

一位八无混子的笔记 – Algèbre Linéaire 2

一位无竞赛、无科研、无学生会、无班委、无社团、
无 npy、无运动、无社交的八无混子

2025 年 9 月 24 日



图 1: 作者

0 引言

0.1 关于本拙作

M.H.Dehon 老师是位非常优秀的数学教育者，学识渊博、水平极高. 不过他的课堂内容深度很大，常常让同学们一时难以消化；板书逻辑严谨，但对我们来说稍显晦涩；教材更是原汁原味，理解起来颇具挑战. 因此，这份笔记的目的，是在本人力所能及的范围内，帮大家更好地理解和消化他课堂上的内容（虽然本人水平有限，但聊胜于无，希望能有所帮助）.

需要说明的是笔记侧重于实用性（即如何理解知识并做题）而非基本概念和定理的证明，即便我认为后者其实更重要.

0.2 注意事项

1. 本人衷心希望您看到这里的时候并非考试前夕，否则我建议您早点睡觉，以便明天以丰沛的精神面对考试；
2. 本作品不可代替教材；
3. 如有问题请立刻向作者指出，本作品允许一切形式的谩骂、攻击、批评.

0.3 关于 Algèbre Linéaire

线性代数是一门非常重要的数学分支，需要我们具备良好的抽象思维能力。如果您认为您已经学明白了线性代数，那多半说明您菌子吃多了，那我建议您再好好想一想。

本笔记将试图以比较简单的方式帮助大家理解线性代数的内容。但依鄙人愚见，线性代数的理解离不开实践。证明定理、做题固然重要，但随手画画矩阵、向量，做做矩阵乘法，探讨极端情况，画画线性映射的几何意义等等，也是不同形式的实践。

1 欧几里得空间上的自同态 (endomorphisms)

自同态实际上就是一种算子，这里我们就不区分这两个名词了。

1.1 线性自同态的伴随算子 (adjoint)

如果哪一天，您想不开要去学量子力学，那请务必学好这一节。

命题. 对于任意自同态 $u \in \mathcal{L}(E)$ ，存在唯一的自同态 $u^* \in \mathcal{L}(E)$ 使得

$$\forall x, y \in E \quad (u^*(x) | y) = (x | u(y)).$$

映射 $u \mapsto u^*$ 是 $\mathcal{L}(E)$ 上的一个对合自同态 (*involutif*，即平方后等于 id)；此外，

$$(\text{id}_E)^* = \text{id}_E, \quad (u \circ v)^* = v^* \circ u^*,$$

对所有 $u, v \in \mathcal{L}(E)$ 都成立。

定义. 对于所有 $u \in \mathcal{L}(E)$ ， E 上的自同态 u^* 被称为 u 的伴随 (*adjoint*)。

现在要证明伴随的存在性和唯一性。先给出一个定理辅助证明。

命题 (Riesz 表示定理). 设 E 是一个欧几里得空间。对于 E 上的每一个线性泛函 $f \in E^*$ ，都可以找到唯一的向量 $v_f \in E$ 使得

$$\forall x \in E \quad f(x) = (v_f | x).$$

称 v_f 为 f 的 *Riesz* 表示。

证明. 唯一性: 若 $v, w \in E$ 满足

$$(v | x) = (w | x), \quad \forall x \in E,$$

则对所有 x 有 $(v - w | x) = 0$ 。取 $x = v - w$ ，得

$$\|v - w\|^2 = (v - w | v - w) = 0,$$

故 $v = w$ ，于是表示向量若存在则唯一。

存在性: 取 E 的一组正交标准基 $\{e_1, \dots, e_n\}$ 。定义

$$v_f := \sum_{i=1}^n f(e_i) e_i \in E.$$

任取 $x \in E$ ，记 $x = \sum_{i=1}^n x_i e_i$ ，则

$$f(x) = f\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i f(e_i).$$

另一方面,

$$(v_f | x) = \left(\sum_{i=1}^n f(e_i) e_i \middle| \sum_{j=1}^n x_j e_j \right) = \sum_{i,j} f(e_i) x_j (e_i | e_j) = \sum_{i=1}^n f(e_i) x_i.$$

两式相等, 因此 $f(x) = (v_f | x)$ 对任意 x 成立. 存在性得证.

有没有更好的方法?

抽象方法 (其实就是书上的方法): 定义线性映射

$$J: E \rightarrow E^*, \quad x \mapsto (x | \cdot)$$

众所周知, E 和 E^* 维数相同, 根据秩定理 (théorème du rang, 中文是我瞎翻译的, 中文教材一般叫秩-零化度定理, 但那个和咱们学的形式不是很一样),

$$\text{rg } J = \dim E - \dim(\ker J)$$

若 J 是单射, 则 $\ker J = \{0\}$, 所以

$$\text{rg } J = \dim E = \dim E^*$$

因此 J 是满射. 从而 J 是同构当且仅当 J 是单射. 下面证明 J 是单射.

取 $v \in \ker J$, 则对所有 $y \in E$ 有

$$0 = J(v)(y) = (y | v).$$

取 $y = v$, 得 $\|v\|^2 = (v | v) = 0$, 所以 $v = 0$. 因此 $\ker J = \{0\}$, J 是单射. 从而 J 是同构.

既然 J 是满射, 则对任意 $f \in E^*$ (这里 f 对应的是 J 定义中的 $(x | \cdot)$), 都存在 $v_f \in E$ (对应 J 定义中的 x) 使得

$$J(v_f) = f,$$

即

$$\forall x \in E \quad f(x) = J(v_f)(x) = (v_f | x).$$

唯一性是因为 J 是单射.

□

我们回到伴随算子的定义上来. 下面证明伴随算子是存在且唯一的.

证明. 对任意 $x \in E$, 定义线性泛函

$$f_x: E \rightarrow \mathbb{R}, \quad y \mapsto (x | u(y)).$$

规定 u^* 的值为 f_x 的 Riesz 表示, 即

$$\forall x, y \in E \quad (u^*(x) | y) = f_x(y) = (x | u(y)).$$

伴随算子的线性请读者自行验证 (或者抄书).

□

注. 我们知道 E s'identifie canoniquement à E^* , 用我们普通话说就是他俩其实是一回事, 只是一个的向量是竖着一条矩阵, 一个是横着一条矩阵. 既然如此, 如果大家还记得转置自同态, 那其实伴随算子和转置自同态基本上就是一回事, 只不过伴随算子是定义在欧几里得空间上的.

好消息, 接下来开始具象了.

命题. 设 b 是 E 的一个正交标准基. 在基 b 下, 线性算子 u 的伴随算子的矩阵是 u 的矩阵的转置:

$$\text{Mat}_b(u^*) = (\text{Mat}_b(u))^{\top}.$$

证明. 设 e_1, \dots, e_n 是基 b 的向量. 矩阵 $\text{Mat}_b(u^*)$ 的第 (i, j) 个元素是 u^*e_j 在基 b 下的第 i 个坐标:

$$(e_i | u^*(e_j)) = (e_j | u(e_i)).$$

因此它正是 $u(e_i)$ 在基 b 下的第 j 个坐标, 也就是说, 就是矩阵 $\text{Mat}_b(u)$ 的第 (j, i) 个元素.

□

接下来来直观理解这一结论.

伴随算子本质上从内积的角度刻画了线性映射. 我们在二次元 (\mathbb{R}^2) 考虑这件事.

取标准正交基 $e_1 = (1, 0)$, $e_2 = (0, 1)$. 设 x 和 y 是 \mathbb{R}^2 中的两个向量. 它们的矩阵形式为:

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

它们的内积定义为:

$$(x | y) = x_1y_1 + x_2y_2 = X^{\top}Y = Y^{\top}X.$$

或者

$$(x | y) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 & y_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

这很好理解. 现在, 设 u 是 \mathbb{R}^2 上的一个线性自同态.

设 u 在基 b 下的矩阵是

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

那么 u 作用在 x 上的结果是

$$u(x) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

因此

$$(u(x) | y) = \begin{pmatrix} y_1 & y_2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

像之前一样, 把 y 立起来, x 躺下. 将后面两项视作一个整体取转置. 注意: $(AB)^\top = B^\top A^\top$

$$(u(x) | y) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

令 u^* 的矩阵为 $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ 那么

$$(u(x) | y) = (x | u^*(y)).$$

这就是伴随算子的矩阵.

命题. 设 $u \in \mathcal{L}(E)$. 则有如下关系:

1. $\ker u^* = (\operatorname{Im} u)^\perp$, $\operatorname{Im} u^* = (\ker u)^\perp$, $\operatorname{rg} u^* = \operatorname{rg} u$
2. 设 F 是 E 的一个向量子空间; F 是 u -稳定的当且仅当 F^\perp 是 u^* -稳定的.

证明. 对于命题 1, 按照定义验证即可:

$$u^*(x) = 0 \iff \forall y \in E \quad (u^*(x) | y) = 0 \iff \forall y \in E \quad (x | u(y)) = 0 \iff x \in (\operatorname{Im} u)^\perp.$$

对于命题 1 的第二条, 只需交换 u 和 u^* (因为 $(u^*)^* = u$). 对于第三条,

$$\operatorname{rg} u^* = \operatorname{codim} \ker u^* = \operatorname{codim} (\operatorname{Im} u)^\perp = \operatorname{rg} u.$$

第一个等号应用了秩定理, 第二个等号是命题 1 的第一条. 至于命题 2, 如果 F 是 u -稳定的, 那么对任意 $x \in F^\perp$,

$$\forall y \in F \quad (u^*(x) | y) = (x | u(y)) = 0,$$

因为 $u(y) \in F$. 反之亦然. □

关于伴随算子, 还有一些比较有趣的理解方法, 但这里太小写不下我们还没学过需要的预备知识.

1.2 正交算子 (orthogonal)

先规定符号, E 是一个欧几里得空间, 维数为 n , u 是 E 上的一个线性自同态.

定义. 称 E 的一个自同态 u 是 **正交的** (*orthogonal*), 如果它不改变内积, 也就是说

$$\forall x, y \in E \quad (u(x) | u(y)) = (x | y).$$

称 u 是 E 的一个 **线性等距映射** (*isométrie linéaire*) (或向量等距映射, ChatGPT 说的, 我不确定), 如果它不改变范数, 也就是说

$$\forall x \in E \quad \|u(x)\| = \|x\|.$$

注. 当您看到两个定义同时出现在 *Dehon* 的书里, 就说明它们是等价的.

命题. 设 $b = (e_1, \dots, e_n)$ 是 E 的一个正交标准基. 以下条件是等价的:

1. 自同态 u 是正交的;
2. 自同态 u 是线性等距映射;
3. 族 $(u(e_1), \dots, u(e_n))$ 是 E 的一个正交标准基.

证明. $1 \implies 3$ 是显然的. 对一个正交标准基, u 既不改变它们的长度, 也不改变它们两两之间的内积 (1 或 0), 因此它们的像仍然是一个正交标准基.

$3 \implies 2$ 来自于在正交标准基中标量积的表达式; 事实上, 在条件 3 下, 由于 u 的线性, 向量 $x \in E$ 在基 (e_1, \dots, e_n) 下的坐标与 $u(x)$ 在基 $(u(e_1), \dots, u(e_n))$ 下的坐标相同. 即如果 $x = x_1 e_1 + \dots + x_n e_n$, 那么 $u(x) = x_1 u(e_1) + \dots + x_n u(e_n)$. 它们的长度只取决于坐标 x_1, \dots, x_n , 因此 $\|u(x)\| = \|x\|$.

$2 \implies 1$ 来自极化公式:

$$(u(x) | u(y)) = \frac{1}{2}(\|u(x+y)\|^2 - \|u(x)\|^2 - \|u(y)\|^2) = \frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2) = (x | y).$$

□

一个典型的例子是之前学过的对称映射. 即如果 p 是一个正交投影算子, 那么 $s = 2p - \text{id}$ 是一个正交算子.

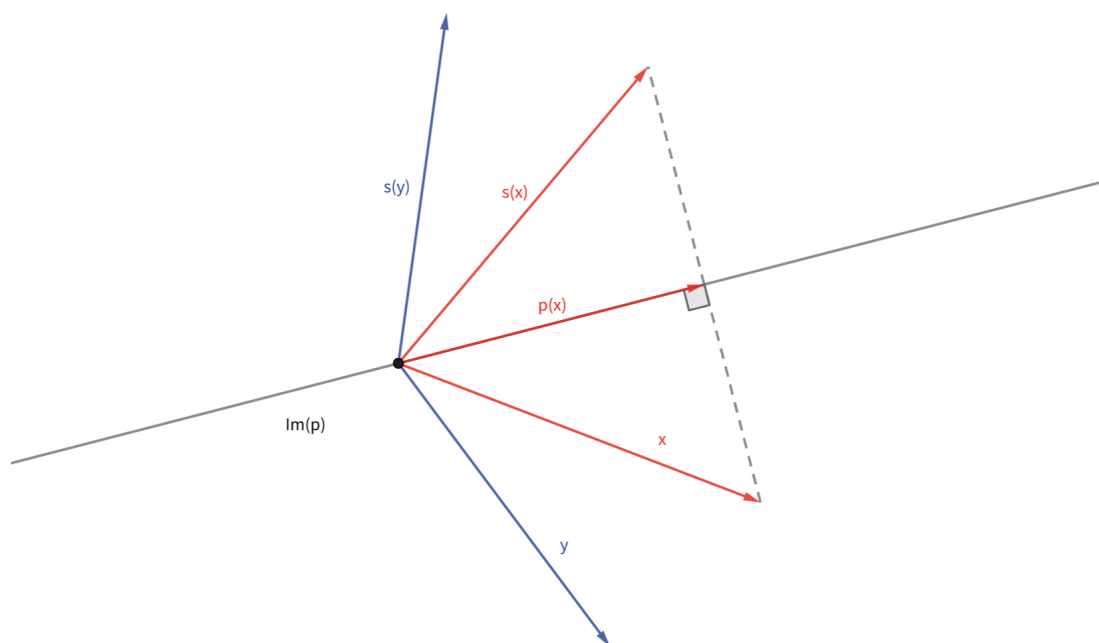


图 2: 对称映射是正交算子

上图是一个典型的二次元正交投影的对称映射. 可以看到, $s(x)$ 和 $s(y)$ 的夹角和 x 与 y 的夹角相同, 且长度也对应相同, 从而 $(s(x) | s(y)) = (x | y)$, 因此对称映射是一个正交算子.

接下来又要开始具象了, 因为涉及到矩阵了.

定义. 设 $A \in M_n(\mathbb{R})$ 是一个实方阵. 称 A 是一个 **正交矩阵** (*matrice orthogonale*), 如果

$$A^T A = I_n.$$

注. 不难看出, 如果 A 是正交矩阵, 那么 A 的各列构成一个正交标准基. 因为 $A^T A$ 的第 (i, j) 个元素是 A^T 的第 i 行与 A 的第 j 列的内积, 也就是 A 的第 i 列与第 j 列的内积. 如果 $A^T A = I_n$, 那么当 $i = j$ 时, 内积为 1; 当 $i \neq j$ 时, 内积为 0. 这正是正交标准基的定义.

$$\begin{pmatrix} 0 & 1 & 0 \\ \frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

命题. 设 b 是 E 的一个正交标准基. 以下条件是等价的:

1. u 是 E 的一个正交自同态;
2. 自同态 u 是可逆的, 并且

$$u^{-1} = u^*;$$

3. u 在基 b 下的矩阵是一个正交矩阵.

证明. 请读者自行验证. □

到现在, 我们不难看出: 正交自同态的作用, 实际上就是就是把坐标轴旋转 (可能还会翻转) 一下, 这一过程不改变向量之间的夹角和长度, 因此不改变内积. 因此, 任两个自同态的复合仍然是一个正交自同态. 聪明的您看到运算的封闭性, 会想到什么呢?

命题. E 中的正交自同态构成一个群, 记作 $O(E)$. 此外, 全体 n 阶正交矩阵也构成一个群.

这个群的单位元是恒等映射; 逆元是伴随算子. 除此之外, 我们之前学过, 1 和 -1 可以构成一个群. 现在, 考虑正交自同态 u 的行列式 $\det(u)$. 设 u 的矩阵为 A , 不难发现:

$$(\det(u))^2 = \det(A) \det(A^T) = \det(I) = 1.$$

命题. 行列式是 $O(E)$ 到 $\{1, -1\}$ 的一个满同态.

定义. 由行列式定义的满同态, 它的核 (注意: 单位元是 1 而非 0!) 被称为 E 的特殊正交群 (*groupe spécial orthogonal*), 记作 $SO(E)$.

命题. 如果 u 是一个正交自同态, 且线性子空间 F 是 u -稳定的, 那么 F^\perp 也是 u -稳定的.

证明. 只需套用之前伴随算子的性质即可. 注意 $u^* = u^{-1}$. □

命题. 正交自同态的特征值是模为 1 的复数. 在实数域上, 正交自同态的特征值只能是 1 或 -1 .

证明. 设 λ 是 u 的一个特征值, $x \in E \setminus \{0\}$ 是对应的特征向量. 那么

$$\|u(x)\| = \|\lambda x\| = |\lambda| \|x\|.$$

另一方面, 由于 u 是一个线性等距映射, $\|u(x)\| = \|x\|$. 因此 $|\lambda| = 1$. □

命题. 考虑二次元的情况.

若 u 是一个正交自同态, 那么:

1. 如果 $u \in SO(E)$, 那么 u 是一个旋转; 具体体现为存在 $\theta \in [0, \pi]$, 使得在一切正交标准基下, u 的矩阵为

$$\begin{pmatrix} \cos \theta & \mp \sin \theta \\ \pm \sin \theta & \cos \theta \end{pmatrix}$$

2. 如果 $u \in O(E) \setminus SO(E)$, 那么 u 是一个反射. 存在一个正交标准基, 使得 u 的矩阵为

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

从直观感受上来说, 旋转就是把坐标轴旋转一个角度, 因此不论是哪个正交标准基下, 矩阵形式都一样. 反射是把坐标翻转, 但并不是所有正交标准基下矩阵形式都一样. 例如标准形式的反射

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

在标准基下是这样的形式, 相当于 x 坐标不变, y 坐标翻转. 也就是关于 x 轴的反射. 但如果我们把 x 轴和 y 轴交换一下, 那么这个反射就相当于 x 坐标翻转, y 坐标不变. x 轴和 y 轴交换的矩阵是:

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

同一线性变换在新基下的矩阵为

$$B = P^{-1}AP = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

表示关于 y 轴的反射. 这就是反射的一个例子.

此外, 我们还注意到 P 的行列式为 -1 , 因此 $P \in O(E) \setminus SO(E)$. 那么在哪个正交标准基下, 反射的矩阵形式是

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

呢? 注意到 P 的实际作用是把 x 坐标和 y 坐标交换了一下, 实质上是关于 $y = x$ 这条直线的反射. 因此, 选取第一个基向量为平行于这条直线的 $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$, 第二个基向量为 $(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$, 就可以了. 读者可以用基变换公式自行验证.

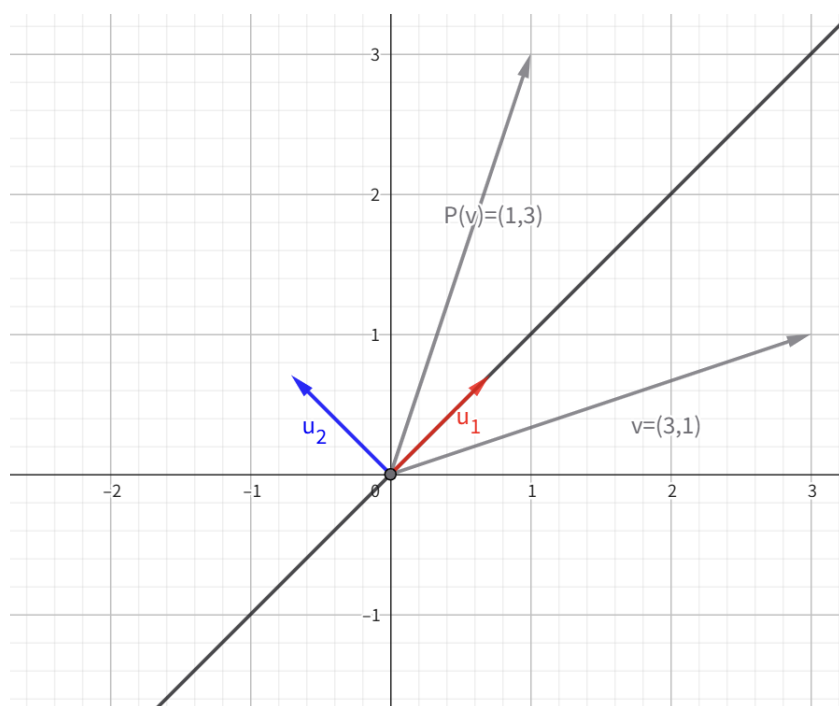


图 3: 以 u_1 和 u_2 为正交标准基, P 的矩阵就是标准形式

换句话说, 二维平面内旋转, 转轴是垂直于这一平面的, 因此不论二维平面的正交标准基怎么选, u 的矩阵形式都一样, 因为转轴的表达式不会变. 然而, 反射需要一个平面内直线作为“镜面”, 因此在不同的正交标准基下, 镜面的表达式是不一样的, 从而 u 的矩阵形式可能会有所不同.

现在我们回到三次元的情况.

命题 (欧拉定理). 假设 E 是三维的. 设 $u \in SO(E) \setminus \{\text{id}_E\}$. 那么, 1 是 u 的一个特征值, 并且与之相关的特征子空间是一个向量张成的直线, 称为 u 的旋转轴. 且存在唯一的实数 $\theta \in]0, \pi]$, 称为 u 的非定向角 (*angle non orienté*), 使得在任意一个正交标准基中, 如果

第一个向量取在 u 的旋转轴上, 那么 u 的矩阵为

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \mp \sin \theta \\ 0 & \pm \sin \theta & \cos \theta \end{pmatrix}.$$

这个定理我们在一门叫理论力学的课程中见到过, 不知道大家还记不记得.

证明. 一定存在一条 u -稳定的向量直线; 否则将不会存在任何 u -稳定的向量平面. (先别管, 当成公理) 设 D 是一条 u -稳定的直线; 则 u 在 D 上诱导的是一个系数为 ± 1 的相似映射:

$$u_D = \pm \text{id}_D \quad (\text{换句话说, } D \text{ 是 } u \text{ 的一个特征向量所张成的, 特征值是 } \pm 1).$$

之前我们证明了平面 $P = D^\perp$ 也是 u -稳定的; u 在 P 上的限制 u_P 是 P 上的一个线性等距映射, 也就是说, 根据前一个命题, 它是 P 的一个旋转或者反射. 换句话说, u 使转轴上的向量保持不变或翻转 (其实可以一定是不变, 待会就知道了), 使垂直于转轴的平面上向量旋转——作者注.

因此, 在适应分解 $E = D \oplus P$ 的某个正交标准基 b 下 (b 的第一个向量在 D 上), u 的矩阵具有如下形式:

$$\begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & & \\ 0 & & C \end{pmatrix},$$

其中 $C \in O(2)$ 是 u_P 在 b 的其余两个向量所构成的基下的矩阵. 由此可得:

$$1 = \det u = \pm 1 \cdot \det C = (\det u_D)(\det u_P).$$

如果 $u_D = -\text{id}_D$, 那么 u_P 是 P 的一个反射, 因此存在 $e \in P$, 使得 $u(e) = u_P(e) = e$. 这里的 e 就是刚才我们说的镜面上的向量. 于是我们可以用直线 $\text{Vect}(e)$ 替换 D , 这样我们就可以回到 $u_D = \text{id}_D$ 的情形. 因此不管怎么说, 都可以让 1 是 u 的一个特征值, 并且我们可以假设直线 D 由与特征值 1 对应的一个特征向量生成.

此时 u_P 是 P 的一个旋转, 且不同于 id_P , 因此 $1 \notin \text{Sp}(u_P)$, 从而 D 就是与特征值 1 相关的特征子空间.

因此平面

$$P = \ker(u - \text{id}_E)^\perp$$

被唯一确定, 并且 u 在适应分解 $E = D \oplus P$ 的某个正交标准基下的矩阵就是 Euler 给出的形式, 其中 θ 是 u_P 的非定向角.

□

总结这个证明的思路: 先声明稳定的直线和与之垂直的平面; 在此基础上, 如果直线对应特征值是 1, 就选为转轴; 否则算子在平面上的限制的特征值就是 -1, 是翻转, 如果这样那就选镜面为转轴. 至于刚才提到的公理...

命题. 任一自同态必有稳定的直线或二维平面.

证明. 考察自同态 u 的矩阵 A . 若自同态有特征值 λ 则存在 $v \in E$, $u(v) = \lambda v$. 矩阵表示就是 $(A - \lambda I)V = 0$. 所以 $\det(A - \lambda I) = 0$ 注意到 $\det(A - \lambda I)$ 是关于 λ 的 n 维多项式, 称为**特征多项式**. 去年我们学过, 实数域高维多项式一定有一次或二次的不可约因式. 如果特征多项式有一次因式, 就有实根, 这个根就是特征值, u 就在对应的特征向量张成的直线上稳定.

反之如果它只有不可约的二次因式, 那么它有一对共轭复根 $\alpha = a_1 + ia_2$ 和 $\bar{\alpha} = a_1 - ia_2$, 在复数域有一个复特征向量 $z = x + iy$, 满足 $u(z) = \alpha z$ 展开, 对比实部和虚部: 由此可得

$$u(x + iy) = (a_1 + ia_2)(x + iy) = (a_1x - a_2y) + i(a_2x + a_1y)$$

比较实部与虚部, 得到

$$u(x) = a_1x - a_2y, \quad u(y) = a_2x + a_1y.$$

因此, 二维平面 $P = \text{Vect}(x, y)$ 关于 u 是稳定的.

□

高维生物如何研究转圈圈和照镜子?

命题. 以下条件是等价的:

1. 自同态 u 是正交的;
2. 存在一族 u -稳定 2 维平面 $(P_i)_{i \in [1, r]}$, 其中整数 r 介于 0 和 $\lfloor n/2 \rfloor$ 之间, 使得 E 可以分解为如下正交直和:

$$E = \ker(u - \text{id}_E) \oplus \ker(u + \text{id}_E) \oplus P_1 \oplus \cdots \oplus P_r,$$

这里任两个子空间互相正交. 并且对于任意 $i \in [1, r]$, 自同态 u 在平面 P_i 上的限制 u_{P_i} 是一个旋转, 而不是 $\pm \text{id}_{P_i}$.

证明. 如果 2 成立, 则 u 是一个线性等距映射, 并且在各个相互正交的子空间上稳定, 那么套用勾股定理即可得证.

反过来, 先假设 u 是一个**没有特征值**的正交自同态, 并对 n 进行归纳. 即假设定理对 $n-1$ 及以前全部成立.

$n=2$ 根据公理很容易得到.

假设 $n > 2$. 自同态 u 必然有一个 u -稳定平面, 记作 P_1 . u_{P_1} 是 P_1 上的一个正交自同态, 并且没有特征值 (因为 u_{P_1} 的任一特征值同时也是 u 的特征值, 我们刚设了 u 没有特征值), 所以它是 P_1 上的一个旋转, 且不同于 $\pm \text{id}_{P_1}$. 另一方面, P_1^\perp 是 E 的一个 $n-2$ 维的 u -稳定子空间; 将归纳假设应用于 u 在 P_1^\perp 上的限制, 就得到如下分解:

$$P_1^\perp = P_2 \oplus \cdots \oplus P_r,$$

其中每个 P_i 都是一个 u -稳定平面, 并且 u 在 P_i 上的限制是一个旋转, 且不同于 $\pm \text{id}_{P_i}$. 因此 2 得证, 并且 n 得是偶数.

上式没有 $\ker(u_{P_1^\perp} - \text{id}_E)$ 和 $\ker(u_{P_1^\perp} + \text{id}_E)$, 因为设了没有特征值. 这两项如果非空则 $u_{P_1^\perp}$ 有特征值 1 或 -1.

在一般情况下 (假设 u 是任意一个正交自同态, 可能有特征值), 设

$$F = \ker(u - \text{id}_E) \oplus \ker(u + \text{id}_E).$$

那么 u_F 是 F 上的一个正交对称映射. 首先这两个子空间 $\ker(u - \text{id}_E)$ 和 $\ker(u + \text{id}_E)$ (可能为空) 是正交的 (请自行验证). 另一方面, F^\perp 是一个 u -稳定子空间, 且 u_{F^\perp} 是 F^\perp 上的一个没有特征值的正交自同态 (因为 u 特征值只能是 ± 1 , 都在 F 里用完了). 对 F^\perp 的讨论, 照搬无特征值的证明过程即可. \square

注. 对于 $n \geq 4$ 的情况, 我无法给出可视化的易于理解的例子. 如果您认识高维生物, 可以请他给大家讲讲.

根据以上内容, 我们不难看出:

命题. 以下条件等价:

1. 线性自同态 u 是正交的;
2. 存在 E 的一组正交标准基, 使得 u 在该基下的矩阵是如下形式的分块对角矩阵:

$$\text{diag}(I_p, -I_q, R(\theta_1), \dots, R(\theta_r)),$$

其中 $p, q, r \in \mathbb{N}$, 满足 $p + q + 2r = n$, 且 $\theta_1, \dots, \theta_r \in (0, \pi)$. 这里

$$R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \in SO(2)$$

表示平面上的旋转矩阵.

很显然, 大矩阵中的 I_p 表示恒等的部分, $-I_q$ 表示翻转的部分, 其余各部分则表示各平面上的旋转. 可见高维生物要想理解旋转也是离不开二次元的. 这一切都是因为不可约的实多项式最高为二次, 大家可以想想为什么.

关于欧拉定理和后面的旋转相关内容, 固然证明方法很值得学习, 但它们在力学、工程学上的应用也是很有价值的. 数学来源于社会实践, 服务于社会实践, 决不能脱离生产而存在 (但可以脱离三次元而存在).

1.3 自伴算子 (auto-adjoint) (或对称算子)

这一节要开始抽象了.

定义. 我们称向量空间 E 上的线性算子 u 是自伴算子或对称算子, 如果 $u^* = u$, 也就是说

$$\forall x, y \in E \quad (u(x) | y) = (x | u(y)).$$

命题. b 是 E 上的正交标准基, 不难得出以下两条是等价的:

1. u 是自伴算子.
2. u 在 b 上的矩阵是对称的.

证明. 这很显然, 因为 u 是自伴算子等价于 u 在 b 上的矩阵的转置等于自己. □

很显然, 对称算子构成 $\mathcal{L}(E)$ 上的一个线性子空间, 记作 $S(E)$. 特别的, 如果 E 是 \mathbb{R}^n , 那么 $S(E)$ 的维数是 $\frac{n(n+1)}{2}$, 因为 n 阶对称矩阵的值完全取决于对角线和对角线上(下)的元素.

命题. 向量空间 E 上的自伴的投影算子就是正交投影算子.

证明. 假设 p 是一个正交投影算子, 那么对于任意 $x, y \in E$, 有

$$x - p(x) \perp p(y),$$

因此

$$(x | p(y)) = (p(x) | p(y)).$$

由对称性(这是个常用技巧, 因为我们没规定 x 和 y 有什么不同, 所以交换他们俩)可得

$$(x | p(y)) = (p(x) | y) = (p(x) | p(y)),$$

于是 $p^* = p$.

反过来, 若 p 是满足 $p^* = p$ 的 E 上的投影算子, 则对于任意 $x \in \ker p$, $y \in \operatorname{Im} p$, 有

$$(x | y) = (x | p(y)) = (p(x) | y) = 0.$$

因此 $\ker p \perp \operatorname{Im} p$, 即 p 是正交投影. □

我很确定这道题我们上学期 TD 讲过, 而且是我讲的.

命题. 假设线性算子 u 是自伴算子. 设 F 是 E 的一个 u -稳定子空间; 那么其正交补 F^\perp 也是 u -稳定的.

不用证了, 每节都有个类似的.

命题. 假设 u 是自伴算子. 那么在 E 中存在一条 u -稳定的直线(向量子空间). 换句话说, 算子 u 的特征值集合非空.

这个命题很重要.

证明. 反证法, 假设在 E 中不存在 u -稳定直线, 也就没有实特征值. 根据那个公理, 可以推出存在一个二维 u -稳定子空间 $F \subset E$. 算子 u 在 F 上的限制 u_F 仍然是自伴算子 (因为 u 是自伴的). 因此, 在 F 的一个正交标准基下, u_F 的矩阵表示为

$$M = \begin{pmatrix} a & c \\ c & b \end{pmatrix}.$$

由于 u_F 没有特征向量 (因为假设了 u 没有), 矩阵 $M - \lambda I_2$ 对所有 $\lambda \in \mathbb{R}$ 都是可逆的. 因此,

$$\forall \lambda \in \mathbb{R}, \quad \det \begin{pmatrix} a - \lambda & c \\ c & b - \lambda \end{pmatrix} = \lambda^2 - (a + b)\lambda + ab - c^2 \neq 0.$$

这对吗? 考察特征多项式

$$X^2 - (a + b)X + ab - c^2$$

它的判别式等于 $(a - b)^2 + 4c^2 > 0$, 因此必有实根. \square

线性代数的本质是数, 是抽象的, 但要想理解线性代数, 必须使用具象的方法. 下面来看另一个证明.

证明. (Boss 进入了祂的二阶段) 设函数

$$q : E \rightarrow \mathbb{R}, \quad x \mapsto (x | u(x)).$$

由于该函数连续, 而单位球面

$$S = \{x \in E \mid \|x\| = 1\}$$

是紧集 (E 是有限维的, S 是闭集且有界), 所以 q 在 S 上面存在最大值和最小值, 也就是存在 $x_0 \in S$, 使得

$$\forall x \in S \quad q(x_0) \leq q(x).$$

我们来证明: x_0 是算子 u 的一个特征向量, 其对应特征值为 $\lambda = q(x_0)$. 只需验证 $u(x_0)$ 属于 $\text{Vect}(x_0)$, 也就是正交于超平面 $(\text{Vect}(x_0))^\perp$ (因为在内积空间, 验证正交性是很容易的, 乘一下等于 0 就行了), 即对于任意单位向量 $y \in E$, 如果 $(x_0 | y) = 0$, 那么有 $(u(x_0) | y) = 0$.

取 $y \in S$, 满足 $(x_0 | y) = 0$. 我们可以定义一个通过 x_0 和 y 的“大圆”:

$$x_t = (\cos t)x_0 + (\sin t)y.$$

考虑函数

$$\varphi_y : \mathbb{R} \rightarrow \mathbb{R}, \quad t \mapsto q(x_t).$$

展开得:

$$\varphi_y(t) = (\cos t)^2 q(x_0) + (\sin 2t)(u(x_0) | y) + (\sin t)^2 q(y).$$

由于 x_0 的定义, φ_y 在 $t = 0$ 处取得极小值, 因此

$$0 = \varphi'_y(0) = 2(u(x_0) | y).$$

于是 $(u(x_0) | y) = 0$, 证毕. \square

既然有了特征值...

命题. 以下条件是等价的:

1. 线性算子 u 是自伴的;
2. 存在 E 的一个正交标准基, 使得 u 在此基下可对角化;
3. u 可对角化, 且它的特征子空间两两正交;
4. 算子 u 有如下的谱分解:

$$u = \lambda_1 p_1 + \cdots + \lambda_s p_s,$$

其中 p_1, \dots, p_s 是正交投影, $\lambda_1, \dots, \lambda_s$ 是两两不同的实数.

证明. 2、3、4 相互是等价的, 这很显然, 只需证明 1 和 2 等价就行了.

如果存在 E 的一个正交标准基 b , 在此基下 u 的矩阵是对角的, 那么由于对角矩阵是对称的, 可知 u 是自伴的.

反过来怎么证? 对维数 n 归纳证明. 若 $n = 1$, 结论显然成立, 没什么可说的. 设 $n > 1$, 并假设任意维数为 $n - 1$ 的欧几里得空间上的自伴算子都能在某个正交标准基下对角化.

取 u 的一个单位特征向量 e_n (它肯定存在, 我们证了 u 有稳定直线). 根据上面那个“不用证了”的命题 (不要藐视任何一个命题的结论, 即便忽视它的证明), 超平面 $H = (\text{Vect}(e_n))^\perp$ 是 u -稳定的. 因此, 限制在 H 上的算子 u_H 在 H 上是自伴的, 那么由归纳假设, 存在一个正交标准基 (e_1, \dots, e_{n-1}) 对角化 u_H . 于是 $(e_1, \dots, e_{n-1}, e_n)$ 就是 E 的一个由特征向量构成的正交标准基. \square

事已至此, 我们最后来看一下对称矩阵的谱定理.

命题. 设 $A \in \mathcal{M}_n(\mathbb{R})$. 以下条件是等价的:

1. 矩阵 A 是对称的;
2. 矩阵 A 可正交对角化, 即存在一个 n 阶正交矩阵 P , 使得

$$P^{-1}AP$$

是对角矩阵.

证明. 设 b_0 是 E 的一个正交标准基, 且 u 是 E 的一个线性算子, 在基 b_0 下的矩阵表示为 A .

说一个基 b 是正交标准基, 等价于说从基 b_0 到基 b 的过渡矩阵

$$P = \text{Mat}_{b_0}(b)$$

是正交的. 根据基变换公式

$$\text{Mat}_b(u) = P^{-1}AP,$$

通过之前那四条等价条件就可以证明这个命题. □

补充一条有用的东西:

命题. 设 $A \in \mathbb{R}^{n \times n}$ 是对称矩阵. 若

$$Ae_i = \lambda_i e_i, \quad i = 1, \dots, n,$$

且 $\{e_1, \dots, e_n\}$ 是由 A 的单位特征向量组成的正交标准基 (即 $e_i^T e_j = \delta_{ij}$), 则

$$A = \sum_{i=1}^n \lambda_i e_i e_i^T.$$

证明. 令

$$Q = [e_1 \ e_2 \ \cdots \ e_n] \in \mathbb{R}^{n \times n},$$

则 Q 正交化 A :

$$Q^T A Q = \text{diag}(\lambda_1, \dots, \lambda_n) =: D.$$

于是有

$$A = Q D Q^T.$$

对任意向量 $x \in \mathbb{R}^n$, 由于 $\{e_i\}$ 是标准正交基, 可以将 x 展开为

$$x = \sum_{i=1}^n (e_i^T x) e_i.$$

于是

$$Ax = A \left(\sum_{i=1}^n (e_i^T x) e_i \right) = \sum_{i=1}^n (e_i^T x) A e_i = \sum_{i=1}^n (e_i^T x) \lambda_i e_i = \sum_{i=1}^n \lambda_i e_i (e_i^T x).$$

把 $(e_i^T x)$ 提出到右侧, 可写成矩阵作用的形式:

$$Ax = \left(\sum_{i=1}^n \lambda_i e_i e_i^T \right) x, \quad \forall x \in \mathbb{R}^n.$$

由线性算子的格等价性 (两个矩阵若对任意向量作用相同则矩阵相等), 得到

$$A = \sum_{i=1}^n \lambda_i e_i e_i^T$$

□

事已至此，我们已经把自伴算子的主要理论内容逐一展开，已经形成了一个严谨的知识体系。这是不是就意味着自伴算子的内容到此为止了呢？恐怕还远远不够。毕竟，一个游戏如果只通关一次就退坑，那未免太可惜了；只有进入二周目，才能在新的角度下发现隐藏的细节和真正的乐趣。同样地，对于自伴算子，如果我们仅仅停留在定理证明和数学形式上，却没有理解它究竟“做什么”“有什么用”“到底在干嘛”，那就等于只玩了一次速通。

(以上是 ChatGPT 作为嘴替在瞎 BB)

1.3' 自伴算子（二周目）

注. 本节仅供辅助理解，不可当做严谨数学证明。

自伴算子起到什么效果？我们回顾一下之前的一个命题：

命题. 以下条件是等价的：

1. 线性算子 u 是自伴的；
2. 存在 E 的一个正交标准基，使得 u 在此基下可对角化；
3. u 可对角化，且它的特征子空间两两正交；
4. 算子 u 有如下的谱分解：

$$u = \lambda_1 p_1 + \cdots + \lambda_s p_s,$$

其中 p_1, \dots, p_s 是正交投影， $\lambda_1, \dots, \lambda_s$ 是两两不同的实数。

对于 E 上的任意向量 x ，假设 $b = (e_1, \dots, e_n)$ 是对角化 u 的正交标准基，令 p_i 为到特征子空间 E_{λ_i} 的正交投影，那么对任意 $x \in E$ ，有

$$u(x) = \lambda_1 p_1(x) + \cdots + \lambda_s p_s(x) = \sum_{i=1}^s \lambda_i p_i(x).$$

这个式子清楚地揭示了自伴算子的几何作用：

首先， $p_i(x)$ 表示将向量 x 投影到特征子空间 E_{λ_i} 上的部分。也就是说， $p_i(x)$ 把 x 分解成在 E_{λ_i} 方向上的分量。算子 u 对这一分量施加作用时，只是把它乘上一个实数因子 λ_i 。换句话说， u 在 E_{λ_i} 内表现为“按比例伸缩”，比例因子就是对应的特征值 λ_i 。

因此，整个向量 x 在 u 的作用下，就是被分解到不同的相互正交的特征子空间里，各部分分别被不同的实数因子拉伸（或压缩，取决于特征值），最后再把这些变换后的分量加起来。

我们来看一个例子，仍然是二次元。考虑一个自伴算子 u ，它在正交标准基下的矩阵是

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

很显然, 这个矩阵每一行各元素的和都是 3, 因此它的一个特征向量是 $(1, 1)$:

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

对应特征值是 3. 此外, 考察垂直于这个特征向量的向量 $(-1, 1)$:

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

这也是特征向量, 对应特征值 1. 因此, u 把向量在 $y = x$ 上的投影伸长为三倍, 对 $y = -x$ 上的投影保持不变. 取 $u_1 = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$, $u_2 = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$, 则 (u_1, u_2) 是对角化 u 的正交标准基.

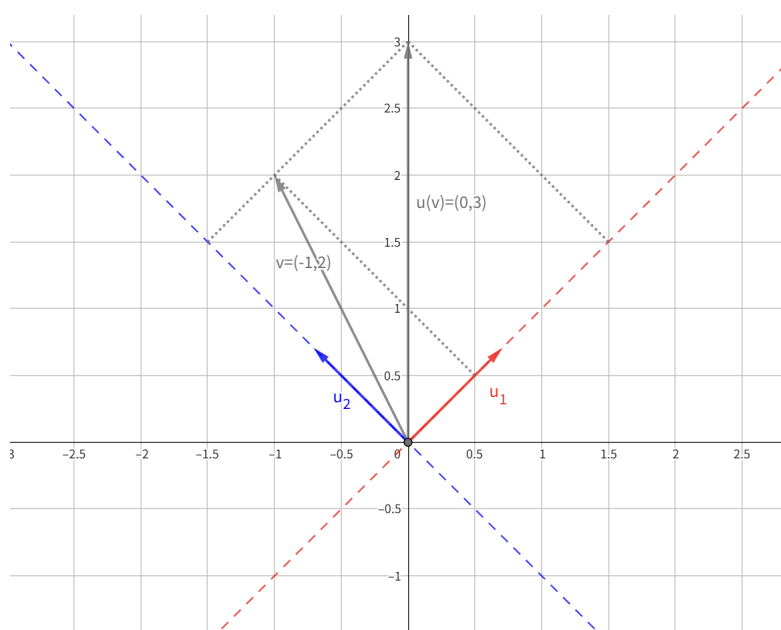


图 4: u 把向量在 u_1 上的投影伸长为三倍, 对 u_2 上的投影保持不变.

注. 有没有感觉这个图很眼熟? 没错, 我为了偷懒再次选了以 $y = x$ 和 $y = -x$ 为子空间的算子, 这样以前的图改一改就行了.

接下来, 我们来回顾这个令人费解的证明:

命题. 假设 u 是自伴算子. 那么在 E 中存在一条 u -稳定的直线 (向量子空间). 换句话说, 算子 u 的特征值集合非空.

证明. 设函数

$$q: E \rightarrow \mathbb{R}, \quad x \mapsto (x | u(x)).$$

由于该函数连续, 而单位球面

$$S = \{x \in E \mid \|x\| = 1\}$$

是紧集 (E 是有限维的, S 是闭集且有界), 所以 q 在 S 上面存在最大值和最小值, 也就是存在 $x_0 \in S$, 使得

$$\forall x \in S \quad q(x_0) \leq q(x).$$

我们来证明: x_0 是算子 u 的一个特征向量, 其对应特征值为 $\lambda = q(x_0)$. 只需验证 $u(x_0)$ 属于 $\text{Vect}(x_0)$, 也就是正交于超平面 $(\text{Vect}(x_0))^\perp$ (因为在内积空间, 验证正交性是很容易的, 乘一下等于 0 就行了), 即对于任意单位向量 $y \in E$, 如果 $(x_0 | y) = 0$, 那么有 $(u(x_0) | y) = 0$.

取 $y \in S$, 满足 $(x_0 | y) = 0$. 我们可以定义一个通过 x_0 和 y 的“大圆”:

$$x_t = (\cos t)x_0 + (\sin t)y.$$

考虑函数

$$\varphi_y : \mathbb{R} \rightarrow \mathbb{R}, \quad t \mapsto q(x_t).$$

展开得:

$$\varphi_y(t) = (\cos t)^2 q(x_0) + (\sin 2t)(u(x_0) | y) + (\sin t)^2 q(y).$$

由于 x_0 的定义, φ_y 在 $t = 0$ 处取得极小值, 因此

$$0 = \varphi'_y(0) = 2(u(x_0) | y).$$

于是 $(u(x_0) | y) = 0$, 证毕. □

这个证明的思路是怎么来的? 为什么要定义一个通过 x_0 和 y 的“大圆”?

思考这样一个问题: 线性算子 u 是自伴的, 它把一个向量在各个相互垂直的子空间上分别拉伸或压缩. 那么一个球在这个线性算子作用下的像应该长什么样?

降维思考, 如果一个线性算子, 就像上文中的 u , 在二维空间把一个向量沿一个轴的分量乘上一个数, 沿垂直于这个轴的轴的分量也乘上另一个数, 那它会把一个圆变成什么样子?

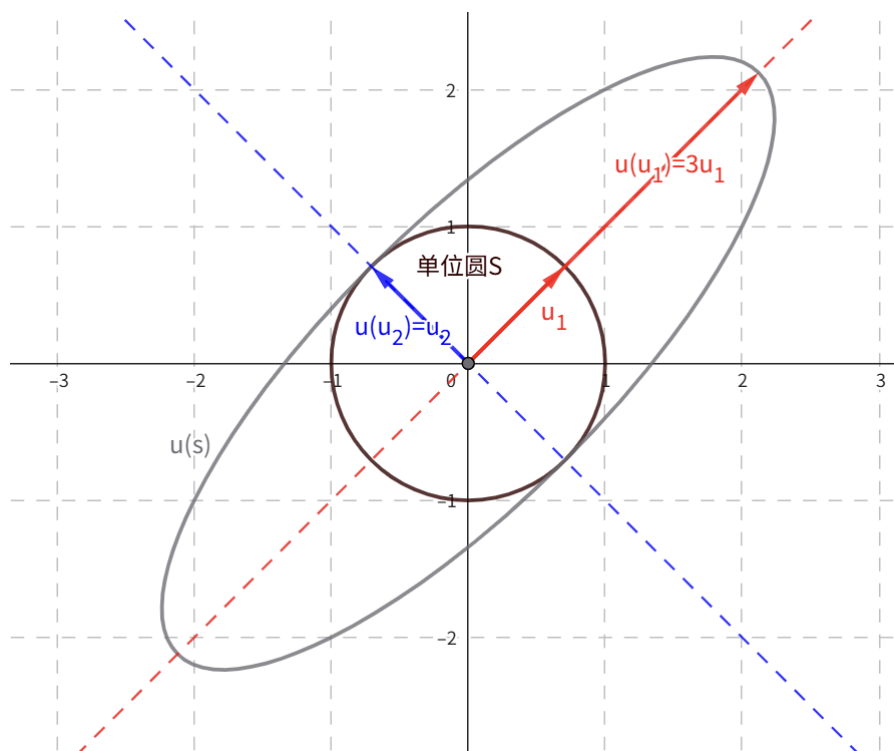


图 5: 猜对了, 就是椭圆!

既然如此, 如图 5, 文中的 x_0 表示 S 上一个点, 使得 $(x_0 | u(x_0))$ 取最小值. 这个点很显然就等于短轴的顶点, 也就是 u_2 或 $-u_2$ 的点. u_2 就是特征向量. 上课的时候, 他的现场发挥和书上相反, 把 x_0 取成使得 $(x_0 | u(x_0))$ 取最大值. 对应的就是长轴的顶点, 也就是 u_1 或 $-u_1$.

注. “这个点很显然就等于短轴的顶点”其实不太显然. 大概原理是定义的函数 $q: x \mapsto (x | u(x))$ 描述了模长为 1 的向量 x 在 u 的作用下, 在它自己的方向上被压缩或拉伸的程度. 因此特征向量会取到极值.

至于为什么要构造通过 x_0 和 y 的“大圆”, 就不能拘泥于二次元了 (见图 6). 我们已经知道了 x_0 使得 q 在单位球 S 上取极值. 在此基础上, x_0 任何沿着球面上的小变动方向的导数都应为 0. 我们需要描述“沿球面微小变动”的方向, 也就是切向.

任取一个单位向量 y 满足 $(x | y) = 0$, 这个 y 就表示从 x_0 出发的切向方向.

沿着 y 的切向构造一个大圆. 这样能够给 x_0 一个微小扰动, 且不改变 $\|x_0\|$. 因为 x_0 是极值点, 沿任意切向的方向导数应为 0, 如果南极点是全世界最冷的点, 那么不论考虑哪个经线圈, 它都应该是这个圈上温度最低的. 所以对于任意与 x_0 垂直的单位向量 y , 有

$$\varphi'_y(0) = 0.$$

求导并代入 $t = 0$, 由于 $\cos 0 = 1$, $\sin 0 = 0$, 得到

$$\varphi'_y(0) = 2(u(x_0) | y).$$

由极值条件 $\varphi'_y(0) = 0$, 于是

$$(u(x_0) | y) = 0.$$

因为上述等式对任意与 x_0 垂直的 y 都成立, 说明 $u(x_0)$ 与切空间 $T_{x_0}S = \{v \in E : (x_0 | v) = 0\}$ 的所有向量都正交. 换言之, $u(x_0)$ 只能和 x_0 平行, 即存在 $\lambda \in \mathbb{R}$ 使

$$u(x_0) = \lambda x_0,$$

这就是证明 x_0 为 u 的特征向量的核心结论.

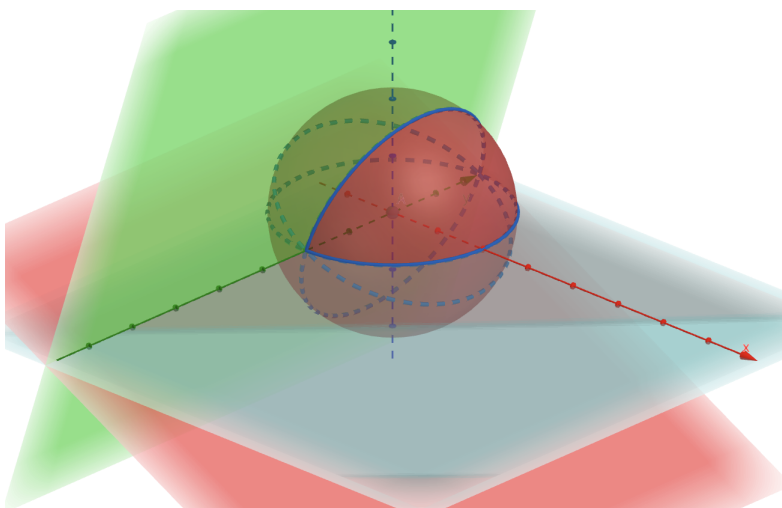


图 6: 二次元的圆只有一个切向, 而我们三维元要考虑的切向就很多了.

再来看这个命题:

命题. 向量空间 E 上的自伴的投影算子就是正交投影算子.

想像这样一个场景: 晴朗的 6 月 10 日 (或者 7 月 4 日) 左右, 小王在深圳平坦的大街上散步. 早晨, 小王能够在地面上看到自己的影子. 这时, 太阳就是一个投影算子, 但显然正交投影算子, 因为阳光是斜着照的. 太阳也不是一个自伴算子. 因为小王站立时在地面上的 (正交) 投影为 0, 影子长度却不为 0, 拉伸了无穷倍; 趴下后, 在地面上的投影为身高, 影子长度也是身高, 拉伸了 1 倍. 拉伸的倍数不是恒定的.

到了正午, 太阳直射深圳市的时候, 就是一个正交投影算子. 此时它也是个自伴算子, 因为小王发现, 当他摔倒在地的过程中, 它的影长和自己在地面上的投影是成正比的, 比值为 1.

回到梦开始的地方:

定义. 我们称向量空间 E 上的线性算子 u 是自伴算子或对称算子, 如果 $u^* = u$, 也就是说

$$\forall x, y \in E \quad (u(x) | y) = (x | u(y)).$$

还是考虑用了不知道多少次的老二次元线性自伴算子: $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ 如图,

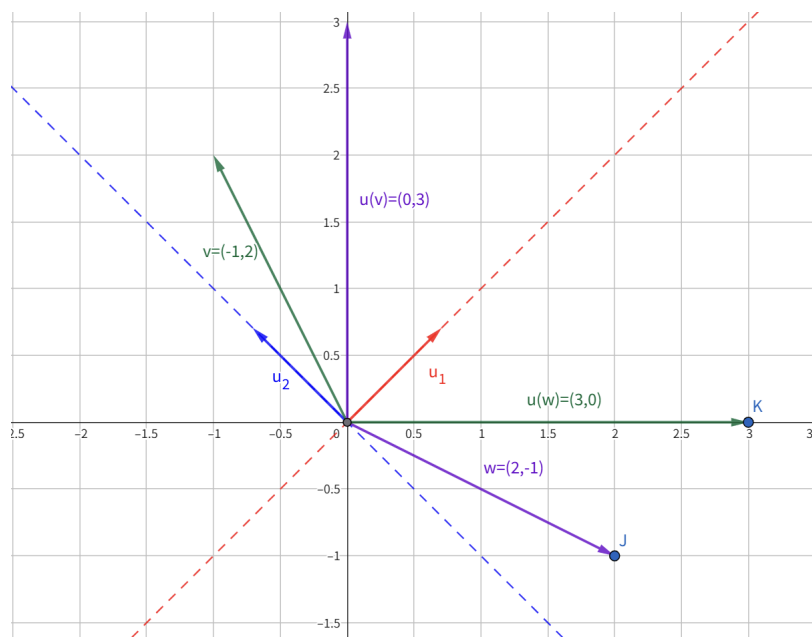


图 7: $(u(v) | w) = (v | u(w))$

啥, 您说没看懂? 那我们换个角度再看一下. 众所周知, 在不同的正交标准基下, 内积是不变的 (还记得正交自同态吗?). 我们换 u_1 和 u_2 的正交标准基, 也就是把这张图顺时针转 45° .

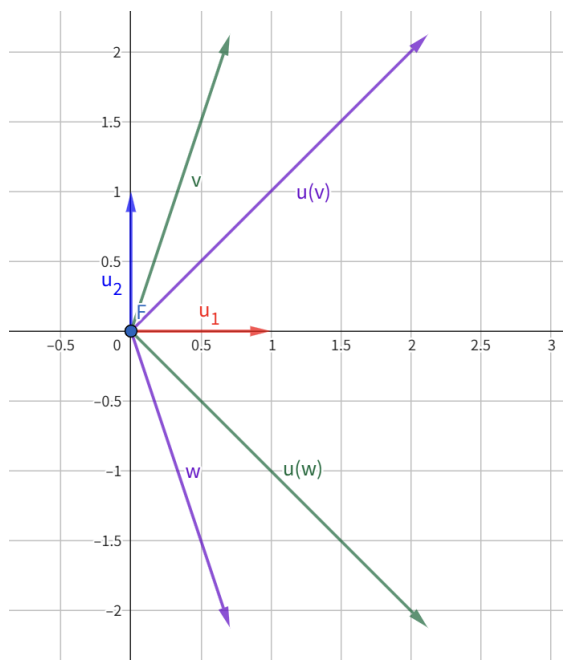


图 8: 孩子们我失算了, 选的矢量过于对称以至于失去了一般性

我想这张图几乎不用解释 $(u(v) | w) = (v | u(w))$ 是怎么看出来的了. 总而言之, 正交

标准基下，内积只取决于两个向量的坐标. 图 8 中两组叉乘的纵坐标对应是一致的，横坐标一个变为 3 倍，一个缩小为 $1/3$ ，所以内积是一样的. 换句话说，自伴算子的性质使得在正交标准基下它以固定倍数放缩向量的坐标，从而它作用到内积的任一项结果是一样的.

至此，您应该对自伴算子有了一些直观形象的理解. 但注意，人的直觉并不是什么时候都靠谱. 因此在考场上还是要从基本的概念、定义和计算出发.

1.4 极分解

讲到了再写. (别忘了后面有例题，白给的 TD 答案)

1.5 例题

例题. 考虑实矩阵

$$A = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}.$$

1. 观察到 A 存在一个明显的特征值. 确定 A 的特征子空间;

2. 计算 A 的谱分解 (特征值分解).

解. 注意到 $A - I = \begin{pmatrix} 1 & 2 & -2 \\ 2 & 4 & -4 \\ -2 & -4 & 4 \end{pmatrix}$. 观察这个矩阵, 好像挺好看的. 把第一列 (也是第一行) 拿出来: 令 $u = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$. 直接计算得

$$uu^T = \begin{pmatrix} 1 & 2 & -2 \\ 2 & 4 & -4 \\ -2 & -4 & 4 \end{pmatrix},$$



对于向量 u 本身:

$$Au = (I + uu^T)u = u + u(u^T u) = (1 + \|u\|^2)u.$$

计算 $\|u\|^2 = 1^2 + 2^2 + (-2)^2 = 1 + 4 + 4 = 9$, 因此

$$Au = 10u.$$

所以 u (或任何与之成比例的向量) 是特征向量, 对应特征值 $\lambda_1 = 10$.

对于任意 $v \in \mathbb{R}^3$ 且 $v \perp u$ (即 $u^T v = 0$), 有

$$Av = (I + uu^T)v = v + u(u^T v) = v.$$

因此所有与 u 正交的向量都是特征向量, 对应特征值 $\lambda_2 = \lambda_3 = 1$ (这里有两个 1, 可通过以下两条得出: 与 u 正交的向量构成二维平面; 矩阵 A 的迹是 $12=10+1+1$).

至于谱分解, 参见“有用的东西”命题.

例题. (TD1.2, 注意做这题之前要先搞明白 TD1.1) 设 E 与 F 分别是维数为 n 与 p 的欧几里得空间, 且 $f: F \rightarrow E$ 为一线性映射 (实内积空间上的线性算子). 记 f^* 为 f 的伴随算子.

1. 证明下列等式:

$$\text{Ker}(f^*f) = \text{Ker } f, \quad \text{Im}(f^*f) = (\text{Ker } f)^\perp, \quad \text{rg}(f^*f) = \text{rg } f.$$

2. 现在假设 $\text{rg } f = p$.

(a) 证明 $n \geq p$ 且 $f^*f \in GL(F)$ (即 f^*f 可逆).

(b) 设 $v \in E$ 且 $x \in F$. 证明: 在 F 中, 方程

$$(f^*f)(x) = f^*(v)$$

的解 x 唯一, 当且仅当 $f(x)$ 是 v 在 $\text{Im } f$ 上的正交投影.

3. 考虑 n 个线性方程、 p 个未知数的实系数线性方程组

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,p}x_p = b_1, \\ \vdots \\ a_{n,1}x_1 + \cdots + a_{n,p}x_p = b_n, \end{cases}$$

将其写成矩阵形式 $AX = B$. 假设 n 严格大于 p (方程数多于未知数), 且该线性系统的秩达到可能的最大值, 即等于 p . 这样的系统 (称为“超定”) (机翻的错了别找我) 一般并不一定有解. 证明: 该系统存在唯一的二次意义下的最佳近似解 (即使残差平方和最小的唯一解), 也就是说存在唯一的 X 使得

$$\sum_{k=1}^n |a_{k,1}x_1 + \cdots + a_{k,p}x_p - b_k|^2$$

最小; 并且该唯一最优解等于 $p \times p$ 的正规方程

$$(A^T A)X = A^T B$$

的唯一精确解.

4. 在 \mathbb{R}^2 上赋予通常的欧几里得结构; 在标准 (基) 坐标下把点的坐标记为 (x, y) . 考虑三条仿射直线

$$y - x = 0, \quad 2y - x - 1 = 0, \quad y + 4x - 6 = 0.$$

(a) 它们是否共点 (是否有公共交点)?

(b) 证明存在唯一一点使得该点到三条直线的距离平方和达到最小, 并求出该点.

解. 第一问略.

2.1 一方面, $E \supset \text{Im } f$, 因此

$$n = \dim E \geq \dim(\text{Im } f) = \text{rg } f = p;$$

另一方面, f^*f 是 F 上的一个自同态, 且其像空间的维数为 $p = \dim F$, 因此 f^*f 是双射.

2.2 既然 f^*f 是双射, 对于任意给定的 $v \in E$, 方程 $(f^*f)(x) = f^*(v)$ 在 F 中都有唯一解 x . 在此基础上, 对于 $v \in E$ 与任意 $x \in F$, 有下列等价关系:

$$\begin{aligned} (f^*f)(x) = f^*(v) &\iff \forall y \in F, (f^*(v) - (f^*f)(x) \mid y)_F = 0 \\ &\iff \forall y \in F, (v - f(x) \mid f(y))_E = 0 \\ &\iff v - f(x) \in (\text{Im } f)^\perp. \end{aligned}$$

也就是说, 点 $f(x) \in \text{Im } f$ 是 v 在子空间 $\text{Im } f$ 上的正交投影; 换句话说, $f(x)$ 是 v 在 $\text{Im } f$ 中的唯一的“最佳近似”.

3 将上述讨论应用到欧几里得空间: 设 $E = \mathbb{R}^n$ 且 $F = \mathbb{R}^p$, 两者都带有标准内积; f 表示在标准基下矩阵为 A 的线性映射, 向量 v 的坐标列记为 B . 我们假设线性系统的秩等于 p (换言之, 矩阵 A 的行向量线性无关), 并且方程的个数多于未知数: $n > p$. 一般情形下方程组没有精确解.

我们打算寻找一个“按二次意义的最佳近似解”, 即确定 x 使得下列量最小:

$$\sum_{k=1}^n |a_{k,1}x_1 + \cdots + a_{k,p}x_p - b_k|^2 = \|f(x) - v\|_E^2.$$

这等价于 $f(x)$ 是 v 在 $\text{Im } f$ 上的正交投影. 根据 2.2, 这样的“解”存在并且唯一:

$$x = (f^*f)^{-1}(f^*(v)).$$

换句话说, 若 $X \in M_{p,1}(\mathbb{R})$ 是表示向量 x 的列向量 (标准基下), 则

$$X = (A^\top A)^{-1} A^\top B.$$

4 直线

$$x - y = 0, \quad x - 2y + 1 = 0, \quad 4x + y - 6 = 0$$

并不共点.

我们考虑平面上三条仿射直线, 其一般形式为

$$\ell_k: a_k x + b_k y + c_k = 0, \quad k = 1, 2, 3,$$

其中 $(a_k, b_k) \neq (0, 0)$. 给定点 $(x, y) \in \mathbb{R}^2$, 该点到直线的欧几里得距离为

$$d_k(x, y) = \frac{|r_k(x, y)|}{\sqrt{a_k^2 + b_k^2}}.$$

因此若我们想要以“点到每条直线的距离的平方和”作为评价函数（即在几何意义上寻找一个对三条直线的“最佳拟合点”），应当最小化的是

$$\sum_{k=1}^3 d_k(x, y)^2 = \sum_{k=1}^3 \frac{r_k(x, y)^2}{a_k^2 + b_k^2}.$$

将每一项除以 $\sqrt{a_k^2 + b_k^2}$ ，定义归一化系数

$$\alpha_k := \frac{1}{\sqrt{a_k^2 + b_k^2}}.$$

则目标函数变为

$$\sum_{k=1}^3 (\alpha_k r_k(x, y))^2 = \sum_{k=1}^3 (\alpha_k a_k x + \alpha_k b_k y + \alpha_k c_k)^2.$$

中间忘了，反正结果是

$$\begin{pmatrix} x \\ y \end{pmatrix} \approx \begin{pmatrix} 1.20 \\ 1.14 \end{pmatrix}.$$

注.