Some Notes For Inner Products

Xiaoyu Chen

Define 1 (Inner Products)**.** *对于一个定义在域 $\mathbb{F}$($\mathbb{F}$ 可以是 $\mathbb{R}, \mathbb{C}$) 上的向量空间 V. 如果函数 $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{F}$ 对于 $\forall x, y, z \in V$ 和 $\forall c \in \mathbb{F}$ 都满足:*

(1) $\langle x, x\rangle \geq 0$ 非负性 *(Nonnegativity)*

(1a) $\langle x, x\rangle = 0$ *iff* $x = 0$ 正性 *(Positivity)*

(2) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ 可加性 *(Additivity)*

(3) $\langle cx, y\rangle = c\langle x, y\rangle$ 齐次性 *(Homogeneity)*

(4) $\langle x, y\rangle = \overline{\langle y, x\rangle}$ 共轭对称性 *(Hermitian Property)*

$\Rightarrow$ 则 $\langle\cdot,\cdot\rangle$ 是**一个内积** *(inner product)*.

$\Rightarrow$ 如果不考虑 *(1a)*, 我们称 $\langle\cdot,\cdot\rangle$ 是**一个半内积** *(semi-inner product)*. 所以内积当然也是半内积.

$\triangle$ 上面这些性质有些时候也被称为内积的**公理**.

Fact 1. $\langle x, y\rangle := y^* x$ 是一个内积

Fact 2. 对于函数 $(\cdot,\cdot) : V \times V \to \mathbb{F}$. $(x, y) := y^* D x$, 其中 $D = \operatorname{diag}(d_1, \cdots, d_n) \in M_n(\mathbb{F})$, 有:

(1) $\Leftarrow$ 如果 D 满足 $d_i \geq 0, \forall i \in [n]$

(1a) $\Leftarrow$ 如果 D 满足 $d_i > 0, \forall i \in [n]$

(2) $\Leftarrow \forall D$

(3) $\Leftarrow \forall D$

(4) $\Leftarrow D = D^*$, *i.e.* D 是 *Hermitian* 阵

特别的, 如果 D 实对角阵, 且对角元全部 > 0, 则 $(\cdot,\cdot)$ 是一个内积.

Property 1. 通过内积的定义, 可以导出如下性质:

(a) $\langle x, cy\rangle = \bar{c}\langle x, y\rangle$

(b) $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$

(c) $\langle ax + by, cw + dz\rangle = a\bar{c}\langle x, w\rangle + a\bar{d}\langle x, z\rangle + b\bar{c}\langle y, w\rangle + b\bar{d}\langle y, z\rangle$

(d) $\langle x, \langle x, y\rangle y\rangle = \langle x, y\rangle\overline{\langle x, y\rangle} = |\langle x, y\rangle|^2$

(e) $\langle x, y\rangle = 0, \forall y \in V$ *iff* $x = 0$

这里只有 *(e)* 的证明需要用到 *(1)* 和 *(1a)*.

Theorem 1 (Cauchy–Schwarz inequality)**.** 对于任意的半内积 $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{F}$, 我们都有:

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle, \quad \forall x, y \in V$$

proof.(半内积). 考虑一个多项式 $p(t)=\langle tx-e^{i\theta}y, tx-e^{i\theta}y\rangle, \forall t,\theta\in\mathbb{R}, x,y\in V$, 这时有:

$$\begin{aligned}
p(t) &= (tx-e^{i\theta}y)^*(tx-e^{i\theta}y)\\
&= (tx^*-e^{-i\theta}y^*)(tx-e^{i\theta}y)\\
&= t^2\langle x,x\rangle - te^{i\theta}\langle y,x\rangle - te^{-i\theta}\langle x,y\rangle + \langle y,y\rangle\\
&= t^2\langle x,x\rangle - \overline{te^{-i\theta}\langle x,y\rangle} - te^{-i\theta}\langle x,y\rangle + \langle y,y\rangle\\
&= t^2\langle x,x\rangle - 2\mathrm{Re}\left[te^{-i\theta}\langle x,y\rangle\right] + \langle y,y\rangle
\end{aligned}$$

这个时候, 我们将 θ 取定, 使其满足 $e^{-i\theta}\langle x,y\rangle = |\langle x,y\rangle|$, 因为 $\langle x,y\rangle\in\mathbb{F}\subset\mathbb{C}$, 所以这样的 θ 一定存在. 所以, 这时我们有 $p(t)=t^2\langle x,x\rangle - 2t|\langle x,y\rangle| + \langle y,y\rangle$.

这个时候, 如果 $\langle x,x\rangle = 0\wedge\langle x,y\rangle\neq 0$, 则 $p(t) = -2t|\langle x,y\rangle| + \langle y,y\rangle$ 那么当 t 足够大的时候, $p(t)<0$, 这不满足**公理 (1)**. 所以当 $\langle x,x\rangle=0$ 时, $\langle x,y\rangle = 0$, $|\langle x,y\rangle|^2\leq\langle x,x\rangle\langle y,y\rangle$ 自然满足.

如果 $\langle x,x\rangle\neq 0$. 则我们可以令 $t_0=\dfrac{|\langle x,y\rangle|}{\langle x,x\rangle}$. 则:

$$\begin{aligned}
p(t_0) \quad &= \frac{|\langle x,y\rangle|^2}{\langle x,x\rangle} - 2\frac{|\langle x,y\rangle|^2}{\langle x,x\rangle} + \langle y,y\rangle\\
&= \quad -\frac{|\langle x,y\rangle|^2}{\langle x,x\rangle} + \langle y,y\rangle \quad \geq 0\\
&\Rightarrow \qquad \langle y,y\rangle \quad \geq \frac{|\langle x,y\rangle|^2}{\langle x,x\rangle}\\
&\qquad \langle x,x\rangle\langle y,y\rangle \quad \geq |\langle x,y\rangle|^2
\end{aligned}$$

□

上面这个证明对于内积和半内积都是有效的. 不过, 如果只关注内积, 还有一个更加简单的证明方法.

proof.(内积). 考虑 $v=\langle y,y\rangle x - \langle x,y\rangle y$. 则

$$\begin{aligned}
0 &\leq \langle v,v\rangle\\
&= \langle\langle y,y\rangle x - \langle x,y\rangle y, \langle y,y\rangle x - \langle x,y\rangle y\rangle\\
&= \langle y,y\rangle^2\langle x,x\rangle - \langle y,y\rangle\overline{\langle x,y\rangle}\langle x,y\rangle \underbrace{-\langle x,y\rangle\langle y,y\rangle\langle y,x\rangle + \langle x,y\rangle\overline{\langle x,y\rangle}\langle y,y\rangle}_{0}\\
&= \langle y,y\rangle^2\langle x,x\rangle - \langle y,y\rangle\overline{\langle x,y\rangle}\langle x,y\rangle\\
&= \langle y,y\rangle(\langle y,y\rangle\langle x,x\rangle - \overline{\langle x,y\rangle}\langle x,y\rangle)
\end{aligned}$$

如果 $\langle y,y\rangle = 0$, 则 $y=0$, 则 $\langle x,y\rangle=0$, 结论显然成立.

另一方面, 如果 $\langle y,y\rangle\neq 0$, 则我们知道:

$$\begin{aligned}
0 \leq \langle y,y\rangle\langle x,x\rangle - \overline{\langle x,y\rangle}\langle x,y\rangle\\
|\langle x,y\rangle|^2 \leq \langle x,x\rangle\langle y,y\rangle
\end{aligned}$$

这种证明方式还告诉我们, **考虑内积时**, 柯西不等式取等的条件是:

$$\begin{aligned}
v &= 0\\
\langle y,y\rangle x &= \langle x,y\rangle y\\
x &= \frac{\langle x,y\rangle}{\langle y,y\rangle}y
\end{aligned}$$

即 x 和 y 线性相关 (平行). □

Define 2 (范数 (norm))**.** *令 V 是 $\mathbb{F}$ ($\mathbb{R}$ 或 $\mathbb{C}$) 上的一个向量空间. 一个函数 $||\cdot||: V \to \mathbb{F}$ 如果对于 $\forall x, y \in V, \forall c \in \mathbb{F}$ 满足:*

(1)	$\|\|x\|\| > 0$	*非负性 (Nonnegativity)*
(1a)	$\|\|x\|\| = 0$ *iff* $x = 0$	*正性 (Positivity)*
(2)	$\|\|cx\|\| = c\|\|x\|\|$	*齐次性 (Homogeneity)*
(3)	$\|\|x+y\|\| \leq \|\|x\|\| + \|\|y\|\|$	*三角形不等式 (Triangle Inequality)*

$\Rightarrow$ *则称 $||\cdot||$ 是一个**范数** (norm).*

$\Rightarrow$ *如果不考虑 (1a), 则称 $||\cdot||$ 是一个 **半范数** (semi-norm).*

Lemma 1. *$||\cdot||$ 是一个在 $\mathbb{F}$ 上的半范数, 则*

$$|\,||x|| - ||y||\,| \leq ||x-y||$$

proof. 因为 $x = y + (x-y)$ 所以根据三角形不等式, 有

$$||x|| \leq ||y|| + ||x-y||$$
$$||x|| - ||y|| \leq ||x-y||$$

同理, 可以得到 $||y|| - ||x|| \leq ||x-y||$. □

Fact 3. *如果有一个 (半) 内积 $\langle\cdot,\cdot\rangle: V \times V \to \mathbb{F}$, 则我们可以通过 $||x|| := \langle x, x\rangle^{1/2}$ 来定义出一个 (半) 范数.*

证明. (1), (1a), (2) 都很好验证, 下面来说明如何验证 (3).

$$\begin{aligned}
||x+y||^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= ||x||^2 + ||y||^2 + \langle x,y\rangle + \langle y,x\rangle \\
&= ||x||^2 + ||y||^2 + 2\mathrm{Re}\left[\langle x,y\rangle\right] \\
&\leq ||x||^2 + ||y||^2 + 2|\langle x,y\rangle| \\
&\leq ||x||^2 + ||y||^2 + 2\langle x,x\rangle^{1/2}\langle y,y\rangle^{1/2} \\
&= ||x||^2 + ||y||^2 + 2||x||\,||y|| \\
&= (||x|| + ||y||)^2
\end{aligned}$$ □

1 Self-Adjoint Operator

Define 3 (adjoint)**.** *Let V and W be real or complex finite dimensional vector spaces with inner products $\langle\cdot,\cdot\rangle_V$ and $\langle\cdot,\cdot\rangle_W$, respectively. Let $L: V \to W$ be linear. If there is a transformation $L^*: W \to V$ for which*

$$\langle Lv, w\rangle_W = \langle v, L^*w\rangle_V \tag{1}$$

holds for every pair of vectors $v \in V$ and $w \in W$, then L^* is said to be the adjoint *of L.*

Remark 1. *这里, $v \in V, L: V \to W$, 所以 $Lv \in W$, 所以 $\langle Lv, w\rangle_W$ 中左右两个变量都在 W 中. 同理, $w \in W, L^*: W \to V$, 所以 $L^*w \in V$, 所以 $\langle v, L^*w >_V$ 中, 左右两个变量都在 V 中.*

在下面的证明中, 我们会用到关于向量和矩阵表示的问题, 这里需要引入基变换的语言.

Define 4 (Change of Basis)**.** *V 中一组相互正交的单位向量 $\mathcal{B}=\{v_1,v_2,\cdots,v_n\}$ 可以作为 V 这个空间的一组基. 对于 $\forall x\in V$, 我们可以通过 $\mathcal{B}$ 来表示出 x. 如果 $x=\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n$, 则*

$$[x]_{\mathcal{B}}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}$$

称为 x 在基 $\mathcal{B}$ 中的表示 ($\mathcal{B}$ basis representation of x).

仅仅有向量的基表示是不够的, 矩阵同样需要有基表示. 假设现在有两个基 $\mathcal{B}_1,\mathcal{B}_2$, 对于任何一个向量 $x\in V$, 我们希望写出 Tx 在 $\mathcal{B}_2$ 中的表示 $[Tx]_{\mathcal{B}_2}$. 不妨 $\mathcal{B}_1=\{v_1,v_2,\cdots,v_n\}$, 且

$$[x]_{\mathcal{B}_1}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}$$

那么我们有

$$\begin{aligned}[Tx]_{\mathcal{B}_2}&=\left[\sum_{j=1}^{n}\alpha_jTv_j\right]_{\mathcal{B}_2}=\sum_{j=1}^{n}\alpha_j[Tv_j]_{\mathcal{B}_2}\\&=\sum_{j=1}^{n}\alpha_j\begin{bmatrix}t_{1j}\\ \vdots\\ t_{nj}\end{bmatrix}=\begin{bmatrix}t_{11}&\cdots&t_{1n}\\ \vdots&\ddots&\vdots\\ t_{n1}&\cdots&t_{nn}\end{bmatrix}\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix}\end{aligned}$$

所以, 可以做如下定义

Define 5.

$${}_{\mathcal{B}_2}[T]_{\mathcal{B}_1}=\begin{bmatrix}t_{11}&\cdots&t_{1n}\\ \vdots&\ddots&\vdots\\ t_{n1}&\cdots&t_{nn}\end{bmatrix}=[[Tv_1]_{\mathcal{B}_2},\cdots,[Tv_n]_{\mathcal{B}_2}]$$

这是因为从上面的观察可以发现 ${}_{\mathcal{B}_2}[T]_{\mathcal{B}_1}[x]_{\mathcal{B}_1}=[Tx]_{\mathcal{B}_2}$. 所以 ${}_{\mathcal{B}_1}[T]_{\mathcal{B}_2}$ 的取值情况和 x 无关. 特别的, 当 $\mathcal{B}_1=\mathcal{B}_2$ 时, 我们称这种形式 ${}_{\mathcal{B}_1}[T]_{\mathcal{B}_1}$ 为 T 在基 $\mathcal{B}_1$ 中的表示 ($\mathcal{B}_1$ basis representation of T).

Fact 4. *基变换, 不改变矩阵的特征值. i.e. $Av=\lambda v\Leftrightarrow_{\mathcal{B}}[A]_{\mathcal{B}}[v]_{\mathcal{B}}=\lambda[v]_{\mathcal{B}}$.*

证明. $\Rightarrow$:

$$\begin{aligned}{}_{\mathcal{B}}[A]_{\mathcal{B}}[v]_{\mathcal{B}}&=[Av]_{\mathcal{B}}\\&=[\lambda v]_{\mathcal{B}}=\lambda[v]_{\mathcal{B}}\end{aligned}$$

$\Leftarrow$:

$$\begin{aligned}[Av]_{\mathcal{B}} &=_{\mathcal{B}} [A]_{\mathcal{B}}[v]_{\mathcal{B}} \\ &= \lambda[v]_{\mathcal{B}} \\ &= [\lambda v]_{\mathcal{B}}\end{aligned}$$

所以 $Av = \lambda v$ (因为 Av 和 λv 在基 $\mathcal{B}$ 下是同一种表示). □

Define 6 (self-adjoint)**.** *对于任意一个 inner product $\langle \cdot, \cdot \rangle_V$, 和矩阵 $L : V \to V$. 如果对任意 $x, y \in V$, 都有:*

$$\langle Lx, y\rangle_V = \langle x, Ly\rangle_V$$

(i.e. $L = L^*$), 则我们说 L 是 self-adjoint 的.

self-adjoint 的矩阵的特征向量有很有用的性质.

1.1 Properties for self-adjoint operator

Remark 2. *这一小节中, 我么讨论的内积 $\langle \cdot, \cdot\rangle$ 都是一般的内积.*

Proposition 1. *Let V be a complex vector space with an inner product. If $L : V \to V$ is a self-adjoint linear transformation, then the eigenvalues of L are real numbers, and eigenvectors of L corresponding to distinct eigenvalues are orthogonal.*

证明. *[1] the eigenvalues of L are all real numbers*:
如果 L 有一个特征值对 (λ, v) (不妨 $\langle v, v\rangle = 1$), 则我们有:

$$\lambda = \lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Lv, v\rangle = \langle v, Lv\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle = \overline{\lambda}$$

所以很容易可以看出来, λ 应该是一个实数.

[2] the eigenvectors of L corresponding to distinct eigenvalues are orthogonal:
考虑 L 的两个不同的特征值对 (λ_1, v_1), (λ_2, v_2), 且 $\lambda_1 \neq \lambda_2$ (注意, 这里由 [1] 可知, $\lambda_1, \lambda_2 \in \mathbb{R}$). 然后就有

$$\lambda_1\langle v_1, v_2\rangle = \langle \lambda_1 v_1, v_2\rangle = \langle Lv_1, v_2\rangle = \langle v_1, Lv_2\rangle = \langle v_1, \lambda_2 v_2\rangle = \overline{\lambda_2}\langle v_1, v_2\rangle = \lambda_2\langle v_1, v_2\rangle$$

容易看出, 因为 $\lambda_1 \neq \lambda_2$, 所以这个式子只在 $\langle v_1, v_2\rangle = 0$ 时成立. □

Fact 5. *If we do not change the definition of matrix multiplication. Then, for two different inner products $\langle \cdot, \cdot\rangle_a, \langle \cdot, \cdot\rangle_b$, defined on the same n-dimensional space V, $\forall A \in V \to V$ which is linear, A has the same eigenvalues on both inner-spaces.*

证明. 要理解这个 Fact, 只需要回顾 eigenvalue 的定义, 通常我们使用下面的方式来定义 A 的 eigenvalues:

$$\{\lambda | v \in V, Av = \lambda v\}$$

所以, 很容易能看出, inner product 的定义对 A 的特征值的取值是不会产生影响的. □

Lemma 2. *Let V be a complex, finite dimensional vector space with dimension ≥ 1, if $L: V \to V$ is linear, then L has at least one eigenvalue.*

证明. 由上面的 Fact 可知, 我们在考虑矩阵的特征值时不需要考虑 inner product 的定义. 而方程 $\det(A-\lambda I) = 0$ 一定会有 $1 \sim n$ 个不同的界. 所以矩阵 L 当然会有至少一个特征值 (这个特征值当然也有对应的特征向量). $\square$

Remark 3. 上面的证明只说明了 L 是一个在 $V \to V$ 上转移的矩阵的情况, 下面我们来证明一个更加强的结论.

Lemma 3. *Let G be a complex, finite dimensional vector space with dimension ≥ 1. If $L: G \to G$ is a linear transformation on G. Then, for a subspace V of G, if $L|_V: V \to V$ (i.e. L is closed on V), then $L|_V$ has at least one eigenvalue on V.*

证明. 不妨 G 是一个 n 维空间, V 是 G 的一个子空间, 有 m 维. $V^\perp$ 当然也是 G 的一个子空间, 有 $n-m$ 维.

这个时候, 我们可以从 V 中选出 m 个正交的单位向量 $\{v_1, v_2, \cdots, v_m\}$. 同样, 我们可以从 $V^\perp$ 中选出 $n-m$ 个正交的单位向量 $\{v_1^\perp, v_2^\perp, \cdots, v_{n-m}^\perp\}$. 显然, 将这些向量放到一起, 我们就得到了 G 的一组基:

$$\mathcal{B} = \{v_1, v_2, \cdots, v_m, v_1^\perp, v_2^\perp, \cdots, v_{n-m}^\perp\}$$

所以, $\forall x \in V$, 有:

$$\begin{aligned}
{}_{\mathcal{B}}[L]_{\mathcal{B}}[x]_{\mathcal{B}} &= [[Lv_1]_{\mathcal{B}}, [Lv_2]_{\mathcal{B}}, \cdots, [Lv_m]_{\mathcal{B}}, [Lv_1^\perp]_{\mathcal{B}}, \cdots [Lv_{n-m}^\perp]_{\mathcal{B}}][x]_{\mathcal{B}} \\
&= \begin{bmatrix} A_{m\times m} & \bigstar_{m\times n-m} \\ 0_{n-m\times m} & \bigstar_{n-m\times n-m} \end{bmatrix} \begin{bmatrix} x_m \\ 0_{n-m} \end{bmatrix} \\
&= \begin{bmatrix} A_{m\times m}x_m \\ 0_{n-m} \end{bmatrix}
\end{aligned}$$

结合前面基变换不改变特征值的结论, 可以知道对于 $x \in V$, L 的特征值, 等于 $A_{m\times m}$ 的特征值 (A 有的特征值 L 都有). 在由前面一个 Lemma 给出的结论, 可以知道, L 在 V 上至少有一个特征值. $\square$

Remark 4. 上面两个 *Lemma* 联合起来, 就证明了一般的情况. 即, 如果 $L: V \to V$, 这里只是说 L 在 V 这个空间封闭, 则 L 一定在这个空间至少有一个特征值 (当然也就有至少一个特征向量).

下面给出 self-adjoint operator 的一个非常重要的性质.

Theorem 2. *Let V be a complex, finite dimensional vector space. If $L: V \to V$ is a self-adjoint linear transformation, then thre is an orthonormal basis for V that is composed of eigenvectors of L. The matrix of L related to this basis is diagonal.*

Remark 5. 用人话说就是: 任意一个内积的 *self-adjoint operator* 的特征向量都可以构成关于这个内积的一组基.

证明. 显然, L 的所有特征向量, 一定可以构成一个子空间的正交基, 计作 $\mathcal{B}$. 我们将 L 的所有特征向量张成的空间记为 $B = \mathrm{Span}(\mathcal{B})$, 然后 $S = B^\perp$. 不难发现, L 对 S 是封闭的, 对于 $\forall v \in \mathcal{B}, u \in S$.

$$\langle v, Lu\rangle = \langle Lv, u\rangle = \langle \lambda v, u\rangle = \lambda\langle v, u\rangle = 0$$

所以 Lu 和 v 依然关于 $\langle\cdot,\cdot\rangle$ 正交. 因为这对于所有的特征向量都成立, 所以有 $Lu \perp B$, 也就是说 $Lu \in S$. 所以, 我们可以说, S 关于 L 是封闭的.

然后利用我们前面的结论可以看出, 如果 $S \neq \{0\}$, 则 L 在 S 中一定存在一个特征值, 也就存在一个特征向量. 但是这样的话, 这个特征向量就也应该存在于 B (根据定义). 这导致 $B \cap S \neq \emptyset$, 就产生了矛盾.

所以我们一定有 $S = \{0\}$. 所以 L 的所有特征向量一定可以关于 $\langle\cdot,\cdot\rangle$ 构成一个 V 的基. □

Remark 6. 这个定理, 会导致本来只在一般 *inner product* 上成立的 *Courant-Fischer Theorem]* 在任意 *self-adjoint* 上成立.