线性代数 (Linear Algebra)

第六章 Orthogonality and Least Squares

§6.1 Inner Product，Length and Orthogonality

内积，长度和正交性

衡 益

2021 年 12 月 21 日，中山大学南校区

Motivation

2

6.1 内积、长度和正交性

设想启动一个巨大的工程，该工程估计持续**10年时间，需要花费大量的人力去构造并计算一个1800000x900000的线性方程组**，这实际上就是美国地质调查局(NGS) 1974年所做的工作，当时准备更新北美地质资料(NAD)个包含268000个仔细测量并标记说明的地点网络，它覆盖整个北美大陆，包括巴拿马地峡、格陵兰岛、夏威夷、维尔京群岛、波多黎哥和其他加勒比海诸岛。

3

6.1 内积、长度和正交性

北美地质资料中记录的经度和纬度范围必须精确到几厘米，其原因是它构成诸如测量，地图，法定边界，国家和区域土地使用计划，像高速公路和公共使用线路等国内工程项目的设计等方面的标准。从上次1927年地名测量调整以来，**至少需要在原始资料中增加200000个新观测点，误差会随时间流逝而不断积累**，并且在某些区域，地球板块以每年5厘米的速度漂移.到1970年，重新检查资料系统的工作已经十分迫切，而且新计划准备选取新的参考观测点。

6.1 内积、长度和正交性

覆盖长达140年的数据资料必须转化为适合计算机运算的格式，且数据本身需要标准化 (例如，地球地壳运动的数学模型，数年前被用于更新测量加利福尼亚的圣安德里亚断层)。此外，测量需要交又检测以确定误差来源是原始数据或是输入计算机的输人错误. 最后的计算包括180万个观测值，每个值的重要性又依赖它的相对精度且涉及一个方程。

5

6.1 内积、长度和正交性

通常意义下，北美地质资料对应的线性方程组没有正常解，只具有最小二乘解. 它用经度和纬度表示参考点以通近180万个实际观测点. 通过对应线性方程组的法方程解出最小二乘解，实际计算包含928735个方程、928735个变量!

6

6.1 内积、长度和正交性

由于得到的法方程对当时的计算机来说规模实在太大.他们通过Helmert分块的技巧将方程组分成小块，将系数矩阵递推分割成越来越小的块，最小块对应的方程组在北美资料中覆盖地理上邻近的500-2000个参考点，图6-1显示美国如何被分成 Helmert块，经过几个中间步骤，小块系统的解最后产生整体928735个变量的所有解。

7

6.1 内积、长度和正交性

图 6-1 美国 Helmert 分块的邻接边界

直到1983年才完成北美资料数据库的更新，三年以后，通过更深入的分析和940小时的计算机计算人类历史上曾经完成的数目最大的最小二乘法问题被完全解出。

8

6.1 内积、长度和正交性

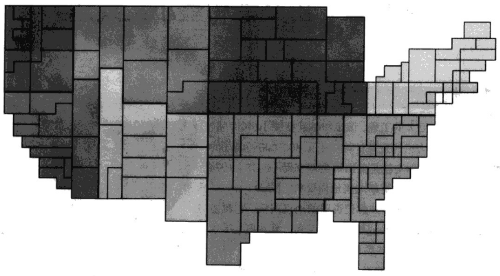


图 6-1 美国 Helmert 分块的邻接边界

实验数据中产生的线性方程组Ax=b通常无解，例如上面的实例。通常可接受的替换解是向量它使得与b之间的距离尽可能小。在线性代数的数值计算中，经常用到矩阵分解的技巧。

9

基本知识

10

内积

定义

设有 n 维向量 $\mathrm{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \mathrm{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$

定义 $[\mathrm{x}, \mathrm{y}] = \mathrm{x}^T \mathrm{y} = x_1 y_1 + x_2 y_2 + \cdots x_n y_n$ 为向量 x 与 y的内积。

$$\mathrm{x}^T \mathrm{y} = (x_1, x_2, \cdots, x_n) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

11

内积

设有 n 维向量 $\mathrm{x}, \mathrm{y}, \mathrm{z} \in \mathbb{R}^n, \lambda \in \mathbb{R}$

(1) $[\mathrm{x}, \mathrm{y}] = [\mathrm{y}, \mathrm{x}]$;

(2) $[\lambda \mathrm{x}, \mathrm{y}] = \lambda [\mathrm{x}, \mathrm{y}]$;

(3) $[\mathrm{x} + \mathrm{y}, \mathrm{z}] = [\mathrm{x}, \mathrm{z}] + [\mathrm{y}, \mathrm{z}]$;

(4) $\mathrm{x} = \mathbf{0}, [\mathrm{x}, \mathrm{x}] = 0; \mathrm{x} \neq \mathbf{0}, [\mathrm{x}, \mathrm{x}] > 0$

施瓦茨 **(Schwarz)** 不等式

$$[\mathrm{x}, \mathrm{y}]^2 \leq [\mathrm{x}, \mathrm{x}][\mathrm{y}, \mathrm{y}]$$

12

长度

定义 $\|\mathbf{x}\| = \sqrt{[\mathbf{x}, \mathbf{x}]} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$,

$\|\mathbf{x}\|$ 称为 n 维向量 $\mathbf{x}$ 的长度（或范数）

某向量

$\mathbf{a} \to \mathbf{x} = \dfrac{\mathbf{a}}{\|\mathbf{a}\|}$

单位化

性质 (1) 非负性　当 $\mathbf{x} \neq 0$ 时，$\|\mathbf{x}\| > 0$;

当 $\mathbf{x} = 0$，$\|\mathbf{x}\| = 0$;

(2) 齐次性　$\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$;

$$-1 \le \frac{[\mathbf{x}, \mathbf{y}]}{\|\mathbf{x}\| \|\mathbf{y}\|} \le 1 \ (\|\mathbf{x}\|, \|\mathbf{y}\| \neq 0 \text{ 时})$$

正交

$[\mathbf{x}, \mathbf{y}] = 0$

$\mathbf{x} \neq 0$ 与 $\mathbf{y} \neq 0$ 的夹角 $\theta = \arccos \dfrac{[\mathbf{x}, \mathbf{y}]}{\|\mathbf{x}\| \|\mathbf{y}\|}$

13

长度

$$\|\mathbf{x} + \mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$$

证明: $\|\mathbf{x} + \mathbf{y}\|^2 = [\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}] = [\mathbf{x}, \mathbf{x}] + 2[\mathbf{x}, \mathbf{y}] + [\mathbf{y}, \mathbf{y}]$

$$= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2[\mathbf{x}, \mathbf{y}]$$

$$\le \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\sqrt{[\mathbf{x}, \mathbf{x}][\mathbf{y}, \mathbf{y}]}$$

$$= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\| \|\mathbf{y}\|$$

$$= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2$$

Schwarz 不等式

$[\mathbf{x}, \mathbf{y}]^2 \le [\mathbf{x}, \mathbf{x}][\mathbf{y}, \mathbf{y}]$

$\Rightarrow \|\mathbf{x} + \mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$ ■

14

向量的距离

定义 设u, v $\in \mathbb{R}^n$，记 dist(u, v)是向量u和v的距离

则, dist(u, v)= $\|u - v\|$

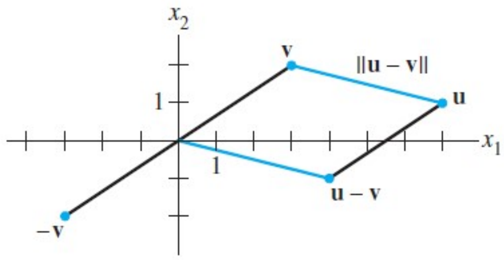


FIGURE 4 The distance between **u** and **v** is the length of **u** − **v**.

15

正交性

16

正交性

定义 设向量$u, v \in \mathbb{R}^n$，若$u \cdot v = 0$，那么向量u, v是正交的。

零向量与$\mathbb{R}^n$中的每个向量v正交，因为$0^T \cdot v = 0$.

定理 两个向量u和v正交，当且仅当$\|u+v\|^2 = \|u\|^2 + \|v\|^2$.

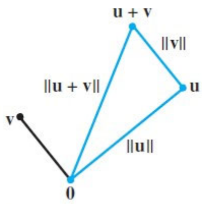


FIGURE 6

17

正交性

定义 若向量z与$\mathbb{R}^n$子空间W 中的所有向量正交，那么称z与W 正交。所有与z具有相同性质的向量组成的集合，称作子空间W 的正交补空间，记作$W^\perp$.

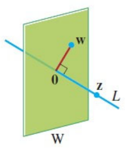


FIGURE 7
A plane and line through **0** as orthogonal complements.

设W是经过原点的平面，L 是过原点且垂直于W 的直线。若$z \in L$，$w \in W$，那么，直线L 上的向量与平面W中的所有向量w正交。L与W 互为正交补，即$L = W^\perp$, 且$W = L^\perp$.

（1）若向量$x \in W^\perp$，当且仅当x 与张成W 的向量集中的每个向量正交；
（2）$W^\perp$是$\mathbb{R}^n$的子空间。

18

正交性

定义 $A_{m\times n}$ 矩阵的列空间 $C(A)$ 包含所有列向量的线性组合，即所有可能的 Ax 向量。

列空间 Column space

线性系统 $Ax = b$ 有解的充要条件是 b 必须在 A 的列空间中。

$A_{m\times n}$ 的列空间是 $\mathbb{R}^m$（不是 $\mathbb{R}^n$）的子空间！

定义 $A_{m\times n}$ 矩阵的零空间 $N(A)$ 包含所有 $Ax = 0$ 的解，包括 $x = 0$，这些向量在 $\mathbb{R}^n$ 中。

消元不改变零空间

零空间 Null space

19

正交性

定义

$A_{m\times n}$ 矩阵的行空间 $C(A^T)$ 包含所有行向量的线性组合，即等同于 A^T 矩阵的列空间，是 $\mathbb{R}^n$ 的子空间。

$A_{m\times n}$ 的行空间是 $\mathbb{R}^n$ 的子空间！

行空间 Row space

定义 $A_{m\times n}$ 矩阵的左零空间是 $N(A^T)$，是 $\mathbb{R}^m$ 的子空间。

左零空间 Left null space

20

正交性

定义 设矩阵$A_{m\times n}$是$m\times n$ 维的，则其行空间 $C(A^T)$ 的正交补空间是零空间$N(A)$，其列空间的正交补空间为A^T的零空间：

$$(1)\ C(A^T)^{\perp}=N(A),\quad (2)\ C(A)^{\perp}=N(A^T).$$

证明

若向量$x\in N(A)$，那么x与A 的每行都正交。由于矩阵A的行空间$C(A^T)$ 由其行向量张成，那么$x\perp C(A^T)$.

相反，若$x\perp C(A^T)$，那么x 与A 的每行都正交，且$Ax=0$. 同理，同理可证式 2。

21

正交性

性质总结

(1) 正交向量有 $v^Tw=0\Rightarrow \|v\|^2+\|w\|^2=\|v+w\|^2=\|v-w\|^2$

(2) 如果对每个 $v\in V$, 每个 $w\in W$ 有 $v^Tw=0$,
则子空间 V 和 W 正交

(3) A 矩阵的行空间与与零空间正交。列空间与左零空间$N(A^T)$正交

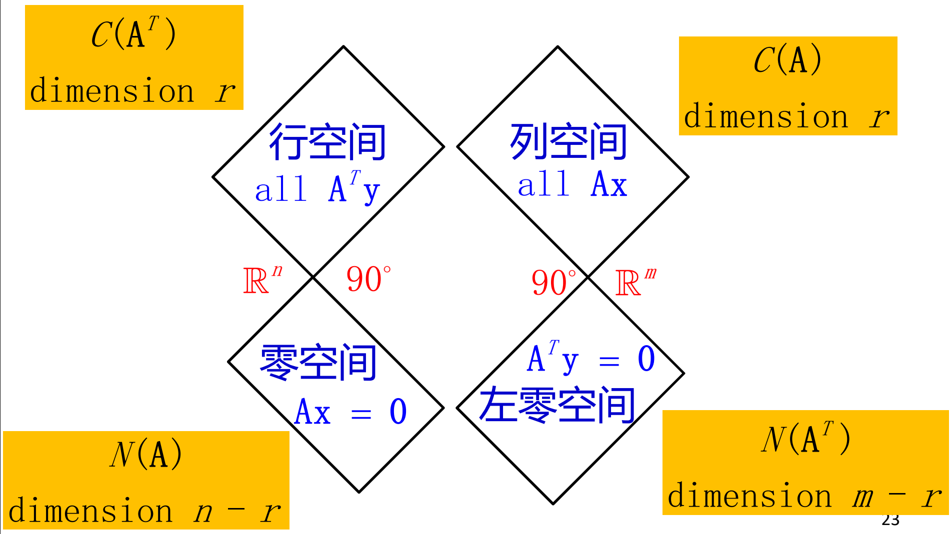
(4) 相应的维数 $r+(n-r)=n,\quad r+(m-r)=m$

$$[x,A^Ty]=x^T(A^Ty)=(Ax)^Ty=0^Ty=0$$
$$\Rightarrow N(A)\perp C(A^T)$$

$$[Ax,y]$$
$$=(Ax)^Ty=x^TA^Ty=x^T0=0$$
$$\Rightarrow C(A)\perp N(A^T)$$

22

正交性



正交性

$$u \cdot v = \|u\| \|v\| \cos \theta$$

内积

向量夹角

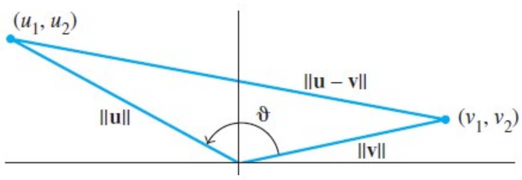


FIGURE 9 The angle between two vectors.

线性代数 (Linear Algebra)

第六章 Orthogonality and Least Squares

§6.2 Orthogonal Sets

正交集

衡 益

2021 年 12 月 21 日，中山大学南校区

正交集

26

6.2 正交集

定义 一组$\mathbb{R}^n$中的向量$\{u_1, \cdots, u_p\}$，如果对任意两个不同向量有$u_i \cdot u_j = 0, i \neq j$，那么这组向量被称为正交集。

例1 证明$\{u_1, u_2, u_3\}$是正交集，其中

$$u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad u_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

27

6.2 正交集

解 考虑三种向量组合，$\{u_1, u_2\}$，$\{u_1, u_3\}$，$\{u_2, u_3\}$.

$$u_1 \cdot u_2 = 3 \times (-1) + 1 \times 2 + 1 \times 1 = 0$$

$$u_1 \cdot u_3 = 3 \times (-\tfrac{1}{2}) + 1 \times (-2) + 1 \times (\tfrac{7}{2}) = 0$$

$$u_2 \cdot u_3 = (-1) \times (-\tfrac{1}{2}) + 2 \times (-2) + 1 \times (\tfrac{7}{2}) = 0$$

集合中任意两个向量正交，因此$\{u_1, u_2, u_3\}$是正交集。

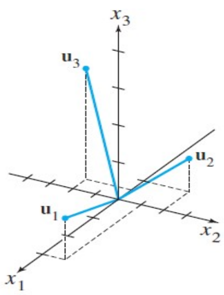


FIGURE 1

28

6.2 正交集

定理 若S= $\{u_1, u_2, \cdots, u_p\}$ 是$\mathbb{R}^n$中非零的正交集，那么S线性无关，并且是由S张成的子空间的基。

证明 设有 $\lambda_1, \lambda_2, \cdots \lambda_p$ 使 $\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_p u_p = 0$

$$\Rightarrow [u_1, (\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_p u_p)] = [u_1, 0] = 0$$

$$\Rightarrow \lambda_1 \|u_1\|^2 = 0 \overset{u_1\text{非零}}{\Rightarrow} \lambda_1 = 0$$

类似可证 $\lambda_2, \lambda_3, \cdots, \lambda_p = 0$

$\Rightarrow$ 向量组 $u_1, u_2, \cdots, u_p$ 线性无关 ■

$\Rightarrow$ 向量组 $u_1, u_2, \cdots, u_p$ 是子空间的基

29

6.2 正交集

定理 若 $\{u_1, u_2, \cdots, u_p\}$ 是$\mathbb{R}^n$子空间W的一组正交基，那么W中的每个向量y都可由$u_1, u_2, \cdots, u_p$唯一表示。即

$$y = c_1 u_1 + \cdots + c_p u_p$$

则 $$c_j = \frac{y \cdot u_j}{u_j \cdot u_j} (j = 1, \cdots, p)$$

证明 由 $\{u_1, u_2, \cdots, u_p\}$ 的正交性可知

$$y \cdot u_j = (c_1 u_1 + c_2 u_2 + \cdots + c_p u_p) \cdot u_j$$

$$\Rightarrow c_j u_j u_j = y \cdot u_j$$

由于$u_j \cdot u_j$非零，上述方程中可解出$c_j = \dfrac{y \cdot u_j}{u_j u_j}$.

30

6.2 正交集

定理 若$\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$是$\mathbb{R}^n$子空间W的一组正交基，那么W中的每个向量y都可由$\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$唯一表示。即

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

则

$$c_j = \frac{\mathbf{y}\cdot\mathbf{u}_j}{\mathbf{u}_j\cdot\mathbf{u}_j}\ (j = 1, \cdots, p)$$

几何意义

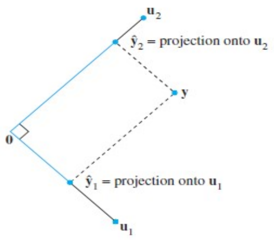


FIGURE 4 A vector decomposed into the sum of two projections.

$$\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2$$

31

6.2 正交集

例2 已知 $\mathbb{R}^3$ 中的两个向量 $\mathbf{a}_1^T = \begin{pmatrix}1 & 1 & 1\end{pmatrix}, \mathbf{a}_2^T = \begin{pmatrix}1 & -2 & 1\end{pmatrix}$ 正交，试求一个非零向量 $\mathbf{a}_3$ ，使 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 两两正交。

解 $A = \begin{pmatrix}\mathbf{a}_1^T \\ \mathbf{a}_2^T\end{pmatrix} = \begin{pmatrix}1 & 1 & 1 \\ 1 & -2 & 1\end{pmatrix}$，$\mathbf{a}_3 = \begin{pmatrix}x_1 \\ x_2 \\ x_3\end{pmatrix}$应满足 $A\mathbf{a}_3 = 0$

$$A = \begin{pmatrix}1 & 1 & 1 \\ 1 & -2 & 1\end{pmatrix} \overset{r_2 - r_1 \to r_2}{\sim} \begin{pmatrix}1 & 1 & 1 \\ 0 & -3 & 0\end{pmatrix} \underset{r_1 - r_2 \to r_1}{\overset{-\frac{1}{3}r_2 \to r_2}{\sim}} \begin{pmatrix}1 & 0 & 1 \\ 0 & 1 & 0\end{pmatrix}$$

基础解系为$\begin{pmatrix}-1 \\ 0 \\ 1\end{pmatrix}$，则 $\mathbf{a}_3 = \begin{pmatrix}-1 \\ 0 \\ 1\end{pmatrix}$ 满足条件

32

6.2 正交集

例3 设$u_1 = \begin{bmatrix} 3 & 1 & 1 \end{bmatrix}$，$u_2 = \begin{bmatrix} -1 & 2 & 1 \end{bmatrix}$，$u_3 = \begin{bmatrix} -1/2 & -2 & 2/7 \end{bmatrix}$
用上述u_1, u_2, u_3的线性组合表示$y=\begin{bmatrix} 6 & 1 & -8 \end{bmatrix}$.

解 首先计算

$y \cdot u_1 = 11$，$y \cdot u_2 = -12$，$y \cdot u_3 = -33$

$u_1 \cdot u_1 = 11$，$u_2 \cdot u_2 = 6$，$u_3 \cdot u_3 = 33/2$

由上述定理可得，

$$y = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2} u_2 + \frac{y \cdot u_3}{u_3 \cdot u_3} u_3$$
$$= \frac{11}{11} u_1 + \frac{-12}{6} u_2 + \frac{-33}{33/2} u_3$$
$$= u_1 - 2u_2 - 2u_3$$

6.2 正交集

力的分解

在力y 的作用下，汽车直线移动
u ：移动方向

将y正交分解！

标准正交基

35

6.2 正交集

定义 设 $e_1, e_2, \cdots, e_r$ 是向量空间 V（$V \subseteq \mathbb{R}^n$）的一个基，如果 $e_1, e_2, \cdots, e_r$ 两两正交，且都是单位向量，则称 $e_1, e_2, \cdots, e_r$ 是V 的一个标准正交基

设 $a \in V, a = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_r e_r$

$$e_i^T a = \lambda_i e_i^T e_i = \lambda_i \Rightarrow \lambda_i = [a, e_i]$$

给向量空间取标准正交基方便计算

36

6.2 正交集

例4 证明 $\{v_1, v_2, v_3\}$ 是$\mathbb{R}^3$中的一组标准正交基。

$$v_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/11 \end{bmatrix},\quad v_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix},\quad v_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

先验证正交 → **再验证长度**

或

先验证长度 → **再验证正交**

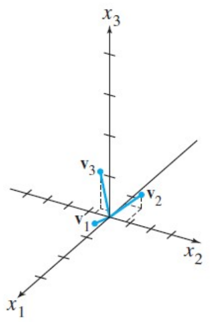


FIGURE 6

37

6.2 正交集

解 计算

$$v_1 \cdot v_2 = -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0$$

$$v_1 \cdot v_3 = -3/\sqrt{726} - 4/\sqrt{726} + 7/\sqrt{726} = 0$$

$$v_2 \cdot v_3 = 1/\sqrt{396} - 8/\sqrt{396} + 7/\sqrt{396} = 0$$

因此 $\{v_1, v_2, v_3\}$ 是正交集合，并且

$$v_1 \cdot v_1 = 9/11 + 1/11 + 1/11 = 1$$

$$v_2 \cdot v_2 = 1/6 + 4/6 + 1/6 = 1$$

$$v_3 \cdot v_3 = 1/66 + 16/66 + 49/66 = 1$$

因此，v_1，v_2，v_3均为单位向量，那么 $\{v_1, v_2, v_3\}$ 是标准正交集合。由于集合线性无关，因此为$\mathbb{R}^3$中的一组基。

38

6.2 正交集

定义 如果 $m \times n$ 阶矩阵 U 的向量是标准正交的 $\Leftrightarrow$ $U^TU = I$.

证明 设$U = [u_1 \quad u_2 \quad \cdots u_n]$，且$u_i \in \mathbb{R}^m$, $i = 1, 2, \ldots, n$.那么，

$$U^TU = \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix} [u_1 \quad u_2 \quad \cdots u_n] = \begin{bmatrix} u_1^Tu_1 & u_1^Tu_2 & u_1^Tu_3 & \\ u_2^Tu_1 & u_2^Tu_2 & u_2^Tu_3 & \\ u_3^Tu_1 & u_3^Tu_2 & u_3^Tu_3 & \\ & & & \ddots \end{bmatrix}$$

U是正交矩阵当且仅当

$u_i^Tu_j = u_j^Tu_i = 0$， $i \neq j$

$u_i^Tu_j = 1$， $i = j$， $\Rightarrow U^TU = I$

6.2 正交集

性质 如果U是 $m \times n$ 阶矩阵，其向量是标准正交的，$x, y \in \mathbb{R}^n$，那么，

(1) $\|Ux\| = \|x\|$ **长度**

(2) $(Ux)\cdot(Uy) = x \cdot y$

(3) $(Ux)\cdot(Uy) = 0$当且仅当$x \cdot y=0$ **正交性**

映射$Ux \mapsto x$保留了其长度和正交性！

40

6.2 正交集

定义 如果 n 阶方阵 A 满足 $A^TA = I$,
即 $A^{-1} = A^T$, 则 A 称之为正交矩阵。

A 为正交矩阵 $\Leftrightarrow$
A 的列向量都是单位向量，且两两正交

$$\begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{pmatrix}(a_1, a_2, \cdots, a_n) = I \Rightarrow a_i^T a_j = \begin{cases} 1, & \text{当 } i = j, \\ 0, & \text{当 } i \neq j \end{cases} \quad (i, j = 1, 2, \cdots, n),$$

单位向量

两两正交

41

6.2 正交集

$$P = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \text{是正交矩阵？}$$

单位向量?

两两正交?

1

2 **或者** $P^TP = I$**?**

42

6.2 正交集

例5 设U= $\begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}$，$x = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$.

注意到U中的向量是标准正交的，验证 $\|Ux\| = \|x\|$.

解

$$Ux = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}\begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$$

$$\|Ux\| = \sqrt{9+1+1} = \sqrt{11}$$

$$\|x\| = \sqrt{2+9} = \sqrt{11}$$

43

6.2 正交集

性质

(1) 若 A 是正交矩阵，则 $A^{-1} = A^T$ 也是正交矩阵，且 $|A| = 1$ 或 -1

(2) 若 A 和 B 都是正交矩阵，则 AB 也是正交矩阵

定义 若 P 为正交矩阵，则线性变换 $y = Px$ 成为正交变换

$$\|y\| = \sqrt{[y,y]} = \sqrt{y^Ty} = \sqrt{(Px)^TPx} = \sqrt{x^TP^TPx} = \sqrt{x^Tx} = \sqrt{[x,x]} = \|x\|$$

经正交变换线段长度保持不变

44

线性代数 (Linear Algebra)

第六章 Orthogonality and Least Squares

§6.3 Orthogonality Projection

正交投影

衡 益

2021 年 12 月 21 日，中山大学南校区

正交投影

46

6.3 正交投影

$$P = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow$$ 投影到 z 轴

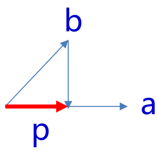
$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow$$ 投影到 xy 平面

$$P = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}, \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

伸缩变换　　旋转变换　　反射

47

6.3 正交投影



将 b 向量投影到 a 向量的方向上

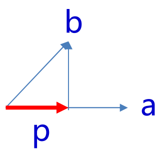
$$p = \frac{a^T b}{a^T a} a$$

系数 λ

48

6.3 正交投影

$$p = \lambda a, \quad e = b - p = b - \lambda a$$



$$< b - \lambda a, a >= 0$$

$$\Rightarrow < b, a > - < \lambda a, a >= 0$$

$$\Rightarrow \lambda = \frac{< b, a >}{< a, a >} = \frac{< a, b >}{< a, a >}$$

$$p = Pb$$

$$e = b - p, \quad b - \frac{< a, b >}{< a, a >} a$$

$$p = a\frac{a^T b}{a^T a} = \frac{aa^T}{a^T a} b$$

6.3 正交投影

对于给定的向量$y \epsilon R^n$, W为子空间，若：

（1）W中有唯一向量$\hat{y}$，使得$y - \hat{y}$与W正交

（2）$\hat{y}$是W中唯一最接近y的向量

上述两个性质提供了**寻找线性系统最小二乘解的关键**。

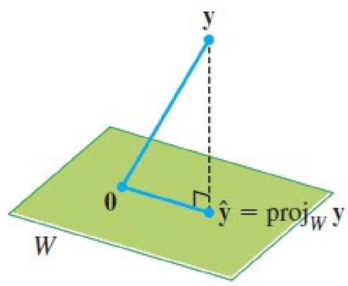


FIGURE 1

50

6.3 正交投影

设y可以表示为$\mathbb{R}^n$ 空间基$u_1 \cdots u_n$的线性组合，y 中和的各项可以分为两个部分，使得 y 写成

$$y = z_1 + z_2$$

此处z_1是其中一些 u_i 的线性组合，z_2 是其余u_j 的线性组合，当$\{u_1 \cdots u_n\}$ 正交时，这个思路很有用 → 回忆6.1节，$W^\perp$ 表示所有与W 正交的向量的集合。

例1 设 $\{u_1, \cdots, u_5\}$ 是$\mathbb{R}^5$ 中的正交基，令

$$y = c_1u_1 + \cdots + c_5u_5$$

考虑子空间W=Span $\{u_1, u_2\}$，那么，y 可由$z_1 \in W$，$z_2 \in W^\perp$表示。

51

6.3 正交投影

例1 设 $\{u_1, \cdots, u_5\}$ 是$\mathbb{R}^5$中的正交基，令

$$y = c_1u_1 + \cdots + c_5u_5$$

考虑子空间W=Span $\{u_1, u_2\}$，那么，y可由$z_1 \in W$，$z_2 \in W^\perp$表示。

解 令

$$y = \underbrace{c_1u_1 + c_2u_2}_{z_1} + \underbrace{c_3u_3 + c_4u_4 + c_5u_5}_{z_2}$$

其中

$$z_1 = c_1u_1 + c_2u_2 \in \text{Span}\{u_1, u_2\}$$

$$z_2 = c_3u_3 + c_4u_4 + c_5u_5 \in \text{Span}\{u_3, u_4, u_5\}$$

52

6.3 正交投影

例1 设 $\{u_1, \cdots, u_5\}$ 是$\mathbb{R}^5$中的正交基，令

$$y = c_1u_1 + \cdots + c_5u_5$$

考虑子空间W=Span $\{u_1, u_2\}$，那么，y可由$z_1 \in W, z_2 \in W^{\perp}$表示。

解 要证$z_2 \in W^{\perp}$，即证z_2与W 的基 $\{u_1, u_2\}$ 正交，利用内积的性质，

$$\begin{aligned} z_2 \cdot u_1 &= (c_3u_3 + c_4u_4 + c_5u_5) \cdot u_1 \\ &= c_3u_3 \cdot u_1 + c_4u_4 \cdot u_1 + c_5u_5 \cdot u_1 \\ &= 0 \end{aligned}$$

由于u_1与u_3，u_4，u_5正交，因此其内积为0. 同理可得$z_2 \cdot u_2$=0.
那么，$z_2 \in W^{\perp}$.

6.3 正交投影

定理 **正交分解定理**

设W 是$\mathbb{R}^n$空间的子空间，那么每个$y \in \mathbb{R}^n$ 都有可以唯一的表示为如下形式

$$y = \hat{y} + z \tag{1}$$

其中$\hat{y} \in W$，$z \in W^{\perp}$.事实上，若 $\{u_1, \cdots, u_p\}$ 是W 的一组正交基，那么

$$\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1}u_1 + \cdots + \frac{y \cdot u_p}{u_p \cdot u_p}u_p \tag{2}$$

且$z = y - \hat{y}$.

6.3 正交投影

注： 其中，(1)式中的$\hat{y}$ 称为y 在W 上的正交投影，常记作 $\mathrm{proj}_W\ y$，

见下图，当W 是一维子空间时，$\hat{y}$ 的公式和之前的公式一致。

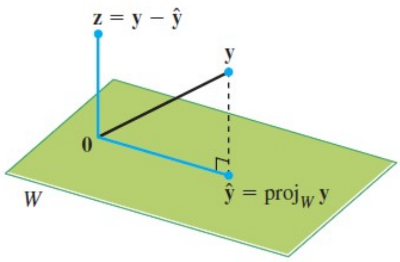


FIGURE 2 The orthogonal projection of **y** onto W.

55

6.3 正交投影

证明

设$\{u_1,\cdots,u_p\}$是W 的正交基，$\hat{y}$ 可以由其线性表示。那么$\hat{y}\in W$因为$\hat{y}$是$u_1,\cdots,u_p$的线性组合。令$z=y-\hat{y}$，由于u_1与$u_2,\cdots,u_p$正交，因此

$$
\begin{aligned}
z\cdot u_1=(y-\hat{y})\cdot u_1 &= y\cdot u_1-\left(\frac{y\cdot u_1}{u_1\cdot u_1}\right)u_1\cdot u_1-0-\cdots-0\\
&= y\cdot u_1-y\cdot u_1\\
&= 0
\end{aligned}
$$

因此 z 正交于 u_1，同理 z 与 $u_j\in W$ 都正交，故$z\in W^{\perp}$.

56

6.3 正交投影

证明

下证 $y = \hat{y} + z$的分解是唯一的。设y 还可以表示为 $y = \hat{y}_1 + z_1, \hat{y}_1 \in W, z_1 \in W^{\perp}$. 那么$\hat{y} + z = \hat{y}_1 + z_1$，则

$$\hat{y} - \hat{y}_1 = z_1 - z$$

上式说明向量$v = \hat{y} - \hat{y}_1$ 同时属于W 和其正交补空间$W^{\perp}$. 因此

$$v \cdot v = 0$$
$$\Rightarrow v = 0$$
$$\Rightarrow \hat{y} = \hat{y}_1 且 z_1 = z.$$

57

6.3 正交投影

例2 设$u_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, u_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.其中$\{u_1, u_2\}$是$W = Span\{u_1, u_2\}$的正交基。将向量y 用W 和$W^{\perp}$ 中的向量分解表示。

解 y在W上的正交投影为

$$\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2} u_2$$

$$= \frac{9}{30}\begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$

58

6.3 正交投影

例2 设$\mathbf{u}_1=\begin{bmatrix}2\\5\\-1\end{bmatrix}$, $\mathbf{u}_2=\begin{bmatrix}-2\\1\\1\end{bmatrix}$, $\mathbf{y}=\begin{bmatrix}1\\2\\3\end{bmatrix}$.其中$\{\mathbf{u}_1,\ \mathbf{u}_2\}$是$W=Span\{\mathbf{u}_1,\ \mathbf{u}_2\}$的正交基。将向量$\mathbf{y}$ 用W 和$W^{\perp}$ 中的向量分解表示。

解 因此

$$\mathbf{y}-\hat{\mathbf{y}}=\begin{bmatrix}1\\2\\3\end{bmatrix}-\begin{bmatrix}-2/5\\2\\1/5\end{bmatrix}=\begin{bmatrix}7/5\\0\\14/5\end{bmatrix}$$

正交分解定理确保$\mathbf{y}-\hat{\mathbf{y}}\in W^{\perp}$.那么$\mathbf{y}$ 可被分解为

$$\mathbf{y}=\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}-2/5\\2\\1/5\end{bmatrix}+\begin{bmatrix}7/5\\0\\14/5\end{bmatrix}$$

59

6.3 正交投影

正交投影的几何意义

W=Span$\{\mathbf{u}_1,\ \mathbf{u}_2\}$

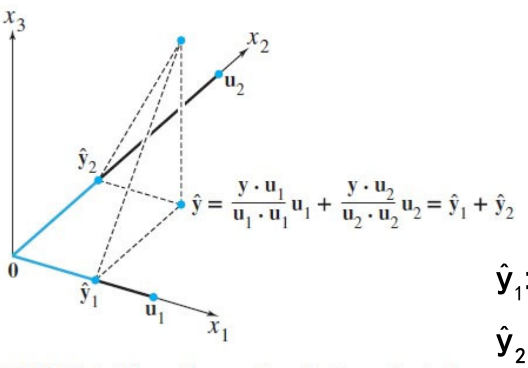


FIGURE 3 The orthogonal projection of **y** is the sum of its projections onto one-dimensional subspaces that are mutually orthogonal.

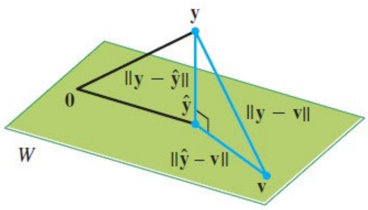
$\hat{\mathbf{y}}_1$: y在Span$\{\mathbf{u}_1\}$ 上的投影

$\hat{\mathbf{y}}_2$: y在Span$\{\mathbf{u}_2\}$ 上的投影

60

6.3 正交投影

性质 若$y \in W = \mathrm{Span}\{u_1, \cdots, u_p\}$，那么 $\mathrm{proj}_W\, y = y$.



定理 最佳逼近定理

设W 为的子空间，y是$\mathbb{R}^n$ 中的任一向量，$\hat{y}$ 为y 在W上的正交投影。那么$\hat{y}$ 是W 中最接近y 的点，即

$$\|y - \hat{y}\| < \|y - v\|$$

对于所有的$v \in W$ 成立，v 不同于$\hat{y}$.

61

6.3 正交投影

证明

设$v \in W$ 是不同于$\hat{y}$，则$\hat{y} - v \in W$. 由正交分解定理可知，$y - \hat{y}$与W 正交。特别的$y - \hat{y}$ 与$\hat{y} - v \in W$ 正交，由于

$$y - v = (y - \hat{y}) + (\hat{y} - v)$$

由勾股定理可知

$$\|y - v\|^2 = \|y - \hat{y}\|^2 + \|\hat{y} - v\|^2$$

由于$\hat{y} - v \neq 0$, 因此$\|\hat{y} - v\|^2 > 0$, 那么不等式

$$\|y - \hat{y}\| < \|y - v\|$$

成立。

62

6.3 正交投影

例3 设$u_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, u_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, W = \text{Span}\{u_1, u_2\}$，

那么W中与y最近的点为

解

$$\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2} u_2 = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$

63

6.3 正交投影

例4 设$y \in \mathbb{R}^n$, $W=Span\{u_1, u_2\}$ 是$\mathbb{R}^n$的一个子空间，那么y与W 的距离由y到W 的最短距离来定义。计算$\text{dist}(y, W)$，

$$y = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, u_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

解 由最佳逼近定理可知，

$$\text{dist}(y, W) = \|y - \hat{y}\|$$

其中$\hat{y} = \text{proj}_W y$. 由于$\{u_1, u_2\}$是W 中的正交基，

$$\hat{y} = \frac{15}{30} u_1 + \frac{-21}{6} u_2 = \frac{1}{2}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} - \frac{7}{2}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}$$

64

6.3 正交投影

例4 设$\mathbf{y} \in \mathbb{R}^n$, W=Span$\{u_1, u_2\}$ 是$\mathbb{R}^n$的一个子空间，那么y与W 的距离由y到W 的最短距离定义。计算dist$(\mathbf{y}, W)$,

$$\mathbf{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

解

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix} - \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

$$\|\mathbf{y} - \hat{\mathbf{y}}\|^2 = 3^2+6^2=45$$

因此

$$\text{dist}(\mathbf{y}, W) = 3\sqrt{5}$$

65

6.3 正交投影

例5 设$\mathbf{y}=\begin{bmatrix} 7 & 6 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 4 & 2 \end{bmatrix}$. 找到一个y在u上的投影。

解

计算

$$\mathbf{y} \cdot \mathbf{u}=\begin{bmatrix} 7 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40, \quad \mathbf{u} \cdot \mathbf{u}=\begin{bmatrix} 4 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20.$$

y在u上的投影为

$$\hat{\mathbf{y}}=\frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u} = \frac{40}{20}\mathbf{u} = 2\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$\mathbf{y} - \hat{\mathbf{y}}$和$\hat{\mathbf{y}}$的和为y

$$\begin{bmatrix} 7 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

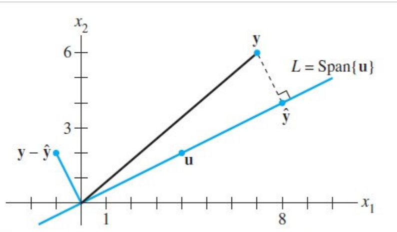


FIGURE 3 The orthogonal projection of **y** onto a line L through the origin.

66

6.3 正交投影

例5 计算下图中y到L的距离。

解 y到L的距离就是y到$\hat{y}$垂直线段的长度，为$y-\hat{y}$. 因此，其距离为：

$$\|y-\hat{y}\|=\sqrt{(-1)^2+2^2}=\sqrt{5}$$

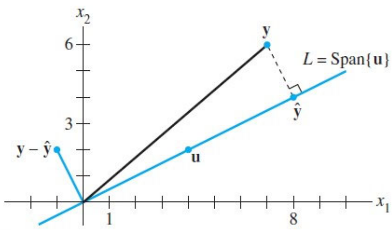


FIGURE 3 The orthogonal projection of y onto a line L through the origin.

67

6.3 正交投影

定理

设$\{u_1,\cdots,u_p\}$是$\mathbb{R}^n$子空间W 的一组标准正交基，那么

$$\mathrm{proj}_W y=(y\cdot u_1)u_1+(y\cdot u_2)u_2+\cdots+(y\cdot u_p)u_p \qquad (3)$$

若$U=\begin{bmatrix}u_1 & u_2 & \cdots & u_p\end{bmatrix}$，那么

$$\mathrm{proj}_W y=UU^T y, \forall y\in\mathbb{R}^n \qquad (4)$$

证明

公式(3)由幻灯片第10页(2)立刻推出，同样(3)说明$\mathrm{proj}_W y$是U 中列的线性组合，且对应权值分别为$y\cdot u_1$ ，$y\cdot u_2$ ，$\cdots$，$y\cdot u_p$，同样可以写成：$u_1^T y, u_2^T y,\cdots,u_p^T y$，说明这些向量可以记为$U^T y$，从而证明了(4).

Q & A