

正交补和最小化

1.正交补

1.1 定义:正交补

若 U 是 V 的子集,那么 U 的正交补,记作 U^\perp ,是与 U 中每个向量都正交的向量构成的集合.即

$$U^\perp = \{v \in V : \forall u \in U, \langle u, v \rangle = 0\}$$

1.2 正交补的性质

- (1) 若 U 是 V 的子集,那么 U^\perp 是 V 的子空间.
- (2) $\{0\}^\perp = V, V^\perp = \{0\}$.
- (3) 若 U 是 V 的子集,那么 $U \cap U^\perp \subseteq \{0\}$.
- (4) 若 U, W 是 V 的子集且 $U \subseteq W$,那么 $W^\perp \subseteq U^\perp$.

下面的结论表明, 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.

1.3 子空间及其正交补的直和

若 U 是 V 的有限维子空间,那么 $V = U \oplus U^\perp$.

Proof.

首先我们证明 $V = U + U^\perp$.

为此,设 $v \in V$,令 e_1, \dots, e_n 是 U 的规范正交基.在Bessel不等式的证明中我们曾做过如下分解

$$v = \underbrace{\sum_{k=1}^n |\langle v, e_k \rangle| e_k}_u + \underbrace{v - \sum_{k=1}^n |\langle v, e_k \rangle| e_k}_w$$

于是 $u \in U$.又对于任意 $1 \leq k \leq n$ 有

$$\langle w, e_k \rangle = \langle v, e_k \rangle - \langle v, e_k \rangle = 0$$

从而 w 与 $\text{span}(e_1, \dots, e_n) = U$ 中的每个向量都正交,这表明 $w \in U^\perp$.于是 $V = U + U^\perp$.

由1.2又可知 $U \cap U^\perp = \{0\}$,于是 $V = U \oplus U^\perp$.

由上式可得 U^\perp 的维数.

1.4 正交补的维数

设 V 是有限维的, U 是 V 的子空间,那么 $\dim U^\perp = \dim V - \dim U$.

这由1.3可以很快地得到.以下结论是1.3的另一重要推论.

1.5 正交补的正交补

设 U 是 V 的一个有限维子空间,那么 $U = (U^\perp)^\perp$.

Proof.

对于任意 $u \in U$ 和任意 $v \in U^\perp$,都有 $\langle u, v \rangle = 0$.于是 $\langle v, u \rangle = 0$,进而 $u \in (U^\perp)^\perp$.于是 $U \subseteq (U^\perp)^\perp$.

对于任意 $v \in (U^\perp)^\perp \subseteq V$,由1.3可知存在唯一 $u \in U, w \in U^\perp$ 使得 $v = u + w$.于是 $v - u = w \in U^\perp$.

由于 $v, u \in (U^\perp)^\perp$,于是 $v - u \in (U^\perp)^\perp$.于是 $v - u \in U^\perp \cap (U^\perp)^\perp$,这表明 $v - u = \mathbf{0}$,从而 $v = u \in U$.因此 $(U^\perp)^\perp \subseteq U$.

结合可知 $U = (U^\perp)^\perp$.

以及我们可以通过正交补的性质反推子空间的性质.

1.6 子空间和其正交补的关联

设 U 是 V 的有限维子空间,那么 $U^\perp = \{\mathbf{0}\} \Leftrightarrow U = V$.

我们现在引入一个重要的算子:正交投影.

1.7 定义: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 $P_U \in \mathcal{L}(V)$:

对于任意 $v \in V$,将其写为 $v = u + w$ 的形式,其中 $u \in U, w \in U^\perp$,定义 $P_U v = u$.

根据它的名字,应该能想到这与高中所说的投影向量的关系.正交投影具有如下性质.

1.8 正交投影的性质

- (1) $P_U \in \mathcal{L}(V)$.
- (2) 对任意 $u \in U$,都有 $P_U u = u$.
- (3) 对任意 $w \in U^\perp$,都有 $P_U w = \mathbf{0}$.
- (4) $\text{range } P_U = U, \text{null } P_U = U^\perp$.
- (5) 对任意 $v \in V$ 都有 $v - P_U v \in U^\perp$.

(6) 对任意 $k \in \mathbb{N}$, 都有 $P_U^k = P_U$.

(7) 对于任意 $v \in V$ 都有 $\|P_U v\| \leq \|v\|$.

(8) 若 e_1, \dots, e_m 是 U 的规范正交基, 那么对于任意 $v \in V$ 有 $P_U v = \langle v, e_1 \rangle e_1 + \dots + \langle v, e_m \rangle e_m$.

你应当自行验证上面所说的每一条性质. 我们将用正交投影再次证明 Riesz 定理.

2.1 Riesz 定理

设 V 是有限维的. 对于任意 $v \in V$, 定义 $\phi_v \in V'$ 为: $\forall u \in V, \phi_v(u) = \langle u, v \rangle$. 那么 $v \mapsto \phi_v$ 是满射.

Proof.

设 $\phi \in V'$. 若 $\phi = \mathbf{0}$, 那么 $\phi = \phi_{\mathbf{0}}$, 从而假设 $\phi \neq \mathbf{0}$.

于是 $\text{null } \phi \neq V$, 这表明 $(\text{null } \phi)^\perp \neq \{\mathbf{0}\}$.

令 $w \in (\text{null } \phi)^\perp$ 且 $w \neq \mathbf{0}$, 令 $v = \frac{\phi(w)}{\|w\|^2} w$, 那么 $v \in (\text{null } \phi)^\perp$.

在 v 的定义两侧取范数可得 $\|v\| = \frac{|\phi(w)|}{\|w\|}$.

在 v 的定义两侧作用 ϕ 可得 $\phi(v) = \frac{\phi(w)}{\|w\|^2} \phi(w) = \frac{|\phi(w)|^2}{\|w\|^2} = \|v\|^2$.

现在设 $u \in V$. 利用上式可得 $u = \left(u - \frac{\phi(u)}{\phi(v)} v\right) + \frac{\phi(u)}{\|v\|^2} v$.

上式中括号内的向量属于 $\text{null } \phi$, 于是与 v 正交. 于是将此式两边与 v 作内积可得

$$\langle u, v \rangle = \frac{\phi(u)}{\|v\|^2} \langle v, v \rangle = \phi(u)$$

因此 $\phi = \phi_v$, 从而 $v \mapsto \phi_v$ 是满射, 因此命题得证.

2. 最小化问题

在学习初等的几何时我们就知道点到直线(或平面)上最短的距离是垂线段的长度. 我们将在向量空间中扩展这一结论.

2.1 到子空间的最短距离

设 U 是 V 的有限维子空间, $v \in V$ 且 $u \in U$. 那么 $\|v - P_U v\| \leq \|v - u\|$, 当且仅当 $u = P_U v$ 时等号成立.

Proof.

我们有一个简洁的证明方法.

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

你可以在二维或三维的欧几里得空间中通过几何直观更清晰地理解上述不等式的证明.

例如,我们可以用这样的方法来逼近正弦函数.假如我们要求出次数不高于5的实系数多项式 u 使得在 $[-\pi, \pi]$ 上逼近 $\sin x$.这里的逼近是指令 $\int_{-\pi}^{\pi} |\sin x - u(x)|^2 dx$ 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构成的实内积空间,其上内积定义为

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

令 $v(x) = \sin x$ 是 $C[-\pi, \pi]$ 上的函数,令 U 表示由次数不高于5的所有实系数多项式构成的 $C[-\pi, \pi]$ 的子空间.现在我们只需求出 $u \in U$ 使得 $\|v - u\|$ 尽可能小.

这样,我们将 U 的标准基 $1, x, \dots, x^5$ 应用Gram-Schmidt过程改写为 U 的规范正交基 e_1, \dots, e_6 .接着利用1.3的(8)计算 $P_U v$,即可得到近似结果 $u(x)$.这结果比同样次数的Taylor多项式在稍远处要精确的多.

3. 伪逆

设 $T \in \mathcal{L}(V, W)$ 且 $w \in W$.考虑这样一个问题:求出 $v \in V$ 使得 $Tv = w$.

如果 T 是可逆的,那么这样的 v 唯一存在.否则,对于特定的 w 很可能是无解或者有无穷多个解.

当 T 不可逆时,我们仍期望对上述方程给出一个较为精确的答案.例如,如果 $Tv = w$ 无解,就找出使得 $\|Tv - w\|$ 尽可能小的 v ;如果 $Tv = w$ 有无穷多个解,就找出使得 $\|v\|$ 最小的那个解.如此,我们将开始伪逆的定义.

3.1 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 $T|_{(\text{null } T)^\perp}$ 是将 $(\text{null } T)^\perp$ 映成 $\text{range } T$ 的单射.

Proof.

设 $v \in (\text{null } T)^\perp$ 且 $T|_{(\text{null } T)^\perp} v = \mathbf{0}$,进而 $Tv = \mathbf{0}$.于是 $v \in \text{null } T$.

又 $(\text{null } T) \cap (\text{null } T)^\perp = \{\mathbf{0}\}$,于是 $v = \mathbf{0}$.进而 $\text{null } T|_{(\text{null } T)^\perp} = \{\mathbf{0}\}$,这表明 $T|_{(\text{null } T)^\perp}$ 是单射.

显然 $\text{range } T|_{(\text{null } T)^\perp} \subseteq \text{range } T$.

设 $w \in \text{range } T$,于是存在 $v \in V$ 使得 $Tv = w$.于是存在 $u \in \text{null } T$ 和 $x \in (\text{null } T)^\perp$ 使得 $v = u + x$.于是

$$T|_{(\text{null } T)^\perp} x = Tx = Tv - Tu = w - \mathbf{0} = w$$

这表明 $w \in \text{range } T|_{(\text{null } T)^\perp}$,因此 $\text{range } T \subseteq \text{range } T|_{(\text{null } T)^\perp}$.

于是 $\text{range } T|_{(\text{null } T)^\perp} = \text{range } T$.

现在,我们可以定义伪逆了.

3.2 定义:伪逆

设 V 是有限维的, $T \in \mathcal{L}(V, W)$. T 的伪逆 $T^\dagger \in \mathcal{L}(W, V)$ 是如下定义的线性映射:

$$\forall w \in W, T^\dagger w = (T|_{(\text{null } T)^\perp})^{-1} P_{\text{range } T} w$$

伪逆具有如下的性质.

3.3 伪逆的性质

设 V 是有限维的且 $T \in \mathcal{L}(V, W)$.

- (1) 若 T 可逆,那么 $T^\dagger = T^{-1}$.
- (2) $TT^\dagger = P_{\text{range } T}$ 是将 W 映成 $\text{range } T$ 的正交投影.
- (3) $T^\dagger T = P_{(\text{null } T)^\perp}$ 是将 V 映成 $(\text{null } T)^\perp$ 的正交投影.

由伪逆出发,我们就可以给出开头所讲的求得线性映射方程的近似解.

3.4 伪逆可给出最佳近似解或最优解

设 V 是有限维的, $T \in \mathcal{L}(V, W)$ 且 $w \in W$.

- (1) 若 $v \in V$,那么

$$\|T(T^\dagger w) - w\| \leq \|Tv - w\|$$

当且仅当 $v \in T^\dagger w + \text{null } T$ 时等号成立.

- (2) 若 $v \in T^\dagger w + \text{null } T$,那么

$$\|T^\dagger w\| \leq \|v\|$$

当且仅当 $v = T^\dagger w$ 时上式成立.

Proof.

- (1) 设 $v \in V$,那么

$$Tv - w = \underbrace{(Tv - TT^\dagger w)}_x + \underbrace{(TT^\dagger w - w)}_y$$

上式中 $x \in \text{range } T$.因为 TT^\dagger 将 W 映成 $\text{range } T$ 的正交投影,于是 $y \in (\text{range } T)^\perp$.

于是 $\langle x, y \rangle = 0$.于是根据勾股定理 $\|Tv - w\|^2 = \|x\|^2 + \|y\|^2 \geq \|y\|^2$,当且仅当 $x = \mathbf{0}$ 时等号成立.

此时 $T(v - T^\dagger w) = \mathbf{0}$,于是 $v - T^\dagger w \in \text{null } T$,即 $v \in T^\dagger w + \text{null } T$.

- (2) 设 $v \in T^\dagger w + \text{null } T$,那么 $v = (v - T^\dagger w) + T^\dagger w$.

而 $T^\dagger w \in (\text{null } T)^\perp, v - T^\dagger w \in \text{null } T$.根据勾股定理可知 $\|T^\dagger w\| \leq \|v\|$,当且仅当 $v = T^\dagger w$ 时等号成立.