

在定义向量空间时，我们推广了 \mathbf{R}^2 和 \mathbf{R}^3 的线性结构（加法和标量乘法），而未顾及几何上的特性，例如长度和角度等概念。内积的概念就蕴含了这些几何层面的想法，它是我们本章研究的主题

每种内积都可诱导出一种范数（你可以把范数看成长度）。范数满足一些重要的性质，例如毕达哥拉斯定理、三角不等式、平行四边形等式和柯西-施瓦兹不等式

在讨论内积空间时，我们将欧几里得几何中的垂直向量这一概念，重命名为正交向量。我们将看到，规范正交基在内积空间中非常有用。格拉姆-施密特过程可构造出这样的基。本章结尾处，我们将综合运用上述工具来解决最小化问题

6A 内积和范数

内积

将 \mathbf{R}^2 和 \mathbf{R}^3 中的向量看成始于原点的箭头。向量 ν 的长度称 σ_ν 的范数，记作 $\|\nu\|$

对 $\nu = (a, b) \in \mathbf{R}^2$ $\|\nu\| = \sqrt{a^2 + b^2}$

对 $\nu = (a, b, c) \in \mathbf{R}^3$ $\|\nu\| = \sqrt{a^2 + b^2 + c^2}$

$x = (x_1, \dots, x_n) \in \mathbf{R}^n$ 的范数 $\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$

范数在 \mathbf{R}^n 上不是线性的。为了把线性引入讨论中，我们引入点积

6.1 点积

$x, y \in \mathbf{R}^n$ 的点积 $x \cdot y = x_1 y_1 + \dots + x_n y_n$ ，其中 $x = (x_1, \dots, x_n)$ ， $y = (y_1, \dots, y_n)$

\mathbf{R}^n 中两个向量的点积是数而不是向量

注意，对所有 $x \in \mathbf{R}^n$ ，有 $x \cdot x = \|x\|^2$ 。此外， \mathbf{R}^n 上的点积还满足下列性质

如果我们把向量视为一个点而不是一个箭头，那么应将 $\|x\|$ 解释为从原点到点 x 的距离

对所有 $x \in \mathbf{R}^n$ ，均有 $x \cdot x \geq 0$

$x \cdot x = 0$ 当且仅当 $x = 0$

对于固定的 $y \in \mathbf{R}^n$ ， \mathbf{R}^n 到 \mathbf{R} 的将 $x \in \mathbf{R}^n$ 对应到 $x \cdot y$ 的映射是线性的对所有 $x, y \in \mathbf{R}^n$ ，均有 $x \cdot y = y \cdot x$

内积是对点积的推广。现在你可能会猜测，内积的定义就是通过将上一段讨论的点积性质抽象化得出的。对于实向量空间，这个猜想没错。然而，为了让我们所作的定义同时适用于实向量空间和复向量空间，我们需要在下定义之前考虑复数的情况

回忆一下，如果 $\lambda = a + bi$ ，其中 $a, b \in \mathbf{R}$ ，那么 λ 的绝对值记作 $|\lambda|$ ，定义为 $|\lambda| = \sqrt{a^2 + b^2}$ ； λ 的复共轭记作 $\bar{\lambda}$ ，定义为 $\bar{\lambda} = a - bi$ ； $|\lambda|^2 = \lambda \bar{\lambda}$

绝对值和复共轭的定义及基本性质见于第4章对于 $z = (z_1, \dots, z_n) \in \mathbf{C}^n$ ，我们定义 z 的范数为

$$\|z\| = \sqrt{|z_1|^2 + \dots + |z_n|^2}$$

上式中需要用到绝对值，因为我们想让 $\|z\|$ 是个非负实数。注意到

$$\|z\|^2 = z_1 \bar{z}_1 + \dots + z_n \bar{z}_n$$

我们想将 $\|z\|^2$ 视为 z 与自身的内积，就如在 \mathbf{R}^n 中那般。于是上式表明，向量 $w = (w_1, \dots, w_n) \in \mathbf{C}^n$ 和 z 的内积应等于

$$w_1 \bar{z}_1 + \dots + w_n \bar{z}_n$$

如果将 w 和 z 的角色互换，上述表达式就会被其复共轭所替代。于是，我们期望内积满足： w 和 z 的内积等于 z 和 w 的内积的复共轭。说明这点动机后，我们现在可以定义任意实向量空间或复向量空间 V 上的内积了

关于下面定义中所用的记号，有一点需要注意：对于 $\lambda \in \mathbf{C}$ ，记号 $\lambda \geq 0$ 意为 λ 是非负实数

6.2 内积

V 上的内积是一个函数，它将由 V 中元素构成的每个有序对 (u, ν) 对应至一个数 $\langle u, \nu \rangle \in \mathbf{F}$ ，并满足如下性质

正性 对于所有 $\nu \in V$ ，均有 $\langle \nu, \nu \rangle \geq 0$

定性 $\langle \nu, \nu \rangle = 0$ 当且仅当 $\nu = 0$

第一位置上的可加性 对所有 $u, \nu, w \in V$ ， $\langle u + \nu, w \rangle = \langle u, w \rangle + \langle \nu, w \rangle$

第一个位置上的齐次性 对所有 $\lambda \in \mathbf{F}$ 和所有 $u, \nu \in V$ ，均有 $\langle \lambda u, \nu \rangle = \lambda \langle u, \nu \rangle$

共轭对称性 对所有 $u, \nu \in V$ ，均有 $\langle u, \nu \rangle = \overline{\langle \nu, u \rangle}$

每个实数都等于其复共轭。因此如果我们讨论的是实向量空间，那么我们可以从上面最后一个条件中省去复共轭，并直接把它表述为：对于所有 $u, \nu \in V$ ，均有 $\langle u, \nu \rangle = \langle \nu, u \rangle$

大部分数学家都把内积定义如上。不过，许多物理学家使用的定义中，要求齐次性在第二个位置上成立，而不是第一个位置

6.3 内积

(a) \mathbf{F}^n 上的欧几里得内积定义为: 对所有 $(w_1, \dots, w_n), (z_1, \dots, z_n) \in \mathbf{F}^n$,

$$\langle (w_1, \dots, w_n), (z_1, \dots, z_n) \rangle = w_1 \bar{z}_1 + \dots + w_n \bar{z}_n$$

(b) 若 c_1, \dots, c_n 是正数, 那么可在 \mathbf{F}^n 上定义内积如下: 对所有 $(w_1, \dots, w_n), (z_1, \dots, z_n) \in \mathbf{F}^n$,

$$\langle (w_1, \dots, w_n), (z_1, \dots, z_n) \rangle = c_1 w_1 \bar{z}_1 + \dots + c_n w_n \bar{z}_n$$

(c) 在由定义在区间 $[-1, 1]$ 上的全体连续实值函数构成的向量空间上, 可定义内积如下: 对所有定义在区间 $[-1, 1]$ 上的连续实值函数 f, g ,

$$\langle f, g \rangle = \int_{-1}^1 fg$$

(d) 在 $\mathcal{P}(\mathbf{R})$ 上, 可定义内积如下: 对所有 $p, q \in \mathcal{P}(\mathbf{R})$

$$\langle p, q \rangle = p(0)q(0) + \int_{-1}^1 p'q'$$

(e) 在 $\mathcal{P}(\mathbf{R})$ 上, 还可定义内积如下: 对所有 $p, q \in \mathcal{P}(\mathbf{R})$

$$\langle p, q \rangle = \int_0^\infty p(x)q(x)e^{-x}dx$$

6.4 内积空间

带有内积的向量空间称为内积空间

内积空间最重要的例子是如上例 (a) 所示欧几里得内积的 \mathbf{F}^n . 当称 \mathbf{F}^n 是内积空间时, 除非另有说明, 你都应假设其上定义的内积是欧几里得内积. 为让我们不用反复重申 V 和 W 是内积空间这个前提条件, 我们作出如下假设

6.5 记号: V, W

在本章的剩余部分和下章中, V 和 W 都指代 \mathbf{F} 上的内积空间

注意, 这里稍微有些滥用术语. 一个内积空间是带有内积的向量空间. 当我们称一向量空间 V 为内积空间时, 我们或是将 V 上的内积隐含于其中, 或是由上下文可明确 V 上的内积如何定义 (又或者, 如果这个向量空间是 \mathbf{F}^n , 那么所用内积就是欧几里得内积)

6.6 内积的基本性质

(a) 对每个固定的 $\nu \in V$, 将 $u \in V$ 对应到 u, ν 的函数都是 V 到 \mathbf{F} 的线性映射

(b) 对每个 $\nu \in V$, 均有 $0, \nu = 0$

(c) 对每个 $\nu \in V$, 均有 $\nu, 0 = 0$

(d) 对所有 $u, \nu, w \in V$, 均有 $\langle u, \nu + w \rangle = \langle u, \nu \rangle + \langle u, w \rangle$

(e) 对所有 $\lambda \in \mathbf{F}$ 和 $u, \nu \in V$, 均有 $\langle u, \lambda \nu \rangle = \bar{\lambda} \langle u, \nu \rangle$

证

(a) 对 $\nu \in V$, 由内积定义中第一个位置上的可加性和齐次性, 即可得 $u \mapsto \langle u, \nu \rangle$ 的线性 (b) 每个线性映射都将 0 对应到 0. 所以由 (a) 即知 (b) 成立

(c) 若 $\nu \in V$, 那么由内积定义中的共轭对称性质和 (b) 即得 $\nu, 0 = \overline{0, \nu} = \overline{0} = 0$

(d) 设 $u, \nu, w \in V$. 则

$$\begin{aligned} u, \nu + w &= \overline{\nu + w, u} \\ &= \overline{\nu, u + w, u} \\ &= \overline{\nu, u} + \overline{w, u} \\ &= u, \nu + u, w \end{aligned}$$

(e) 设 $\lambda \in \mathbf{F}$ 及 $u, \nu \in V$. 则

$$\begin{aligned}\langle u, \lambda \nu \rangle &= \overline{\langle \lambda \nu, u \rangle} \\ &= \overline{\lambda \langle \nu, u \rangle} \\ &= \overline{\lambda} \overline{\langle \nu, u \rangle} \\ &= \overline{\lambda} \langle u, \nu \rangle\end{aligned}$$

范数

我们定义内积的动机最初来源于 \mathbf{R}^2 和 \mathbf{R}^3 上向量的范数. 现在我们就会看到, 每种内积都能确定一种范数

6.7 范数

对 $\nu \in V$, ν 的范数记作 $\|\nu\|$, 定义为 $\|\nu\| = \sqrt{\langle \nu, \nu \rangle}$

6.8 范数

(a) 若 $(z_1, \dots, z_n) \in \mathbf{F}^n$ (其上定义了欧几里得内积), 那么

$$\|(z_1, \dots, z_n)\| = \sqrt{|z_1|^2 + \dots + |z_n|^2}$$

(b) 定义在区间 $[-1, 1]$ 上的连续实值函数所构成的向量空间【其上内积定义如6.3 (c)】中的元素 f 满足

$$\|f\| = \sqrt{\int_{-1}^1 f^2}$$

6.9 范数的基本性质

$\nu \in V$

(a) $\|\nu\| = 0$ 当且仅当 $\nu = 0$ (b) 对所有 $\lambda \in \mathbf{F}$, 均有 $\|\lambda \nu\| = |\lambda| \|\nu\|$

证

(a) 因为 $\langle \nu, \nu \rangle = 0$ 当且仅当 $\nu = 0$, 所以原命题成立

(b) 设 $\lambda \in \mathbf{F}$. 那么

$$\begin{aligned}\|\lambda \nu\|^2 &= \langle \lambda \nu, \lambda \nu \rangle \\ &= \lambda \langle \nu, \lambda \nu \rangle \\ &= \lambda \overline{\lambda} \langle \nu, \nu \rangle \\ &= |\lambda|^2 \|\nu\|^2\end{aligned}$$

两边开平方根即得我们欲证的等式

上述结论中(b)的证明形象说明了一个普遍适用的道理: 处理范数的平方通常比直接处理范数来得容易

6.10 正交

两向量 $u, \nu \in V$ 是正交的, 若 $\langle u, \nu \rangle = 0$

在上述定义中, 两个向量的顺序是无关紧要的, 因为 $\langle u, \nu \rangle = 0$ 当且仅当 $\langle \nu, u \rangle = 0$ 除了说“ u 和 ν 是正交的”, 我们有时也说“ u 正交于 ν ” “orthogonal” 这个词来自希腊语中的 “or-thogonios”, 后者意为“直角的”

习题 15 要求你证明若 u, ν 是 \mathbf{R}^2 中的非零向量, 那么

$$\langle u, v \rangle = \|u\| \|v\| \cos \theta,$$

其中 θ 是 u 和 v 间的夹角（将 u 和 v 看成始于原点的箭头）。于是 \mathbf{R}^2 中两个非零向量（在欧几里得内积下）是正交的，当且仅当它们之间夹角的余弦值是 0。这又等价于通常讨论平面几何时所说的两向量垂直。于是，你可将正交这个词看成表达垂直（perpendicular）之意的一个更炫的说法。我们从一个简单的结论开始研究正交性。

6.11 正交性和 0

- (a) 0 与 V 中每个向量都正交
- (b) 0 是 V 中唯一与自身正交的向量

6.11 正交性和 0

- (a) 由 6.6 (b)，对每个 $v \in V$ ， $v \perp 0$
- (b) 若 $v \in V$ 且 $\langle v, v \rangle = 0$ ，那么 $v = 0$ （由内积的定义可得）

下面的定理在 $V = \mathbf{R}^2$ 时的特殊情形，早在 3500 多年前就已经为古巴比伦人所知，在 2500 余年前又被希腊人重新发现并证明。当然，下面的证明用的不是原始的证法。

6.12 毕达哥拉斯定理

设 $u, v \in V$ 。若 u 和 v 是正交的，那么

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

证

设 $u, v \perp$ 。那么

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle \\ &= \|u\|^2 + \|v\|^2 \end{aligned}$$

设 $u, v \in V$ 且 $v \neq 0$ 。我们想将 u 写成 v 的标量倍加上正交于 v 的向量 w 的形式，如下图所示。

找不到“images/53e03eab13f7d039a0c179840f1322c89cd41411de855b84db3fce897b1b86dd.jpg”。

正交分解：将 u 表示成 v 的标量倍加上正交于 v 的向量的形式。为探究如何将 u 写成 v 的标量倍加上正交于 v 的向量的形式，令 $c \in \mathbf{F}$ 表示一标量。则

$$u = cv + (u - cv)$$

于是我们需选取 c 使得 v 与 $u - cv$ 正交。因此我们希望有

$$0 = \langle u - cv, v \rangle = \langle u, v \rangle - c\|v\|^2.$$

上式表明应将 c 取为 $\langle u, v \rangle / \|v\|^2$ 。取定该 c 后，我们就可写出

$$u = \frac{\langle u, v \rangle}{\|v\|^2} v + \left(u - \frac{\langle u, v \rangle}{\|v\|^2} v \right)$$

可验证上述等式将 u 显式地写成了 v 的标量倍加上正交于 v 的向量的形式。于是我们就证明了下面这个关键的结论。

6.13 一种正交分解

设 $u, \nu \in V$ ，且 $\nu \neq 0$ 。取 $c = \frac{\langle u, \nu \rangle}{\|\nu\|^2}$ 及 $w = u - \frac{\langle u, \nu \rangle}{\|\nu\|^2} \nu$ 那么

我们将利用正交分解 6.13 来证明下面的柯西-施瓦兹不等式，它是数学中最重要的不等式之一

6.14 柯西-施瓦兹不等式

设 $u, \nu \in V$ 。那么

$$|\langle u, \nu \rangle| \leq \|u\| \|\nu\|$$

当且仅当 u, ν 成标量倍数关系时，上述不等式取得等号

证

若 $\nu = 0$ ，那么欲证不等式的两端都等于 0，即得证。于是我们可假设 $\nu \neq 0$ 。考察 6.13 给出的正交分解式

$$u = \frac{\langle u, \nu \rangle}{\|\nu\|^2} \nu + w$$

其中 w 正交于 ν 。由毕达哥拉斯定理，

$$\begin{aligned} \|u\|^2 &= \left(\frac{\langle u, \nu \rangle}{\|\nu\|^2} \nu \right)^2 + \|w\|^2 \\ &= \frac{|\langle u, \nu \rangle|^2}{\|\nu\|^2} + \|w\|^2 \\ &\geq \frac{|\langle u, \nu \rangle|^2}{\|\nu\|^2} \end{aligned}$$

1821 年，奥古斯汀·路易·柯西 (Augustin-Louis Cauchy, 1789-1857) 证明了 6.16 (a) 1859 年，柯西的学生维克多·布尼亚科夫斯基 (Viktor Bunyakovsky, 1804-1889) 证明了类似于 6.16 (b) 的积分不等式。几十年后，赫尔曼·施瓦兹 (Hermann Schwarz, 1843-1921) 发现了类似的结论并引发了更多的关注，柯西-施瓦兹不等式由此得名
将上式两边都乘以 $\|\nu\|^2$ 并同时开平方根即得欲证的不等式

由上段的证明可得柯西-施瓦兹不等式取得等号当且仅当式 (6.15) 取得等号，这又等价于 $w = 0$ 。而 $w = 0$ 当且仅当 u 是 ν 的标量倍 (见 6.13)。于是柯西-施瓦兹不等式取得等号，当且仅当 u 是 ν 的标量倍或 ν 是 u 的标量倍 (或互成标量倍数关系)。这样说可以涵盖 u 或 ν 等于 0 这个特殊情形

6.16 柯西-施瓦兹不等式

(a) 若 $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbf{R}$ ，那么

$$(x_1 y_1 + \dots + x_n y_n)^2 \leq (x_1^2 + \dots + x_n^2)(y_1^2 + \dots + y_n^2)$$

这就是将柯西-施瓦兹不等式应用于向量 $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbf{R}^n$ 的结果 (所用内积是通常的欧几里得内积)

(b) 若 f, g 是定义在 $[1, 1]$ 上的连续实值函数，那么

$$\left| \int_{-1}^1 fg \right|^2 \leq \left(\int_{-1}^1 f^2 \right) \left(\int_{-1}^1 g^2 \right)$$

这就是将柯西-施瓦兹不等式应用于例 6.3 (c) 的结果

接下来的结论被称为三角不等式，从几何角度解释，就是三角形任意一边的长度都小于另外两边的长度之和。三角不等式表明，连接两点的折线路径中，最短的是一条线段 (可视作由共线的线段构成的折线)

找不到“images/4b048b34fa9be49b8ff99d4f7f3b2f73cecb9c4b1555ae3f4ce1f4756c8e8fe.jpg”。

这个三角形里 $u + \nu$ 的长度小于 u 的长度加上 ν 的长度

6.17 三角不等式

$u, v \in V$. 那么 $\|u + v\| \leq \|u\| + \|v\|$, 取等当且仅当 u, v 中任意一者是另一者的非负实数倍

证

$$\begin{aligned}\|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \langle u, u \rangle + \langle v, v \rangle + \langle u, v \rangle + \langle v, u \rangle \\ &= \|u\|^2 + \|v\|^2 + 2\langle u, v \rangle \\ &\leq \|u\|^2 + \|v\|^2 + 2|\langle u, v \rangle| \\ &\leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| \\ &= (\|u\| + \|v\|)^2\end{aligned}$$

其中式(6.19) 源自柯西-施瓦兹不等式 (6.14) 同时开平方根得证

上述证明表明, 三角不等式取得等号, 当且仅当式(6.18) 和(6.19) 取得等号. 于是三角不等式取得等号等价于

$$\langle u, v \rangle = \|u\|\|v\|$$

如果 u, v 之一是另一者的非负实数倍, 那么式 (6.20) 成立. 反之, 设式 (6.20) 成立. 那么柯西-施瓦兹不等式 (6.14) 取等号的条件就表明, u, v 必成标量倍数关系. 由式(6.20) 又可知该标量必须是非负实数, 这就完成了证明

反向三角不等式见于本节习题20

下面结果因其几何解释而被称作平行四边形等式: 在每个平行四边形中, 对角线的长度平方之和等于四条边的长度平方之和. 注意, 此处的证明比欧几里得几何中的证明更直截了当

找不到“images/a8eb0e0ba5c4ac1a6a7b7455f9a142924a8ce768c5fb42a17241c90675343e7d.jpg”。

图示平行四边形的对角线是 $u + v$ 和 $u - v$

6.21 平行四边形等式

设 $u, v \in V$. 那么

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$$

证

我们有

$$\begin{aligned}\|u + v\|^2 + \|u - v\|^2 &= \langle u + v, u + v \rangle + \langle u - v, u - v \rangle \\ &= \|u\|^2 + \|v\|^2 + \langle u, v \rangle + \langle v, u \rangle \\ &\quad + \|u\|^2 + \|v\|^2 - \langle u, v \rangle - \langle v, u \rangle \\ &= 2(\|u\|^2 + \|v\|^2)\end{aligned}$$

命题得证

6B 规范正交基

规范正交组和格拉姆-施密特过程

6.22 规范正交

如果一个向量组中所有向量的范数都是 1, 且每个向量与其他向量都正交, 则称该向量组是规范正交的

换言之, V 中向量组 e_1, \dots, e_m 是规范正交的, 若对所有 $j, k \in \{1, \dots\}$,

$$\langle e_j, e_k \rangle = \begin{cases} 1, \\ 0, \end{cases}$$

6.23 规范正交组

(a) \mathbf{F}^n 的标准基是一个规范正交组(b) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$ 是 \mathbf{F}^3 中的规范正交组. (c) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$ 是 \mathbf{F}^3 中的规范正交组(d) 设 n 是正整数. 那么, 正如习题4 要求你验证的那样,

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots, \frac{\sin nx}{\sqrt{\pi}}$$

是 $C[-\pi, \pi]$ 中的规范正交向量组. $C[-\pi, \pi]$ 是由定义在区间 $[-\pi, \pi]$ 上的连续实值函数所构成的向量空间, 且其上内积定义为

$$\langle f, g \rangle = \int_{-\pi}^{\pi} fg$$

上述规范正交组经常用于建立潮汐等周期现象的数学模型

(e) 在 $\mathcal{P}_2(\mathbf{R})$ 上定义内积为: 对于所有 $p, q \in \mathcal{P}_2(\mathbf{R})$,

$$\langle p, q \rangle = \int_{-1}^1 pq.$$

这样 $\mathcal{P}_2(\mathbf{R})$ 便成为内积空间. $\mathcal{P}_2(\mathbf{R})$ 的标准基 $1, x, x^2$ 不是规范正交组, 因为该组中向量的范数不是1. 将其中每个向量都除以各自的范数, 可得组 $1/\sqrt{2}, \sqrt{3}/2x, \sqrt{5}/2x^2$, 其中每个向量的范数都是1, 并且第二个向量与第一个、第三个向量正交. 然而, 第一个向量和第三个向量不正交, 从而该组不是规范正交组. 很快我们就会知道如何由标准基 $1, x, x^2$ 构造一个规范正交组 (见例 6.34)

处理规范正交组很容易, 如下面结论所示

6.24 规范正交组线性组合的范数

设 e_1, \dots, e_m 是 V 中的规范正交向量组. 那么对所有 $a_1, \dots, a_m \in \mathbf{F}$, 有

$$\|a_1 e_1 + \dots + a_m e_m\|^2 = |a_1|^2 + \dots + |a_m|^2$$

证

因各 e_k 的范数都是1, 所以只要反复利用毕达哥拉斯定理 (6.12) 得证

下面结论是上述结论的一个重要推论

6.25 规范正交组是线性无关的

每个规范正交向量组都是线性无关的

证

设 e_1, \dots, e_m 是 V 中的规范正交向量组, 且 $a_1, \dots, a_m \in \mathbf{F}$ 满足

$$a_1 e_1 + \dots + a_m e_m = 0$$

那么 $|a_1|^2 + \dots + |a_m|^2 = 0$ (由 6.24), 这表明各 a_k 都等于0. 于是 e_1, \dots, e_m 是线性无关的

现在我们给出一个重要的不等式

6.26 贝塞尔不等式

设 e_1, \dots, e_m 是 V 中的规范正交向量组. 若 $\nu \in V$, 那么

$$|\langle \nu, e_1 \rangle|^2 + \dots + |\langle \nu, e_m \rangle|^2 \leq \|\nu\|^2$$

证

设 $\nu \in V$. 那么

$$\nu = \underbrace{\langle \nu, e_1 \rangle e_1 + \dots + \langle \nu, e_m \rangle e_m}_u + \underbrace{\nu - \langle \nu, e_1 \rangle e_1 - \dots - \langle \nu, e_m \rangle e_m}_w$$

令 u 和 w 按上式定义. 若 $k \in \{1, \dots, m\}$, 那么 $\langle w, e_k \rangle = \langle \nu, e_k \rangle - \langle \nu, e_k \rangle \langle e_k, e_k \rangle = 0$. 这意味着 $\langle w, u \rangle = 0$. 现由毕达哥拉斯定理

$$\begin{aligned} \|\nu\|^2 &= \|u\|^2 + \|w\|^2 \\ &\geq \|u\|^2 \\ &= |\langle \nu, e_1 \rangle|^2 + \dots + |\langle \nu, e_m \rangle|^2 \end{aligned}$$

其中最后一行来自 6.24

6.27 规范正交基

V 的一个规范正交基, 就是 V 中既是规范正交组又是基的向量组

如标准基是 \mathbf{F}^n 的一个规范正交基

6.28 长度恰好的规范正交组是规范正交基

有限维 V 中每长为 $\dim V$ 的规范正交向量组都是 V 的规范正交基

证

由 6.25, V 中每个规范正交向量组都是线性无关的. 于是每个这样的组, 只要长度恰好, 就是一个基——见 2.38

上面提到, 标准基是 \mathbf{F}^4 的规范正交基. 我们现在说明

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

同样是 \mathbf{F}^4 的规范正交基
我们有

$$\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) = \sqrt{\frac{1}{2^2} + \frac{1}{2^2} + \frac{1}{2^2} + \frac{1}{2^2}} = 1$$

类似地, 上述组中其余三个向量的范数也为 1
注意到

$$\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \left(-\frac{1}{2}\right) + \frac{1}{2} \cdot \left(-\frac{1}{2}\right) = 0$$

类似地, 上述组中任意两个不同向量的内积都等于 0

于是上述组是规范正交的. 因为该组是 4 维向量空间 \mathbf{F}^4 中长度为 4 的规范正交组, 所以由 6.28 可得该组是 \mathbf{F}^4 的规范正交基
一般来说, 给定 V 的基 e_1, \dots, e_n 与向量 $\nu \in V$, 我们知道存在一组标量 $a_1, \dots, a_n \in \mathbf{F}$ 满足

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

对于 V 的一个任取的基来说, 计算出满足上述等式的数 a_1, \dots, a_n 可能很麻烦. 然而, 下面结论表明, 若取的是规范正交基, 各系数的计算就很容易了——只需取 $a_k = \langle \nu, e_k \rangle$

下面结论中, $\langle \nu, e_1 \rangle, \dots, \langle \nu, e_n \rangle$ 充当 \mathbf{F}^n 中向量坐标的角色, 从而使得维数为 n 的内积空间看起来就像 \mathbf{F}^n 一样. 请注意观察下面结论是如何体现这点的. 下面关于 $\|\nu\|$ 的公式被称为帕塞瓦尔恒等式 (Parseval's identity), 它于 1799 年发表在研究傅里叶级数 (Fourier series) 的文章中

6.30 将向量写成规范正交基的线性组合

e_1, \dots, e_n 是 V 的规范正交基且 $u, \nu \in V$, 那么

(a) $\nu = \langle \nu, e_1 \rangle e_1 + \dots + \langle \nu, e_n \rangle e_n$

(b) $\|\nu\|^2 = |\langle \nu, e_1 \rangle|^2 + \dots + |\langle \nu, e_n \rangle|^2$

(c) $\langle u, \nu \rangle = \langle u, e_1 \rangle \overline{\langle \nu, e_1 \rangle} + \dots + \langle u, e_n \rangle \overline{\langle \nu, e_n \rangle}$

证

因为 e_1, \dots, e_n 是 V 的一个基, 所以存在标量 a_1, \dots, a_n 满足

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

因为 e_1, \dots, e_n 是规范正交的, 所以将上式两端同时与 e_k 作内积, 可得 $\langle \nu, e_k \rangle = a_k$. 于是 (a) 成立

由 (a) 和 6.24 立刻可知 (b) 成立

将 u 同时与式(a) 两侧作内积, 并利用内积在第二个位置上的共轭线性【6.6 (d) 和 6.6 (e)】即可得 (c)

6.31 求出一个线性组合的系数

设我们想将向量 $(1, 2, 4, 7) \in \mathbf{F}^4$ 写成 \mathbf{F}^4 中规范正交基 (见例 6.29)

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

的线性组合. 如果我们用的是非规范正交基, 我们一般需要解含有四个方程、四个未知数的方程组来得到系数. 而此处我们用的是规范正交基, 所以只需四个内积并利用 6.30 (a), 即可得 $(1, 2, 4, 7)$ 等于

$$7\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) - 4\left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right) + \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right) + 2\left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)$$

既然我们懂得了规范正交基很有用, 那么我们怎么求出它们呢? 例如, 带有例 6.3 (c) 中内积的 $\mathcal{P}_m(\mathbf{R})$ 具有规范正交基吗? 接下来这个结论会将我们引向这些问题的答案

下面证明中所用的算法称作格拉姆-施密特过程. 它给出了将一个线性无关组转化成与之张成空间相同的规范正交组的方法

约根·格拉姆 (Jørgen Gram, 1850-1916) 和埃哈德·施密特 (Erhard Schmidt, 1876-1959) 推广了这个构造规范正交组的算法

6.32 格拉姆-施密特过程

设 ν_1, \dots, ν_m 是 V 中的线性无关向量组. 令 $f_1 = \nu_1$. 对 $k = 2, \dots, m$, 依次定义 f_k 为

$$f_k = \nu_k - \frac{\langle \nu_k, f_1 \rangle}{\|f_1\|^2} f_1 - \dots - \frac{\langle \nu_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1}.$$

对每个 $k = 1, \dots, m$, $e_k = \frac{f_k}{\|f_k\|}$ 那么 e_1, \dots, e_m 是 V 中的规范正交向量组, 且对每个 $k = 1, \dots, m$ 满足

$$\text{span}(\nu_1, \dots, \nu_k) = \text{span}(e_1, \dots, e_k).$$

证

对 k 用归纳法

$k = 1$

注意因为 $e_1 = f_1/\|f_1\|$ 所以我们有 $\|e_1\| = 1$; 另外, $\text{span}(\nu_1) = \text{span}(e_1)$ 因为 ν_1 是 ν_1 的非零倍

$1 < k \leq m$

设由 6.32 构造出的组 e_1, \dots, e_{k-1} 是规范正交组且满足下式:

$$\text{span}(\nu_1, \dots, \nu_{k-1}) = \text{span}(e_1, \dots, e_{k-1}).$$

因为 ν_1, \dots, ν_m 是线性无关的, 所以我们有 $\nu_k \notin \text{span}(\nu_1, \dots, \nu_{k-1})$. 于是 $\nu_k \notin \text{span}(e_1, \dots, e_{k-1}) = \text{span}(f_1, \dots, f_{k-1})$, 这意味着 $f_k \neq 0$. 于是 6.32 中定义 e_k 时就不会出现除以零的问题. 将一个向量除以其范数会得到范数为 1 的向量, 因而 $\|e_k\| = 1$. 令 $j \in \{1, \dots, k-1\}$. 那么

$$\begin{aligned}\langle e_k, e_j \rangle &= \frac{1}{\|f_k\| \|f_j\|} \langle f_k, f_j \rangle \\ &= \frac{1}{\|f_k\| \|f_j\|} \nu_k - \frac{\langle \nu_k, f_1 \rangle}{\|f_1\|^2} f_1 - \dots - \frac{\langle \nu_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1}, f_j \\ &= \frac{1}{\|f_k\| \|f_j\|} (\langle \nu_k, f_j \rangle - \langle \nu_k, f_j \rangle) \\ &= 0\end{aligned}$$

于是 e_1, \dots, e_k 是规范正交组

由 6.32 给出的 e_k 定义, 可见 $\nu_k \in \text{span}(e_1, \dots, e_k)$. 将其与式(6.33)相结合可得

$$\text{span}(\nu_1, \dots, \nu_k) \subseteq \text{span}(e_1, \dots, e_k).$$

上式中的两个组都是线性无关的 (各 ν 线性无关是由于前提条件, 各 e 线性无关是由于它们规范正交和 6.25). 于是这两个子空间的维数都为 k , 因此它们相等, 这就完成了归纳步骤的证明, 由此完成了整个证明

6.34 $\mathcal{P}_2(\mathbf{R})$ 的一个规范正交基

$\mathcal{P}_2(\mathbf{R})$ 上定义内积为: 对于所有 $p, q \in \mathcal{P}_2(\mathbf{R})$,

$$\langle p, q \rangle = \int_{-1}^1 pq.$$

这样 $\mathcal{P}_2(\mathbf{R})$ 便成为内积空间. 我们知道 $1, x, x^2$ 是 $\mathcal{P}_2(\mathbf{R})$ 的基, 但不是规范正交基. 我们将通过把格拉姆-施密特过程应用在 $\nu_1 = 1, \nu_2 = x, \nu_3 = x^2$ 上来求出 $\mathcal{P}_2(\mathbf{R})$ 的一个规范正交基

首先, 取 $f_1 = \nu_1 = 1$. 于是 $\|f_1\|^2 = \int_{-1}^1 1 = 2$. 则由 6.32 中的公式得

$$f_2 = \nu_2 - \frac{\langle \nu_2, f_1 \rangle}{\|f_1\|^2} f_1 = x - \frac{\langle x, 1 \rangle}{\|f_1\|^2} = x$$

式中最后一个等号成立是由于 $\langle x, 1 \rangle = \int_{-1}^1 t dt = 0$

由上述 f_2 的表达式得 $\|f_2\|^2 = \int_{-1}^1 t^2 dt = \frac{2}{3}$. 现在由 6.32 中的公式得

$$f_3 = \nu_3 - \frac{\nu_3, f_1}{\|f_1\|^2} f_1 - \frac{\nu_3, f_2}{\|f_2\|^2} f_2 = x^2 - \frac{1}{2} \langle x^2, 1 \rangle - \frac{3}{2} \langle x^2, x \rangle x = x^2 - \frac{1}{3}$$

由上述 f_3 的表达式

$$\|f_3\|^2 = \int_{-1}^1 \left(t^2 - \frac{1}{3}\right)^2 dt = \int_{-1}^1 \left(t^4 - \frac{2}{3}t^2 + \frac{1}{9}\right) dt = \frac{8}{45}$$

现在将 f_1, f_2, f_3 都除以其范数, 可得规范正交组

$$\sqrt{\frac{1}{2}}, \sqrt{\frac{3}{2}}x, \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right)$$

这个规范正交组的长度是 3, 恰等于 $\mathcal{P}_2(\mathbf{R})$ 的维数. 因此该规范正交组就是 $\mathcal{P}_2(\mathbf{R})$ 的规范正交基 (由 6.28)

现在我们可以回答有关规范正交基存在性的问题了

6.35 规范正交基的存在性

每个有限维内积空间都有规范正交基

证

设 V 是有限维的. 选取 V 的一个基, 对其应用格拉姆-施密特过程 (6.32), 可得长为 $\dim V$ 的规范正交组. 由 6.28, 这个规范正交组就是 V 的一个规范正交基

有时我们不仅要知道规范正交基存在, 还要知道每个规范正交组都能被扩充成一个规范正交基. 在下面的推论中, 格拉姆-施密特过程表明, 这样的扩充总是可以完成的

6.36 每个规范正交组都可被扩充为规范正交基

设 V 是有限维的. 那么 V 中每个规范正交向量组都能被扩充为 V 的一个规范正交基

证

设 e_1, \dots, e_m 是 V 中的一个规范正交向量组. 那么 e_1, \dots, e_m 是线性无关的 (由 6.25) 因此该组可被扩充为 V 的一个基 $e_1, \dots, e_m, \nu_1, \dots, \nu_n$ (见 2.32). 现在对 $e_1, \dots, e_m, \nu_1, \dots, \nu_n$ 应用格拉姆-施密特过程 (6.32), 可得规范正交组

$$e_1, \dots, e_m, f_1, \dots, f_n$$

格拉姆-施密特过程所用的公式保持组中前 m 个向量不变, 因为这些向量已经是规范正交的了由 6.28 可知上述组是 V 的规范正交基

上章中, 我们给出了一个算子关于某个基具有上三角矩阵的充要条件 (见 5.44). 既然我们现在研究内积空间, 我们就想知道是否存在一个规范正交基, 使得算子关于它存在上三角矩阵. 下面结论说明, 算子关于某规范正交基具有上三角矩阵的条件, 和关于一个任取的基具有上三角矩阵的条件是一样的

6.37 关于某个规范正交基有上三角矩阵

V 有限维, $T \in \mathcal{L}(V)$ 关于 V 的某个规范正交基有上三角矩阵当且仅当 T 的最小多项式等于 $(z - \lambda_1) \cdots (z - \lambda_m)$, 其中 $\lambda_1, \dots, \lambda_m \in \mathbf{F}$

证

设 T 关于 V 的某个基 ν_1, \dots, ν_n 具有上三角矩阵. 于是对各 $k = 1, \dots, n$, $\text{span}(\nu_1, \dots, \nu_k)$ 在 T 下是不变的 (见 5.39)

对 ν_1, \dots, ν_n 应用格拉姆-施密特过程, 可得 V 的一个规范正交基 e_1, \dots, e_n . 因为

$$\text{span}(e_1, \dots, e_k) = \text{span}(\nu_1, \dots, \nu_k)$$

对每个 k 都成立 (见 6.32), 所以我们可推断出, 对各 $k = 1, \dots, n$, $\text{span}(e_1, \dots, e_k)$ 在 T 下是不变的. 从而由 5.39, T 关于规范正交基 e_1, \dots, e_n 具有上三角矩阵. 现在利用 5.44 即可完成证明

对复向量空间, 下面结论是上述结论的重要应用. 舒尔定理适用于多个算子的版本见于习题 20
伊赛·舒尔 (Issai Schur, 1875-1941) 于 1909 年发表了下述结论的证明

6.38 舒尔定理

有限维复内积空间上的每个算子都关于某个规范正交基有上三角矩阵

证

由代数基本定理的版本二 (4.13) 和 6.37 得证

内积空间上的线性泛函

因为映射至标量域 \mathbf{F} 的线性映射扮演特殊的角色，所以我们在 3F 节中，为它们以及它们构成的向量空间起了特殊的名称。因为你可能跳过 3F 节，所以下面重述一下这两个定义

6.39 线性泛函，对偶空间

V 上的一个线性泛函是一个从 V 到 \mathbf{F} 的线性映射

V 的对偶空间记作 V' ，是 V 上全体线性泛函所构成的向量空间。换言之， $V' = \mathcal{L}(V, \mathbf{F})$

6.40 \mathbf{F}^3 上的线性泛函

定义为

$$\varphi(z_1, z_2, z_3) = 2z_1 - 5z_2 + z_3$$

的函数 $\varphi: \mathbf{F}^3 \rightarrow \mathbf{F}$ 是 \mathbf{F}^3 上的线性泛函。我们可将该泛函写成如下形式：对每个 $z \in \mathbf{F}^3$ ，

$$\varphi(z) = \langle z, w \rangle$$

其中 $w = (2, -5, 1)$

6.41 $\mathcal{P}_5(\mathbf{R})$ 上的线性泛函

定义为

$$\varphi(p) = \int_{-1}^1 p(t) (\cos(\pi t)) dt$$

的函数 $\varphi: \mathcal{P}_5(\mathbf{R}) \rightarrow \mathbf{R}$ 是 $\mathcal{P}_5(\mathbf{R})$ 上的线性泛函

若 $\nu \in V$ ，那么将 u 对应至 u, ν 的映射是 V 上的线性泛函。下面结论说明， V 上每个线性泛函都可写成该形式。例如，对于例 6.40，我们可取 $\nu = (2, -5, 1)$ ，

为纪念弗里杰什·里斯 (Frigyes Riesz, 1880-1956)，下面结论用他的名字命名。里斯在 20 世纪早期证明了几个类似于下述结论的定理

设我们通过定义 $\langle p, q \rangle = \int_{-1}^1 pq$ 使向量空间 $\mathcal{P}_5(\mathbf{R})$ 成为内积空间。令 φ 定义如例 6.41。我们没法一眼看出存在 $q \in \mathcal{P}_5(\mathbf{R})$ 使得

$$\int_{-1}^1 p(t) (\cos(\pi t)) dt = \langle p, q \rangle$$

对每个 $p \in \mathcal{P}_5(\mathbf{R})$ 都成立【我们不能取 $q(t) = \cos(\pi t)$ ，因为该 q 不属于 $\mathcal{P}_5(\mathbf{R})$ 】。下面结论却能给出一个惊人的结果，那就是的确存在多项式 $q \in \mathcal{P}_5(\mathbf{R})$ 使得上式对于所有 $p \in \mathcal{P}_5(\mathbf{R})$ 都成立

6.42 里斯表示定理

V 是有限维， φ 是 V 上的线性泛函。那么存在唯一的向量 $\nu \in V$ ，使得对每个 $u \in V$ 都有

$$\varphi(u) = \langle u, \nu \rangle$$

证

(存在性)

设 e_1, \dots, e_n 是 V 的规范正交基。那么对每个 $u \in V$ 有

$$\begin{aligned} \varphi(u) &= \varphi(\langle u, e_1 \rangle e_1 + \dots + \langle u, e_n \rangle e_n) \\ &= \langle u, e_1 \rangle \varphi(e_1) + \dots + \langle u, e_n \rangle \varphi(e_n) \\ &= \left\langle u, \overline{\varphi(e_1)} e_1 + \dots + \overline{\varphi(e_n)} e_n \right\rangle \end{aligned}$$

其中第一个等号源于 6.30 (a)

从而设

$$\nu = \overline{\varphi(e_1)}e_1 + \cdots + \overline{\varphi(e_n)}e_n$$

故对每个 $u \in V$ ，都有 $\varphi(u) = \langle u, \nu \rangle$

(唯一性)

设 $\nu_1, \nu_2 \in V$ 使得对于每个 $u \in V$ ，有

$$\varphi(u) = \langle u, \nu_1 \rangle = \langle u, \nu_2 \rangle.$$

那么对每个 $u \in V$ ，有

$$0 = \langle u, \nu_1 \rangle - \langle u, \nu_2 \rangle = \langle u, \nu_1 - \nu_2 \rangle.$$

取 $u = \nu_1 - \nu_2$ 就可得 $\nu_1 - \nu_2 = 0$ 。于是 $\nu_1 = \nu_2$

6.44 里斯表示定理的计算演示

设我们想求出多项式 $q \in \mathcal{P}_2(\mathbf{R})$ 使得

$$\int_{-1}^1 p(t)(\cos(\pi t))dt = \int_{-1}^1 pq$$

对所有多项式 $p \in \mathcal{P}_2(\mathbf{R})$ 都成立。为此，我们定义对于 $p, q \in \mathcal{P}_2(\mathbf{R})$ ， $\langle p, q \rangle$ 等于上式右侧，从而使 $\mathcal{P}_2(\mathbf{R})$ 成为内积空间。注意到上式左侧并不等于 p 与函数 $t \mapsto \cos(\pi t)$ 在 $\mathcal{P}_2(\mathbf{R})$ 上的内积，因为后者并不是多项式

定义 $\mathcal{P}_2(\mathbf{R})$ 上的线性泛函 φ 使得对于每个 $p \in \mathcal{P}_2(\mathbf{R})$ 都有

$$\varphi(p) = \int_{-1}^1 p(t)(\cos(\pi t))dt$$

现在，选用例 6.34 中的规范正交基并运用里斯表示定理证明过程中的式 (6.43)，我们就能得出，若 $p \in \mathcal{P}_2(\mathbf{R})$ ，那么 $\varphi(p) = \langle p, q \rangle$ ，其中

$$\begin{aligned} q(x) = & \left(\int_{-1}^1 \sqrt{\frac{1}{2}} \cos(\pi t) dt \right) \sqrt{\frac{1}{2}} + \left(\int_{-1}^1 \sqrt{\frac{3}{2}} t \cos(\pi t) dt \right) \sqrt{\frac{3}{2}} x \\ & + \left(\int_{-1}^1 \sqrt{\frac{45}{8}} \left(t^2 - \frac{1}{3} \right) \cos(\pi t) dt \right) \sqrt{\frac{45}{8}} \left(x^2 - \frac{1}{3} \right) \end{aligned}$$

运用一些微积分的知识，即可解出

$$q(x) = \frac{15}{2\pi^2} (1 - 3x^2)$$

运用相同的步骤，即可说明如果想让 $q \in \mathcal{P}_5(\mathbf{R})$ 使得式 (6.45) 对所有 $p \in \mathcal{P}_5(\mathbf{R})$ 都成立，那么应取

$$q(x) = \frac{105}{8\pi^4} \left((27 - 2\pi^2) + (24\pi^2 - 270)x^2 + (315 - 30\pi^2)x^4 \right)$$

设 V 是有限维的且 φ 是 V 上的线性泛函。那么式(6.43) 给出了使得

$$\varphi(u) = \langle u, \nu \rangle$$

对所有 $u \in V$ 都成立的向量 ν 的公式。具体而言，我们有

$$\nu = \overline{\varphi(e_1)}e_1 + \cdots + \overline{\varphi(e_n)}e_n$$

上式右侧似乎同时依赖于规范正交基 e_1, \dots, e_n 与 φ 。然而，6.42 告诉我们， ν 由 φ 唯一确定于是上式右端同样与 V 的规范正交基 e_1, \dots, e_n 的选取无关

里斯表示定理的另外两种不同证法见于6.58 和6C 节习题13

6C 正交补和最小化问题

正交补

6.46 正交补

V 的子集 U 的正交补 U^\perp 是与 U 中的每个向量都正交的所有 V 中向量所构成的集合

正交补 U 同时依赖于 V 和 U 。然而，我们总会由上下文明确得知内积空间 V 的选取，于是我们可从记号中省去它

6.47 正交补

若 $V = \mathbf{R}^3$, U 是仅包含点 $(2, 3, 5)$ 的 V 的子集，那么 U 就是平面 $\{(x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = 0\}$

若 $V = \mathbf{R}^3$ ， U 是平面 $\{(x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = 0\}$ ，那么 U 就是直线 $\{(2t, 3t, 5t) : t \in \mathbf{R}\}$ 更一般地说，若 U 是 \mathbf{R}^3 中过原点的平面，那么 U 是过原点且垂直于 U 的直线

若 U 是 \mathbf{R}^3 中过原点的直线，那么 U 是过原点且垂直于 U 的平面

若 $V = \mathbf{F}^5$ 且 $U = \{(a, b, 0, 0, 0) \in \mathbf{F}^5 : a, b \in \mathbf{F}\}$ ，那么

$$U = \{(0, 0, x, y, z) \in \mathbf{F}^5 : x, y, z \in \mathbf{F}\}$$

若 $e_1, \dots, e_m, f_1, \dots, f_n$ 是 V 的规范正交基，那么

$$(\text{span}(e_1, \dots, e_m))^\perp = \text{span}(f_1, \dots, f_n).$$

我们首先介绍一些由定义即可推出的简单结论

6.48 正交补的性质

- (a) 若 U 是 V 的子集，那么 U 是 V 的子空间
- (b) $\{0\}^\perp = V$
- (c) $V = \{0\}^\perp$
- (d) 若 U 是 V 的子集，那么 $U \cap U^\perp \subseteq \{0\}$
- (e) 若 U 和 W 是 V 的子集且 $U \subseteq W$ ，那么 $W^\perp \subseteq U^\perp$

证

(a) 设 U 是 V 的子集。那么对每个 $u \in U$ 有 $\langle u, 0 \rangle = 0$ ，从而 $0 \in U^\perp$ 。设 $\nu, w \in U^\perp$ 。若 $u \in U$ ，那么

$$\langle u, \nu + w \rangle = \langle u, \nu \rangle + \langle u, w \rangle = 0 + 0 = 0$$

于是 $\nu + w \in U^\perp$ ，表明 U^\perp 对于加法封闭

类似地，设 $\lambda \in \mathbf{F}$ 且 $\nu \in U^\perp$ 。若 $u \in U$ ，那么

$$\langle u, \lambda \nu \rangle = \overline{\lambda} \langle u, \nu \rangle = \overline{\lambda} \cdot 0 = 0$$

于是 $\lambda \nu \in U^\perp$ ，表明 U^\perp 对于标量乘法封闭。于是 U^\perp 是 V 的子空间

(b) 设 $\nu \in V$ 。那么 $\langle 0, \nu \rangle = 0$ ，这表明 $\nu \in \{0\}^\perp$ 。于是 $\{0\}^\perp = V$

(c) 设 $\nu \in V$ 。那么 $\langle \nu, \nu \rangle = 0$ ，这表明 $\nu = 0$ 。于是 $V = \{0\}^\perp$

(d) 设 U 是 V 的子集，且 $u \in U \cap U^\perp$ 。那么 $\langle u, u \rangle = 0$ ，这表明 $u = 0$ 。于是 $U \cap U^\perp \subseteq \{0\}$

(e) 设 U 和 W 是 V 的子集且 $U \subseteq W$ 。设 $\nu \in W^\perp$ 。那么对每个 $u \in U$ 有 $\langle u, \nu \rangle = 0$ ，这表明对每个 $u \in U$ 也有 $\langle u, \nu \rangle = 0$ 。因而 $\nu \in U^\perp$ 。于是 $W^\perp \subseteq U^\perp$

回忆一下，若 U 和 W 是 V 的子空间，那么如果 V 中每个元素都能被唯一地写成 U 中的一个向量加上 W 中的一个向量的形式，那么 V 就是 U 和 W 的直和（记为 $V = U \oplus W$ ）（见 1.41）。这还等价于 $V = U + W$ 且 $U \cap W = \{0\}$ （见 1.46）

下面结论表明， V 的每个有限维子空间都能引出 V 的一个很自然的直和分解。习题 16 举例说明了如果没有子空间 U 是有限维这个前提条件，下面的结论可能没法成立

6.49 子空间及其正交补的直和

U 是 V 的有限维子空间。那么

$$V = U \oplus U^\perp.$$

证

(+)

设 $\nu \in V$ ，令 e_1, \dots, e_m 是 U 的规范正交基。我们将 ν 写成 U 中的向量与正交于 U 的向量之和

我们有

$$\nu = \underbrace{\langle \nu, 1 \rangle_1 + \cdots + \langle \nu, m \rangle_m}_u + \underbrace{\nu - \langle \nu, 1 \rangle_1 - \cdots - \langle \nu, m \rangle_m}_w$$

令 u 和 w 定义如上 (6.26 的证明中也曾这样做). 因各 $k \in U$, $u \in U$

因 e_1, \dots, e_m 是规范正交组, 所以对各 $k = 1, \dots, m$

$$\begin{aligned}\langle w, e_k \rangle &= \langle \nu, e_k \rangle - \langle \nu, e_k \rangle \\ &= 0\end{aligned}$$

从而 w 与 $\text{span}(e_1, \dots, e_m)$ 中的每个向量都正交, 这表明 $w \in U^\perp$. 因此我们可写出 $\nu = u + w$, 其中 $u \in U$ 且 $w \in U^\perp$, 这就完成了 $V = U + U^\perp$ 的证明

($\cap = \{0\}$)

由 6.48 (d) 得

现在我们可知如何由 $\dim U$ 计算出 $\dim U^\perp$ 了

6.51 正交补的维数

设 V 是有限维的, U 是 V 的子空间. 那么

$$\dim U^\perp = \dim V - \dim U$$

证

由 6.49 和 3.94 可得

下面结论是6.49的一个重要推论

6.52 正交补的正交补

设 U 是 V 的一个有限维子空间, 则

$$U^{\perp\perp} = U$$

证

(\subseteq)

为此, 设 $u \in U^\perp$. 那么对每个 $w \in U$ 都有 $\langle u, w \rangle = 0$ (由 U^\perp 的定义即得). 因为 u 与 U 中的每个向量都正交, 所以我们有 $u \in (U^\perp)^\perp$, 这就证明了式(6.53)

(另一方向)

设 $\nu \in (U^\perp)^\perp$. 由6.49, 我们可写出 $\nu = u + w$, 其中 $u \in U$ 且 $w \in U^\perp$, 则有 $\nu - u = w \in U^\perp$. 因为 $\nu \in (U^\perp)^\perp$ 且 $u \in U$ 【由式(6.53)】, 故可得 $\nu - u \in (U^\perp)^\perp$. 于是 $\nu - u \in U \cap (U^\perp)^\perp$, 这表明 $\nu - u = 0$ 【由 6.48 (d)】, 进而表明 $\nu = u$, 因此 $\nu \in U$. 因此 $(U^\perp)^\perp \subseteq U$, 结合式(6.53)得证

设 U 是 V 的一个子空间, 且我们想证明 $U^{\perp\perp} = U$. 在某些情况下, 最简单的证法是先证习题 16 (a) 表明, 去掉 U 是有限维这个前提条件后, 下面结论就不成立了

明与 U 正交的向量只有0, 然后运用下述结论. 例如, 下述结论在习题4中就很用处

6.54 定理

U 是 V 的有限维子空间, 则

$$U = \{0\} \iff U = V$$

证

先设 $U = \{0\}$. 那么由 6.52, $U = (U) = \{0\} = V$, 得证反之, 若 $U = V$, 则由 6.48 (c) 知 $U = V = \{0\}$ 我们现在为 V 的每个有限维子空间 U 定义算子 P_U

6.55 正交投影

设 U 是 V 的一个有限维子空间. 将 V 映成 U 的正交投影是定义如下的算子 $P_U \in \mathcal{L}(V)$: 对每个 $\nu \in V$, 将其写成 $\nu = u + w$, 其中 $u \in U$ 且 $w \in U^\perp$, 然后令 $P_U \nu = u$

6.49 给出的直和分解式 $V = U \oplus U^\perp$ 表明, $\nu \in V$ 可以被唯一表示为 $\nu = u + w$ (其中 $u \in U$ 且 $w \in U^\perp$) 的形式. 由此可见 $P_U \nu$ 的定义是完善的. 请看 6.61 的证明中所附的示意图——你应该将这张描述 $P_U \nu$ 的图像牢记于心

6.56 映成一维子空间的正交投影

设 $u \in V$ 且 $u \neq 0$, U 是 V 的一维子空间, 定义为 $U = \text{span}(u)$ 若 $\nu \in V$, 那么

$$\nu = \frac{\langle \nu, u \rangle}{\|u\|^2} u + \left(\nu - \frac{\langle \nu, u \rangle}{\|u\|^2} u \right)$$

式中右侧第一项属于 $\text{span}(u)$ (从而属于 U), 第二项正交于 u (从而属于 U^\perp). 于是 $P_U \nu$ 等于上式右侧第一项. 换言之, 对每个 $\nu \in V$,

$$P_U \nu = \frac{\langle \nu, u \rangle}{\|u\|^2} u$$

取 $\nu = u$, 上式变为 $P_U u = u$; 取 $\nu \in \{u\}^\perp$, 上式则变为 $P_U \nu = 0$. 这是下面结果中(b)与(c)的具体例子

6.57 正交投影 P_U 的性质

U 是 V 的有限维子空间, 则

- (a) $P_U \in \mathcal{L}(V)$;
- (b) 对每个 $u \in U$, 都有 $P_U u = u$;
- (c) 对每个 $w \in U^\perp$, 都有 $P_U w = 0$;
- (d) $\text{range } P_U = U$;
- (e) $\text{null } P_U = U^\perp$;
- (f) 对每个 $\nu \in V$, 都有 $\nu - P_U \nu \in U^\perp$;
- (g) $P_U^2 = P_U$;
- (h) 对每个 $\nu \in V$, 都有 $\|P_U \nu\| \leq \|\nu\|$;
- (i) 若 e_1, \dots, e_m 是 U 的一个规范正交基且 $\nu \in V$, 那么

$$P_U \nu = \langle \nu, e_1 \rangle e_1 + \dots + \langle \nu, e_m \rangle e_m$$

证

(a) 为证明 P_U 是 V 上的线性映射, 设 $\nu_1, \nu_2 \in V$. 将 ν_1, ν_2 写成

$$\nu_1 = u_1 + w_1 \quad \nu_2 = u_2 + w_2,$$

其中 $u_1, u_2 \in U$ 且 $w_1, w_2 \in U^\perp$. 从而 $P_U \nu_1 = u_1$ 且 $P_U \nu_2 = u_2$. 则有

$$\nu_1 + \nu_2 = (u_1 + u_2) + (w_1 + w_2),$$

其中 $u_1 + u_2 \in U$ 且 $w_1 + w_2 \in U^\perp$. 于是

$$P_U(\nu_1 + \nu_2) = u_1 + u_2 = P_U\nu_1 + P_U\nu_2.$$

类似地, 设 $\lambda \in \mathbf{F}$ 且 $\nu \in V$. 将 ν 写成 $\nu = u + w$, 其中 $u \in U$ 且 $w \in U^\perp$. 那么 $\lambda\nu = \lambda u + \lambda w$, 其中 $\lambda u \in U$, $\lambda w \in U^\perp$. 因此 $P_U(\lambda\nu) = \lambda u = \lambda P_U\nu$

因此 P_U 是从 V 到 V 的线性映射

(b) 设 $u \in U$. 我们可将 u 写为 $u = u + 0$, 其中 $u \in U$ 且 $0 \in U^\perp$. 于是 $P_U u = u$

(c) 设 $w \in U^\perp$. 我们可将 w 写为 $w = 0 + w$, 其中 $0 \in U$ 且 $w \in U^\perp$. 于是 $P_U w = 0$

(d) 由 P_U 的定义得 $\text{range } P_U \subseteq U$. 又由 (b) 可知 $U \subseteq \text{range } P_U$. 于是 $\text{range } P_U = U$

(e) 由 (c) 可知 $U^\perp \subseteq \text{null } P_U$. 为证明这个包含关系反过来也成立, 注意到若 $\nu \in \text{null } P_U$, 那么 6.49 给出的分解式必为 $\nu = 0 + \nu$, 其中 $0 \in U$ 且 $\nu \in U^\perp$. 于是 $\text{null } P_U \subseteq U^\perp$

(f) 若 $\nu \in V$, 且有 $\nu = u + w$ (其中 $u \in U$ 且 $w \in U^\perp$), 那么

$$\nu - P_U\nu = \nu - u = w \in U^\perp.$$

(g) 若 $\nu \in V$, 且有 $\nu = u + w$ (其中 $u \in U$ 且 $w \in U^\perp$), 那么

$$(P_U^2)\nu = P_U(P_U\nu) = P_U u = u = P_U\nu.$$

(h) 若 $\nu \in V$, 且有 $\nu = u + w$ (其中 $u \in U$ 且 $w \in U^\perp$), 那么

$$\|P_U\nu\|^2 = \|u\|^2 \leq \|u\|^2 + \|w\|^2 = \|\nu\|^2,$$

式中最后一个等号源于毕达哥拉斯定理(i) 由 6.49 证明中的式 (6.50) 即可得 $P_U\nu$ 的表达式

在上节中我们证明了里斯表示定理 (6.42), 它的关键内容是, 有限维内积空间上的每个线性泛函, 都可由与某个固定向量的内积式来表示. 了解不同的证明往往能带给我们新的理解于是, 我们现在改用正交补 (而非之前所用的规范正交基) 来重新证明里斯表示定理

下面重述的里斯表示定理使我们能将 V 与 V' 等同起来看. 我们只证明下述结论的“映成”部分, 因为“一对一”部分的证明很常规——按 6.42 的证法即可这个新证法背后的直观想法是: 若 $\varphi \in V'$, $\nu \in V$ 且 $\varphi(u) = \langle u, \nu \rangle$ 对所有 $u \in V$ 都成立, 那么 $\nu \in (\text{null } \varphi)$. 而由 6.51 和 3.21, $(\text{null } \varphi)$ 是 V 的一维子空间 (除 $\varphi = 0$ 的平凡情形外). 于是我们可通过取 $(\text{null } \varphi)$ 中的任一非零元素并将其与一适当的标量相乘而得到 ν . 下面证明正是这样做的

6.58 里斯表示定理再讨论

设 V 是有限维的. 对每个 $\nu \in V$, 定义 $\varphi_\nu \in V'$ 为: 对每个 $u \in V$,

$$\varphi_\nu(u) = \langle u, \nu \rangle.$$

那么 $\nu \mapsto \varphi_\nu$ 是将 V 映成 V' 的一对一函数

证

为证 $\nu \mapsto \varphi_\nu$ 是满射, 设 $\varphi \in V'$. 若 $\varphi = 0$, 那么 $\varphi = \varphi_0$. 从而假设 $\varphi \neq 0$. 因此 $\text{null } \varphi \neq V$, 这表明 $(\text{null } \varphi) \neq \{0\}$ (由 6.49, 其中取 $U = \text{null } \varphi$)

令 $w \in (\text{null } \varphi)$ 且 $w \neq 0$. 令注意: 若 $\mathbf{F} = \mathbf{R}$, 则函数 $\nu \mapsto \varphi_\nu$ 是从 V 到 V' 的线性映射. 然而, 若 $\mathbf{F} = \mathbf{C}$, 则该函数不是线性的, 因为若 $\lambda \in \mathbf{C}$, 则 $\varphi_{\lambda\nu} = \bar{\lambda}\varphi_\nu$

$$\nu = \frac{\overline{\varphi(w)}}{\|w\|^2} w.$$

那么 $\nu \in (\text{null } \varphi)$, 并且 $\nu \neq 0$ (因为 $w \notin \text{null } \varphi$)

在式(6.59) 两侧同取范数

$$\|\nu\| = \frac{|\varphi(w)|}{\|w\|}$$

将 φ 同时作用于式 (6.59) 两端并利用式 (6.60), 可得

$$\varphi(\nu) = \frac{|\varphi(w)|^2}{\|w\|^2} = \|\nu\|^2$$

现在设 $u \in V$. 利用上式

$$u = \left(u - \frac{\varphi(u)}{\varphi(\nu)} \nu \right) + \frac{\varphi(u)}{\|\nu\|^2} \nu$$

式中带括号的一项属于 $\text{null } \varphi$, 因此与 ν 正交. 于是将此式两边都与 ν 作内积可得

$$\langle u, \nu \rangle = \frac{\varphi(u)}{\|\nu\|^2} \langle \nu, \nu \rangle = \varphi(u)$$

因此 $\varphi = \varphi_\nu$, 表明 $\nu \mapsto \varphi_\nu$ 是满射, 则原命题得证

里斯表示定理的另一种不同证法见于习题13

最小化问题

常遇到下面问题: 给定 V 的子空

这个最小化问题的解特别简洁, 由此产生了间 U 与一点 $\nu \in V$, 求出一 $u \in U$ 使得 $\| \nu - u \|$ 尽可能小. 下面结论表明, $u = P_U \nu$ 是该最小化问题的唯一解

6.61 到子空间的最短距离

U 是 V 的有限维子空间, $\nu \in V$ 且 $u \in U$. 那么 $\| \nu - P_U \nu \| \leq \| \nu - u \|$, 取等当且仅当 $u = P_U \nu$

证

找不到“images/b9b50fdb89101d1b8ffac61422e5ffd56868e94b36e1d74e8b8da083fcf0ef81.jpg”。

$$\begin{aligned} \| \nu - P_U \nu \|^2 &\leq \| \nu - P_U \nu \|^2 + \| P_U \nu - u \|^2 \\ &= \| (\nu - P_U \nu) + (P_U \nu - u) \|^2 \\ &= \| \nu - u \|^2, \end{aligned}$$

其中第一行成立是因为 $0 \leq \| P_U \nu - u \|^2$, 第二行源于毕达哥拉斯定理【可以运用它是因为由6.57 (f) 知 $\nu - P_U \nu \in U$, 又有 $P_U \nu - u \in U$ 】, 简单计算即知第三行成立. 两边开平方根就得到了我们欲证的不等式

$P_U \nu$ 是 U 中距离 ν 最近的点

上面证明的不等式取得等号当且仅当式 (6.62) 取得等号, 这等价于 $\| P_U \nu - u \| = 0$, 又等价于 $u = P_U \nu$

我们常将上述结论与公式6.57 (i) 相结合来计算出最小化问题的显式解, 如下例所示

6.63 利用线性代数来逼近正弦函数

假如我们要求出次数不高于5的实系数多项式 u , 使其在区间 $[-\pi, \pi]$ 上尽可能逼近正弦函数, 这里的“逼近”指的是使

$$\int_{-\pi}^{\pi} |\sin x - u(x)|^2 dx$$

尽可能小

令 $C[-\pi, \pi]$ 表示定义在 $[-\pi, \pi]$ 上的全体连续实值函数所构成的实内积空间, 其上内积定义为

$$\langle f, g \rangle = \int_{-\pi}^{\pi} fg.$$

令 $\nu \in C[-\pi, \pi]$ 是定义为 $\nu(x) = \sin x$ 的函数. 令 U 表示由次数不高于5的所有实系数多项式所构成的 $C[-\pi, \pi]$ 的子空间. 现在可将问题重新表述为:

求出 $u \in U$ 使得 $\|\nu - u\|$ 尽可能小

为了计算出这个逼近问题的解，首先对

一台可计算积分的电脑对此很有用处

U 的基 $1, x, x^2, x^3, x^4, x^5$ 用格拉姆-施密特过程【所用内积定义如式 (6.64)】，得到 U 的一个规范正交基 $e_1, e_2, e_3, e_4, e_5, e_6$. 接着，再采用式(6.64) 给出的内积并利用6.57 (i) 来计算 $P_U \nu$ (其中 $m = 6$) . 计算得 $P_U \nu$ 就是定义如下的函数 u :

$$u(x) = 0.987862x - 0.155271x^3 + 0.00564312x^5,$$

这里我们把精确解中出现的含 π 系数用足够准确的十进制近似值来代替. 由 6.61, 上述多项式 u 是利用不超过5 次的多项式在区间 $[-\pi, \pi]$ 上对正弦函数的最佳逼近 (此处“最佳逼近”是就 $\int_{-\pi}^{\pi} |\sin x - u(x)|^2 dx$ 最小而言的)

为了观察逼近程度有多高，我们在下图中同时绘出正弦函数和式 (6.65) 给出的逼近函数 u 在区间 $[-\pi, \pi]$ 上的图像

找不到“images/dfc4bc405dcf9da2f8484346759f3c677d44459845b10de8c53b20123f26c6f1.jpg”。

$[-\pi, \pi]$ 区间上正弦函数 (红线) 和式(6.65) 给出的最佳逼近五次多项式 u (蓝线) 的图像

逼近式 (6.65) 准确到使得两函数的图像几乎一致——我们用肉眼只能看到一个图像! 此处红色图像几乎覆盖掉了蓝色图像. 如果你在电子设备上查看这张图，你可以在 π 或 $-\pi$ 附近将它放大至原先的 400% , 以便观察两函数图像的细微差别

用五次多项式来逼近正弦函数的另一种著名方法，就是利用定义如下的泰勒多项式 (Tay-lor polynomial) p :

$$p(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!}.$$

为了观察逼近程度有多高，我们在下图中同时绘出正弦函数和泰勒多项式 p 在区间 $[-\pi, \pi]$ 上的图像

找不到“images/216cece12e45c12a0c92b3fe06ea607834bdee344225ff29ce9d6e680a1b74ad.jpg”。

$[-\pi, \pi]$ 区间上正弦函数 (红线) 和式(6.66) 所示的泰勒多项式 (蓝线) 的图像

在 x 接近 0 时，泰勒多项式能很好地逼近 $\sin x$. 然而上述图像表明，对于 $|x| > 2$, 泰勒多项式就不那么准确了，特别是与式(6.65) 相比. 例如，取 $x = 3$, 逼近式 (6.65) 对于 $\sin 3$ 的估计的误差约为 0.001, 而泰勒多项式 (6.66) 对于 $\sin 3$ 的估计的误差约是 0.4 . 于是在 $x = 3$ 处，泰勒多项式的误差比式 (6.65) 误差大数百倍. 线性代数帮助我们找到了正弦函数的新的逼近方法，这个方法改进了我们在微积分中学过的方法!

伪逆

设 $T \in \mathcal{L}(V, W)$ 且 $w \in W$

考虑问题：求出 $\nu \in V$, $T\nu = w$

如若 $V = \mathbf{F}^n$ 且 $W = \mathbf{F}^m$, 那么上式可表示含有 m 个线性方程， n 个未知数 ν_1, \dots, ν_n 的方程组，其中 $\nu = (\nu_1, \dots, \nu_n)$

若 T 可逆，那么上述方程的唯一解就是 $\nu = T^{-1}w$. 然而，若 T 不可逆，那么对于某些 $w \in W$, 上述方程可能无解；对于另外一些 $w \in W$, 上述方程可能有无穷多解

当 T 不可逆时，我们仍可尝试尽可能好地处理上述方程. 例如，若上述方程无解，那么我们就求解方程 $T\nu - w = 0$, 而是尝试解出使 $\|T\nu - w\|$ 尽可能小的 $\nu \in V$. 又例如，如果有无穷多个 $\nu \in V$ 满足上述方程，那么我们可以选出这些解中使得 $\|\nu\|$ 最小的那个

伪逆就为我们尽可能好地求解上述方程提供了有力工具，即便 T 并不可逆. 为定义伪逆，我们需要下面的结论

在接下来的两个证明中，我们将不加注解地使用这个结论：若 V 是有限维的，且 $T \in \mathcal{L}(V, W)$, 那么 $\text{null } T$, $(\text{null } T)$ 与 $\text{range } T$ 都是有限维的

6.67 限制线性映射以获得既单又满的映射

设 V 是有限维的，且 $T \in \mathcal{L}(V, W)$. 那么 $T|_{(\text{null } T)}$ 是将 $(\text{null } T)$ 映成 $\text{range } T$ 的单射

证

设 $\nu \in (\text{null } T)$ 且 $T|_{(\text{null } T)}\nu = 0$. 因此 $T\nu = 0$, 进而 $\nu \in (\text{null } T) \cap (\text{null } T)$, 从而 $\nu = 0$ 【由 6.48 (d)】. 因此 $\text{null } T|_{(\text{null } T)} = \{0\}$, 这表明 $T|_{(\text{null } T)}$ 是单射

显然, $\text{range } T|_{(\text{null } T)} \subseteq \text{range } T$. 为证明这个包含关系反过来也成立, 设 $w \in \text{range } T$. 因此存在 $\nu \in V$ 满足 $w = T\nu$. 由6.49, 存在 $u \in \text{null } T$ 和 $x \in (\text{null } T)$ 使得 $\nu = u + x$. 则

$$T|_{(\text{null } T)} x = Tx = T\nu - Tu = w - 0 = w$$

这表明 $w \in \text{range } T|_{(\text{null } T)}$. 因此 $\text{range } T \subseteq \text{range } T|_{(\text{null } T)}$, 这就证明了 $\text{range } T|_{(\text{null } T)} = \text{range } T$

下面定义 (乃至之后的章节) 中, 将 $T|_{(\text{null } T)}$ 看成把 $(\text{null } T)$ 映成 $\text{range } T$ 的可逆线性映射 (上面证明了这一点)

6.68 伪逆

设 V 是有限维的, $T \in \mathcal{L}(V, W)$. T 的伪逆 $T \in \mathcal{L}(W, V)$ 是定义如下的从 W 到 V 的线性映射: 对每个 $w \in W$, $Tw = (T|_{(\text{null } T)})^{-1} P_{\text{range } T} w$

回忆一下, 若 $w \in (\text{range } T)$, 则 $P_{\text{range } T} w = w$; 若 $w \in \text{range } T$, 则 $P_{\text{range } T} w = w$. 于是, 如果 $w \in (\text{range } T)$, 则 $Tw = 0$; 若 $w \in \text{range } T$, 则 Tw 是 $(\text{null } T)$ 中唯一满足 $T(Tw) = w$ 的元素
我们将看到, 伪逆看上去很像一般的逆

6.69 伪逆的代数性质

V 有限维, $T \in \mathcal{L}(V, W)$

(a) 若 T 可逆, 则 $T = T^{-1}$

(b) $TT = P_{\text{range } T}$ 是将 W 映成 $\text{range } T$ 的正交投影

(c) $TT = P_{(\text{null } T)}$ 是将 V 映成 $(\text{null } T)$ 的正交投影

证

(a) 设 T 可逆. 那么 $(\text{null } T) = V$, 且 $\text{range } T = W$. 从而 $T|_{(\text{null } T)} = T$ 且 $P_{\text{range } T}$ 是 W 上的恒等算子. 因此 $T = T^{-1}$

(b) 设 $w \in \text{range } T$. 于是

$$TTw = T(T|_{(\text{null } T)})^{-1} w = w = P_{\text{range } T} w.$$

若 $w \in (\text{range } T)$, 那么 $Tw = 0$. 因此 $TTw = 0 = P_{\text{range } T} w$. 于是, 在 $\text{range } T$ 和 $(\text{range } T)$ 上, TT 和 $P_{\text{range } T}$ 都相同. 因此这两个线性映射相等 (由6.49)

(c) 设 $\nu \in (\text{null } T)$. 因为 $T\nu \in \text{range } T$, 所以由 T 定义可知

$$T(T\nu) = (T|_{(\text{null } T)})^{-1}(T\nu) = \nu = P_{(\text{null } T)} \nu.$$

若 $\nu \in \text{null } T$, 那么 $TT\nu = 0 = P_{(\text{null } T)} \nu$. 于是, 在 $(\text{null } T)$ 和 $\text{null } T$ 上, TT 和 $P_{(\text{null } T)}$ 都相同. 因此这两个线性映射相等 (由6.49)

设 $T \in \mathcal{L}(V, W)$. 如果 T 是满射, 那么由上面结论的 (b) 可得, TT 是 W 上的恒等算子伪逆也被称作摩尔-彭罗斯逆 (Moore-Penrose inverse) 子. 如果 T 是单射, 那么由上面结论的(c) 可得, TT 是 V 上的恒等算子. 伪逆的更多代数性质见于习题 19 至23

对于 $T \in \mathcal{L}(V, W)$ 和 $w \in W$, 我们现在回到求出满足方程 $T\nu = w$ 的 $\nu \in V$ 这个问题上来前面提过, 若 T 是可逆的, 那么 $\nu = T^{-1}w$ 就是方程的唯一解, 但若 T 不可逆, 那么 T^{-1} 没有定义. 然而, 伪逆 T 是有定义的. 下面结论的 (a) 就会告诉我们, 取 $\nu = Tw$ 可使 $T\nu$ 尽可能接近 w . 于是伪逆给出了上述方程的所谓最佳拟合 (best fit)

结合下面的 (b) 以及 (a) 中不等式取等号的条件可知, 在所有使得 $T\nu$ 尽可能接近 w 的向量 $\nu \in V$ 中, 向量 Tw 具有最小的范数

6.70 伪逆可给出最佳近似解或最优解

V 有限维, $T \in \mathcal{L}(V, W)$, $w \in W$

(a) $\nu \in V$, 则

$$\|T(Tw) - w\| \leq \|T\nu - w\|$$

当且仅当 $\nu \in Tw + \text{null } T$ 时, 取等

(b) $\nu \in Tw + \text{null } T$, 则

$$Tw \leq \|\nu\|$$

当且仅当 $\nu = Tw$ 时, 上式取得等号

证

(a) 设 $\nu \in V$. 那么

$$T\nu - w = (T\nu - TT w) + (TT w - w).$$

上式第一个括号中的项属于 $\text{range } T$. 因为算子 TT 是将 W 映成 $\text{range } T$ 的正交投影【由6.69 (b)】, 所以上式第二个括号中的项属于 $(\text{range } T)^\perp$ 【见 6.57 (f)】

于是, 由毕达哥拉斯定理可知, 上式第二个括号中的项的范数小于或等于 $\|T\nu - w\|$, 并且当且仅当上式第一个括号中的项等于0 时可取得等号. 因此上述不等式取得等号, 当且仅当 $\nu - Tw \in \text{null } T$, 这等价于 $\nu \in Tw + \text{null } T$. 这就完成了(a)的证明

(b) 设 $\nu \in Tw + \text{null } T$. 因此 $\nu - Tw \in \text{null } T$. 我们有

$$\nu = (\nu - Tw) + Tw.$$

T 的定义表明, $Tw \in (\text{null } T)^\perp$. 于是由毕达哥拉斯定理得 $Tw \leq \nu$, 并且当且仅当 $\nu = Tw$ 时取等

下章将给出 T 的一个计算公式 (见7.78)

6.71 \mathbf{F}^4 到 \mathbf{F}^3 的线性映射的伪逆

设 $T \in \mathcal{L}(\mathbf{F}^4, \mathbf{F}^3)$ 定义为

$$T(a, b, c, d) = (a + b + c, 2c + d, 0).$$

该线性映射既不是单射也不是满射, 但我们仍可计算其伪逆. 为此, 首先注意到 $\text{range } T = \{(x, y, 0) : x, y \in \mathbf{F}\}$. 从而对每个 $(x, y, z) \in \mathbf{F}^3$,

$$P_{\text{range } T}(x, y, z) = (x, y, 0).$$

并且,

$$\text{null } T = \{(a, b, c, d) \in \mathbf{F}^4 : a + b + c = 0\}$$

由 $\text{null } T$ 中的两个向量构成的组 $(-1, 1, 0, 0), (-1, 0, 1, -2)$ 张成 $\text{null } T$, 因为若 $(a, b, c, d) \in \text{null } T$, 那么

$$(a, b, c, d) = b(-1, 1, 0, 0) + c(-1, 0, 1, -2).$$

又由于组 $(-1, 1, 0, 0), (-1, 0, 1, -2)$ 线性无关, 所以该组是 $\text{null } T$ 的一个基

现设 $(x, y, z) \in \mathbf{F}^3$. 那么

$$T(x, y, z) = (T|_{(\text{null } T)})^{-1} P_{\text{range } T}(x, y, z) = (T|_{(\text{null } T)})^{-1}(x, y, 0).$$

上式右侧就是满足 $T(a, b, c, d) = (x, y, 0)$ 和 $(a, b, c, d) \in (\text{null } T)$ 的向量 $(a, b, c, d) \in \mathbf{F}^4$. 换言之, a, b, c, d 必满足下列方程:

$$\begin{aligned} a + b + c &= x \\ 2c + d &= y \\ -a + b &= 0 \\ -a + c - 2d &= 0, \end{aligned}$$

其中前两个方程等价于方程 $T(a, b, c, d) = (x, y, 0)$, 后两个方程则是由于 (a, b, c, d) 正交于 $\text{null } T$ 的基向量 $(-1, 1, 0, 0), (-1, 0, 1, -2)$. 将 x, y 视为常数, 并将 a, b, c, d 视为未知量, 上述方程组便含有四个方程、四个未知数, 解之得

$$a = \frac{1}{11}(5x - 2y), b = \frac{1}{11}(5x - 2y), c = \frac{1}{11}(x + 4y), d = \frac{1}{11}(-2x + 3y).$$

因此由式 (6.72),

$$T(x, y, z) = \frac{1}{11}(5x - 2y, 5x - 2y, x + 4y, -2x + 3y).$$

由上述 T 的计算公式可得，对所有 $(x, y, z) \in \mathbf{F}^3$ ，有 $TT(x, y, z) = (x, y, 0)$ ，这形象阐释了6.69 (b) 中的 $TT = P_{\text{range}T}$ 这个式子