

第六章 内积空间和正交性

2021 年 12 月 29 日



在线性空间中增加了一个额外的结构, 从而我们可以谈论向量的长度, 两个向量之间的距离, 角度, 正交关系及投影...

度和几何引进来了



内容

- 1 内积, 长度和正交性
- 2 正交集与标准正交基
- 3 正交投影
- 4 直和与正交补
- 5 正交分解与最佳逼近



标准内积

\mathbb{R}^n 中向量的点乘: $\vec{u} \cdot \vec{v} := \vec{u}^T \vec{v}$ (标准内积)

标准内积的性质:

- 对称性

$$\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$$

- 双线性

$$(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}, \quad (c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (c\vec{v})$$

- 正定性

$$\vec{u} \cdot \vec{u} \geq 0, \text{ 且 } \vec{u} \cdot \vec{u} = 0 \text{ 成立的充要条件是 } \vec{u} = \vec{0}$$



定义 (内积的一般定义)

二元函数 $\langle, \rangle : V \times V \rightarrow \mathbb{R}$, 满足

- 对称性

$$\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$$

- 双线性

$$\langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle, \quad \langle c\vec{u}, \vec{v} \rangle = c\langle \vec{u}, \vec{v} \rangle$$

- 正定性

$$\langle \vec{u}, \vec{u} \rangle \geq 0, \text{ 且 } \langle \vec{u}, \vec{u} \rangle = 0 \text{ 成立的充要条件是 } \vec{u} = \vec{0}$$

定义了内积的向量 (线性) 空间称为内积空间; 特别地, 定义了内积的实线性空间通常称为欧氏空间.



非标准内积的例子

- $\vec{u}, \vec{v} \in \mathbb{R}^2$, $\langle \vec{u}, \vec{v} \rangle = 5u_1v_1 + 6u_2v_2$.
- $\vec{p}, \vec{q} \in \mathbb{P}_n$, 其中 \mathbb{P}_n 是次数最高为 n 的实多项式集合. \mathbb{P}_n 对于通常多项式的加法和数乘构成了一个线性空间. 给定互异的 $n+1$ 个实数 t_0, \dots, t_n , 可以定义 \mathbb{P}_n 中的内积为

$$\langle \vec{p}, \vec{q} \rangle = \sum_{i=0}^n p(t_i)q(t_i).$$

- 考虑 $[a, b]$ 区间上的所有实连续函数构成的线性空间, 记为 $C[a, b]$, 内积可以定义为

$$\langle f, g \rangle = \int_a^b f(t)g(t) dt, \quad f, g \in C[a, b].$$



向量的长度: 范数

- 向量 \vec{v} 的长度 (范数, norm) 定义为 $\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$ (或 $\sqrt{\langle \vec{v}, \vec{v} \rangle}$).
- 向量的单位化: $\vec{v} \rightarrow \vec{v}/\|\vec{v}\|$.
- \mathbb{R}^n 中两个向量的距离: $\text{dist}(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\|$.
(关于距离的约定 & 三角形不等式)
- 两个向量之间的夹角: $\langle \vec{u}, \vec{v} \rangle = \|\vec{u}\| \|\vec{v}\| \cos \theta$.
(Cauchy-Schwarz 不等式)
- 正交向量: 如果 $\langle \vec{u}, \vec{v} \rangle = 0$, 则称 \vec{u} 和 \vec{v} 是相互正交的: $\vec{u} \perp \vec{v}$.
- (勾股定理) \vec{u}, \vec{v} 正交的充要条件是 $\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$.



两个不等式

- 柯西-施瓦茨 (Cauchy-Schwarz) 不等式

$$|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \|\vec{v}\|$$

证明:

$$0 \leq \langle \vec{u} + \lambda \vec{v}, \vec{u} + \lambda \vec{v} \rangle = \|\vec{u}\|^2 + \lambda^2 \|\vec{v}\|^2 + 2\lambda \langle \vec{u}, \vec{v} \rangle \implies$$

$$\Delta = 4\langle \vec{u}, \vec{v} \rangle^2 - 4\|\vec{u}\|^2 \|\vec{v}\|^2 \leq 0 \implies |\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \|\vec{v}\|$$

等号成立当且仅当 \vec{u}, \vec{v} 线性相关.

- 三角不等式

$$\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$$



例

- $$a_1 b_1 + \cdots + a_n b_n \leq \sqrt{a_1^2 + \cdots + a_n^2} \sqrt{b_1^2 + \cdots + b_n^2}$$

- $$\int_a^b f(x)g(x) \, dx \leq \sqrt{\int_a^b [f(x)]^2 \, dx} \sqrt{\int_a^b [g(x)]^2 \, dx}$$



内容

- 1 内积, 长度和正交性
- 2 正交集与标准正交基
- 3 正交投影
- 4 直和与正交补
- 5 正交分解与最佳逼近



定义

若向量组里的向量两两正交, 则该向量组称为一个正交集.

- (正交集向量的线性无关性) 由非零向量构成的正交集线性无关.
- 正交基: 如果这个基是正交集的话.
- 标准正交基 (orthonormal basis): 由单位向量构成的正交基.



向量在正交基中的展开

设 $\{\vec{u}_1, \dots, \vec{u}_p\}$ 是线性空间 V 的一个正交基, 把向量 $\vec{y} \in V$ 在这个基下展开:

$$\vec{y} = c_1 \vec{u}_1 + \dots + c_p \vec{u}_p,$$

则展开系数

$$c_i = \frac{\langle \vec{y}, \vec{u}_i \rangle}{\langle \vec{u}_i, \vec{u}_i \rangle}.$$

进一步的, 若 $\{\vec{u}_1, \dots, \vec{u}_p\}$ 是标准正交基, 则直接有 $c_i = \langle \vec{y}, \vec{u}_i \rangle$.



具有单位正交列的矩阵

一个 $m \times n$ 的矩阵 $U = [\vec{u}_1 \cdots \vec{u}_n]$, 若 $\langle \vec{u}_i, \vec{u}_j \rangle = \delta_{ij}$, 则我们称矩阵 U 具有单位正交列.

- 显然首先应当有 $m \geq n$.
- U 具有单位正交列的充要条件是 $U^T U = I_n$.
- U 定义了从一个从 \mathbb{R}^n 到 \mathbb{R}^m 的线性变换: $U: \mathbb{R}^n \rightarrow \mathbb{R}^m, \vec{x} \mapsto U\vec{x}$.
该变换具有如下性质:
 - 1 保长度, 即 $\|U\vec{x}\| = \|\vec{x}\|$.
 - 2 保内积, 即 $\langle U\vec{x}, U\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle$.
 - 3 保夹角, 特别的, 保持正交性: $\langle U\vec{x}, U\vec{y} \rangle = 0 \iff \langle \vec{x}, \vec{y} \rangle = 0$.

注: 其实单由保长度就可以推出其余的两个性质.



正交矩阵

如果 U 是方阵, 则我们称这种矩阵为**正交矩阵**. 它所定义的变换称为**正交变换**.

- 因为 U 是方阵, 所以我们有 $U^T U = I = U U^T$, U^T 是 U 的逆矩阵.
- 正交矩阵的列和行都是单位正交的.



线性无关集的正交化方法

把一个线性无关集 $\{\vec{x}_1, \dots, \vec{x}_n\}$ 变成标准正交集的步骤: 格拉姆-施密特 (Gram-Schmidt) 方法

① 正交化过程

$$\vec{v}_1 = \vec{x}_1, \quad \vec{v}_2 = \vec{x}_2 - \frac{\langle \vec{x}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \vec{v}_1, \quad \dots,$$

$$\vec{v}_n = \vec{x}_n - \sum_{i=1}^{n-1} \frac{\langle \vec{x}_n, \vec{v}_i \rangle}{\langle \vec{v}_i, \vec{v}_i \rangle} \vec{v}_i.$$

② 标准化

在正交化过程中可以同时进行归一化 (标准化): $\vec{u}_i = \vec{v}_i / \|\vec{v}_i\|$.

从上面的步骤可以看到, 生成的标准正交基 $\{\vec{u}_1, \dots, \vec{u}_p\}$ 与原来的基 $\{\vec{x}_1, \dots, \vec{x}_p\}$ 之间的过渡矩阵为一个上三角可逆矩阵. \Rightarrow QR 分解



QR 分解

从上面的正交化过程可知,

$$\begin{aligned} [\vec{x}_1 \ \vec{x}_2 \ \vec{x}_3 \ \cdots] &= [\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3 \ \cdots] \begin{bmatrix} 1 & \frac{\langle \vec{x}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} & \frac{\langle \vec{x}_3, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} & \\ 0 & 1 & \frac{\langle \vec{x}_3, \vec{v}_2 \rangle}{\langle \vec{v}_2, \vec{v}_2 \rangle} & \\ 0 & 0 & 1 & \\ & & & \ddots \end{bmatrix} \\ &= [\vec{u}_1 \ \vec{u}_2 \ \vec{u}_3 \ \cdots] \begin{bmatrix} \|\vec{v}_1\| & & & \\ & \|\vec{v}_2\| & & \\ & & \|\vec{v}_3\| & \\ & & & \ddots \end{bmatrix} \begin{bmatrix} 1 & \frac{\langle \vec{x}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} & \frac{\langle \vec{x}_3, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} & \\ 0 & 1 & \frac{\langle \vec{x}_3, \vec{v}_2 \rangle}{\langle \vec{v}_2, \vec{v}_2 \rangle} & \\ 0 & 0 & 1 & \\ & & & \ddots \end{bmatrix} \end{aligned}$$

QR 分解: 对于一个**列线性无关**的矩阵 $A_{m \times n}$ ($m \geq n$), 有 $A = QR$, 其中 $m \times n$ 矩阵 Q 满足 $Q^T Q = I_n$, R 是 n 阶上三角矩阵且对角元 > 0 .



内容

- 1 内积, 长度和正交性
- 2 正交集与标准正交基
- 3 正交投影**
- 4 直和与正交补
- 5 正交分解与最佳逼近



沿给定向量的正交投影

Q: 把一个向量 \vec{y} 沿着给定向量 \vec{u} 及 $\perp \vec{u}$ 的方向展开. 即

$$\vec{y} = \vec{y}_{\parallel} + \vec{y}_{\perp}, \quad \text{要求 } \vec{y}_{\parallel} = \alpha \vec{u}, \vec{y}_{\perp} \perp \vec{u}.$$

A: 由 $\langle \vec{y}_{\perp}, \vec{u} \rangle = \langle \vec{y} - \vec{y}_{\parallel}, \vec{u} \rangle = 0 \implies \alpha = \langle \vec{y}, \vec{u} \rangle / \langle \vec{u}, \vec{u} \rangle$. \vec{y}_{\parallel} 称为 \vec{y} 在 \vec{u} 上的正交投影. $\vec{y}_{\parallel} = \frac{\langle \vec{y}, \vec{u} \rangle}{\langle \vec{u}, \vec{u} \rangle} \vec{u} \equiv \text{proj}_{\vec{u}} \vec{y}$.

投影算符 (矩阵)

$$\text{proj}_{\vec{u}} = \vec{u} \frac{\langle \vec{u}, \cdot \rangle}{\langle \vec{u}, \vec{u} \rangle} = \frac{\vec{u} \vec{u}^T}{\vec{u} \cdot \vec{u}}$$



在正交基下作投影

向量 \vec{y} 在正交基下的展开实际上就是将 \vec{y} 分解为在一系列一维子空间上的正交投影之和, 相应的投影矩阵为 $\frac{\vec{u}_i \vec{u}_i^T}{\vec{u}_i \cdot \vec{u}_i}$.

若 W 的一个正交基为 $\{\vec{u}_1, \dots, \vec{u}_p\}$ ($\dim W = p$), 则

$$\begin{aligned}\vec{y}_{\parallel} &= \text{proj}_W \vec{y} = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \dots + \frac{\vec{y} \cdot \vec{u}_p}{\vec{u}_p \cdot \vec{u}_p} \vec{u}_p \\ &= \left(\frac{\vec{u}_1 \vec{u}_1^T}{\vec{u}_1 \cdot \vec{u}_1} + \dots + \frac{\vec{u}_p \vec{u}_p^T}{\vec{u}_p \cdot \vec{u}_p} \right) \vec{y}, \\ \vec{y}_{\perp} &= \vec{y} - \vec{y}_{\parallel}.\end{aligned}$$

容易证明 $\vec{y}_{\perp} \cdot \vec{u}_i = 0$ ($i = 1, \dots, p$), 即 $\vec{y}_{\perp} \perp W$.



内容

- 1 内积, 长度和正交性
- 2 正交集与标准正交基
- 3 正交投影
- 4 直和与正交补**
- 5 正交分解与最佳逼近



子空间的直和

设 V_1, V_2 是线性空间 V 的两个子空间, 且 $V_1 \cap V_2 = \{\vec{0}\}$. 定义两个子空间的直和为

定义

$$V_1 \oplus V_2 = \{\vec{v} : \vec{v} = \vec{v}_1 + \vec{v}_2, \vec{v}_1 \in V_1, \vec{v}_2 \in V_2\}$$

- 由 $V_1 \cap V_2 = \{\vec{0}\}$ 可知, 该分解 $\vec{v} = \vec{v}_1 + \vec{v}_2, \vec{v}_1 \in V_1, \vec{v}_2 \in V_2$ 是唯一的.
- $\dim(V_1 \oplus V_2) = \dim V_1 + \dim V_2$



设 W 是 \mathbb{R}^n (或某线性空间 V) 的一个子空间,

- 向量与子空间的正交

如果向量 \vec{z} 与 W 中的任意向量都正交, 则称 \vec{z} 正交于 W , 可以记为 $\vec{z} \perp W$.

- 正交补 W^\perp

与 W 正交的向量全体组成的集合称为 W 的正交补.

- W^\perp 也是一个线性空间, 且 $W \cap W^\perp = \{\vec{0}\}$, $V = W \oplus W^\perp$.

- $\dim V = \dim W + \dim W^\perp$



例

- $ax + by + cz = 0$ 确定了一个平面 P , 且有
$$P^\perp = \text{Span}\{\vec{v} = (a, b, c)\}.$$
- $\text{Nul } A = (\text{Row } A)^\perp, \text{Nul } A^T = (\text{Col } A)^\perp$
- $A\vec{x} = \vec{b}$ 有解 $\Leftrightarrow \vec{b} \in \text{Col } A \Leftrightarrow \vec{b} \perp \text{Nul } A^T$
- $\text{Nul } A = \text{Nul } A^T A$



- 如果矩阵 A 各列线性无关, 则 $A^T A$ 可逆. 反之亦成立.

证明.

$A = [\vec{v}_1 \ \cdots \ \vec{v}_n]$, $A^T A = [A^T \vec{v}_1 \ \cdots \ A^T \vec{v}_n]$. $A^T A$ 可逆当且仅当 $A^T \vec{v}_1, \dots, A^T \vec{v}_n$ 线性无关. 考虑 $c_1 A^T \vec{v}_1 + \cdots + c_n A^T \vec{v}_n = \vec{0}$, 则有 $A^T (c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n) = \vec{0}$, 即 $c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n \in \text{Nul } A^T = (\text{Col } A)^\perp$. 而同时 $c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n \in \text{Col } A$, 所以只可能有 $c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n = \vec{0}$. 再由 A 各列线性无关, 因此 $c_1 = \cdots = c_n = 0$. 所以我们从 A 线性无关可以推出 $A^T A$ 可逆.

反过来, 由 $A^T A$ 可逆, 可以很容易推出 A 线性无关. 因为如果 A 线性相关的话, 则向量组 $\{A^T \vec{v}_1, \dots, A^T \vec{v}_n\}$ 一定线性相关. □

另外的一种证明利用 ($\text{Nul } A = \text{Nul } A^T A$), 可以参看 6.5 节习题的 19-22 题.



内容

- 1 内积, 长度和正交性
- 2 正交集与标准正交基
- 3 正交投影
- 4 直和与正交补
- 5 正交分解与最佳逼近



正交分解定理

定理 (正交分解的存在性与唯一性定理)

设 W 是 \mathbb{R}^n 的一个子空间, W^\perp 是 W 的正交补, 则 $\mathbb{R}^n = W \oplus W^\perp$.

对定理表述的解释

$\forall \vec{y} \in \mathbb{R}^n$, 有 $\vec{y} = \vec{y}_\parallel + \vec{y}_\perp$, 其中 $\vec{y}_\parallel \in W$, $\vec{y}_\perp \in W^\perp$, 且分解是唯一的.

我们称 \vec{y}_\parallel 是 \vec{y} 在 W 子空间上的正交投影.



最佳逼近

定理 (最佳逼近定理)

W 是 V 的一个子空间, $\vec{y} \in V$, 则 $\vec{y}_{\parallel} = \text{proj}_W \vec{y}$ 是 W 中最接近 \vec{y} 的一点.

证明.

按照正交分解定理, 我们有 $\vec{y} = \vec{y}_{\parallel} + \vec{y}_{\perp}$, 其中 $\vec{y}_{\parallel} \in W$, $\vec{y}_{\perp} \in W^{\perp}$. 对于 $\forall \vec{v} \in W$, 有

$$\|\vec{y} - \vec{v}\|^2 = \|\vec{y}_{\parallel} + \vec{y}_{\perp} - \vec{v}\|^2 = \|\vec{y}_{\parallel} - \vec{v}\|^2 + \|\vec{y}_{\perp}\|^2 \geq \|\vec{y}_{\perp}\|^2.$$

则当且仅当 $\vec{v} = \vec{y}_{\parallel}$ 时, $\|\vec{y} - \vec{v}\|$ 最小. (上面的证明用到了勾股定理)



应用: 最小二乘问题

Q: 当 $A\vec{x} = \vec{b}$ 不相容 (无解) 时, 找到合适的 \vec{x} , 使得 $\|A\vec{x} - \vec{b}\|$ 最小.

A: 实际上就是要在 $\text{Col } A$ 中找到最接近 \vec{b} 的向量. 由最佳逼近定理, 我们知道这个向量就是 $\vec{b}_{\parallel} = \text{proj}_{\text{Col } A} \vec{b}$.

设 $A\vec{x} = \vec{b}_{\parallel}$, 则 $\vec{b} - A\vec{x} \in (\text{Col } A)^{\perp} = \text{Nul } A^T$, 所以

$$A^T(\vec{b} - A\vec{x}) = \vec{0} \quad \text{或} \quad A^T A \vec{x} = A^T \vec{b}$$

- 上面的方程也可以从要求 $\|A\vec{x} - \vec{b}\|$ 取极值得到.
- 该方程一定有解, 因为 \vec{b} 中 $\perp \text{Col } A$ 的部分对 $A^T \vec{b}$ 没有贡献.
- A 是列线性无关时有唯一解. 这时若对 A 作 QR 分解, 有 $\vec{b}_{\parallel} = QQ^T \vec{b}$. 特别的, 若 A 中的列向量已是单位正交的, 则有 $\vec{b}_{\parallel} = AA^T \vec{b}$. 这时 AA^T (或是 QQ^T) 是关于 $\text{Col } A$ 的投影算符.



本章作业

- 6.1 节: 1, 3, 15, 20, 27
- 6.2 节: 5, 9, 15, 31
- 6.3 节: 1, 9, 15, 23
- 6.4 节: 3, 13, 17
- 6.5 节: 3, 7, 11
- 6.7 节: 9, 10, 13, 15



课堂练习

- 6.1 节: 5, 7, 17, 19, 23
- 6.2 节: 3, 7, 13, 29
- 6.3 节: 3, 7, 11, 19,
- 6.4 节: 1, 11, 15
- 6.5 节: 1, 13, 17
- 6.7 节: 1, 3, 11, 17

