) null $T = \mathrm{null} T^*$
(b) range $T =$ range T^*
(c) $V = \mathrm{null} T \oplus \mathrm{range} T$
(d) 对任一 $\boldsymbol{\lambda} \in \mathbf{F}$， $T - \lambda I$ 都是正规的
(e) 如果 $\nu \in V$ 且 $\boldsymbol{\lambda} \in \mathbf{F}$， 则 $T\nu = \lambda\nu$ 当且仅当 $T^*\nu = \bar{\lambda}\nu$

证

(a) 设 $\nu \in V$， 那么有

$$\nu \in \mathrm{null} T \iff \|T\nu\| = 0 \iff \|T^*\nu\| = 0 \iff \nu \in \mathrm{null} T^*$$

以上等价关系的中间那个由7.20 得到．因此， $\mathrm{null} T = \mathrm{null} T^*$
(b)

$$\mathrm{range} T = (\mathrm{null} T^*)^\perp = (\mathrm{null} T)^\perp = \mathrm{range} T^*$$

其中，第一个相等关系来自 7.6 (d)，第二个相等关系来自本结果的 (a)，第三个相等关系来自 7.6 (b)

$$V = (\mathrm{null} T) \oplus (\mathrm{null} T)^\perp = \mathrm{null} T \oplus \mathrm{range} T^* = \mathrm{null} T \oplus \mathrm{range} T.$$

其中，第一个相等关系来自6.49，第二个相等关系来自7.6 (b)，第三个相等关系来自本结果的 (b)
(d) 设 $\boldsymbol{\lambda} \in \mathbf{F}$， 那么有

$$\begin{aligned}(T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)(T^* - \bar{\lambda} I) \\ &= TT^* - \bar{\lambda} T - \lambda T^* + |\lambda|^2 I \\ &= T^*T - \bar{\lambda} T - \lambda T^* + |\lambda|^2 I \\ &= (T^* - \bar{\lambda} I)(T - \lambda I) \\ &= (T - \lambda I)^*(T - \lambda I).\end{aligned}$$

因此， $T - \lambda I$ 和它的伴随可交换．于是 $T - \lambda I$ 是正规的
(e) 设 $\nu \in V$ 且 $\boldsymbol{\lambda} \in \mathbf{F}$． 那么 (d) 和 7.20 蕴涵

$$\|(T - \lambda I)\nu\| = \|(T - \lambda I)^*\nu\| = \|(T^* - \bar{\lambda} I)\nu\|$$

因此，$\|(T-\lambda I)\nu\|=0$ 当且仅当 $\|(T^*-\bar{\lambda}I)\nu\|=0$．于是 $T\nu=\lambda\nu$ 当且仅当 $T^*\nu=\bar{\lambda}\nu$ 因为每个自伴算子都是正规的，所以下一条结果也适用于自伴算子

> 7.22 正规算子的正交特征向量
>
> $T\in\mathcal{L}(V)$ 是正规的，则 T 的对应于不同特征值的特征向量正交

证

设 α,β 是 T 的不同特征值，对应的特征向量是 u,ν.．因此 $Tu=\alpha u$ 且 $T\nu=\beta\nu$．根据 7.21(e)，有 $T^*\nu=\bar{\beta}\nu$，故

$$\begin{aligned}(\alpha-\beta)\langle u,\nu\rangle&=\langle\alpha u,\nu\rangle-\langle u,\bar{\beta}\nu\rangle\\&=\langle Tu,\nu\rangle-\langle u,T^*\nu\rangle\\&=0\end{aligned}$$

因为 $\alpha\neq\beta$，所以上式表明 $u,\nu=0$．因此 u 和 ν 正交，命题得证

如下所述，下一条结果仅在 $\mathbf{F}=\mathbf{C}$ 条件下有意义．然而在本节习题 12 中，可以看到一个不论 $\mathbf{F}=\mathbf{C}$ 还是 $\mathbf{F}=\mathbf{R}$ 都有意义的版本
设 $\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$．将 $\mathcal{L}(V)$ 类比为 $\mathbf{c}$，$\mathcal{L}(V)$ 上的伴随就如同C 上的复共轭，那么由式 (7.24) 定义的算子 A 和 B 也就相当于 T 的“实部”和“虚部”．这样一来，以下结果的非正式标题才有了意义

> 7.23 T 是正规的 T 的实部和虚部可交换
>
> 设 $\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$，则 T 是正规的当且仅当存在可交换的自伴算子 A 和 B 使得$T=A+\mathrm{i}B$

证

设 T 是正规的，令

$$A=\frac{T+T^*}{2}$$

$$B=\frac{T-T^*}{2\mathrm{i}}$$

则 A 和 B 是自伴的且 $T=A+\mathrm{i}B$．可以很快计算得

$$AB-BA=\frac{T^*T-TT^*}{2\mathrm{i}}$$

因为 T 是正规的，所以上式右侧等于0．因此，算子 A 和 B 可交换，此方向得证
为了证明另一方向的蕴涵关系，现在假设存在可交换的自伴算子 A 和 B 使得 $T=A+\mathrm{i}B$，那么 $T^*=A-\mathrm{i}B$．将前面这两个式子相加再除以2，得到式(7.24) 中 A 的式子；将前面这两个式子相减再除以 2i，得到式 (7.24) 中 B 的式子．这样一来，式 (7.24) 就蕴涵式 (7.25)．因为 B 和 A 可交换，所以式 (7.25) 蕴涵 T 是正规的，此方向得证

7B 谱定理

回顾对角矩阵是对角线外处处为0 的方阵；V 上的算子若关于 V 的某个基有对角矩阵就被称为可对角化的，而这种情况当且仅当 V 有由该算子的特征向量组成的基时才会发生（见 5.55）
V 上最好的算子，就是关于 V 中的某个规范正交基有对角矩阵的算子．也正是这些算子，其特征向量可用来组成 V 的规范正交基．本节我们的目标是，证明用以刻画这些算子的谱定理——它指出，当 $\mathbf{F}=\mathbf{R}$ 时这些算子即自伴算子，当 $\mathbf{F}=\mathbf{C}$ 时这些算子即正规算子
在对内积空间上的算子的研究中，谱定理也许是最有用的工具．它在特定的无限维内积空间上的推广——例如，见作者的《测度、积分与实分析》（Measure, Integration & Real Analysis）一书的章节10D——在泛函分析中起到了关键作用
由于谱定理的结论依赖于 F，所以我们将谱定理划成两块，分别称作实谱定理和复谱定理

实谱定理

证实谱定理需两个引理．它们在实内积空间和复内积空间上都成立，但是在复谱定理的证明中不需要用到
对于接下来这条结果，你只要想到实系数二次多项式，就能猜到它是正确的，甚至还能给出它的证明．具体而言，设 $\boldsymbol{b}, \boldsymbol{c} \in \mathbf{R}$ 且 $b^2 < 4c$，令 x 是一实数，从而有

$$x^2 + bx + c = \left(x + \frac{b}{2}\right)^2 + \left(c - \frac{b^2}{4}\right) > 0$$

特别地，$x^2 + bx + c$ 是一可逆实数（这不过是把"非 ″ 往复杂了说）．将实数 x 替换为自伴算子（回想一下实数和自伴算子之间的类比），就引出了下面这条结果

> **7.26 可逆二次表达式**
>
> $T \in \mathcal{L}(V)$ 是自伴的，$\boldsymbol{b}, \boldsymbol{c} \in \mathbf{R}$ 使得 $b^2 < 4c$，则
>
> $$T^2 + bT + cI$$
>
> 是可逆算子

证

令 ν 是 V 中的非零向量，那么

$$\begin{aligned}
&\langle (T^2 + bT + cI)\nu, \nu \rangle = \langle T^2\nu, \nu \rangle + b\langle T\nu, \nu \rangle + c\langle \nu, \nu \rangle \\
&= \langle T\nu, T\nu \rangle + b\langle T\nu, \nu \rangle + c\|\nu\|^2 \\
&\geq \|T\nu\|^2 - |b|\|T\nu\|\|\nu\| + c\|\nu\|^2 \\
&= \left(\|T\nu\| - \frac{|b|\|\nu\|}{2}\right)^2 + \left(c - \frac{b^2}{4}\right)\|\nu\|^2 \\
&> 0
\end{aligned}$$

其中第三行成立是由柯西-施瓦兹不等式（6.14）得到的．最后一个不等号意味着 $(T^2 + bT + cI)\nu \neq 0$．因此 $T^2 + bT + cI$ 是单射，这意味着它是可逆的（见3.65）

接下来这条结果将是我们证明实谱定理所用的关键工具

> **7.27 自伴算子的最小多项式**
>
> $T \in \mathcal{L}(V)$ 是自伴的，则 T 的最小多项式等于 $(z - \lambda_1) \cdots (z - \lambda_m)$，其中 $\lambda_1, \ldots, \lambda_m \in \mathbf{R}$

证

先设 $\mathbf{F} = \mathbf{C}$． T 的最小多项式的零点即 T 的特征值【根据 5.27 (a)】．而 T 的所有特征值都是实的（根据 7.12）．因此代数基本定理的版本二（见4.13）告诉我们 T 的最小多项式具有待证的形式
设 $\mathbf{F} = \mathbf{R}$．由 $\mathbf{R}$ 上的多项式分解（见4.16）可得，存在 $\lambda_1, \ldots, \lambda_m \in \mathbf{R}$ 和 $b_1, \ldots, b_N$,$c_1, \ldots, c_N \in \mathbf{R}$ （对任一 k 都有 $b_k^2 < 4c_k$）使得 T 的最小多项式等于

$$(z - \lambda_1) \cdots (z - \lambda_m)(z^2 + b_1 z + c_1) \cdots (z^2 + b_N z + c_N).$$

其中 m 或 N 可能等于0，也就是没有相应形式的项．这样一来，就有

$$(T - \lambda_1 I) \cdots (T - \lambda_m I)(T^2 + b_1 T + c_1 I) \cdots (T^2 + b_N T + c_N I) = 0$$

如果 $N > 0$，那么我们可以对上式两边同时乘以 $T^2 + b_N T + c_N I$ （由 7.26 可知其为可逆算子）的逆，这样得到的 T 的多项式仍等于 0，但其次数比式 (7.28) 低 2，这和最小多项式"次数最小"的性质相违背．所以必然有 $N = 0$，这意味着式 (7.28) 中的最小多项式具有的形式是 $(z - \lambda_1) \cdots (z - \lambda_m)$，命题得证

以上结果连同 5.27 (a) 蕴涵了"每个自伴算子都有特征值"这一命题．事实上，我们在接下来的结果中还会看到，自伴算子有足够的特征向量以形成一个基
接下来这条结果，是线性代数中的主要定理之一．它全面地描述了实内积空间上的自伴算子．

7.29 实谱定理

$\mathbf{F}=\mathbf{R}$ 且 $T\in\mathcal{L}(V)$. 下列等价:
(a) T 是自伴的
(b) T 关于 V 的某个规范正交基有对角矩阵
(c) V 有由 T 的特征向量构成的规范正交基

证

设 (a) 成立，那么 T 是自伴的．根据 6.37 和 7.27 这两条有关最小多项式的结果，可得 T 关于 V 的某个规范正交基有上三角矩阵．关于这一规范正交基， T^* 的矩阵是 T 的矩阵的转置．然而， $T^*=T$ ．因此， T 的矩阵的转置等于 T 的矩阵．又因为 T 的矩阵是上三角的，所以这意味着该矩阵对角线上方和下方的元素都为 0．于是， T 关于这一规范正交基的矩阵是对角矩阵．因此，(a) 蕴涵(b)
设(b) 成立，那么 T 关于 V 的某个规范正交基有对角矩阵．这个对角矩阵和它的转置相等．所以关于那个基， T^* 的矩阵等于 T 的矩阵．于是， $T^*=T$ ， 证得 (b) 蕴涵 (a)
(b) 和(c) 之间的等价关系由定义得【或者参见 5.55 (a) 和(b) 等价的证明】

7.30 一算子的特征向量所成的规范正交基

考虑 $\mathbf{R}^3$ 上的算子 T ，其矩阵（关于标准基）为

$$\begin{pmatrix} 14 & -13 & 8 \\ -13 & 14 & 8 \\ 8 & 8 & -7 \end{pmatrix}$$

该矩阵等于自身的转置且其元素都为«实数，所以 T 是自伴的．可验证

$$\frac{(1,-1,0)}{\sqrt{2}},\frac{(1,1,1)}{\sqrt{3}},\frac{(1,1,-2)}{\sqrt{6}}$$

是 $\mathbf{R}^3$ 的规范正交基且由 T 的特征向量组成．关于这个基， T 的矩阵是对角矩阵

$$\begin{pmatrix} 27 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -15 \end{pmatrix}$$

实谱定理同时应用于不止一个算子的版本，见本节习题17

复谱定理

接下来这条结果，全面地描述了复内积空间上的正规算子

7.31 复谱定理

设 $\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$. 那么下列等价:
(a) T 是正规的
(b) T 关于 V 的某个规范正交基有对角矩阵
(c) V 有由 T 的特征向量构成的规范正交基

证

设 (a) 成立，那么 T 是正规的．由舒尔定理（6.38)，存在 V 的一规范正交基$e_1,\ldots,e_n$ ，使得 T 关于它有上三角矩阵．于是我们可以写出

$$M(T,(e_1,\ldots,e_n))=\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \ddots & \vdots \\ 0 & & a_{n,n} \end{pmatrix}$$

我们将证明这个矩阵实际上是对角矩阵
我们从以上这个矩阵可以得到

$$\|Te_1\|^2 = |a_{1,1}|^2,$$
$$\|T^*e_1\|^2 = |a_{1,1}|^2 + |a_{1,2}|^2 + \cdots + |a_{1,n}|^2.$$

因为 T 是正规的，所以 $\|Te_1\| = \|T^*e_1\|$ （见 7.20）. 因此由以上两个等式可得，式 (7.32)中的矩阵第一行除了第一个元素 $a_{1,1}$ 可能不为0 外都为0
这样一来，式(7.32) 蕴涵了

$$\|Te_2\|^2 = |a_{2,2}|^2$$

（因为我们在上一段证明了 $a_{1,2} = 0$ ）和

$$\|T^*e_2\|^2 = |a_{2,2}|^2 + |a_{2,3}|^2 + \cdots + |a_{2,n}|^2$$

因为 T 是正规的，所以 $\|Te_2\| = \|T^*e_2\|$. 因此由以上两个等式可得，式(7.32) 中的矩阵第二行除了对角线元素 $a_{2,2}$ 可能不为0 外都为0
按这种方式继续下去，我们可以得到，矩阵(7.32) 的所有非对角线元素都等于0．因此(b)成立，也就完成了(a) 蕴涵(b) 的证明
现在假设(b) 成立，那么 T 关于 V 的某个规范正交基有一对角矩阵．取 T 的矩阵的共轭转置，可以得到 T^* （关于同一个基）的矩阵；于是 T^* 也有对角矩阵．任意两个对角矩阵都可交换，从而 T 和 T^* 可交换，意即 T 是正规的．换句话说，(a) 成立．这也就完成了(b) 蕴涵(a) 的证明
(b) 和(c) 的等价关系，由定义可得（也见5.55）
关于前述结果中的(a) 蕴涵(b)，可在本节习题13 和20 找到其他方式的证明

本节习题 14 和 15 通过将定义空间表达成特征空间的正交直和，来解释实谱定理和复谱定理
复谱定理同时应用于不止一个算子的版本，见习题16
复谱定理的主要结论是：有限维复内积空间上的每个正规算子，都可由规范正交基对角化．接下来的例子呈现了这一点

7.33 一算子的特征向量所成的规范正交基

$T \in \mathcal{L}(\mathbf{C}^2)$，其定义为： $T(w,z) = (2w-3z, 3w+2z)$． T （关于标准基）的矩阵为

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}$$

正如我们在例 7.19 中所见， T 是正规算子
可以验证

$$\frac{1}{\sqrt{2}}(\mathrm{i},1), \frac{1}{\sqrt{2}}(-\mathrm{i},1)$$

是 $\mathbf{C}^2$ 的规范正交基，且由 T 的特征向量组成．关于这个基， T 的矩阵是对角矩阵

$$\begin{pmatrix} 2+3\mathrm{i} & 0 \\ 0 & 2-3\mathrm{i} \end{pmatrix}$$

7C 正算子

7.34 正算子

$T \in \mathcal{L}(V)$ 称为正的，若 T 是自伴的且对所有 $\nu \in V$ 有

$$\langle T\nu, \nu \rangle \geq 0$$

如果 V 是复向量空间，那么以上定义中“ T 是自伴的”这一条件可以去掉（根据7.14）

7.35 正算子

(a) 令算子 $T \in \mathcal{L}(\mathbf{F}^2)$ 的矩阵（关于标准基）是 $\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$. 那么 T 是自伴的，而且
$T(w,z),(w,z) = 2|w|^2 - 2\mathrm{Re}(w\bar{z}) + |z|^2 = |w-z|^2 + |w|^2 \geq 0$ 对所有 $(w,z) \in \mathbf{F}^2$ 成立．因此 T 是正算子

(b) 如果 U 是 V 的子空间，那么正交投影 P_U 是正算子．你应该验证一下
(c) 如果 $T \in \mathcal{L}(V)$ 是自伴的，且 $\boldsymbol{b}, \boldsymbol{c} \in \mathbf{R}$ 使得 $b^2 < 4c$ ，那么 $T^2 + bT + cI$ 是正算子，如7.26的证明所示

7.36 平方根

算子 R 称为算子 T 的平方根，如果 $R^2 = T$

7.37 算子的平方根

$T \in \mathcal{L}(\mathbf{F}^3)$ 定义为 $T(z_1, z_2, z_3) = (z_3, 0, 0)$ ，那么定义为 $R(z_1, z_2, z_3) = (z_2, z_3, 0)$ 的算子 $R \in \mathcal{L}(\mathbf{F}^3)$ 是 T 的一个平方根，因为 $R^2 = T$ （你可以验证）

接下来这条结果中的正算子的特性，对应于 C 中非负数的特性．具体而言， $z \in \mathbf{C}$ 非负当且仅当它有非负平方根，这对应于条件 (d)；z 非负当且仅当它有实平方根，这对应于条件 (e)；z 非负当且仅当存在 $w \in \mathbf{C}$ 使因为正算子对应于非负数，所以更好的术语应该是"非负算子"．然而，算子理论家们一贯称之为正算子，所以我们也遵循习惯．一些数学家用的是"正半定算子"（positive semidef-inite operator)，意思和"正算子"一样得 $z = \overline{w}w$ ，这对应于条件(f)．等价于"某算子是正算子"的又一条件，见本节习题20

7.38 正算子的特性

令 $T \in \mathcal{L}(V)$ ．那么下列等价：
(a) T 是正算子
(b) T 自伴且所有特征值非负
(c) 关于 V 的某个规范正交基， T 的矩阵是对角矩阵且对角线上仅有非负数
(d) T 有正平方根
(e) T 有自伴平方根
(f) 存在某个 $R \in \mathcal{L}(V)$ 使得 $T = R^*R$

证

将证：($\mathrm{a}) \implies (\mathrm{b}) \implies (\mathrm{c}) \implies (\mathrm{d}) \implies (\mathrm{e}) \implies () \implies (\mathrm{a})$
首先假设 (a) 成立，那么 T 是正的，也就意味着 T 是自伴的（根据正算子的定义）．为证明(b) 中的另一条件，设 λ 是 T 的特征值．令 ν 是 T 的对应于 λ 的特征向量．那么

$$0 \leq T\nu, \nu = \lambda\nu, \nu = \lambda\,\nu, \nu.$$

因此， λ 是非负数．于是，(b) 成立，也就表明(a) 蕴涵(b)
现在假设(b) 成立，那么 T 自伴且所有特征值非负．由谱定理（7.29 和7.31）， V 有一规范正交基 $e_1, \ldots, e_n$ 由 T 的特征向量组成．令 $\lambda_1, \ldots, \lambda_n$ 是 T 的对应于 $e_1, \ldots, e_n$ 的特征值，那么每个 λ_k 都是非负数． T 关于 $e_1, \ldots, e_n$ 的矩阵是以 $\lambda_1, \ldots, \lambda_n$ 为对角线元素的对角矩阵，这表明 (b) 蕴涵 (c)
现在假设 (c) 成立．设 $e_1, \ldots, e_n$ 是 V 的规范正交基，使得 T 关于这个基的矩阵是对角矩阵，且对角线上是非负数 $\lambda_1, \ldots, \lambda_n$ ．由线性映射引理（3.4）可得，存在 $R \in \mathcal{L}(V)$ 使得

$$Re_k = \sqrt{\lambda_k} e_k$$

对任一 $k = 1, \ldots, n$ 都成立．你应自行验证， R 是正算子．此外，有 $R^2 e_k = \lambda_k e_k = Te_k$ 对任一k 都成立．这意味着 $R^2 = T$ ．因此， R 是 T 的正平方根．于是，(d) 成立，这表明(c) 蕴涵(d)
每个正算子都是自伴的（根据正算子的定义），于是(d) 蕴涵(e)
现在假设(e) 成立，意即 V 上存在一自伴算子 R 使得 $T = R^2$ ．那么 $T = R^*R$ （因为 $R^* = R$ ）于是 (e) 蕴涵 (f)
最后，假设 (f) 成立．令 $R \in \mathcal{L}(V)$ 使得 $T = R^*R$ ．那么 $T^* = (R^*R)^* = R^*(R^*)^* = R^*R = T$ 于是 T 自伴．离证明(a) 成立就差一步——注意到

$$\langle T\nu, \nu \rangle = \langle R^*R\nu, \nu \rangle = \langle R\nu, R\nu \rangle \geq 0$$

对任一 $\nu \in V$ 都成立．因此 T 是正的，这表明(f) 蕴涵(a)

每个非负数都有唯一非负平方根．接下来这条结果表明正算子也有类似的性质

7.39 每个正算子都只有一个正平方根

V 上的每个正算子都有唯一正平方根

证

设 $T \in \mathcal{L}(V)$ 是正的．设 $\nu \in V$ 是 T 的特征向量．于是存在一实数 $\lambda \geq 0$ 使得 $T\nu = \lambda\nu$
令 R 是 T 的正平方根．我们将证明 $R\nu = \sqrt{\lambda}\nu$，这意味着 R 对 T 的特征向量的作用是唯一确定的．因为 V 中存在由 T 的特征向量组成的基（根据谱定理），所以这意味着 R 是唯一确定的
一个正算子可以有无穷多个平方根（但是它们当中只能有一个是正的）．例如，如果$\dim V > 1$ 的话， V 上的恒等算子就有无穷多个平方根
为了证明 $R\nu = \sqrt{\lambda}\nu$，注意到，谱定理指出 V 中存在由 R 的特征向量组成的规范正交基 $e_1, \dots, e_n$．因为 R 是正算子，所以它的所有特征值非负．因此存在非负数 $\lambda_1, \dots, \lambda_n$ 使得$Re_k = \sqrt{\lambda_k}e_k$ 对任一 $k = 1, \dots, n$ 成立
因为 $e_1, \dots, e_n$ 是 V 的基，所以我们可以写出

$$\nu = a_1 e_1 + \cdots + a_n e_n$$

其中 $a_1, \dots, a_n \in \mathbf{F}$．因此

$$R\nu = a_1\sqrt{\lambda_1}e_1 + \cdots + a_n\sqrt{\lambda_n}e_n$$

于是

$$\lambda\nu = T\nu = R^2\nu = a_1\lambda_1 e_1 + \cdots + a_n\lambda_n e_n$$

上式表明

$$a_1\lambda e_1 + \cdots + a_n\lambda e_n = a_1\lambda_1 e_1 + \cdots + a_n\lambda_n e_n$$

故 $a_k(\lambda - \lambda_k) = 0$ 对任一 $k = 1, \dots, n$ 都成立．于是

$$\nu = \sum_{\{k:\lambda_k=\lambda\}} a_k e_k$$

故

$$R\nu = \sum_{\{k:\lambda_k=\lambda\}} a_k\sqrt{\lambda}e_k = \sqrt{\lambda}\nu$$

命题得证

由于上面这条结果，下面定义的这个记号就有了意义

7.40 记号：$\sqrt{T}$

对正算子 T， $\sqrt{T}$ 表示 T 的唯一正平方根

7.41 正算子的平方根

在（具有通常的欧几里得内积的） $\mathbf{R}^2$ 上定义算子 s, T 为
那么，关于 $\mathbf{R}^2$ 的标准基，我们有

$$M(S) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \overrightarrow{\#} \quad M(T) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

这两个矩阵都等于自身的转置，因此 s 和 T 都是自伴的
如果 $(x, y) \in \mathbf{R}^2$，那么

$$\langle S(x,y), (x,y)\rangle = x^2 + 2y^2 \geq 0$$

且

$$\langle T(x,y), (x,y)\rangle = x^2 + 2xy + y^2 = (x+y)^2 \geq 0$$

因此 S 和 T 都是正算子
$\mathbf{R}^2$ 的标准基是由 s 的特征向量组成的规范正交基．注意到

$$\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right),\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)$$

是 T 的特征向量所成的规范正交基，其中第一个特征向量对应的特征值为 2，第二个对应的特征值为0．因此 $\sqrt{T}$ 的特征向量与 T 的相同，而特征值为 $\sqrt{2}$ 和 0
通过证明以下这些矩阵的平方就是式 (7.42) 中的矩阵，并且以下每个矩阵都是正算子关于标准基的矩阵，可验证

$$\mathcal{M}\left(\sqrt{S}\right)=\begin{pmatrix}1 & 0\\ 0 & \sqrt{2}\end{pmatrix}\quad \mathcal{M}\left(\sqrt{T}\right)=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}$$

接下来这条结果的表述并不涉及平方根，但是这个简洁的证明很好地利用了正算子的平方根

7.43 定理

设 T 是 V 上的正算子且 $\nu\in V$ 使得 $T\nu,\nu=0$，那么 $T\nu=0$

证

$$0=\langle T\nu,\nu\rangle=\sqrt{T}\sqrt{T}\nu,\nu=\sqrt{T}\nu,\sqrt{T}\nu=\left\|\sqrt{T}\nu\right\|^2$$

故$\sqrt{T}\nu=0$，故$T\nu=\sqrt{T}\left(\sqrt{T}\nu\right)=0$

7D 等距映射、幺正算子和矩阵分解

等距映射

保持范数的线性映射值得有个名字

7.44 等距映射

线性映射 $S\in\mathcal{L}(V,W)$ 称等距映射若$\|S\nu\|=\|\nu\|$对任 $\nu\in V$ 都成立

如果 $S\in\mathcal{L}(V,W)$ 是等距映射且 $\nu\in V$ 使得 $S\nu=0$，那么 $\|\nu\|=\|S\nu\|=\|0\|=0$，这意味着 $\nu=0$．因此每个等距映射都是单射
希腊语单词 isos 意为相等、metron 意为度量，所以 isometry 的字面意思就是度量相等．

7.45 将规范正交基映射到规范正交组 $\Longrightarrow$ 等距映射

设 $e_1,\ldots,e_n$ 是 V 的规范正交基，$g_1,\ldots,g_n$ 是 W 中的规范正交组．令 $S\in\mathcal{L}(V,W)$ 是使得对任一 $k=1,\ldots,n$ 都有 $Se_k=g_k$ 的线性映射．下面证明 s 是等距映射．设 $\nu\in V$．那么有

$$\nu=\nu,e_1e_1+\cdots+\nu,e_ne_n$$

以及

$$\|\nu\|^2=|\langle\nu,e_1\rangle|^2+\cdots+|\langle\nu,e_n\rangle|^2$$

其中我们用到了6.30 (b)．将 s 作用于式(7.46) 的两侧，得到

$$S\nu=\nu,e_1Se_1+\cdots+\nu,e_nSe_n=\nu,e_1g_1+\cdots+\nu,e_ng_n$$

因此

$$\|S\nu\|^2 = |\langle \nu, e_1\rangle|^2 + \cdots + |\langle \nu, e_n\rangle|^2$$

对比式 (7.47) 和式 (7.48)，可知 $\|\nu\| = \|S\nu\|$. 因此， s 是等距映射

下一条结果给出了“某线性映射是等距映射”的等价条件．(a) 和(c) 的等价关系表明，线性映射是等距映射当且仅当它是保持内积的．(a) 和(d) 的等价关系表明，线性映射是等距映射当且仅当它将某个规范正交基映射为规范正交组．因此例7.45 涵盖了所有等距映射．更进一步还有，线性映射是等距映射，当且仅当它将每个规范正交基都映射为规范正交组——因为 (a)的成立与否并不依赖于基 $e_1, \ldots, e_n$ 的选取

接下来这条结果中(a) 和(e) 的等价关系表明，线性映射是等距映射当且仅当它（关于任一规范正交基）的矩阵的列形成一规范正交组．这里我们将 $m \times n$ 矩阵的列等同于 $\mathbf{F}^m$ 的元素，然后应用 $\mathbf{F}^m$ 上的欧几里得内积

7.49 等距映射的性质

设 $S \in \mathcal{L}(V, W)$. 设 $e_1, \ldots, e_n$ 是 V 的规范正交基， $f_1, \ldots, f_m$ 是 W 的规范正交基．那么下列等价：

(a) s 是等距映射

(b) $S^*S = I$

(c) $\langle Su, S\nu\rangle = \langle u, \nu\rangle$ 对所有 $u, \nu \in V$ 成立

(d) $Se_1, \ldots, Se_n$ 是 W 中的规范正交组

(e) $M(S, (e_1, \ldots, e_n), (f_1, \ldots, f_m))$ 的列形成 $\mathbf{F}^m$ 中关于欧几里得内积的规范正交组

证

设(a) 成立，那么 s 是等距映射．如果 $\nu \in V$ ，那么

$$\langle (I - S^*S)\nu, \nu\rangle = \langle \nu, \nu\rangle - \langle S^*S\nu, \nu\rangle = \|\nu\|^2 - \langle S\nu, S\nu\rangle = \|\nu\|^2 - \|S\nu\|^2 = 0$$

于是自伴算子 $I - S^*S$ 等于0（根据7.16）．因此 $S^*S = I$ ，证明了 (a) 蕴涵 (b)

设(b) 成立，那么 $S^*S = I$. 如果 $u, \nu \in V$ ，那么

$$\langle Su, S\nu\rangle = \langle S^*Su, \nu\rangle = \langle Iu, \nu\rangle = \langle u, \nu\rangle$$

证明了 (b) 蕴涵 (c)

设(c) 成立，那么 $\langle Su, S\nu\rangle = \langle u, \nu\rangle$ 对所有 $u, \nu \in V$ 成立．因此，如果 $j, k \in \{1, \ldots, n\}$ ，那么

$$Se_j, Se_k = e_j, e_k$$

于是， $Se_1, \ldots, Se_n$ 是 W 中的规范正交组，证明了(c) 蕴涵(d)

设(d) 成立，那么 $Se_1, \ldots, Se_n$ 是 W 中的规范正交组．令 $A = \mathcal{M}\big(S, (e_1, \ldots, e_n), (f_1, \ldots, f_m)\big)$. 如果 $k, r \in \{1, \ldots, n\}$ ，那么

$$\sum_{j=1}^{m} A_{j,k}\overline{A_{j,r}} = \sum_{j=1}^{m} A_{j,k}f_j, \sum_{j=1}^{m} A_{j,r}f_j = \langle Se_k, Se_r\rangle = \begin{cases} 1 & -k \\ & -k \end{cases}$$

式 (7.50) 的左侧是 A 的第 k 列和第 $\boldsymbol{r}$ 列在 $\mathbf{F}^m$ 中的内积．因此 A 的□列形成了 $\mathbf{F}^m$ 中的规范正交组，证明了 (d) 蕴涵 (e)

现在假设(e) 成立，那么上一段定义的矩阵 A 的列形成 $\mathbf{F}^m$ 中的规范正交组．那么式(7.50) 表明 $Se_1, \ldots, Se_n$ 是W 中的规范正交组．所以，根据例 7.45 ：(：$Se_1, \ldots, Se_n$ 相当于其中的 $g_1, \ldots, g_n$) ，可得 s 是等距映射，证明了(e) 蕴涵(a)

“某线性映射是等距映射”的更多等价条件，见本节习题第1、11 题

幺正算子

7.51 幺正算子

算子 $S \in \mathcal{L}(V)$ 被称为幺正的，如果 S 是可逆等距映射

前面提到每个等距映射都是单射，而有限维向量空间上的每个单射算子都是可逆的（见 3.65），“V 是有限维内积空间”又是本章总成立的一个假设，因此我们可以从上面的定义中移去“可逆”一词，而不改变其含虽然对于有限维内积空间上的算子而言，幺正（unitary）和等距（isometry）是同一个意思；但记得幺正算子将向量空间映射到自身，而等距映射将向量空间映射到另一（可能不同的）向量空间

义．但这里保留“可逆”这个不必要的词，是为了定义的一致性，毕竟读者在学习不一定有限维的内积空间时，遇到的应该还是上面这个定义

7.52 $\mathbf{R}^2$ 的旋转

$\theta \in \mathbf{R}$ 且 s 是 $\mathbf{F}^2$ 上的算子，其关于 $\mathbf{F}^2$ 的标准基的矩阵是

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

该矩阵的这两列形成了 $\mathbf{F}^2$ 中的规范正交组，于是 s 是等距映射【根据 7.49 中 (a) 和 (e) 的等价关系】. 因此， s 是幺正算子
如果 $\mathbf{F}=\mathbf{R}$ ，那么 s 是“绕 $\mathbf{R}^2$ 的原点逆时针旋转 θ 弧度”这一算子．观察到这一点，能让我们用另一种方式来思考 s 为什么是等距映射——因为绕 $\mathbf{R}^2$ 的原点旋转是保持范数的

下一条结果（7.53）列出了“某算子是幺正算子”的几个等价条件．7.49 中所有等价于“某线性映射是等距映射”的条件都应该纳入其中；除此之外，7.53 中出现了另一些条件，这是因为我们将讨论的范围限制在从向量空间到自身的线性映射当中．例如，7.49 表明线性映射$S \in \mathcal{L}(V,W)$ 是等距映射当且仅当 $S^*S=I$ ，而 7.53 表明算子 $S \in \mathcal{L}(V)$ 是幺正算子当且仅当$S^*S=SS^*=I$
7.49 与 7.53 的差异还有：7.49 (d) 提到的是规范正交组，而 7.53 (d) 提到的是规范正交基；
7.49 (e) 提到的是 $\nabla\mathcal{M}(T)$ 的列，而 7.53 (e) 提到的是 $\nabla\mathcal{M}(T)$ 的行；7.49 (e) 中的 $\nabla\mathcal{M}(T)$ 是关于 V
中一个规范正交基和W 中一个规范正交基的，而 7.53 (e) 中的 $\nabla\mathcal{M}(T)$ 是仅关于 V 中一个基的—这个基“身兼二职”

7.53 幺正算子的特性

$S \in \mathcal{L}(V)$ ， $e_1,\dots,e_n$ 是 V 的规范正交基。下列等价：
(a) s 是幺正算子
(b) $S^*S=SS^*=I.$
(c) s 可逆且 $S^{-1}=S^*$
(d) $Se_1,\dots,Se_n$ 是 V 的规范正交基
(e) $M(S,(e_1,\dots,e_n))$ 的行形成 $\mathbf{F}^n$ 中关于欧几里得内积的规范正交基
(f) S^* 是幺正算子

证

设(a) 成立，那么 s 是幺正算子．于是，根据7.49 中(a) 和(b) 的等价关系，有

$$S^*S=I$$

等式两侧同时右乘 S^{-1} 可得 $S^*=S^{-1}$. 因此 $SS^*=SS^{-1}=I$ ，待证等式得证，进而 (a) 蕴涵 (b) 得证
可逆性和逆的定义表明，(b) 蕴涵(c)
设 (c) 成立，那么 s 可逆且 $S^{-1}=S^*$. 因此， $S^*S=I$. 于是，根据 7.49 中 (b) 和(d) 的等价关系，可得 $Se_1,\dots,Se_n$ 是 V 中的规范正交组．这个组的长度又等于 $\dim V$ ，因此$Se_1,\dots,Se_n$ 是 V 的规范正交基，证明了(c) 蕴涵(d)
设(d) 成立，那么 $Se_1,\dots,Se_n$ 是 V 的规范正交基．7.49 中(a) 和(d) 的等价关系表明 s 是幺正算子．因此

$$(S^*)^*S^*=SS^*=I$$

其中最后一个等号成立，是因为我们已经证明了本结果中的 (a) 蕴涵 (b)．上式和 7.49 中 (a)和(b) 的等价关系表明 S^* 是等距映射．因此，$M\left(S^*,(e_1,\dots,e_n)\right)$ 的列形成了 $\mathbf{F}^n$ 的规范正交基【根据 7.49 中 (a) 和 (e) 的等价关系】. $M(S,(e_1,\dots,e_n))$ 的行是 $M\left(S^*,(e_1,\dots\dots,e_n)\right)$ 的列的复共轭．因此， $M(S,(e_1,\dots,e_n))$ 的行形成 $\mathbf{F}^n$ 的规范正交基，证明了(d) 蕴涵(e)
设(e) 成立，那么 $M\left(S^*,(e_1,\dots,e_n)\right)$ 的列形成 $\mathbf{F}^n$ 的规范正交基．7.49 中 (a) 和 (e)之间的等价关系表明， S^* 是等距映射，证明了(e) 蕴涵(f)
设(f) 成立，那么 S^* 是幺正算子．在这一结果中，我们已经证明了的蕴涵关系链表明 (a) 蕴涵 (f)．对 S^* 应用这一结果，可证明 $(S^*)^*$ 是幺正算子，也就证明了(f) 蕴涵(a)
我们证明了 $\text{(a)} \implies \text{(b)} \implies \text{(c)} \implies \text{(d)} \implies \text{(e)} \implies \text{()} \implies$ (a)，从而完成了证明

回顾一下我们所做的 $\mathbf{c}$ 与 $\mathcal{L}(V)$ 之间的类比．在这一类比下，复数 z 对应于算子 $S \in \mathcal{L}(V)$ ，z 对应于 S^* ．实数（ $z=\bar{z}$ ）对应于自伴算子（ $S=S^*$ ），而非负数对应于（没取对名字的）正算子
C 的另一个重要的子集是单位圆（unit circle)，由满足 $|z|=1$ 的复数 z 构成． $|z|=1$ 这一条件等价于 $\bar{z}z=1$ 这一条件．在我们的类比下，这对应于 $S^*S=I$ 这一条件，等价于 S 是幺正算子．于是，这一类比表明，C 中的单位圆对应于幺正算子构成的集合．在接下来的两条结果中，这一类比体现在了幺正算子的特征值上．这一类比的另一例子见习题 15

7.54 幺正算子的特征值绝对值是1

设 λ 是幺正算子的特征值，那么 $|\lambda| = 1$

证

设 $S \in \mathcal{L}(V)$ 是幺正算子且 λ 是 s 的特征值．令 $\nu \in V$ （$\nu \neq 0$）使得 $S\nu = \lambda\nu$．那么有

$$|\lambda| \, \| \nu \| = \|\lambda\nu\| = \|S\nu\| = \| \nu \|.$$

因此，$|\lambda| = 1$，命题得证

接下来这条结果，以复谱定理为主要工具，刻画了有限维复内积空间上的幺正算子

7.55 对复内积空间上的幺正算子的描述

$\mathbf{F} = \mathbf{C}$ 且 $S \in \mathcal{L}(V)$，下列等价：
(a) s 是幺正算子
(b) 存在 V 的一规范正交基由 s 的特征向量组成，其对应的特征值的绝对值都是1

证

设(a) 成立，那么 s 是幺正算子．7.53 中(a) 和(b) 的等价关系表明 S 是正规的．因此，由复谱定理（7.31）可得，存在 V 的一规范正交基由 s 的特征向量组成．而 S 的每个特征值的绝对值都是1（根据 7.54），也就完成了(a) 蕴涵(b) 的证明
设 (b) 成立．令 $e_1, \ldots, e_n$ 是 V 的规范正交基，由 s 的特征向量组成，其对应特征值 $\lambda_1, \ldots, \lambda_n$ 的绝对值都为1．那么 $Se_1, \ldots, Se_n$ 也是 V 的规范正交基，这是因为

$$Se_j, Se_k = \lambda_j e_j, \lambda_k e_k = \lambda_j \overline{\lambda_k}\, e_j, e_k = \begin{cases} 0 \\ 1 \end{cases}$$

对所有 $j, k = 1, \ldots, n$ 成立．因此，由7.53 中(a) 和(d) 的等价关系可得，S 是幺正的，证明了(b) 蕴涵 (a)

QR 分解

本小节将视角由算子转向方阵．这有助于你培养将算子等同于方阵的思维方式（在选取了算子的定义空间的基之后）．你还应该更加熟练地在算子和方阵之间，将种种概念和结果来回转换
当我们从 $n \times n$ 矩阵而非算子出发时，除非另有说明，否则假设相关的算子都在（具有欧几里得内积的） $\mathbf{F}^n$ 上，并且它们的矩阵都是关于 $\mathbf{F}^n$ 的标准基得到的
我们首先做如下定义，将幺正算子的概念转换为幺正矩阵

7.56 幺正矩阵

一 $n \times n$ 矩阵被称为幺正的，如果它的列形成 $\mathbf{F}^n$ 中的规范正交组

在以上定义中，我们可以把 $\mathbf{F}^n$ 中的规范正交组”替换为“ $\mathbf{F}^n$ 的规范正交基”，因为 n 维内积空间中的长度为 n 的每个规范正交组都是规范正交基．如果 $S \in \mathcal{L}(V)$ 且 $e_1, \ldots, e_n$ 和 $f_1, \ldots, f_n$ 是 V 的规范正交基，那么 s 是幺正算子当且仅当 $\mathcal{M}(S, (e_1, \ldots, e_n), (f_1, \ldots, f_n))$ 是幺正矩阵，这可由 7.49 中 (a) 和 (e) 的等价关系得到．还要注意，我们可以将定义中的“列”替换为“行”，这用到了7.53 中(a) 和(e) 的等价关系
接下来这条结果，给出了“某方阵是幺正矩阵”的一些等价条件，其证明留作读者的习题在 (c) 中， ν 表示 和 ν 的矩阵乘积，这里将 $\mathbf{F}^n$ 的元素等同于 $n \times 1$ 矩阵（有时也称为列向量）．下面的 (c) 中的范数是 $\mathbf{F}^n$ 上通常的欧几里得范数，其来自欧几里得内积．在 (d) 中， * 表示矩阵 的共轭转置，对应于其所关联的算子的伴随

7.57 幺正矩阵的特性

是 $n \times n$ 矩阵．那么下列等价：
(a) 是幺正矩阵
(b) 的行形成 $\mathbf{F}^n$ 中的规范正交组
(c) $\|\nu\| = \|\nu\|$ 对任一 $b\nu \in \mathbf{F}^n$ 都成立
(d) $^* = {}^* = I$， I 是对角线上为1、其余处处为0 的 $n \times n$ 矩阵

下面陈述并证明的 QR 分解，是被广泛使用的 QR 算法（此处不作讨论）中的主要工具，该算法用于求出方阵特征值和特征向量的良好近似．在下面的结果中，如果矩阵 A 属于 $\mathbf{F}^{n,n}$，那么矩阵 和 R 也属于 $\mathbf{F}^{n,n}$

7.58 QR 分解

A 是各列线性无关的方阵．那么存在唯一矩阵 和 R，其中 是幺正的，而 R 是上三角的且对角线上仅有正数，使

$$A = R$$

证

令 $\nu_1, \ldots, \nu_n$ 表示 A 的列，并将其视为 $\mathbf{F}^n$ 的元素．对 $\nu_1, \ldots, \nu_n$ 应用格拉姆-施密特过程（6.32），得到 $\mathbf{F}^n$ 的规范正交基 $e_1, \ldots, e_n$，使

$$\operatorname{span}(\nu_1, \ldots, \nu_k) = \operatorname{span}(e_1, \ldots, e_k)$$

对任一 $k = 1, \ldots, n$ 都成立．令 R 是 $n \times n$ 矩阵，定义为

$$R_{j,k} = \nu_k, e_j$$

其中 $R_{j,k}$ 表示 R 的第 j 行第 k 列的元素．如果 $j > k$，那么 $\boldsymbol{e}_j$ 正交于 $\operatorname{span}(e_1, \ldots, e_k)$ 从而正交于 $|\ \nu_k$ 【根据式(7.59)】．换句话说，如果 $j > k$，那么 $\nu_k, e_j = 0$． 因此 R 是上三角矩阵
令 是各列为 $e_1, \ldots, e_n$ 的幺正矩阵．如果 $k \in \{1, \ldots, n\}$，那么 R 的第 k 列等于 各列的线性组合，其系数来自 R 的第 k 列——见 3.51 (a)．于是 R 的第 k 列等于

$$\nu_k, e_1 e_1 + \cdots + \nu_k, e_k e_k$$

也就等于 ν_k 【根据 6.30 (a)】，即 A 的第 k 列．因此， $A = R$，即待证等式
格拉姆-施密特过程的定义式（见6.32）表明，每个 ν_k 都等于 $\boldsymbol{e}_k$ 的正数倍加上 $e_1, \ldots, e_{k-1}$ 的线性组合．因此 $\langle \nu_k, e_k \rangle$ 是正数，于是 R 的对角线上所有元素都是正数，这一点也得证
最后，为证明 和 R 是唯一的，假设我们还有 $A = \ R$，其中 是幺正的，而 R 是上三角的且对角线上仅有正数．令 $q_1, \ldots, q_n$ 表示 的列．像上面那样看待矩阵乘法，可以得到，每个 ν_k 都是 $q_1, \ldots, q_k$ 的线性组合，其系数来自 R 的第 k 列．这意味着 $\operatorname{span}(\nu_1, \ldots, \nu_k) = \operatorname{span}(q_1, \ldots, q_k)$ 且 $\nu_k, q_k > 0$． 而满足这些条件的规范正交组具有唯一性（见6B 节习题10），这表明 $q_k = e_k$ 对任一 $k = 1, \ldots, n$ 都成立．于是， $=$ ，也就意味着 $R = R$，从而完成了唯一性的证明

QR 分解的证明表明，将格拉姆-施密特过程应用于待分解矩阵的列，可以得到其分解式中幺正矩阵的列．接下来这个例子，基于我们刚刚完成的证明，呈现了QR 分解的计算过程

7.60 QR 分解

下面求矩阵

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -4 \\ 0 & 3 & 2 \end{pmatrix}$$

的QR 分解．我们按照 7.58 的证明，令 ν_1, ν_2, ν_3 等于 A 的列：

$$\nu_1 = (1, 0, 0), \quad \nu_2 = (2, 1, 3), \quad \nu_3 = (1, -4, 2)$$

对 ν_1, ν_2, ν_3 应用格拉姆-施密特过程，得规范正交组

$$e_1 = (1, 0, 0), \quad e_2 = \left(0, \frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}}\right), \quad e_3 = \left(0, -\frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right)$$

顺着 7.58 的证明，令 是以 e_1, e_2, e_3 为列的幺正矩阵：

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ 0 & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix}$$

再按 7.58 的证明，令 R 是 3×3 矩阵，«第 j 行第 k 列的¬元素为 ν_k, e_j ，得到

$$R = \begin{pmatrix} 1 & 2 & 1 \\ 0 & \sqrt{10} & \frac{\sqrt{10}}{5} \\ 0 & 0 & \frac{7\sqrt{10}}{5} \end{pmatrix}$$

注意到 R 确实是对角线上仅有正数的上«三角矩阵，这正是QR 分解所保证的
现在，做矩阵乘法即可验证 $A = R$ 就是待求的分解

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ 0 & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & \sqrt{10} & \frac{\sqrt{10}}{5} \\ 0 & 0 & \frac{7\sqrt{10}}{5} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -4 \\ 0 & 3 & 2 \end{pmatrix} = A$$

因此， $A = R$ ，正如我们预期的

在下一小节科列斯基分解（7.63）的证明中，QR 分解将是主要工具．QR 分解的另一个很好的应用，见阿达马不等式（9.66）的证明
若一矩阵有QR 分解，那么求解该矩阵对应的线性方程组就可以用这一分解式而不用高斯消元法．具体来说，设 A 是各列线性无关的 $n \times n$ 方阵， $b \in \mathbf{F}^n$ ，我们想要求解方程 $Ax = b$ ，其中 $x = (x_1, \ldots, x_n) \in \mathbf{F}^n$ （我们还是照常把 $\mathbf{F}^n$ 的元素等同于 $n \times 1$ 列向量）
设 $A = R$ ，其中 是幺正的，而 R 是上三角的且对角线上仅有正数（ 和 R 可以用格拉姆-施密特过程由 A 得到，如7.58 的证明所示）．方程 $Ax = b$ 等价于 $Rx = b$ ．等式两边同时左乘 * ，再用上 7.57 (d)，就可得到方程$Rx = {}^*b$
因为矩阵 * 是 的共轭转置，所以计算 $^*b$ 是很简单的．因为 R 是对角线上为正数的上三角矩阵，所以以上方程所表示的线性方程组，求解起来很快——先求 x_n ，然后求 x_{n-1} ，依次类推

科列斯基分解

我们以从内积角度对可逆正算子的刻画作为本小节的开始

7.61 可逆正算子

自伴算子 $T \in \mathcal{L}(V)$ 是可逆正算子当且仅当 $\langle T\nu, \nu \rangle > 0$ 对任意非零 $\nu \in V$ 都成立

证

设 T 是可逆正算子，若 $\nu \in V$ 且 $\nu \neq 0$ ，由 T 可逆， $T\nu \neq 0$ ，故$\langle T\nu, \nu \rangle \neq 0$ （由 7.43），故 $\langle T\nu, \nu \rangle > 0$
设 $\langle T\nu, \nu \rangle > 0$ 对任意非零 $\nu \in V$ 都成立，则$T\nu \neq 0$ 对任意非零 $\nu \in V$ 都成立，故 T 单射，故 T 可逆

接下来的定义将以上结果转换为矩阵语言．这里我们用的是 $\mathbf{F}^n$ 上通常的欧几里得内积，并且将 $\mathbf{F}^n$ 中的元素等同于 $n \times 1$ 列向量

7.62 正定

矩阵 $B \in \mathbf{F}^{n,n}$ 称为正定的，如果 $B^* = B$ 且

$$\langle Bx, x \rangle > 0$$

对任意非零 $x \in \mathbf{F}^n$ 都成立

矩阵是上三角的当且仅当它的共轭转置是下三角的（意即对角线上方所有元素为 0）．下面的这个分解，将正定矩阵写成下三角矩阵和它的共轭转置的乘积，它在数值线性代数4 中具有重要意义
接下来这条结果完全是关于矩阵的，不过它的证明也利用了算子和矩阵两者结论之间的等同关系．在下面这条结果中，如果矩阵 B 属于 $\mathbf{F}^{n,n}$ ，那么矩阵 R 也属于 $\mathbf{F}^{n,n}$

7.63 科列斯基分解

B是正定矩阵．那么存在唯一一对角线上仅含正数的上三角矩阵 R 使

$$B = R^*R$$

证

因为 B 是正定的，所以存在与 B 大小相同的可逆矩阵 A 使得 $B = A^*A$ 【根据 7.38 中 (a)和(f) 之间的等价关系】
令 $A = R$ 是 A 的QR 分解（见7.58），其中 是幺正的，而 R 是上三角的且对角线上仅有正数．那么， $A^* = R^{**}$
故

$$B = A^*A = R^{**}R = R^*R$$

即待证等式
为证明该结果的唯一性部分，设 S 是对角线上仅含正数的上三角矩阵，且 $B = S^*S$ ．因为矩阵 B 可逆，所以矩阵 S 可逆（见 3D 节的习题 11）．
在等式 $B = S^*S$ 两侧右乘 S^{-1} 可得$BS^{-1} = S^*$
令 A 是该证明第一段的那个矩阵 A ，那么有

$$(AS^{-1})^*(AS^{-1}) = (S^*)^{-1}A^*AS^{-1}$$

$$= (S^*)^{-1}BS^{-1}$$

$$= (S^*)^{-1}S^*$$

$$= I$$

故 AS^{-1} 是幺正的
故 $A = (AS^{-1})S$ 这一分解，将 A 化为一个幺正矩阵和对角线上仅含正数的一个上三角矩阵的乘积．而正如7.58 中所述，QR 分解具有唯一性，故 $S = R$

以上证明的第一段中，我们本可以将 A 选定为 B 的平方根当中唯一的正定矩阵（见7.39）然而这一证明选择了更一般的 A ，这是因为对于特定的正定矩阵 B ，求出不那么特殊的 A 可能更容易

7E 奇异值分解

奇异值

本节我们将用到以下这条结果

7.64 $T*T$ 的性质

$T \in \mathcal{L}(V, W)$ ，则
(a) T^*T 是 V 上的正算子；
(b) null $T^*T = \text{null}T$ ；
(c) range $T^*T = \text{range}T^*$ ；
(d) dim range $T =$ dim range $T^* =$ dim range T^*T

证

(a) 由

$$(T^*T)^* = T^*(T^*)^* = T^*T,$$

故 T^*T 是自伴的
若 $\nu \in V$

$$(T^*T)\nu, \nu = T^*(T\nu), \nu = T\nu, T\nu = T\nu^2 \geq 0$$

故 T^*T 是正算子
(b) 设 $\nu \in \operatorname{null}T^*T$ ，则

$$\|T\nu\|^2 = T\nu, T\nu = T^*T\nu, \nu = 0, \nu = 0$$

故 $T\nu = 0$ ，证明了 $\operatorname{null}T^*T \subseteq \operatorname{null}T$
另一方向的包含关系是显然的，因为如果 $\nu \in V$ 且 $T\nu = 0$ 就有 $T^*T\nu = 0$ ，故 $\operatorname{null}T^*T = \operatorname{null}T$
(c) 由(a) T^*T 是自伴的，则

其中第一个和最后一个相等关系来自7.6，而第二个相等关系来自(b)
(d) 注意到

$$T = \dim(\operatorname{null}T^*)^{\perp} = \dim W - \dim \operatorname{null}T^* = \dim \operatorname{range}T^*$$

其中第一个相等关系来自 7.6 (d)，第二个来自 6.51，最后一个来自线性映射基本定理（3.21），从而验证了(d) 中的第一个等式
等式 dim range $T^* = \operatorname{dimrange}T^*T$ 由 (c) 可得

算子的特征值可以反映算子的一些性质．还有一种被称为“奇异值”的数，也是很有用的下面提到的特征空间和记号 E （在示例中会用到）已在5.52 中定义

7.65 奇异值

$T \in \mathcal{L}(V, W)$ ． T 的奇异值是 T^*T 的特征值的非负平方根，按降序排列，而且每个奇异值的出现次数，等于 T^*T 对应特征空间的维数

7.66 $\mathbf{F}^4$ 上一算子的奇异值

$T \in \mathcal{L}(\mathbf{F}^4)$ ， $T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4)$，

$$T^*T(z_1, z_2, z_3, z_4) = (9z_1, 4z_2, 0, 9z_4)$$

用 $\mathbf{F}^4$ 的标准基对角化 T^*T ，可得 T^*T 的特征值是 9、4、0．以及，对应于这些特征值的特征空间维数是

$$\dim E(9, T^*T) = 2 \qquad \dim E(4, T^*T) = 1 \qquad \dim E(0, T^*T) = 1$$

取 T^*T 的这些特征值的非负平方根，并用到以上的维数信息，我们可以得出结论： T 的奇异值是3、3、2、0
T 仅有的特征值是 -3 和0．所以在这个例子中，特征值并没有包含出现在 T 的定义中的2，而奇异值中却包含了2

7.67 从 $\mathbf{F}^4$ 到 $\mathbf{F}^3$ 的一线性映射的奇异值

$T \in \mathcal{L}(\mathbf{F}^4, \mathbf{F}^3)$ （关于标准基）的矩阵是

$$\begin{pmatrix} 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

T^*T 的矩阵是

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 25 \end{pmatrix}$$

特征值是 25、2、0；而 $\dim E(25, T^*T) = 1$ ， $\lim E(2, T^*T) = 1$ ， $\dim E(0, T^*T) = 2$ ，故 T 的奇异值是5、 $\sqrt{2}$, 0, 0

本节习题2 给出了正奇异值的一种刻画

7.68 正奇异值的作用

$T \in \mathcal{L}(V, W)$ ，则
(a) T 是单射 $\Longleftrightarrow 0$ 不是 T 的奇异值

(b) T 的正奇异值个数等于 dim range T
(c) T 是满射 $\iff$ T 的正奇异值个数等于 $\dim W$

证

线性映射 T 是单射当且仅当 ull$T=\{0\}$，当且仅当 null$T^*T=\{0\}$ 【根据 7.64 (b)】，当且仅当0 不是 T^*T 的特征值，当且仅当0 不是 T 的奇异值，这就完成了(a) 的证明
将谱定理应用于 T^*T 可得 dim range T^*T 等于 T^*T 的正特征值个数（考虑重复）．那么，由 7.64 (d) 可得 dim range T 等于 T 的正奇异值个数，证明了(b)
利用 (b) 和 2.39 可证明 (c) 成立

以下表格对比了特征值和奇异值

由特征值构成的组	由奇异值构成的组
背景：向量空间	背景：内积空间
仅对从向量空间到自身的线性映射有定义	对于从内积空间到可能不同的内积空间的线性映 射都有定义
组中元素可以是任意实数（若F=R）或复数（若 F=C）	组中元素必为非负数
如果F=R，可以为空组	组的长度等于定义空间维数
包含0算子不可逆	包含0线性映射不是单射
没有标准顺序，特别是当F=C时	总是降序排列

证

S 是等距映射 $\iff$ $S^*S=I$ \$ \ s ^ { } s\$ *的所有特征值都等于1*$\iff$ s *的所有奇异值都等于1*
其中第一个等价关系来自7.49，第二个由谱定理（7.29 或7.31）应用于自伴算子 \$S ^ { } S\$ 可得

线性映射和矩阵的奇异值分解

接下来的结果表明，对于从 V 到W 的每个线性映射，都可以用它的奇异值以及 V 和W 中的规范正交组，给出非常干净利落的描述．在下一节，我们将看到奇异值分解（也经常称作SVD）的一些重要应用
奇异值分解在数值线性代数中很有用，因为有很好的工具可用于逼近正算子的特征值和特征向量，而正算子 T^*T 的特征值和特征向量又能引出奇异值分解

7.70 奇异值分解

设 $T\in\mathcal{L}(V,W)$ 且 T 的正奇异值是 $s_1,\ldots,s_m$．那么存在 V 中的规范正交组 $e_1,\ldots,e_m$ 和W 中的规范正交组 $f_1,\ldots,f_m$ 使

$$T\nu = s_1\nu, e_1f_1+\cdots+s_m\nu, e_mf_m$$

对任一 $\nu\in V$ 都成立

证

令 $s_1,\ldots,s_n$ 表示 T 的奇异值（则 $n=\dim V$）．因为 T^*T 是正算子【见 7.64 (a)】，由谱定理可得，存在 V 中的规范正交基 $e_1,\ldots,e_n$ 使得

$$T^*Te_k=s_k^2e_k$$

对任一 $k=1,\ldots,n$ 都成立
对于任一 $k=1,\ldots,m$，令

$$f_k = \frac{Te_k}{s_k}.$$

如果 $j, k \in \{1, \dots, m\}$，那么

$$f_j, f_k = \frac{1}{s_j s_k} Te_j, Te_k = \frac{1}{s_j s_k} e_j, T^*Te_k = \frac{s_k}{s_j} e_j, e_k = \begin{cases} 0 \\ 1 \end{cases}$$

那么，$f_1, \dots, f_m$ 是 W 中的规范正交组
如果 $k \in \{1, \dots, n\}$ 且 $k > m$，那么 $s_k = 0$，从而 $T^*Te_k = 0$（根据 7.72），也就意味着$Te_k = 0$【根据 7.64 (b)】
设 $\nu \in V$，那么

$$\begin{aligned} T\nu &= T\big(\langle \nu, e_1 \rangle e_1 + \cdots + \langle \nu, e_n \rangle e_n\big) \\ &= \langle \nu, e_1 \rangle Te_1 + \cdots + \langle \nu, e_m \rangle Te_m \\ &= s_1 \langle \nu, e_1 \rangle f_1 + \cdots + s_m \langle \nu, e_m \rangle f_m. \end{aligned}$$

其中，最后一个下标从第一行的 n 变为第二行的 m 是因为如果 $k > m$ 就有 $Te_k = 0$（见上一段），而第三行是根据式(7.73) 得到的．上式就是我们预期的结果

设 $T \in \mathcal{L}(V, W)$，T 的正奇异值是 $s_1, \dots, s_m$，而 $e_1, \dots, e_m$ 和 $f_1, \dots, f_m$ 是奇异值分解（7.70）中的那两组向量．规范正交组 $e_1, \dots, e_m$ 可以扩充为 V 的规范正交基 $e_1, \dots, e_{\dim V}$，规范正交组 $f_1, \dots, f_m$ 可以扩充为 W 的规范正交基 $f_1, \dots, f_{\dim W}$．公式 (7.71) 表明

$$Te_k = \begin{cases} s_k f_k & 1 \le k \le m, \\ 0 & m < k \le \dim V. \end{cases}$$

那么 T 关于规范正交基 $(e_1, \dots, e_{\dim V})$ 和 $(f_1, \dots, f_{\dim W})$ 的矩阵具有如下简单形式：

$$\boldsymbol{\mathcal{M}}\big(T, (e_1, \dots, e_{\dim V}), (f_1, \dots, f_{\dim W})\big)_{j,\boldsymbol{k}} = \begin{cases} k & \frac{\text{ddag } s}{\sqrt{\pi}} \\ & \frac{\text{ddag } s}{\sqrt{\pi}/\sqrt{\pi}} \end{cases}$$

如果 $\dim V = \dim W$（比如，如果 $W = V$ 就是这样），那么上一段描述的矩阵就是对角矩阵．如果我们像下面这样推广对角矩阵的定义，使其适用于那些不一定是方阵的矩阵，那么我们就证明了一个极好的结果——每个从 V 到W 的线性映射关于恰当的规范正交基都有对角矩阵

7.74 对角矩阵

$M \times N$ 矩阵 A 被称为对角矩阵，如果除了 $A_{k,k}$（$k = 1, \dots, \min\{M, N\}$）可能不为0以外，所有元素都为0

以下表格对比了谱定理（7.29 和7.31）和奇异值分解（7.70）

谱定理	奇异值分解
只描述自伴算子（F=R的情形下）或正规算子（F=C的情形下）	描述任意从内积空间到可能不同的内积空间的线 性映射
得到一个规范正交基	得到两个规范正交组，一个是定义空间的，一个是 值域的，而且即使在值域等于定义空间的情况下也 不一定是同一个
取决于是F=R还是F=C而有不同的证明	不管是F=R还是F=C，证明都是一样的

7.75 伴随和伪逆的奇异值分解

设 $T \in \mathcal{L}(V, W)$ 且 T 的正奇异值是 $s_1, \dots, s_m$．设 $e_1, \dots, e_m$ 和 $f_1, \dots, f_m$ 是 V 和 W 中的规范正交组，使得对任一 $\nu \in V$ 都有

$$T\nu = s_1 \, \nu, e_1 f_1 + \cdots + s_m \, \nu, e_m f_m.$$

那么，对任一 $w \in W$ 都有

$$T^*w = s_1 \, w, f_1 e_1 + \cdots + s_m \, w, f_m e_m$$

和

$$T^{\dagger}w=\frac{\langle w,f_1\rangle}{s_1}e_1+\cdots+\frac{\langle w,f_m\rangle}{s_m}e_m.$$

证

如果 $\nu\in V$ 且 $w\in W$ ，那么

$$\begin{aligned}T\nu,w&=s_1\,\nu,e_1f_1+\cdots+s_m\,\nu,e_mf_m,w\\&=s_1\,\nu,e_1\,f_1,w+\cdots+s_m\,\nu,e_m\,f_m,w\\&=\nu,s_1\,w,f_1e_1+\cdots+s_m\,w,f_me_m.\end{aligned}$$

这意味着

$$T^*w=s_1\,w,f_1e_1+\cdots+s_m\,w,f_me_m,$$

证明了式 (7.77)
为证式 (7.78)，设 $w\in W$. 令

$$\nu=\frac{\langle w,f_1\rangle}{s_1}e_1+\cdots+\frac{\langle w,f_m\rangle}{s_m}e_m.$$

将 T 作用于上式两侧，得到

$$\begin{aligned}T\nu&=\tfrac{w,f_1}{s_1}Te_1+\cdots+\tfrac{w,f_m}{s_m}Te_m\\&=w,f_1f_1+\cdots+w,f_mf_m\\&=P_{\mathrm{range}T}w.\end{aligned}$$

其中，第二行成立是因为式 (7.76) 蕴涵了若 $k=1,\ldots,m$ 则 $Te_k=s_kf_k$ ， 最后一行成立是因为式 (7.76) 蕴涵了 $f_1,\ldots,f_m$ 张成 range T 从而是 range T 的规范正交基【这样就可以用到6.57 (i)了】. 有了上式，并且观察到 $\nu\in(\mathrm{null}T)^{\perp}$ 【见本节习题8 (b)】，再结合 $T^{\dagger}w$ 的定义（见6.68），可得 $\nu=T^{\dagger}w$ ， 证明了式 (7.78)

7.79 求奇异值分解

$T\in\mathcal{L}(\mathbf{F}^4,\mathbf{F}^3)$ ， $T(x_1,x_2,x_3,x_4)=(-5x_4,0,x_1+x_2)$
想求 T 的奇异值分解. T （关于标准基）的矩阵是

$$\begin{pmatrix}0&0&0&-5\\0&0&0&0\\1&1&0&0\end{pmatrix}$$

T^*T 的矩阵是

$$\begin{pmatrix}1&1&0&0\\1&1&0&0\\0&0&0&0\\0&0&0&25\end{pmatrix}$$

正特征值是 25 和 2，而 $\dim E(25,T^*T)=1$ ， $\dim E(2,T^*T)=1$
故T 的正奇异值是 5 和$\sqrt{2}$
为求出 T 的奇异值分解，我们必须求出 $\mathbf{F}^4$ 中的一规范正交组 $\boldsymbol{e}_1,\boldsymbol{e}_2$ 和 $\mathbf{F}^3$ 中的一规范正交组 f_1,f_2 使得

$$T\nu=5\,\nu,e_1f_1+\sqrt{2}\,\nu,e_2f_2$$

对所有 $\nu\in\mathbf{F}^4$ 成立
$E(25,T^*T)$ 的一规范正交基是向量 $(0,0,0,1)$ ， $E(2,T^*T)$ 的一规范正交基是向量$\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0,0\right)$. 所以，根据7.70 的证明，我们可以取

$$\boldsymbol{e}_1=(0,0,0,1)$$

$$e_2=\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0,0\right)$$

以及

$$f_1=\frac{Te_1}{5}=(-1,0,0)\qquad f_2=\frac{Te_2}{\sqrt{2}}=(0,0,1)$$

不出所料，可以看到 e_1, e_2 就是 $\mathbf{F}^4$ 中的规范正交组， f_1, f_2 就是 $\mathbf{F}^3$ 中的规范正交组，且

$$T\nu = 5\,\nu, e_1 f_1 + \sqrt{2}\,\nu, e_2 f_2$$

对所有 $\nu \in \mathbf{F}^4$ 成立．那么，我们就求出了 T 的奇异值分解

接下来的结果，将奇异值分解从线性映射背景下迁移到矩阵背景下．具体而言，它给出了对任意矩阵的分解，将其化为三个性质良好的矩阵的乘积．这里的证明过程利用线性映射奇异值分解的语言给出了这三个矩阵的具体构造方法

接下来这条结果中，“列规范正交”一词应该解释为，各列关于标准欧几里得内积是规范正交的

7.80 奇异值分解（SVD）的矩阵版本

设 A 是 $p \times n$ 矩阵（秩 $m \geq 1$ ）．那么，存在列规范正交的 $p \times m$ 矩阵 B ，对角线上为正数的 $m \times m$ 对角矩阵 D 以及列规范正交的 $n \times m$ 矩阵 c 使

$$A = BDC^*$$

证

令线性映射 $T : \mathbf{F}^n \mathbf{F}^p$ 关于标准基的矩阵等于 A ．那么，dim range $T = m$ （根据 3.78）令

$$T\nu = s_1\,\nu, e_1 f_1 + \cdots + s_m\,\nu, e_m f_m$$

是 T 的奇异值分解．令

令 $u_1, \ldots, u_m$ 表示 $\mathbf{F}^m$ 的标准基．如果 $k \in \{1, \ldots\ldots, m\}$ ，那么

$$(AC - BD)u_k = Ae_k - B(s_k u_k) = s_k f_k - s_k f_k = 0.$$

故 $AC = BD$

两侧右乘 C^* （ C 的共轭转置），得到

$$ACC^* = BDC^*$$

注意到 C^* 的行是 $e_1, \ldots, e_m$ 的复共轭．因此，如果 $k \in \{1, \ldots, m\}$ ，那么由矩阵乘法的定义可得 $C^* e_k = u_k$ ，于是 $CC^* e_k = e_k$ ．因此，$ACC^*\nu = A\nu$ 对所有 $\nu \in \operatorname{span}(e_1, \ldots, e_m)$ 成立

如果 $\nu \in \left(\operatorname{span}(e_1, \ldots, e_m)\right)^\perp$ ，那么（由 7.81 可得） $A\nu = 0$ ，（由矩阵乘法的定义可得）$C^*\nu = 0$ ．于是， $ACC^*\nu = A\nu$ 对所有 $\nu \in \left(\operatorname{span}(e_1, \ldots, e_m)\right)^\perp$ 成立

因为 ACC^* 和 A 在 $\operatorname{span}(e_1, \ldots, e_m)$ 和 $(\operatorname{span}(e_1, \ldots, e_m))^\perp$ 上都是一致的，我们可以得出结论 $ACC^* = A$ ．因此前面那个行间公式就变成了

$$A = BDC^*$$

注意以上结果中的矩阵 A 有 pn 个元素．相比之下，以上的矩阵 B, D, C 总共有

$$m(p + m + n)$$

个元素．因此，如果 p 和 n 是很大的数而秩 ∇_m 比 p 和 n 小得多，那么为了表示 A 而必须存储在电脑中的元素个数就比 pn 小得多

7F 奇异值分解的推论

线性映射的范数

奇异值分解引出了接下来的 $\|T\nu\|$ 的上界

7.82 $T\nu$ 的上界

设 $T \in \mathcal{L}(V, W)$ ．令 s_1 是 T 的最大奇异值．那么

$$\|T\nu\| \leq s_1\,\|\nu\|$$

对所有 $\nu \in V$ 成立

证

令 $s1, \ldots, sm$ 表示 T 的正奇异值，令 $//T\nu//$ 的一个下界，见 7E 节的习题 14$e_1, \ldots, e_m$ 是 V 中的规范正交组， $f_1, \ldots, f_m$ 是 W 中的规范正交组，且由此可给出 T 的奇异值分解．那么

$$T\nu = s_1 \nu, e_1 f_1 + \cdots + s_m \nu, e_m f_m$$

对所有 $\nu \in V$ 成立．于是，若 $\nu \in V$ 则有

$$\begin{aligned}\|T\nu\|^2 &= s_1^2|\langle\nu, e_1\rangle|^2 + \cdots + s_m^2|\langle\nu, e_m\rangle|^2 \\ &\le s_1^2\left(|\langle\nu, e_1\rangle|^2 + \cdots + |\langle\nu, e_m\rangle|^2\right) \\ &\le s_1^2 \, \| \nu \| \, ^2.\end{aligned}$$

其中最后一个不等关系由贝塞尔不等式（6.26）得到．不等式两侧取平方根，可得 $\|T\nu\| \le s_1 \, \| \nu \|$

设 $T \in \mathcal{L}(V, W)$ 且 s_1 是 T 的最大奇异值．以上结果表明

在式 (7.83) 中，取 $\nu = e_1$ ，可得 $Te_1 = s_1 f_1$ ．因为 $\|f_1\| = 1$ ，所以这意味着 $\|Te_1\| = s_1$ ．那么，由于 $\|e_1\| = 1$ ，7.84 中的不等式可以引出等式

$$\max\{\|T\nu\| : \nu \in V \ \ \|\nu\| \le 1\} = s_1.$$

上式促使我们直接将 T 的范数定义为上式的左侧，而不需要提到（右侧的）奇异值或者奇异值分解，如下所示

7.86 线性映射的范数

设 $T \in \mathcal{L}(V, W)$ ．那么， T 的范数，记为 $\|T\|$ ，定义为

$$\|T\| = \max\{\|T\nu\| : \nu \in V\}$$

一般地，由非负数构成的无限集不一定存在最大值．然而，7.86 前的讨论表明，从 V 到W的线性映射 T 的范数定义中的最大值确实存在（且等于 T 的最大奇异值）
关于“范数”一词和 $//\cdot//$ 这一记号，我们现在有两种不同用法：第一种用法跟 V 上的内积相关，我们定义的是 $\|\nu\| = \sqrt{\nu, \nu}$ 对任一 $\nu \in V$ 都成立；第二种用法，则是用我们刚刚为
$T \in \mathcal{L}(V, W)$ 所作的 $\|T\|$ 的定义． $T \in \mathcal{L}(V, W)$ 的范数 $\|T\|$ 通常不是通过对 T 和其本身取内积得到的（见习题21）．从上下文和所用的符号中，你应该能够判断出其中的范数是何种含义
$\mathcal{L}(V, W)$ 上的范数性质列出如下，和内积空间上的范数性质（见 6.9 和 6.17）看起来是一样的．(d) 中的不等式称为三角不等式，这就和我们对 V 上的范数用的术语是一样的．至于反向三角不等式，见本节习题1

7.87 线性映射范数的基本性质

设 $T \in \mathcal{L}(V, W)$ ，那么
(a) $\|\boldsymbol{T}\| \ge 0$;
(b) $\|T\| = 0 \iff T = 0$;
(c) $\|\lambda T\| = |\lambda|\|T\|$ 对所有 $\boldsymbol{\lambda} \in \mathbf{F}$ 成立；
(d) $\|S + T\| \le \|S\| + \|T\|$ 对所有 $S \in \mathcal{L}(V, W)$ 成立

证

(a) 因为 $\|T\nu\| \ge 0$ 对任一 $\nu \in V$ 都成立，所以由 T 的定义可得 $\|\boldsymbol{T}\| \ge 0$
(b) 设 $\|T\| = 0$ ．那么 $T\nu = 0$ 对所有满足 $\| \nu \| \le 1$ 的 $\nu \in V$ 成立．如果 $u \in V$ （ $u \ne 0$ ），那么

$$Tu = \| u \| T\left(\frac{u}{\| u \|}\right) = 0$$

其中最后一个相等关系成立是因为 $u/\|u\|$ 范数为 1．因为 $Tu = 0$ 对所有 $u \in V$ 成立，所以我们有 $T = 0$

反之如果 $T=0$，那么 $T\nu=0$ 对所有 $\nu\in V$ 成立，于是 $\|T\|=0$ ·
(c) 设 $\boldsymbol{\lambda}\in\mathbf{F}$，那么

$$\begin{aligned}\|\lambda T\| &= \max\{\|\lambda T\nu\|:\nu\in V\}\|\nu\|\le 1\\ &= |\lambda|\max\{\|T\nu\|:\nu\in V\}\|\nu\|\le 1\\ &= |\lambda|\|T\|\end{aligned}$$

(d) 设 $S\in\mathcal{L}(V,W)$．由 $\|S+T\|$ 的定义可知，存在 $\nu\in V$ 使得 $\|\nu\|\le 1$ 且 $\|S+T\|=\|(S+T)\nu\|$ 这样一来，

$$\|S+T\|=\|(S+T)\nu\|=\|S\nu+T\nu\|\le\|S\nu\|+\|T\nu\|\le\|S\|+\|T\|$$

完成了(d) 的证明

对于 $S,T\in\mathcal{L}(V,W),\|S-T\|$ 这个量常常被称为 s 和 T 之间的距离．不正式地说， $\|S-T\|$ 这个数比较小就意味着 s 和 T 比较接近．例如本节习题 9 中指出，对于任一 $T\in\mathcal{L}(V)$，总可以找到与 T 任意接近的可逆算子

7.88 T 的多种表达式

设 $T\in\mathcal{L}(V,W)$，则
(a) $\|T\|=T$ 的最大奇异值
(b) $\|T\|=\max\{\|T\nu\|:\nu\in V$ 且 $\|\nu\|=1\}$
(c) $\|T\|=$ 使得 $\|T\nu\|\le c\|\nu\|$ 对所有 $\nu\in V$ 成立的最小数 $\boldsymbol{c}$

证

(a) 见 7.85
(b) 令 $\nu\in V$， $0<\|\nu\|\le 1$．令 $u=\nu/\|\nu\|$，则

$$\|u\|=\left\|\frac{\nu}{\|\nu\|}\right\|=1$$

$$Tu=T\left(\frac{\nu}{\nu}\right)=\frac{T\nu}{\nu}\ge T\nu$$

故求 $\|T\nu\|$ 在 $\|\nu\|\le 1$ 条件下的最大值时可只关注 V 中范数为1 的那些向量
(c) 设 $\nu\in V$ 且 $\nu\ne 0$．那么由 $\|T\|$ 的定义

$$\left\|T\left(\frac{\nu}{\|\nu\|}\right)\right\|\le\|T\|$$

故

$$\|T\nu\|\le\|T\|\ \|\nu\|$$

设 $c\ge 0$ 且 $\|T\nu\|\le c\,\|\nu\|$ 对所有 $\nu\in V$ 成立，故

$$||T\nu||\le c$$

对所有 $\|\nu\|\le 1$ 的 $\nu\in V$ 成立．遍历所有满足 $\|\nu\|\le 1$ 的 $\nu\in V$，取得的以上不等式左侧最大值 $||T||\le c$．因此， $\|T\|$ 是使得 $\|T\nu\|\le c\|\nu\|$ 对所有 $\nu\in V$ 成立的最小数 c

在处理线性映射的范数时，你可能会经常用到不等式(7.89)
给定一线性映射 T 关于某个规范正交基的矩阵，要计算 T 的范数近似值，7.88 (a) 可能是最有用的．用矩阵乘法可以很快算出 T^*T 的矩阵，用计算机可以求出 T^*T 的最大特征值的近似值（为此已有极好的数值算法)，然后开平方根，再用 7.88 (a) 便可得到 T 的范数的近似值（这通常无法确切计算）
你应该自行验证下面例子中的所有结论

7.90 范数

如果 I 表示 V 上通常的恒等算子，那么 $\|I\|=1$ 如果 $T\in\mathcal{L}(\mathbf{F}^n)$ 且 T 关于 $\mathbf{F}^n$ 的标准基的矩阵各元素全为1，那么 $\|T\|=n$.如果 $T\in\mathcal{L}(V)$ 且 V 有一规范正交基由 T 的对应于特征值 $\lambda_1,\ldots,\lambda_n$ 的特征向量组成，那么$\|T\|$ 是 $|\lambda_1|,\ldots,|\lambda_n|$ 中的最大值设算子 $T\in\mathcal{L}(\mathbf{R}^5)$ （关于标准基）的 5×5

矩阵第 j 行第 k 列的元素是 $1/(j^2+k)$．通用的数学软件显示， T 的最大奇异值约等于0.8，而最小的约等于 10^{-6}．因此， $\|\boldsymbol{T}\|\ 0.8$ 以及（用7E节的习题10 所述结论可得） $\|T^{-1}\|\ 10^6$．而求出这些范数的确切表达式是不可能的接下来这条结论表明，线性映射和它的伴随范数相同

7.91 伴随的范数

设 $T\in\mathcal{L}(V,W)$，那么 $T^*=T$

证

设 $w\in W$，那么

$$\|T^*w\|^2=\langle T^*w,T^*w\rangle=\langle TT^*w,w\rangle\leq\|TT^*w\|\|w\|\leq\|T\|\|T^*w\|\|w\|.$$

以上不等式意味着

$$\|T^*w\|\leq\|T\|\|w\|,$$

结合 7.88 (c) 可得 $\|T^*\|\leq\|T\|$
将 $\|T^*\|\leq\|T\|$ 中的 T 替换成 T^*，再用上 $(T^*)^*=T$，可得 $\|\boldsymbol{T}\|\leq\|\boldsymbol{T}^*\|$．因此， $\|T^*\|=\|T\|$

你可能还想用 7E 节的习题 9 构造以上结果的另一证明．那道习题指出，线性映射和它的伴随的正奇异值相同

用具有较低维值域的线性映射进行逼近

接下来的结果是奇异值分解的精彩应用．它说的是，削去奇异值分解中前 k 项之后的项，从而用值域维数至多为 k 的线性映射，得到线性映射的最佳逼近．具体而言，接下来这条结果中，线性映射 T_k 满足这样的性质——dim range $T_k=k$，并且在值域维数至多为 k 的所有线性映射中， T_k 到 T 的距离最小．这条结果引出了用以压缩巨大矩阵并同时保留其重要信息的算法

7.92 用值域维数至多为 k 的线性映射得到的最佳逼近

$T\in\mathcal{L}(V,W)$ 且 $s_1\geq\cdots\geq s_m$ 是 T 的正奇异值．设 $1\leq k<m$，那么有
进一步，如果

$$T\nu=s_1\,\nu,e_1f_1+\cdots+s_m\,\nu,e_mf_m$$

是 T 的奇异值分解，而 $T_k\in\mathcal{L}(V,W)$ 定义为

$$T_k\nu=s_1\,\nu,e_1f_1+\cdots+s_k\,\nu,e_kf_k$$

对任一 $\nu\in V$ 都成立，那么 dim range $T_k=k$ 且 $\|\boldsymbol{T}-\boldsymbol{T}_k\|=s_{k+1}$

证

若 $\nu\in V$，那么

$$\begin{aligned}\|(T-T_k)\nu\|^2&=\|s_{k+1}\,\nu,e_{k+1}f_{k+1}+\cdots+s_m\,\nu,e_mf_m\|^2\\&=s_{k+1}^2|\langle\nu,e_{k+1}\rangle|^2+\cdots+s_m^2|\langle\nu,e_m\rangle|^2\\&\leq s_{k+1}^2\left(|\langle\nu,e_{k+1}\rangle|^2+\cdots+|\langle\nu,e_m\rangle|^2\right)\\&\leq s_{k+1}^2\|\nu\|^2\end{aligned}$$

故 $\|T-T_k\|\leq s_{k+1}$．又由 $(T-T_k)e_{k+1}=s_{k+1}f_{k+1}$，可得 $\|\boldsymbol{T}-\boldsymbol{T}_k\|=s_{k+1}$
设 $S\in\mathcal{L}(V,W)$ 且 dim range $S\leq k$．则长度为 $k+1$ 的组 $Se_1,\ldots,Se_{k+1}$ 线性相关．因此，存在不全为0 的 $a_1,\ldots,a_{k+1}\in\mathbf{F}$ 使

$$a_1Se_1+\cdots+a_{k+1}Se_{k+1}=0$$

因为 $a_1,\ldots,a_{k+1}$ 不全为0，故 $a_1e_1+\cdots+a_{k+1}e_{k+1}\neq0$，故

$$\begin{aligned}\|(T-S)(a_1e_1+\cdots+a_{k+1}e_{k+1})\|^2 &= \|T(a_1e_1+\cdots+a_{k+1}e_{k+1})\|^2\\ &= \|s_1a_1f_1+\cdots+s_{k+1}a_{k+1}f_{k+1}\|^2\\ &= s_1^2|a_1|^2+\cdots+s_{k+1}^2|a_{k+1}|^2\\ &\geq s_{k+1}^2\left(|a_1|^2+\cdots+|a_{k+1}|^2\right)\\ &= s_{k+1}^2\|a_1e_1+\cdots+a_{k+1}e_{k+1}\|^2\end{aligned}$$

因为 $a_1e_1+\cdots+a_{k+1}e_{k+1}\neq 0$

$$\|T-S\|\geq s_{k+1}.$$

因此，在 dim range $S\leq k$ 的 $S\in\mathcal{L}(V,W)$ 当中， $S=T_k$ 使得 $\|T-S\|$ 最小

奇异值分解用于最佳逼近的其他例子，见本节习题 22、27．前一题是求给定维数的子空间，使得限制在其上的线性映射尽可能小；而后一题是求一幺正算子，使其跟给定的算子尽可能近

极分解

回顾一下，我们在 7.54 前讨论了满足 $|z|=1$ 的复数 z 和幺正算子之间的类比．我们继续用这个类比，并注意到除0 以外每个复数 z 都可以写成这样的形式：

$$z=\left(\frac{z}{|z|}\right)|z|$$

$$=\left(\frac{z}{|z|}\right)\sqrt{\bar{z}z},$$

其中第一个因子，也就是 $z/|z|$ ，绝对值为1
根据上段，我们这样猜测：每个算子 $T\in\mathcal{L}(V)$ 都可以写成一幺正算子乘以 $\sqrt{T^*T}$ 的形式这样猜确实是对的．相应的结果就称为极分解，它很漂亮地描述了 V 上的任意算子
注意若 $T\in\mathcal{L}(V)$ ，那么 T^*T 是正算子【如 7.64 (a) 中所示】．因此，算子 $\sqrt{T^*T}$ 是有意义的，而且明确定义为 V 上的正算子
我们将要陈述并证明的极分解，说的是 V 上的每个算子都是一幺正算子和一正算子的乘积．这样，我们就能将 V 上的任意算子写成两个性质良好的算子的乘积，并且这两种算子我们既能全面地描述也能很清楚地理解．幺正算子由 $\mathbf{F}=\mathbf{C}$ 情形下的 7.55 描述，而正算子由实谱定理（7.29）和复谱定理（7.31）描述
具体而言，考虑 $\mathbf{F}=\mathbf{C}$ 的情形，设

$$T=S\sqrt{T^*T}$$

是算子 $T\in\mathcal{L}(V)$ 的极分解，其中 s 是一幺正算子．那么就存在 V 的规范正交基使得 S 关于其有对角矩阵，也存在 V 的规范正交基使得 $\sqrt{T^*T}$ 关于其有对角矩阵．注意：可能并不存在一个规范正交基，能同时将 s 和 $\sqrt{T^*T}$ 的矩阵都化为对角矩阵这样的好形式——可能S 需要一个规范正交基，而 $\sqrt{T^*T}$ 需要另一个规范正交基
不过（仍然假设 $\mathbf{F}=\mathbf{C}$ ），如果 T 是正规的，那就可以选出一个规范正交基，使得 s 和$\sqrt{T^*T}$ 关于这一个基都有对角矩阵——见本节习题 31．逆命题也是对的：如果 $T\in\mathcal{L}(V)$ 且$T=S\sqrt{T^*T}$ ，其中 $S\in\mathcal{L}(V)$ 是幺正算子且 S 和 $\sqrt{T^*T}$ 关于同一规范正交基有对角矩阵，那么 T 是正规的．具体而言，在这一逆命题的条件下，可得 T 关于同一规范正交基有对角矩阵，也就可得 T 是正规的【根据7.31 中(c) 和(a) 的等价关系】，即结论成立
下面的极分解在实内积空间和复内积空间上都成立，对这些空间上的所有算子都适用

7.93 极分解

设 $T\in\mathcal{L}(V)$ ，那么存在幺正算子 $S\in\mathcal{L}(V)$ 使

$$T=S\sqrt{T^*T}$$

证

令 $s_1,\ldots,s_m$ 是 T 的正奇异值，令 $e_1,\ldots,e_m$ 和 $f_1,\ldots,f_m$ 是 V 中的规范正交组，使得

$$T\nu=s_1\nu,e_1f_1+\cdots+s_m\nu,e_mf_m$$

对任一 $\nu\in V$ 都成立．将 $e_1,\ldots,e_m$ 和 $f_1,\ldots,f_m$ 扩充为 V 的规范正交基 $e_1,\ldots,e_n$ 和 $f_1,\ldots,f_n$
定义 $S\in\mathcal{L}(V)$ 为

$$S\nu=\nu,e_1f_1+\cdots+\nu,e_nf_n$$

对任一 $\nu \in V$ 都成立．那么

$$\begin{aligned}\|S\nu\|^2 &= \|\langle\nu, e_1\rangle f_1 + \cdots + \langle\nu, e_n\rangle f_n\|^2 \\ &= |\langle\nu, e_1\rangle|^2 + \cdots + |\langle\nu, e_n\rangle|^2 \\ &= \|\nu\|^2\end{aligned}$$

因此， s 是幺正算子
将 T^* 作用于式 (7.94) 两侧，再用由式 (7.77) 给出的 T^* 的表达式，可得

$$T^*T\nu = s_1^2\,\nu, e_1e_1 + \cdots + s_m^2\,\nu, e_me_m$$

对任一 $\nu \in V$ 都成立．那么，如果 $\nu \in V$ 就有

$$\sqrt{T^*T}\nu = s_1\,\nu, e_1e_1 + \cdots + s_m\,\nu, e_me_m$$

这是因为将 σ_ν 映成上式右侧这一表达式的算子为正算子且其平方就等于 T^*T． 这样一来，就有

$$\begin{aligned}S\sqrt{T^*T}\nu &= S\left(s_1\,\nu, e_1e_1 + \cdots + s_m\,\nu, e_me_m\right) \\ &= s_1\,\nu, e_1f_1 + \cdots + s_m\,\nu, e_mf_m \\ &= T\nu\end{aligned}$$

其中最后一个等式由式(7.94) 得到

本节习题27 表明，以上这一证明中得到的幺正算子 s ， 是最接近 T 的幺正算子

极分解还有其他的证法，其中直接使用了谱定理而回避了奇异值分解．然而，以上的证明看起来比那些证明更简洁

作用于椭球和平行体的算子

7.95 球

V 中半径为1、以0 为心的球（ball)，记为 B ，定义为

$$B = \{\nu \in V : \|\nu\| \leq 1\}$$

如果 $\dim V = 2$ ，那么我们有时候用的词是“圆盘”（disk）而不是“球”然而，在所有维度中都用“球”这个词可以避免混乱．类似地，如果 $\dim V = 2$ ，那么我们有时候用的词是“椭圆”（ellipse）而不是我们将要定义的“椭球”同样，在所有维度中都用“椭球”这个词可以避免混乱

以下定义的椭球可以视为将球 B 沿每个 f_k 轴伸缩至 s_k 倍得到的

找不到“images/d96998e18688f4d21d54fba7359635e8e4443de4f79de7095359a2cb5b62eec7.jpg”。

$\mathbf{R}^2$ 中的球 B

7.96 椭球、主轴

$f_1, \ldots, f_n$ 是 V 的规范正交基， $s_1, \ldots, s_n$ 是正数．主轴为 $s_1f_1, \ldots, s_nf_n$ 的椭球

$$E(s_1f_1, \ldots, s_nf_n) = \left\{\nu \in V : \frac{|\langle\nu, f_1\rangle|^2}{s_1^2} + \cdots + \frac{|\langle\nu, f_n\rangle|^2}{s_n^2} < 1\right\}.$$

椭球的记号 $E(s_1f_1, \ldots, s_nf_n)$ 并没有显式地包含内积空间 V ，尽管以上的定义依赖于 V ．然而，从上下文以及“ $f_1, \ldots, f_n$ 是 V 的规范正交基”这一要求来看，内积空间 V 应该是明确的

7.97 椭球

找不到“images/d64e232094e949a75614c5b214db6e603cc77cdbbec68ea9735da333dedebdce.jpg”。

$\mathbf{R}^2$ 中的椭球 $E(2f_1, f_2)$，其中 f_1, f_2 是$\mathbf{R}^2$ 的标准基

找不到"images/ff36ae0e7b9f17639ec9630343a89223b52edb22cc2fd7d5edafea49725caeb8.jpg"。

找不到"images/bab59beb2d0b552a20beb7a938b00359471881b82cfe1f45a6c76f3aeb33a0f1.jpg"。

$\mathbf{R}^2$ 中的椭球 $E(2f_1, f_2)$，其中 $f_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), f_2 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$

$\mathbf{R}^3$ 中的椭球 $E(4f_1, 3f_2, 2f_3)$，其中f_1, f_2, f_3 是 $\mathbf{R}^3$ 的标准基

对于 V 的任一规范正交基 $f_1, \dots, f_n$，椭球 $E(f_1, \dots, f_n)$ 都等于 V 中的球 B 【根据帕塞瓦尔恒等式 6.30 (b)】

7.98 T (ω)

对定义在 V 上的函数 T 和 $\subseteq V$

$$T() = \{T\nu : \nu \in \}$$

那么，如果 T 是定义在 V 上的函数，就有 $T(V) = \mathrm{range} T$
接下来这条结果指出，每个可逆算子 $T \in \mathcal{L}(V)$ 都将 V 中的球 B 映成 V 中的椭球．其证明表明这个椭球的主轴就由 T 的奇异值分解得到

7.99 可逆算子化球为椭球

$T \in \mathcal{L}(V)$ 可逆，则 T 将 V 中的球 B 映成 V 中的椭球

证

设 T 有如下奇异值分解：

$$T\nu = s_1 \nu, e_1 f_1 + \dots + s_n \nu, e_n f_n$$

对所有 $\nu \in V$ 成立，其中 $s_1, \dots, s_n$ 是 T 的奇异值，而 $e_1, \dots, e_n$ 和 $f_1, \dots, f_n$ 都是 V 的规范正交基．我们将证明 $T(B) = E(s_1 f_1, \dots, s_n f_n)$.
设 $\nu \in B$．因为 T 可逆，所以奇异值 $s_1, \dots, s_n$ 都不等于 0（见 7.68）．那么，由式(7.100) 可得

$$\frac{|\langle T\nu, f_1\rangle|^2}{s_1^2} + \dots + \frac{|\langle T\nu, f_n\rangle|^2}{s_n^2} = |\langle \nu, e_1\rangle|^2 + \dots + |\langle \nu, e_n\rangle|^2 < 1.$$

因此，$T\nu \in E(s_1 f_1, \dots, s_n f_n)$．于是，$T(B) \subseteq E(s_1 f_1, \dots, s_n f_n)$
为证明另一方向的包含关系，现在设 $w \in E(s_1 f_1, \dots, s_n f_n)$，令

$$\nu = \frac{\langle w, f_1\rangle}{s_1} e_1 + \dots + \frac{\langle w, f_n\rangle}{s_n} e_n.$$

那么，由 $\|\nu\| < 1$ 和式 (7.100)，可得 $T\nu = w, f_1 f_1 + \dots + w, f_n f_n = w$．于是，$T(B) \supseteq E(s_1 f_1, \dots, s_n f_n)$

我们现在用前一条结果证明，可逆算子将所有椭球——不只是半径为 1 的球——都映成椭球

7.101 可逆算子化椭球为椭球

$T \in \mathcal{L}(V)$ 可逆，且 E 是 V 中的椭球．那么 $T(E)$ 是 V 中的椭球

证

存在 V 的规范正交基 $f_1,\ldots,f_n$ 和正数 $s_1,\ldots,s_n$ 使得 $E=E(s_1f_1,\ldots,s_nf_n)\$$. 定义$\$S\in\mathcal{L}(V)$ 为

$$S(a_1f_1+\cdots+a_nf_n)=a_1s_1f_1+\cdots+a_ns_nf_n$$

那么， s 将 V 中的球 B 映成 E （你可以验证），故

$$T(E)=T\big(S(B)\big)=(TS)(B).$$

结合上式并将7.99 应用于 TS 得 $T(E)$ 是 V 中的椭球

回顾一下（见3.95），如果 $u\in V$， $\subseteq V$，那么 $u+$ 定义为

$$u+\;=\{u+w:w\in\}$$

几何上，集合 和 $u+$ 看起来一样，但是位置不同

7.102 平行体

$\nu_1,\ldots,\nu_n$ 是 V 的基，令

$$P(\nu_1,\ldots,\nu_n)=\{a_1\nu_1+\cdots+a_n\nu_n:a_1,\ldots,a_n\in(0,1)\}$$

平行体是形如 $u+P(\nu_1,\ldots,\nu_n)$ 的集合，其中 $u\in V$． 向量 $\nu_1,\ldots,\nu_n$ 称为该平行体的边

如果 $\dim V=2$，那么我们经常用“平行四边形”一词而不是“平行体”

7.103 平行体

$\mathbf{R}^2$ 中的平行体 $(0.3,0.5)+P\big((1,0),(1,1)\big)$

找不到“images/855371f152a3df0b26aac79b5565de12a0344787c98d9b03c759d7189c171b58.jpg”。

$\mathbf{R}^3$ 中的一个平行体

找不到“images/6d64e646737beee49282caae2fba480f62a8da5bcfb643eeb120efe4e4327444.jpg”。

7.104 可逆算子化平行体为平行体

设 $u\in V$ 且 $\nu_1,\ldots,\nu_n$ 是 V 的基．设 $T\in\mathcal{L}(V)$ 可逆，那么

$$T\big(u+P(\nu_1,\ldots,\nu_n)\big)=Tu+P(T\nu_1,\ldots,T\nu_n)$$

证

由 T 可逆，组 $T\nu_1,\ldots,T\nu_n$ 是 V 的基
由T 是线性的

$$T(u+a_1\nu_1+\cdots+a_n\nu_n)=Tu+a_1T\nu_1+\cdots+a_nT\nu_n$$

对所有 $a_1,\ldots,a_n\in(0,1)$ 成立．因此， $T\big(u+P(\nu_1,\ldots,\nu_n)\big)=Tu+P(T\nu_1,\ldots,T\nu_n)$

正如我们把矩形从 $\mathbf{R}^2$ 中的平行四边形里特别区分出来一样，我们也给 V 中定义边

7.105 长方体

V 中的长方体是形如

$$u+P(r_1e_1,\ldots,r_ne_n)$$

的集合，其中 $u \in V$， $r_1,\ldots,r_n$ 是正数， $e_1,\ldots,e_n$ 是 V 的规范正交基

$\mathbf{R}^2$ 中长方体就是矩形，但长方体这一术语适用于所有维度

7.106 长方体

找不到“images/2e02f4ff71070d2fae9b9e82f07fd6e474a85ab49e6794d450eca0da90dec834.jpg”。

找不到“images/b6ed4638ac9b0bf776d38a5046fd39710207cdffac9a6a82d69e201fbb2da3f8.jpg”。

长方体 $(1,0)+P(\sqrt{2}e_1,\sqrt{2}e_2)$，其中$e_1=\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right),e_2=\left(-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)$
长方体 $P(e_1,2e_2,e_3)$，其中 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 是$\mathbf{R}^3$ 的标准基

设 $T\in\mathcal{L}(V)$ 可逆．那么 T 将 V 中每个平行体都映成 V 中的平行体（根据7.104）．特别地，T 将 V 中每个长方体都映成 V 中的平行体．这就提出一个问题： T 是否把 V 中的某些长方体映成 V 中的长方体？接下来这条结果借助奇异值分解回答了这个问题

7.107 每个可逆算子都将某些长方体化成长方体

$T\in\mathcal{L}(V)$ 可逆．设 T 有奇异值分解

$$T\nu=s_1\,\nu,e_1f_1+\cdots+s_n\,\nu,e_nf_n$$

其中， $s_1,\ldots,s_n$ 是 T 的奇异值， $e_1,\ldots,e_n$ 和 $f_1,\ldots,f_n$ 是 V 的规范正交基，上式对所有$\nu\in V$ 成立．那么，对于所有正数 $r_1,\ldots,r_n$ 和所有 $u\in V$， T 将长方体 $u+P(r_1e_1,\ldots,r_ne_n)$ 映成长方体 $Tu+P(r_1s_1f_1,\ldots,r_ns_nf_n)$

证

如果 $a_1,\ldots,a_n\in(0,1)$， $r_1,\ldots,r_n$ 是正数， $u\in V$，那么

$$T(u+a_1r_1e_1+\cdots+a_nr_ne_n)=Tu+a_1r_1s_1f_1+\cdots+a_nr_ns_nf_n.$$

用奇异值计算体积

本小节我们的目标是，理解算子如何改变其定义域子集的体积．因为体积的概念属于分析学而非线性代数，所以我们仅仅讨论体积的直觉概念．而在分析学的体系之下，我们用来处理体积的直观方法可以转化为适当的正确定义、正确陈述和正确证明
我们关于体积的直觉，在实内积空间上最有效．因此在本小节的剩余部分，经常会出现$\mathbf{F}=\mathbf{R}$ 这一假设
如果 $\dim V=n$， 那么我们说的“体积”指的是 n 维体积．你应该很熟悉 $\mathbf{R}^3$ 中的这一概念当 $n=2$ 时，我们通常称之为“面积”而不是“体积”，但是为了保持一致，我们在所有维度上都用“体积”这个词．关于体积最基本的直觉就是，长方体（根据定义，其定义边互相正交）的体积是所有定义边长度的乘积

7.108 长方体的体积

设 $\mathbf{F}=\mathbf{R}$．如果 $u\in V$， $r_1,\ldots,r_n$ 是正数， $e_1,\ldots,e_n$ 是 V 的规范正交基，那么

$$\text{olume}\,(u+P(r_1e_1,\ldots,r_ne_n))=r_1\times\cdots\times r_n.$$

以上定义跟常见的 $\mathbf{R}^2$ 中矩形面积（这里我们称为体积）公式和 $\mathbf{R}^3$ 中长方体的体积公式是一致的．例如，例7.106 中第一个长方体具有二维体积（或面积），其值为2，因为其定义边的长度为 $\sqrt{2}$ 和 $\sqrt{2}$；第二个长方体具有三维体积，其值为2，因为其定义边的长度为1、2、1
为定义 V 的子集的体积，我们用有限个互不相交的长方体来逼近这一子集，并且将用于逼近的这些长方体的体积加起来．我们对越多的不相交长方体取并集，就能越准确地逼近 V 的一子集，对其体积的逼近也越好
这些想法应该让你想起了黎曼积分是怎样定义的——用若干互不相交的矩形来逼近曲线下的面积．上面的讨论引出了以下不严谨但合乎直觉的定义

找不到"images/9156dea2aeb006bfd0eb46278953c05c173ae274a8e1295d4b15eeae568096cf.jpg"。

该球的体积 五个长方体的体积之和

7.109 体积

设 $\mathbf{F} = \mathbf{R}$， $\subseteq V$．那么 的体积，记为 volume ，约等于逼近 的若干个不相交长方体的体积之和

我们用一种直观的方法处理体积，如此就忽略了许多值得推敲的问题．例如，如果我们用关于一个基的长方体逼近 ，那么我们再用关于另一个基的长方体逼近 又能否得到相同的体积？如果 $_1$ 和 $_2$ 是 V 的不相交子集，那么是否有volume $(_1 \cup _2) =$ volume $_1+$ volume $_2$？假如我们仅考虑 V 的性质相当好的子集，那么借助分析学的手段可以说明，这两个问题都有肯定的回答，和我们对体积的直觉一致

7.110 线性映射改变体积

$T \in \mathcal{L}(\mathbf{R}^2)$， $T\nu = 2\,\nu, e_1 e_1 + \nu, e_2 e_2$，其中 $\boldsymbol{e}_1, \boldsymbol{e}_2$ 是 $\mathbf{R}^2$ 的标准基．这一线性映射使向量沿 $\boldsymbol{e}_1$ 轴拉伸至2 倍而沿 e_2 轴不变．上面那个由五个长方体逼近的球，由 T 映射为此处展示的椭球．原始图片中的五个长方体，每一个都被映射为宽度为原来的两倍而高度跟原来相同的长方体．于是每个长方体都被映射为体积（面积）为原来两倍的长方体．这五个新长方体的体积总和就近似为这个椭球的体积．所以， T 将球的体积改变为原来的2 倍

找不到"images/91bb1043e07637b806b7641302adabc99fde35f4a5c813d1e135ee2003116cb3.jpg"。

此图中每个长方体的宽度都是上一张图中的两倍，而其高度不变

上例 T 将关于基 $\boldsymbol{e}_1, \boldsymbol{e}_2$ 的长方体映射为关于同一基的长方体，我们可以从中看出 T 是怎样改变体积的．一般来说，算子将长方体映射为不是长方体的平行体．然而，如果我们选取适当的基（来自奇异值分解的基！），那么关于该基的长方体就能被映射为关于可能不同的一个基的长方体，如7.107 所示．观察这点可得下条结果的自然证明

7.111 体积变化倍数是奇异值的乘积

$\mathbf{F} = \mathbf{R}$， $T \in \mathcal{L}(V)$ 可逆，且 $\subseteq V$．那么volume $T() = T$ 的奇异值的乘积 volume

证

设 T 有奇异值分解

$$T\nu = s_1\,\nu, e_1 f_1 + \cdots + s_n\,\nu, e_n f_n$$

对所有 $\nu \in V$ 成立，其中 $e_1, \ldots, e_n$ 和 $f_1, \ldots, f_n$ 是 V 的规范正交基
用形如 $u + P(r_1 e_1, \ldots, r_n e_n)$ 的长方体逼近 ，这些长方体的体积是 $r_1 \times \cdots \times r_n$．算子 T 将每个长方体 $u + P(r_1 e_1, \ldots, r_n e_n)$ 都映成长方体 $Tu + P(r_1 s_1 f_1, \ldots, r_n s_n f_n)$，其体积为 $(s_1 \times \cdots \times s_n)(r_1 \times \cdots \times r_n)$
算子 T 将用于逼近 的那些长方体映成用于逼近 $T()$ 的那些长方体．因为 T 改变了用于逼近 的每个长方体的体积，对应倍数为 $s_1 \times \cdots \times s_n$，所以线性映射 T 按同样的倍数改变 的体积

设 $T \in \mathcal{L}(V)$．等我们讨论到行列式的时候，我们就会看到， T 的奇异值的乘积就等于$|\det T|$，见 9.60 和 9.61

取决于特征值的算子性质

我们以下面这个表格结束本章．注意，这个表格描述的是有限维复内积空间中的情形．表格的第一列显示的是正规算子在有限维复内积空间上可能具有的性质；表格的第二列显示的是相应的 C 的子集——当且仅当一算子的所有特征值都属于这一特定子集，该算子才具有左边那列中相应的性质．例如，表格的第一行指出，正规算子可逆当且仅当它的所有特征值非零（这第一行是表格中唯一不需要假设算子正规的）

你得确保自己可以解释表格里所有结果为什么都成立．例如，表格的最后一行成立，是因为算子的范数等于它的最大奇异值（根据 7.85），而正规算子的奇异值——假定 $\mathbf{F} = \mathbf{C}$. ——等于特征值的绝对值（根据 7E 节习题7）

正规算子的性质	特征值属于
可逆	C\{0}
自伴	R
斜	{∈C:Reλ=0}
正交投影	{0,1}
正	[0,8)
幺正	{∈C:||=1}
范数小于1	{λ ∈ C: || <1}