

内积空间上的算子

Eliauk

2025 年 12 月 12 日

目录

1 对偶空间和伴随映射	1
2 正规算子	4
2.1 复正规算子的酉相似	6
2.2 实正规算子的正交相似	8
3 酉算子和正交算子	12
4 自伴随算子和反自伴随算子	15
5 半正定算子	18
6 奇异值分解和极分解	20
7 Moore–Penrose 广义逆和最小二乘	25
8 QR 分解和 Cholesky 分解	32

1 对偶空间和伴随映射

本文的域 \mathbb{F} 均指代 \mathbb{R} 或者 \mathbb{C} . 对于 \mathbb{F} 上的 n 向量空间 V , 记对偶空间为 $V^* = \text{Hom}(V, \mathbb{F})$, 即 $V \rightarrow \mathbb{F}$ 的线性映射全体构成的向量空间. 我们有以下重要的结果:

SECTION 1 对偶空间和伴随映射

定理 1.1. 对于 n 维向量空间 V , 其对偶空间 V^* 也是 n 维向量空间. 并且对于 V 的一组基 e_1, \dots, e_n , 记线性映射 $e_i^* : V \rightarrow \mathbb{F}$ 满足

$$e_i^*(e_j) = \delta_{ij},$$

那么 e_1^*, \dots, e_n^* 是 V^* 的一组基, 被称为 e_1, \dots, e_n 的对偶基.

Proof. 我们有 $\dim V^* = (\dim V)(\dim \mathbb{F}) = n$. 所以实际上我们只需要说明 e_1^*, \dots, e_n^* 线性无关即可, 但是为了强调对偶基的作用, 这里仍然验证 e_1^*, \dots, e_n^* 线性无关且张成 V^* . 任取 $f \in V^*$, 对于任意 $v \in V$, 设 $v = \sum v_i e_i$, 那么

$$f(v) = f\left(\sum v_i e_i\right) = \sum v_i f(e_i) = \sum f(e_i) e_i^*(v),$$

所以 $f = \sum f(e_i) e_i^*$, 故 e_1^*, \dots, e_n^* 张成 V^* .

令 $a_1, \dots, a_n \in \mathbb{F}$ 使得

$$a_1 e_1^* + \cdots + a_n e_n^* = 0,$$

那么

$$a_i = a_1 e_1^*(e_i) + \cdots + a_n e_n^*(e_i) = 0,$$

所以 e_1^*, \dots, e_n^* 线性无关. \square

下面假设 V 是内积空间, 即 Euclid 空间 (实内积空间) 或者酉空间 (复内积空间). 内积空间上的线性变换我们一般称之为算子. 假设 e_1, \dots, e_n 是 V 的标准正交基. 那么对偶基满足

$$e_i^*(u) = u_i = \langle u, e_i \rangle \quad \forall u = \sum u_i e_i \in V.$$

任取 $f \in V^*$, 设 $f = \sum a_i e_i^*$, 那么

$$f(u) = \sum a_i e_i^*(u) = \sum a_i \langle u, e_i \rangle = \left\langle u, \sum \bar{a}_i e_i \right\rangle,$$

即存在 $v = \sum \bar{a}_i e_i$ 使得

$$f(u) = \langle u, v \rangle.$$

此外, 这样的 v 还是唯一的: 假设 $w \in V$ 也使得 $f(u) = \langle u, w \rangle$, 那么 $\langle u, v \rangle = \langle u, w \rangle$, 即 $\langle u, v - w \rangle = 0$, 取 $u = v - w$, 就得出 $v - w = 0$, 所以 $w = v$. 事实上, 上面的结果被称为 **Riesz 表示定理**. 这直接引出了伴随算子的概念.

设 $\varphi : V \rightarrow W$ 是 n 维内积空间 V 到 m 维内积空间 W 的线性映射, 那么对于任意的 $w \in W$, 我们考虑映射 $f : V \rightarrow \mathbb{F}$ 为 $f(v) = \langle \varphi(v), w \rangle$. 容易验证 f 是线性映射, 所以 $f \in V^*$. 根据上

SECTION 1 对偶空间和伴随映射

面的叙述, 存在唯一的 $u \in V$ 使得 $\langle \varphi(v), w \rangle = f(v) = \langle v, u \rangle$. 如果我们定义 $\varphi^* : W \rightarrow V$ 为 $\varphi^*(w) = u$, 那么就有

$$\langle \varphi(v), w \rangle = \langle v, \varphi^*(w) \rangle.$$

由于 u 的唯一性, 所以 $\varphi^* : W \rightarrow V$ 是良定义的. 这个 φ^* 被称为 φ 的**伴随映射**. 如果 $W = V$, 那么 φ^* 被称为 φ 的**伴随算子**.

命题 1.2. φ 的伴随映射 φ^* 是 $W \rightarrow V$ 的线性映射.

Proof. 任取 $a_1, a_2 \in \mathbb{F}$, $w_1, w_2 \in W$, 那么

$$\begin{aligned} \langle u, \varphi^*(a_1 w_1 + a_2 w_2) \rangle &= \langle \varphi(u), a_1 w_1 + a_2 w_2 \rangle = \bar{a}_1 \langle \varphi(u), w_1 \rangle + \bar{a}_2 \langle \varphi(u), w_2 \rangle \\ &= \bar{a}_1 \langle u, \varphi^*(w_1) \rangle + \bar{a}_2 \langle u, \varphi^*(w_2) \rangle \\ &= \langle u, a_1 \varphi^*(w_1) + a_2 \varphi^*(w_2) \rangle, \end{aligned}$$

由 u 的任意性, 所以 $\varphi^*(a_1 w_1 + a_2 w_2) = a_1 \varphi^*(w_1) + a_2 \varphi^*(w_2)$, 即 φ^* 是 $W \rightarrow V$ 的线性映射. \square
接下来我们介绍伴随映射的一些重要的性质.

命题 1.3 (伴随映射的性质). 设 $\varphi : V \rightarrow W$ 是内积空间之间的线性映射, 那么

- (1) 对于任意的线性映射 $\psi : V \rightarrow W$, 有 $(\varphi + \psi)^* = \varphi^* + \psi^*$;
- (2) 对于任意 $\lambda \in \mathbb{F}$, 有 $(\lambda \varphi)^* = \bar{\lambda} \varphi^*$;
- (3) $(\varphi^*)^* = \varphi$;
- (4) 对于任意的线性映射 $\psi : V \rightarrow W$, 有 $(\varphi \psi)^* = \psi^* \varphi^*$;
- (5) 对于恒等变换 $\mathbb{1}_V$, 有 $\mathbb{1}_V^* = \mathbb{1}_V$;
- (6) 若 φ 可逆, 那么 φ^* 可逆, 并且 $(\varphi^*)^{-1} = (\varphi^{-1})^*$.

Proof. (1) 任取 $w \in W$, 那么

$$\begin{aligned} \langle u, (\varphi + \psi)^*(w) \rangle &= \langle (\varphi + \psi)(u), w \rangle = \langle \varphi(u) + \psi(u), w \rangle \\ &= \langle \varphi(u), w \rangle + \langle \psi(u), w \rangle = \langle u, \varphi^*(w) \rangle + \langle u, \psi^*(w) \rangle \\ &= \langle u, \varphi^*(w) + \psi^*(w) \rangle = \langle u, (\varphi^* + \psi^*)(w) \rangle, \end{aligned}$$

根据 u 的任意性, 所以 $(\varphi + \psi)^* = \varphi^* + \psi^*$.

(2) 任取 $w \in W$, 那么

$$\langle u, (\lambda \varphi)^*(w) \rangle = \langle (\lambda \varphi)(u), w \rangle = \lambda \langle \varphi(u), w \rangle = \lambda \langle u, \varphi^*(w) \rangle = \langle u, (\bar{\lambda} \varphi^*)(w) \rangle,$$

根据 u 的任意性, 所以 $(\lambda \varphi)^* = \bar{\lambda} \varphi^*$.

(3) 和 (4) 的证明和上面完全类似.

(5) 任取 $v \in V$, 那么 $\langle u, \mathbb{1}_V^*(v) \rangle = \langle u, v \rangle$, 所以 $\mathbb{1}_V^* = \mathbb{1}_V$.

SECTION 2 正规算子

(6) 由于 $\varphi\varphi^{-1} = \varphi^{-1}\varphi = \mathbb{1}$, 利用 (4) 和 (5), 所以

$$(\varphi^{-1})^*\varphi^* = \varphi^*(\varphi^{-1})^* = \mathbb{1},$$

故 φ^* 可逆, 且 $(\varphi^*)^{-1} = (\varphi^{-1})^*$. \square

命题 1.4. 设线性映射 $\varphi : V \rightarrow W$ 在 n 维内积空间 V 的一组标准正交基 e_1, \dots, e_n 和 m 维内积空间 W 的一组标准正交基 f_1, \dots, f_m 下的表示矩阵为 $m \times n$ 矩阵 A , 那么伴随算子 φ^* 在这组基下的表示矩阵为 $n \times m$ 矩阵 A^* . 这里 A^* 表示共轭转置.

Proof. 设 $A = (a_{ij})$, 那么 $\varphi(e_j) = \sum_i a_{ij} f_i$, 所以

$$\langle e_i, \varphi^*(f_j) \rangle = \langle \varphi(e_i), f_j \rangle = \left\langle \sum_k a_{ki} f_k, f_j \right\rangle = \sum_k a_{ki} \langle f_k, f_j \rangle = a_{ji}.$$

于是我们有

$$\langle u, \varphi^*(f_j) \rangle = \sum_i u_i \langle e_i, \varphi^*(f_j) \rangle = \sum_i u_i a_{ji} = \sum_i a_{ji} \langle u, e_i \rangle = \left\langle u, \sum_i \bar{a}_{ji} e_i \right\rangle,$$

根据 u 的任意性, 所以 $\varphi^*(f_j) = \sum_i \bar{a}_{ji} e_i$, 即 φ^* 在这组基下的表示矩阵为 A^* . \square

命题 1.5. 设内积空间 V 上的算子 φ 有不变子空间 U , 那么正交补 U^\perp 是 φ^* 的不变子空间.

Proof. 任取 $w \in U^\perp$, $u \in U$, 那么 $\varphi(u) \in U$, 所以

$$\langle u, \varphi^*(w) \rangle = \langle \varphi(u), w \rangle = 0,$$

所以 $\varphi^*(w) \in U^\perp$, 即 U^\perp 是 φ^* 的不变子空间. \square

命题 1.6. 设 φ 是内积空间 V 上的算子, 那么

- (1) $\ker \varphi^* = (\operatorname{im} \varphi)^\perp$;
- (2) $\operatorname{im} \varphi^* = (\ker \varphi)^\perp$.

Proof. (1) 任取 $v \in \ker \varphi^*$, 即 $\varphi^*(v) = 0$, 当且仅当对于任意的 $u \in V$ 有 $\langle u, \varphi^*(v) \rangle = 0$, 当且仅当 $\langle \varphi(u), v \rangle = 0$, 当且仅当 $v \in (\operatorname{im} \varphi)^\perp$.

(2) 由 (1) 我们有 $\operatorname{im} \varphi = (\ker \varphi)^{\perp\perp}$, 用 φ^* 替换 φ 即得 $\operatorname{im} \varphi^* = (\ker \varphi)^\perp$. \square

2 正规算子

线性代数的主要研究内容是在各种意义下对线性变换进行分类, 如相抵、相似等, 对于内积空间, 我们主要研究线性变换在标准正交基下的形态, 即正交相似或者酉相似, 其中正交对角化

SECTION 2 正规算子

和酉对角化又是主要的研究目标, 在研究这两类对角化问题的时候, 人们发现有一类算子发挥着主要的作用(实际上是酉对角化的充要条件), 即正规算子.

对于 n 维内积空间 V 上的算子 φ , 如果 φ 和 φ^* 可交换, 即

$$\varphi\varphi^* = \varphi^*\varphi,$$

那么我们说 φ 是正规算子. 根据 命题 1.4, 正规算子 φ 的表示矩阵 A 满足

$$AA^* = A^*A,$$

所以这样的矩阵我们称之为正规矩阵.

命题 2.1. 设 φ 是内积空间 V 上的算子, 那么下面的说法等价:

- (1) φ 是正规算子;
- (2) 任取 $v \in V$, 有 $\|\varphi(v)\| = \|\varphi^*(v)\|$;
- (3) φ 在 V 的标准正交基下的表示矩阵为正规矩阵.

Proof. (1) \Rightarrow (2) 任取 $v \in V$, 那么

$$\langle \varphi(v), \varphi(v) \rangle = \langle v, \varphi^*\varphi(v) \rangle = \langle v, \varphi\varphi^*(v) \rangle = \langle \varphi^*(v), \varphi^*(v) \rangle,$$

所以 $\|\varphi(v)\| = \|\varphi^*(v)\|$.

(2) \Rightarrow (1) 任取 $v \in V$, 有

$$\langle v, \varphi^*\varphi(v) \rangle = \langle \varphi(v), \varphi(v) \rangle = \langle \varphi^*(v), \varphi^*(v) \rangle = \langle v, \varphi\varphi^*(v) \rangle,$$

故 $\langle v, (\varphi^*\varphi - \varphi\varphi^*)(v) \rangle = 0$, 而 $\varphi^*\varphi - \varphi\varphi^*$ 是自伴随算子, 根据 定理 4.2, 所以 $\varphi^*\varphi - \varphi\varphi^* = 0$.

(3) \Rightarrow (1) 设 e_1, \dots, e_n 是 V 的标准正交基, φ 在这组基下的表示矩阵为 A . 那么

$$\varphi(e_1, \dots, e_n) = (e_1, \dots, e_n)A, \quad \varphi^*(e_1, \dots, e_n) = (e_1, \dots, e_n)A^*,$$

所以

$$\varphi\varphi^*(e_1, \dots, e_n) = (e_1, \dots, e_n)AA^*, \quad \varphi^*\varphi(e_1, \dots, e_n) = (e_1, \dots, e_n)A^*A,$$

$AA^* = A^*A$ 表明 $\varphi\varphi^* = \varphi^*\varphi$. \square

命题 2.2. 正规算子 φ 满足 $(\text{im } \varphi)^\perp = \ker \varphi$.

Proof. 根据 命题 1.6, 我们知道 $(\text{im } \varphi)^\perp = \ker \varphi^*$, 根据 命题 2.1 的 (2), 有 $\ker \varphi^* = \ker \varphi$, 所以 $(\text{im } \varphi)^\perp = \ker \varphi$. \square

命题 2.3. 正规算子 φ 的不同特征值 λ, μ 对应的特征子空间互相正交.

SECTION 2 正规算子

Proof. 任取非零的 $v \in V_\lambda$, $w \in V_\mu$, 由于 $(\varphi - \mu \mathbb{1})(w) = 0$, 所以 $(\varphi^* - \bar{\mu} \mathbb{1})(w) = 0$, 所以 $\varphi^*(w) = \bar{\mu}w$, 所以

$$\lambda \langle v, w \rangle = \langle \lambda v, w \rangle = \langle \varphi(v), w \rangle = \langle v, \varphi^*(w) \rangle = \langle v, \bar{\mu}w \rangle = \mu \langle v, w \rangle,$$

$\lambda \neq \mu$ 表明 $\langle v, w \rangle = 0$. \square

2.1 复正规算子的酉相似

首先我们利用 Jordan 标准型说明有限维酉空间上的算子都能在某组标准正交基下被上三角化, 这一结论被称为 Schur 定理.

定理 2.4. 令 V 是域 \mathbb{F} 上的 n 维向量空间, $\varphi : V \rightarrow V$ 是线性变换, 那么存在 V 的一组基 e_1, \dots, e_n 使得 φ 在这组基下的表示矩阵为上三角矩阵当且仅当 φ 的最小多项式在 \mathbb{F} 上能够完全分裂为一次多项式.

Proof. (\Rightarrow) 若 φ 在 e_1, \dots, e_n 下的表示矩阵为上三角矩阵

$$A = \begin{pmatrix} a_{11} & & & * \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix},$$

那么 φ 的特征多项式显然为

$$c(x) = (x - a_{11})(x - a_{22}) \cdots (x - a_{nn}).$$

由于最小多项式 $m(x)$ 整除 $c(x)$, 所以 $m(x)$ 在 \mathbb{F} 上完全分裂为一次多项式.

(\Leftarrow) 若 φ 的最小多项式 $m(x)$ 在 \mathbb{F} 上完全分裂为一次多项式. 那么 φ 在 \mathbb{F} 上有 Jordan 标准型, 是上三角矩阵. \square

推论 2.5. 令 V 是域 \mathbb{F} 上的 n 维内积空间, $\varphi : V \rightarrow V$ 是线性变换, 那么存在 V 的一组标准正交基 e_1, \dots, e_n 使得 φ 在这组基下的表示矩阵为上三角矩阵当且仅当 φ 的最小多项式在 \mathbb{F} 上能够完全分裂为一次多项式.

Proof. (\Rightarrow) 显然.

(\Leftarrow) 若 φ 的最小多项式 $m(x)$ 在 \mathbb{F} 上完全分裂为一次多项式, 那么存在 V 的一组基 f_1, \dots, f_n 使得 φ 在这组基下的表示矩阵为上三角矩阵, 等价地说, 对于 $1 \leq k \leq n$, $\text{span}(f_1, \dots, f_k)$ 是 φ 的不变子空间. 对 f_1, \dots, f_n 使用 Gram-Schmidt 正交化, 得到一组标准正交基 e_1, \dots, e_n , 并且满足

$$\text{span}(e_1, \dots, e_k) = \text{span}(f_1, \dots, f_k) \quad 1 \leq k \leq n.$$

SECTION 2 正规算子

由于 $\text{span}(e_1, \dots, e_k)$ 是 φ 的不变子空间, 所以 φ 在标准正交基 e_1, \dots, e_n 下的表示矩阵为上三角矩阵. \square

推论 2.6 (Schur 定理). V 是有限维酉空间, $\varphi : V \rightarrow V$ 是线性变换, 那么存在 V 的一组标准正交基 e_1, \dots, e_n 使得 φ 在这组基下的表示矩阵为上三角矩阵.

Proof. 由于 \mathbb{C} 上的任意多项式都能完全分裂为一次多项式的乘积, 再根据 **推论 2.5** 即得. \square

现在我们便可以证明复谱定理, 即复矩阵能够酉对角化的充分必要条件是其为正规矩阵.

定理 2.7 (复谱定理). 对于 n 维酉空间 V , $\varphi : V \rightarrow V$ 是线性变换, 那么存在 V 的一组标准正交基 e_1, \dots, e_n 使得 φ 在这组基下的表示矩阵为对角阵当且仅当 φ 是正规算子.

Proof. (\Rightarrow) 若存在 V 的一组标准正交基 e_1, \dots, e_n 使得 φ 在这组基下的表示矩阵为对角阵

$$A = \text{diag}(a_1, \dots, a_n),$$

那么 φ^* 在这组基下的表示矩阵为

$$A^* = \text{diag}(\bar{a}_1, \dots, \bar{a}_n),$$

又因为

$$AA^* = \text{diag}(a_1\bar{a}_1, \dots, a_n\bar{a}_n) = A^*A,$$

所以 $\varphi\varphi^* = \varphi^*\varphi$, 即 φ 是正规算子.

(\Leftarrow) 若 φ 是正规算子, 根据 Schur 定理, 存在 V 的一组标准正交基 e_1, \dots, e_n 使得 φ 在这组基下的表示矩阵为上三角阵

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \ddots & & \vdots \\ & & & a_{nn} \end{pmatrix},$$

那么 φ^* 在这组基下的表示矩阵为 A^* . φ 正规表明 $AA^* = A^*A$, 即

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \ddots & & \vdots \\ & & & a_{nn} \end{pmatrix} \begin{pmatrix} \bar{a}_{11} & & & \\ \bar{a}_{12} & \bar{a}_{22} & & \\ \vdots & \vdots & \ddots & \\ \bar{a}_{1n} & \bar{a}_{2n} & \cdots & \bar{a}_{nn} \end{pmatrix} = \begin{pmatrix} \bar{a}_{11} & & & \\ \bar{a}_{12} & \bar{a}_{22} & & \\ \vdots & \vdots & \ddots & \\ \bar{a}_{1n} & \bar{a}_{2n} & \cdots & \bar{a}_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \ddots & & \vdots \\ & & & a_{nn} \end{pmatrix}.$$

对比两边的 (1, 1)-元得到 $\|a_{11}\|^2 + \|a_{12}\|^2 + \cdots + \|a_{1n}\|^2 = \|a_{11}\|^2$, 故 $\|a_{12}\|^2 + \cdots + \|a_{1n}\|^2 = 0$, 这表明 $a_{12} = \cdots = a_{1n} = 0$. 然后对比 (2, 2)-元可得 $a_{23} = \cdots = a_{2n} = 0$, 以此类推可得 $i \neq j$ 时有 $a_{ij} = 0$, 即 A 为对角阵. \square

推论 2.8. 两个复正规矩阵酉相似的充分必要条件是它们有相同的特征值 (计重数).

2.2 实正规算子的正交相似

对于实正规算子而言，我们想要仿照复正规算子的方法，但是 \mathbb{R} 上的不可约多项式最高为二次多项式，所以实正规算子的最小多项式可能不会完全分裂为一次多项式的乘积，故我们首先研究最小多项式为二次不可约多项式的实正规算子。

定理 2.9. 设 φ 是 n 维 Euclid 空间 V 上的正规算子，其最小多项式为 $m(x) = (x - a)^2 + b^2$ ，其中 $a, b \in \mathbb{R}$ 且 $b > 0$ ，即 $m(x)$ 是二次不可约多项式。那么 $n = 2k$ 是偶数，并且存在 φ 的 2 维不变子空间 V_1, V_2, \dots, V_k 使得

- (1) $i \neq j$ 时 $V_i \perp V_j$.
- (2) $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$.
- (3) 对于 $1 \leq i \leq k$ ，存在 V_i 的标准正交基 e_{2i-1}, e_{2i} 使得 $\varphi|_{V_i}$ 在这组基下的表示矩阵为

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Proof. 如果 n 是奇数，那么 φ 的特征多项式 $c(x)$ 是奇数次多项式，所以有实根 λ_0 ，而最小多项式 $m(x)$ 和 $c(x)$ 的根相同，所以 λ_0 是 $m(x)$ 的根，这与 $m(x)$ 在 \mathbb{R} 上不可约矛盾，所以 n 必须为偶数。设 $n = 2k$ ，我们对 k 进行归纳。

$k = 1$ 时，即 $n = 2$ ，我们说明存在 V 的一组标准正交基使得 φ 的表示矩阵如上。令

$$\psi = \frac{1}{b}(\varphi - a\mathbb{1}),$$

那么 ψ 也是正规算子，并且

$$\psi^2 + \mathbb{1} = \frac{1}{b^2}(\varphi - a\mathbb{1})^2 + \mathbb{1} = 0,$$

所以 ψ 不是零映射且最小多项式为 $x^2 + 1$ 。取单位向量 e_1 使得 $\psi(e_1) \neq 0$ ，记 $e_2 = -\psi(e_1)$ ，那么 $\psi(e_2) = -\psi^2(e_1) = e_1$ 。 ψ 是正规算子表明 $\|\psi(v)\| = \|\psi^*(v)\|$ ，故

$$\begin{aligned} & \|\psi^*(e_1) - e_2\|^2 + \|\psi^*(e_2) + e_1\|^2 \\ &= \|\psi^*(e_1)\|^2 - 2\langle\psi^*(e_1), e_2\rangle + \|e_2\|^2 + \|\psi^*(e_2)\|^2 + 2\langle\psi^*(e_2), e_1\rangle + \|e_1\|^2 \\ &= \|\psi(e_1)\|^2 - 2\langle e_1, \psi(e_2)\rangle + \|e_2\|^2 + \|\psi(e_2)\|^2 + 2\langle e_2, \psi(e_1)\rangle + \|e_1\|^2 \\ &= \|\psi(e_2)\|^2 - 2\|\psi(e_2), e_1\| + \|e_1\|^2 + \|\psi(e_1)\|^2 + 2\langle\psi(e_1), e_2\rangle + \|e_2\|^2 \\ &= \|\psi(e_2) - e_1\|^2 + \|\psi(e_1) + e_2\|^2 = 0, \end{aligned}$$

所以 $\psi^*(e_1) = e_2$ 以及 $\psi^*(e_2) = -e_1$ 。这就表明

$$\langle e_1, e_2 \rangle = \langle \psi(e_2), e_2 \rangle = \langle e_2, \psi^*(e_2) \rangle = \langle e_2, -e_1 \rangle = -\langle e_1, e_2 \rangle,$$

SECTION 2 正规算子

所以 $\langle e_1, e_2 \rangle = 0$, 即 e_1, e_2 正交. 又因为

$$\|e_2\|^2 = \langle e_2, e_2 \rangle = \langle -\psi(e_1), e_2 \rangle = -\langle e_1, \psi^*(e_2) \rangle = \langle e_1, e_1 \rangle = \|e_1\|^2 = 1,$$

所以 e_2 是单位向量. 故 e_1, e_2 是 V 的标准正交基且 ψ 在这组基下的表示矩阵为

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

那么 φ 在这组基下的表示矩阵为

$$bA + a = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

下面假设结论对 $k - 1$ 成立. 同样记

$$\psi = \frac{1}{b}(\varphi - a\mathbb{1}),$$

那么 ψ 不是零映射且最小多项式为 $x^2 + 1$. 取单位向量 e_1 使得 $\psi(e_1) \neq 0$, 记 $e_2 = -\psi(e_1)$, 那么 $\psi(e_2) = e_1$. 重复上面的证明, 可知 e_1, e_2 正交且均为单位向量. 记 $V_1 = \text{span}(e_1, e_2)$. 那么 V_1 是 φ 的不变子空间, 且 $\varphi|_{V_1}$ 在 e_1, e_2 下的表示矩阵为 $(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix})$. 此外, 我们有 $V = V_1 \oplus V_1^\perp$. 注意到

$$\begin{aligned} \varphi^*(e_1) &= (b\psi^* + a\mathbb{1})(e_1) = be_2 + ae_1 \in V_1, \\ \varphi^*(e_2) &= (b\psi^* + a\mathbb{1})(e_2) = ae_2 - be_1 \in V_1, \end{aligned}$$

所以 V_1 是 φ^* 的不变子空间, 根据 命题 1.5, 所以 V_1^\perp 是 $(\varphi^*)^* = \varphi$ 的不变子空间. 那么我们可以考虑限制 $\varphi|_{V_1^\perp}$. 显然 $m(x)$ 仍然是 $\varphi|_{V_1^\perp}$ 的零化多项式, 而 $m(x)$ 不可约, 所以 $\varphi|_{V_1^\perp}$ 的最小多项式还是 $m(x)$. 那么根据归纳假设, 存在 $\varphi|_{V_1^\perp}$ 的 2 维不变子空间 V_2, \dots, V_k 使得

$$V_1^\perp = V_2 \oplus \cdots \oplus V_k,$$

其中 $i \neq j$ 时有 $V_i \perp V_j$, 并且 V_i 有标准正交基 e_{2i-1}, e_{2i} 使得 $\varphi|_{V_i^\perp}$ 限制在 V_i 上的表示矩阵为 $(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix})$. 显然 V_2, \dots, V_k 也是 φ 的不变子空间, 这就完成了证明. \square

下面我们就可以研究 Euclid 空间上的一般正规算子了.

定理 2.10. 设 φ 是 Euclid 空间 V 上的正规算子, 其最小多项式 $m(x)$ 的所有不可约因子为

$$(x - a_1)^2 + b_1^2, \dots, (x - a_s)^2 + b_s^2, x - a_{2s+1}, \dots, x - a_{2s+t},$$

其中 $a_1, \dots, a_s, a_{2s+1}, \dots, a_{2s+t} \in \mathbb{R}, b_1, \dots, b_s > 0$. 那么

(1) 每个不可约因子的幂次都是一次, 即

$$m(x) = ((x - a_1)^2 + b_1^2) \cdots ((x - a_s)^2 + b_s^2)(x - a_{2s+1}) \cdots (x - a_{2s+t}).$$

SECTION 2 正规算子

(2) 存在 φ 的两两正交的不变子空间 $W_1, \dots, W_s, W_{2s+1}, \dots, W_{2s+t}$, 使得 $\varphi|_{W_i}$ 的最小多项式 $m_i(x)$ 为

$$m_i(x) = \begin{cases} (x - a_i)^2 + b_i^2 & 1 \leq i \leq s, \\ x - a_i & 2s + 1 \leq i \leq 2s + t. \end{cases}$$

(3) $V = W_1 \oplus \dots \oplus W_s \oplus W_{2s+1} \oplus \dots \oplus W_{2s+t}$.

Proof. 假设

$$m(x) = ((x - a_1)^2 + b_1^2)^{r_1} \cdots ((x - a_s)^2 + b_s^2)^{r_s} (x - a_{2s+1})^{r_{2s+1}} \cdots (x - a_{2s+t})^{r_{2s+t}},$$

其中 $r_1, \dots, r_s, r_{2s+1}, \dots, r_{2s+t} \geq 1$. 假设某个 $r_i \geq 2$, 记这个不可约因子为 $p_i(x)$. 设 $g_i(x)$ 满足 $m(x) = p_i^2(x)g_i(x)$. 那么 $m(\varphi) = p_i^2(\varphi)g_i(\varphi) = 0$. 任取 $v \in V$, 那么 $p_i^2(\varphi)g_i(\varphi)(v) = 0$, 所以 $p_i(\varphi)g_i(\varphi)(v) \in \ker p_i(\varphi)$, 同时注意到 $p_i(\varphi)g_i(\varphi)(v) \in \operatorname{im} p_i(\varphi)$. 由于 $p_i(\varphi)$ 也是正规算子, 根据命题 2.2, 所以 $p_i(\varphi)g_i(\varphi)(v) = 0$, 故 $p_i(\varphi)g_i(\varphi) = 0$. 这表明 $p_i(x)g_i(x)$ 是 φ 的零化多项式, 与 $m(x)$ 是最小多项式矛盾. 所以 $r_i = 1$. 这就证明了(1).

记号不变, 那么对于每个 i , 设 $m(x) = p_i(x)f_i(x)$, 其中 $p_i \nmid f_i$. 令 $W_i = \ker p_i(\varphi)$. 任取 $w \in W_i$, 那么 $p_i(\varphi)(w) = 0$, 那么 $p_i(\varphi)(\varphi(w)) = \varphi(p_i(\varphi)(w)) = 0$, 所以 $\varphi(w) \in W_i$, 故 W_i 是 φ 的不变子空间. 记 $\varphi_i = \varphi|_{W_i}$, 那么

$$p_i(\varphi_i)(w) = p_i(\varphi)(w) = 0 \quad w \in W_i,$$

这表明 $p_i(x)$ 是 φ_i 的零化多项式. 又因为 p_i 不可约, 所以 $p_i(x)$ 就是 φ_i 的最小多项式. $i \neq j$ 时, 任取 $w_i \in W_i$, $w_j \in W_j$. 此时 $p_i(x)$ 和 $p_j(x)$ 互素, 所以存在 $q_i(x)$ 和 $q_j(x)$ 使得 $p_i(x)q_i(x) + p_j(x)q_j(x) = 1$, 故 $p_i(\varphi)q_i(\varphi) + p_j(\varphi)q_j(\varphi) = 1$. 那么

$$w_i = p_j(\varphi)q_j(\varphi)(w_i), \quad w_j = p_i(\varphi)q_i(\varphi)(w_j),$$

注意 $p_j(\varphi)$ 也是正规算子, 所以 $\ker p_j(\varphi)^* = \ker p_j(\varphi)$, 所以

$$\begin{aligned} \langle w_i, w_j \rangle &= \langle p_j(\varphi)q_j(\varphi)(w_i), p_i(\varphi)q_i(\varphi)(w_j) \rangle \\ &= \langle q_j(\varphi)(w_i), p_i(\varphi)q_i(\varphi)p_j(\varphi)^*(w_j) \rangle \\ &= 0. \end{aligned}$$

这就表明 $W_1, \dots, W_s, W_{2s+1}, \dots, W_{2s+t}$ 两两正交. 所以(2)成立.

由于多项式 $f_i(x)$ 互素, 所以存在多项式 $u_i(x)$ 使得

$$\sum_i u_i(x) f_i(x) = 1,$$

SECTION 2 正规算子

即 $\sum_i u_i(\varphi) f_i(\varphi) = \mathbb{1}$. 任取 $v \in V$, 那么

$$v = \sum_i u_i(\varphi) f_i(\varphi)(v),$$

由于

$$p_i(\varphi) u_i(\varphi) f_i(\varphi)(v) = u_i(\varphi) m(\varphi)(v) = 0,$$

所以 $u_i(\varphi) f_i(\varphi)(v) \in \ker p_i(\varphi) = W_i$, 所以

$$V = W_1 + \cdots + W_s + W_{2s+1} + \cdots + W_{2s+t}.$$

W_i 两两正交表明这是直和, 这就证明了(3). \square

定理 2.11 (实正规算子的正交相似标准型). 设 φ 是 n 维 Euclid 空间 V 上的正规算子, 那么存在 2 维不变子空间 V_1, \dots, V_s 和 1 维不变子空间 V_{2s+1}, \dots, V_n , 它们两两正交, 并且

$$V = V_1 \oplus \cdots \oplus V_s \oplus V_{2s+1} \oplus \cdots \oplus V_n.$$

此外, 对于 $1 \leq i \leq s$, 存在 V_i 的标准正交基 e_{2i-1}, e_{2i} , 对于 $2s+1 \leq j \leq n$, 存在 V_j 的单位向量 e_j , 使得 φ 在标准正交基 $e_1, e_2, \dots, e_{2s-1}, e_{2s}, e_{2s+1}, \dots, e_n$ 下的表示矩阵为下列分块对角阵:

$$\text{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \dots, \begin{pmatrix} a_s & b_s \\ -b_s & a_s \end{pmatrix}, a_{2s+1}, \dots, a_n\right),$$

其中 $a_1, \dots, a_s, a_{2s+1}, \dots, a_n$ 是实数, $b_1, \dots, b_s > 0$. 此时容易看出, a_{2s+1}, \dots, a_n 是 φ 的 $n-2s$ 个实特征值, $a_1 \pm b_1i, \dots, a_s \pm b_si$ 是 φ 的 $2s$ 个复特征值.

Proof. 根据定理 2.10, 可设 φ 的最小多项式为

$$m(x) = ((x - a_1)^2 + b_1^2) \cdots ((x - a_\ell)^2 + b_\ell^2)(x - a_{2\ell+1}) \cdots (x - a_{2\ell+t}),$$

并且

$$V = W_1 \oplus \cdots \oplus W_\ell \oplus W_{2\ell+1} \oplus \cdots \oplus W_{2\ell+t},$$

其中 W_j 两两正交且为 φ 的不变子空间. 此时 $\varphi|_{W_j}$ 是 W_j 上的正规算子.

当 $1 \leq j \leq \ell$ 的时候, $\varphi|_{W_j}$ 的最小多项式为 $(x - a_j)^2 + b_j^2$, 根据定理 2.9, 有

$$W_j = V_{j,1} \oplus \cdots \oplus V_{j,k_j},$$

其中 $V_{j,1}, \dots, V_{j,k_j}$ 是两两正交的 $\varphi|_{W_j}$ 的不变子空间, 对于每个 $V_{j,r}$, 存在标准正交基 $e_{2r-1}^{(j)}, e_{2r}^{(j)}$ 使得 $\varphi|_{W_j}$ 在这组基下的表示矩阵为

$$\begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix},$$

SECTION 3 酉算子和正交算子

那么 $\varphi|_{W_j}$ 在 W_j 的标准正交基 $e_1^{(j)}, e_2^{(j)}, \dots, e_{2k_j-1}^{(j)}, e_{2k_j}^{(j)}$ 下的表示矩阵为

$$\text{diag} \left(\begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}, \dots, \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix} \right),$$

其中有 k_j 个分块.

当 $2\ell + 1 \leq j \leq 2\ell + t$ 的时候, $\varphi|_{W_j}$ 的最小多项式为 $x - a_j$, 因此 $\varphi|_{W_j} = a_j \mathbb{1}_{W_j}$, 此时任取 W_j 的一组标准正交基 $e_1^{(j)}, \dots, e_{k_j}^{(j)}$, 记 $V_{j,r} = \text{span}(e_r^{(j)})$, 所以

$$W_j = V_{j,1} \oplus \dots \oplus V_{j,k_j},$$

并且 $\varphi|_{W_j}$ 在这组基下的表示矩阵为 $a_j I_{k_j}$.

综上, 我们有

$$V = \bigoplus_{j=1}^{\ell} \bigoplus_{r=1}^{k_j} V_{j,r} \oplus \bigoplus_{j=2\ell+1}^{2\ell+t} \bigoplus_{r=1}^{k_j} V_{j,r},$$

此时 $e_1^{(1)}, \dots, e_{k_1}^{(1)}, \dots, e_1^{(\ell)}, \dots, e_{k_\ell}^{(\ell)}, e_1^{(2\ell+1)}, \dots, e_{k_{2\ell+1}}^{(2\ell+1)}, \dots, e_1^{2\ell+t}, \dots, e_{k_{2\ell+t}}^{(2\ell+t)}$ 是 V 的标准正交基, 并且 φ 在这组基下的表示矩阵即为所求. \square

推论 2.12. 两个实正规矩阵正交相似当且仅当它们有相同的实特征值和复特征值 (计重数).

3 酉算子和正交算子

对于酉空间 V 上的算子 φ , 如果满足

$$\|\varphi(v)\| = \|v\| \quad \forall v \in V,$$

那么我们说 φ 是**酉算子**. 对于 Euclid 空间而言, 我们说 φ 是**正交算子**.

命题 3.1. 设 φ 是酉空间 V 上的算子, 那么下面的说法等价:

- (1) φ 是酉算子;
- (2) 任取 $u, v \in V$ 有 $\langle \varphi(u), \varphi(v) \rangle = \langle u, v \rangle$;
- (3) φ 将标准正交基变为标准正交基, 即若 e_1, \dots, e_n 是 V 的标准正交基, 那么 $\varphi(e_1), \dots, \varphi(e_n)$ 是 V 的标准正交基;
- (4) φ 在任意标准正交基下的表示矩阵是酉矩阵.
- (5) φ 是正规算子, 且 $\varphi\varphi^* = \varphi^*\varphi = \mathbb{1}$.

SECTION 3 酉算子和正交算子

Proof. (1) \Rightarrow (2) 任取 $u, v \in V$, 那么

$$\begin{aligned}\langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle &= \langle u + v, u + v \rangle \\&= \langle \varphi(u + v), \varphi(u + v) \rangle \\&= \langle \varphi(u), \varphi(u) \rangle + 2\langle \varphi(u), \varphi(v) \rangle + \langle \varphi(v), \varphi(v) \rangle \\&= \langle u, u \rangle + 2\langle \varphi(u), \varphi(v) \rangle + \langle v, v \rangle,\end{aligned}$$

这就表明 $\langle \varphi(u), \varphi(v) \rangle = \langle u, v \rangle$.

(2) \Rightarrow (3) 注意到

$$\langle \varphi(e_i), \varphi(e_j) \rangle = \langle e_i, e_j \rangle = \delta_{ij},$$

故 $\varphi(e_1), \dots, \varphi(e_n)$ 是 V 的标准正交基.

(3) \Rightarrow (4) 设 φ 在标准正交基 e_1, \dots, e_n 下的表示矩阵为 $A = (a_{ij})$, 即

$$\varphi(e_j) = \sum_{k=1}^n a_{kj} e_k.$$

那么

$$\begin{aligned}\langle \varphi(e_i), \varphi(e_j) \rangle &= \left\langle \sum_{k=1}^n a_{ki} e_k, \sum_{\ell=1}^n a_{\ell j} e_\ell \right\rangle = \sum_{k=1}^n \sum_{\ell=1}^n a_{ki} \bar{a}_{\ell j} \langle e_k, e_\ell \rangle \\&= \sum_{k=1}^n \sum_{\ell=1}^n a_{ki} \bar{a}_{kj} \delta_{k\ell} = \sum_{k=1}^n a_{ki} \bar{a}_{kj},\end{aligned}$$

由于 $\varphi(e_1), \dots, \varphi(e_n)$ 也是标准正交基, 所以

$$\sum_{k=1}^n a_{ki} \bar{a}_{kj} = \delta_{ij},$$

这就表明 $A^* A = I_n$, 故 A 是酉矩阵.

(4) \Rightarrow (5) 此时 φ 在标准正交基下的表示矩阵是酉矩阵 A , 所以 $\varphi \varphi^*$ 在这组基下的表示矩阵为 $AA^* = I_n$, 故 φ 正规并且 $\varphi \varphi^* = \varphi^* \varphi = 1$.

(5) \Rightarrow (1) 任取 $v \in V$, 那么

$$\|\varphi(v)\|^2 = \langle \varphi(v), \varphi(v) \rangle = \langle v, \varphi^* \varphi(v) \rangle = \langle v, v \rangle = \|v\|^2,$$

所以 φ 是酉算子. \square

SECTION 3 酉算子和正交算子

推论 3.2. 设 $\varphi : V \rightarrow V$ 是酉算子, 那么 φ 可逆, 并且 $\varphi^{-1} = \varphi^*$ 也是酉算子. 此外, 若 φ, ψ 都是酉算子, 那么 $\varphi\psi$ 也是酉算子.

Proof. 由于

$$\varphi^*(\varphi^*)^* = \varphi^*\varphi = \mathbb{1},$$

所以 φ^* 也是酉算子. 类似地,

$$(\varphi\psi)^*(\varphi\psi) = \psi^*\varphi^*\varphi\psi = \psi^*\psi = \mathbb{1},$$

所以 $\varphi\psi$ 也是酉算子. \square

显然上述结论对于 Euclid 空间上的正交算子也都是成立的. **推论 3.2** 告诉我们酉空间 V 上所有酉算子的集合 $O_n(\mathbb{C})$ 构成群, 我们称为**酉群**. 类似地, Euclid 空间上的正交算子集合 $O_n(\mathbb{R})$ 也构成群, 我们称为**正交群**. 由于酉算子是正规算子, 这允许我们使用上一节的结果.

定理 3.3. 若 φ 是酉算子, 那么 φ 的特征值 λ_0 的模长为 1. 特别的, 若 φ 是正交算子, 那么 $\lambda_0 = \pm 1$.

Proof. 设 $v \in V$ 是 λ_0 的特征向量, 即 $\varphi(v) = \lambda_0 v$, 那么

$$|\lambda_0| \|v\| = \|\lambda_0 v\| = \|\varphi(v)\| = \|v\|,$$

由于 $\|v\| \neq 0$, 所以 $|\lambda_0| = 1$. \square

定理 3.4 (酉算子的酉相似). 设 φ 是 n 维酉空间 V 上的酉算子, 其特征值为

$$e^{i\theta_1}, \dots, e^{i\theta_n},$$

那么存在 V 的一组标准正交基 e_1, \dots, e_n , 使得 φ 在这组基下的表示矩阵为对角阵

$$\text{diag}(e^{i\theta_1}, \dots, e^{i\theta_n}).$$

Proof. 由 **定理 2.7** 立即得到. \square

定理 3.5 (正交算子的正交相似). 设 φ 是 n 维 Euclid 空间 V 上的正交算子, 其特征值为

$$e^{i\theta_1}, e^{-i\theta_1}, \dots, e^{i\theta_s}, e^{-i\theta_s}, \underbrace{1, \dots, 1}_{t \text{ times}}, \underbrace{-1, \dots, -1}_{n-2s-t \text{ times}},$$

其中 $0 < \theta_1 \leq \dots \leq \theta_s < \pi$. 那么存在 V 的一组标准正交基 e_1, \dots, e_n , 使得 φ 在这组基下的表示矩阵为分块对角阵

$$\text{diag} \left(\begin{pmatrix} \cos \theta_1 & \sin \theta_1 \\ -\sin \theta_1 & \cos \theta_1 \end{pmatrix}, \dots, \begin{pmatrix} \cos \theta_s & \sin \theta_s \\ -\sin \theta_s & \cos \theta_s \end{pmatrix}, \underbrace{1, \dots, 1}_{t \text{ times}}, \underbrace{-1, \dots, -1}_{n-2s-t \text{ times}} \right).$$

Proof. 由 **定理 2.11** 立即得到. \square

4 自伴随算子和反自伴随算子

如果内积空间 V 上的算子 φ 满足 $\varphi = \varphi^*$, 那么我们说 φ 是**自伴随算子**. 如果 $\varphi = -\varphi^*$, 那么我们说 φ 是**反自伴随算子**.

设 φ 在某组标准正交基下的表示矩阵为 A , 那么 φ^* 在这组基下的表示矩阵为 A^* , 所以自伴随算子对应 $A = A^*$ 的矩阵, 这类矩阵被称为**Hermite 矩阵**, 对于 Euclid 空间的情况, 则对应 $A = A^T$ 的矩阵, 被称为**实对称矩阵**. 类似地, 反自伴随算子对应 $A = -A^*$ 的矩阵, 被称为**反 Hermite 矩阵**, 对于 Euclid 空间的情况, 则对应 $A = -A^T$ 的矩阵, 被称为**反对称矩阵**.

定理 4.1. 自伴随算子 φ 的特征值必为实数.

Proof. 设 λ_0 是 φ 的特征值, 那么任取 λ_0 的特征向量 v , 有

$$\lambda_0 \langle v, v \rangle = \langle \lambda_0 v, v \rangle = \langle \varphi(v), v \rangle = \langle v, \varphi(v) \rangle = \bar{\lambda}_0 \langle v, v \rangle,$$

$\langle v, v \rangle \neq 0$ 表明 $\lambda_0 = \bar{\lambda}_0$, 所以 λ_0 为实数. \square

定理 4.2. φ 是内积空间 V 上的自伴随算子, 如果对于任意的 $v \in V$ 都有 $\langle \varphi(v), v \rangle = 0$, 那么 $\varphi = 0$.

Proof. 首先假设 V 是 Euclid 空间, 任取 $u, v \in V$, 此时有恒等式 (实际上就是仿照 $ab = ((a+b)^2 - (a-b)^2)/4$)

$$\langle \varphi(u), v \rangle = \frac{\langle \varphi(u+v), u+v \rangle - \langle \varphi(u-v), u-v \rangle}{4},$$

这表明 $\langle \varphi(u), v \rangle = 0$, 令 $v = \varphi(u)$, 所以 $\varphi(u) = 0$, 这就表明 $\varphi = 0$.

现在假设 V 是酉空间, 任取 $u, v \in V$, 此时有恒等式

$$\begin{aligned} \langle \varphi(u), v \rangle &= \frac{\langle \varphi(u+v), u+v \rangle - \langle \varphi(u-v), u-v \rangle}{4} \\ &\quad + \frac{\langle \varphi(u+iv), u+iv \rangle - \langle \varphi(u-iv), u-iv \rangle}{4}i, \end{aligned}$$

同样表明 $\langle \varphi(u), v \rangle = 0$, 令 $v = \varphi(u)$, 所以 $\varphi(u) = 0$, 这就表明 $\varphi = 0$. \square

定理 4.3. φ 是酉空间 V 上的算子, 那么 φ 是自伴随算子当且仅当对于任意的 $v \in V$, $\langle \varphi(v), v \rangle$ 都是实数.

Proof. 若 φ 是自伴随算子, 那么

$$\langle \varphi(v), v \rangle = \langle v, \varphi(v) \rangle = \overline{\langle \varphi(v), v \rangle},$$

SECTION 4 自伴随算子和反自伴随算子

即 $\langle \varphi(v), v \rangle$ 是实数.

若对于任意的 $v \in V$, $\langle \varphi(v), v \rangle$ 都是实数. 那么

$$\langle (\varphi - \varphi^*)(v), v \rangle = \langle \varphi(v), v \rangle - \overline{\langle v, \varphi^*(v) \rangle} = \langle \varphi(v), v \rangle - \overline{\langle \varphi(v), v \rangle} = 0,$$

所以 $\varphi = \varphi^*$, 即 φ 是自伴随算子. \square

定理 4.4 (自伴随算子的酉相似). 设 φ 是 n 维酉空间 V 上的自伴随算子, 其特征值为

$$\lambda_1, \dots, \lambda_n,$$

其中 $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, 那么存在 V 的一组标准正交基 e_1, \dots, e_n 使得 φ 在这组基下的表示矩阵为对角阵

$$\text{diag}(\lambda_1, \dots, \lambda_n).$$

Proof. 由 [定理 2.7](#) 立即得到. \square

定理 4.5 (自伴随算子的正交相似, 实谱定理). 设 φ 是 n 维 Euclid 空间 V 上的自伴随算子, 其特征值为

$$\lambda_1, \dots, \lambda_n,$$

其中 $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, 那么存在 V 的一组标准正交基 e_1, \dots, e_n 使得 φ 在这组基下的表示矩阵为对角阵

$$\text{diag}(\lambda_1, \dots, \lambda_n).$$

Proof. 由 [定理 2.11](#) 立即得到. \square

推论 4.6. 对于 n 维 Euclid 空间 V , $\varphi : V \rightarrow V$ 是线性变换, 那么存在 V 的一组标准正交基 e_1, \dots, e_n 使得 φ 在这组基下的表示矩阵为对角阵当且仅当 φ 是自伴随算子.

Proof. 充分性由 [定理 4.5](#) 保证. 由于 φ^* 的表示矩阵是 φ 表示矩阵的转置, 对角阵的转置还是本身, 所以必要性显然. \square

例 4.7. 给定实对称矩阵

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix},$$

求正交矩阵 P 使得 $P^{-1}AP$ 为对角阵.

Solution. 根据谱定理, A 可以正交对角化, 此时 A 的特征向量可以组成 \mathbb{R}^3 的一组标准正交基, 故我们只需要求出 A 的一组正交的特征向量然后按列拼成的矩阵就是 P 矩阵.

SECTION 4 自伴随算子和反自伴随算子

计算特征多项式为

$$\det(xI_3 - A) = (x - 8)(x - 2)^2,$$

所以 A 的特征值为 $\lambda_1 = 8$, $\lambda_2 = \lambda_3 = 2$. 这表明特征值 8 的特征子空间只有一维, 求解方程组 $(8I_3 - A)x = 0$ 可得解空间的基为 $(1, 1, 1)^T$. 特征值 2 的特征子空间为二维, 求解方程组 $(2I_3 - A)x = 0$ 可得解空间的基为 $(-1, 1, 0)^T, (-1, 0, 1)^T$. 命题 2.3 表明不同特征值对应的特征向量一定正交. 所以只需要 Gram-Schmidt 正交化特征值 2 的两个特征向量, 得到 $(-1, 1, 0)^T, (-1/2, -1/2, 1)^T$. 然后将上述三个向量单位化便得到

$$P = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}, \quad P^{-1}AP = \begin{pmatrix} 8 & & \\ & 2 & \\ & & 2 \end{pmatrix}. \quad \blacksquare$$

下面我们研究反自伴随算子.

定理 4.8. 反自伴随算子 φ 的特征值必为纯虚数.

Proof. 设 λ_0 是 φ 的特征值, 那么任取 λ_0 的特征向量 v , 有

$$\lambda_0 \langle v, v \rangle = \langle \lambda_0 v, v \rangle = \langle \varphi(v), v \rangle = \langle v, -\varphi(v) \rangle = -\bar{\lambda}_0 \langle v, v \rangle,$$

$\langle v, v \rangle \neq 0$ 表明 $\lambda_0 = -\bar{\lambda}_0$, 所以 λ_0 为纯虚数. \square

定理 4.9 (反自伴随算子的酉相似). 设 φ 是 n 维酉空间 V 上的反自伴随算子, 其特征值为

$$\pm i b_1, \dots, \pm i b_s, \underbrace{0, \dots, 0}_{n-2s \text{ times}},$$

其中 $b_1, \dots, b_s > 0$, 那么存在 V 的一组标准正交基 e_1, \dots, e_n 使得 φ 在这组基下的表示矩阵为对角阵

$$\text{diag} \left(i b_1, -i b_1, \dots, i b_s, -i b_s, \underbrace{0, \dots, 0}_{n-2s \text{ times}} \right).$$

Proof. 由 定理 2.7 立即得到. \square

定理 4.10 (反自伴随算子的正交相似). 设 φ 是 n 维 Euclid 空间 V 上的反自伴随算子, 其特征值为

$$\pm i b_1, \dots, \pm i b_s, \underbrace{0, \dots, 0}_{n-2s \text{ times}},$$

其中 $b_1, \dots, b_s > 0$, 那么存在 V 的一组标准正交基 e_1, \dots, e_n 使得 φ 在这组基下的表示矩阵为对角阵

$$\text{diag} \left(\begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & b_s \\ -b_s & 0 \end{pmatrix}, \underbrace{0, \dots, 0}_{n-2s \text{ times}} \right).$$

Proof. 由 [定理 2.11](#) 立即得到. \square

5 半正定算子

内积空间 V 上的自伴随算子 φ 如果满足

$$\langle \varphi(v), v \rangle \geq 0 \quad \forall v \in V,$$

那么我们说 φ 是半正定算子. 注意, 对于酉空间而言, 由于 [定理 4.3](#), 所以上述定义是有意义的. 对于一个算子 φ , 如果存在算子 ψ 使得 $\varphi = \psi^2$, 那么我们说 ψ 是 φ 的平方根.

现在我们将上述定义翻译成矩阵的语言. 设自伴随算子 φ 在标准正交基 e_1, \dots, e_n 下的表示矩阵为 Hermite (实对称) 矩阵 P , 那么对于任意 $v \in V$, 设 v 在这组基下的坐标为 $x = (x_1, \dots, x_n) \in \mathbb{F}^n$, 那么

$$\langle \varphi(v), v \rangle = (Px)^* x = x^* P^* x = x^* Px.$$

所以对于 Hermite (实对称) 矩阵 P , 如果任意 $x \in \mathbb{F}^n$ 使得

$$x^* Px \geq 0,$$

那么我们说 P 是半正定矩阵. 同样的, 如果存在矩阵 S 使得 $P = S^2$, 那么我们说 S 是 P 的平方根.

定理 5.1 (半正定算子的刻画). 令 φ 是内积空间 V 上的算子, 那么下面的说法等价:

- (1) φ 是半正定算子;
- (2) φ 是自伴随算子并且 φ 的所有特征值非负;
- (3) φ 有半正定的平方根;
- (4) φ 有自伴随的平方根;
- (5) 存在算子 ψ 使得 $\varphi = \psi^* \psi$.

Proof. (1) \Rightarrow (2) 设 λ_0 是 φ 的特征值, [定理 4.1](#) 表明 λ_0 为实数. 任取 λ_0 的特征向量 v , 那么

$$0 \leq \langle \varphi(v), v \rangle = \lambda_0 \langle v, v \rangle,$$

由于 $\langle v, v \rangle > 0$, 所以 $\lambda_0 \geq 0$.

SECTION 5 半正定算子

(2) \Rightarrow (3) 根据谱定理, 我们知道存在 V 的一组标准正交基 e_1, \dots, e_n 使得 φ 在这组基下的表示矩阵为对角阵, 对角线为 φ 的所有特征值 $\lambda_1, \dots, \lambda_n$, 由于 $\lambda_i \geq 0$, 所以可以令 $\psi : V \rightarrow V$ 满足

$$\psi(e_i) = \sqrt{\lambda_i}e_i,$$

此时对于任意 $v \in V$, 设 v 在这组基下的坐标为 $x = (x_1, \dots, x_n) \in \mathbb{F}^n$, 那么

$$\langle \psi(v), v \rangle = x^* \operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) x = \sqrt{\lambda_1}|x_1|^2 + \dots + \sqrt{\lambda_n}|x_n|^2 \geq 0,$$

所以 ψ 是半正定算子. 此外 $\psi^2(e_i) = \lambda_i e_i = \varphi(e_i)$, 所以 $\varphi = \psi^2$.

(3) \Rightarrow (4) 显然.

(4) \Rightarrow (5) 设 $\varphi = \psi^2$, ψ 是自伴随算子, 那么 $\varphi = \psi^2 = \psi^* \psi$.

(5) \Rightarrow (1) 如果存在算子 ψ 使得 $\varphi = \psi^* \psi$, 那么 $\varphi^* = \psi^* \psi^* = \varphi$, 所以 φ 是自伴随算子. 任取 $v \in V$, 那么

$$\langle \varphi(v), v \rangle = \langle \psi^* \psi(v), v \rangle = \langle \psi(v), \psi(v) \rangle \geq 0,$$

所以 φ 是半正定算子. \square

定理 5.2. 内积空间 V 上的半正定算子的半正定平方根是唯一的.

Proof. 设 φ 是半正定算子, $\lambda_0 \geq 0$ 是 φ 的特征值, v 是 λ_0 的特征向量. 令 ψ 是 φ 的半正定平方根, 我们首先证明 ψ 必须满足 $\psi(v) = \sqrt{\lambda_0}v$. 根据谱定理, 存在 V 的一组标准正交基 e_1, \dots, e_n 使得 ψ 在这组基下的表示矩阵为对角阵, 对角线为 ψ 的特征值 μ_1, \dots, μ_n . 设 $v = x_1e_1 + \dots + x_ne_n$, 那么

$$\psi(v) = x_1\mu_1e_1 + \dots + x_n\mu_ne_n,$$

于是

$$\lambda_0v = \varphi(v) = \psi^2(v) = x_1\mu_1^2e_1 + \dots + x_n\mu_n^2e_n,$$

于是

$$x_1(\lambda_0 - \mu_1^2)e_1 + \dots + x_n(\lambda_0 - \mu_n^2)e_n = 0,$$

所以 $x_i(\lambda_0 - \mu_i^2) = 0$, 故 $x_i = 0$ 或者 $\lambda_0 = \mu_i^2$. 这表明

$$v = \sum_{\{k | \mu_k^2 = \lambda_0\}} x_k e_k,$$

所以

$$\psi(v) = \sum_{\{k | \mu_k^2 = \lambda_0\}} x_k \mu_k e_k = \sum_{\{k | \mu_k^2 = \lambda_0\}} x_k \sqrt{\lambda_0} e_k = \sqrt{\lambda_0}v.$$

SECTION 6 奇异值分解和极分解

根据谱定理, 存在 V 的一组标准正交基 f_1, \dots, f_n 使得 φ 在这组基下的表示矩阵为对角阵, 对角线为 φ 的特征值, 所以此时 f_1, \dots, f_n 为 φ 的特征向量, 而上面的叙述表明 ψ 在 f_i 上的作用是唯一确定的, 所以 ψ 是唯一的. \square

定理 5.2 表明我们可以良好的定义一个半正定算子 φ 的半正定平方根, 此时我们把这个半正定平方根记为 $\sqrt{\varphi}$.

6 奇异值分解和极分解

命题 6.1. 令 $\varphi : V \rightarrow W$ 是内积空间之间的线性映射, 那么

- (1) $\varphi^* \varphi$ 是 V 上的半正定算子;
- (2) $\ker \varphi^* \varphi = \ker \varphi$;
- (3) $\operatorname{im} \varphi^* \varphi = \operatorname{im} \varphi^*$;
- (4) $\dim \operatorname{im} \varphi = \dim \operatorname{im} \varphi^* = \dim \operatorname{im} \varphi^* \varphi$.

Proof. (1) 显然 $\varphi^* \varphi$ 是自伴随算子. 任取 $v \in V$ 有

$$\langle \varphi^* \varphi(v), v \rangle = \langle \varphi(v), \varphi(v) \rangle \geq 0,$$

所以 $\varphi^* \varphi$ 是半正定算子.

(2) 显然 $\ker \varphi \subseteq \ker \varphi^* \varphi$. 任取 $v \in \ker \varphi^* \varphi$, 那么

$$\langle \varphi(v), \varphi(v) \rangle = \langle \varphi^* \varphi(v), v \rangle = 0,$$

所以 $\varphi(v) = 0$, 所以 $\ker \varphi^* \varphi \subseteq \ker \varphi$.

(3) 我们有

$$\operatorname{im} \varphi^* \varphi = (\ker \varphi^* \varphi)^\perp = (\ker \varphi)^\perp = \operatorname{im} \varphi^*.$$

(4) 我们有

$$\dim \operatorname{im} \varphi = \dim V - \dim \ker \varphi = \dim V - \dim (\operatorname{im} \varphi^*)^\perp = \dim \operatorname{im} \varphi^*. \quad \square$$

对于线性映射 $\varphi : V \rightarrow W$, 我们将 $\varphi^* \varphi$ 的所有特征值的算数平方根从大到小排列, 称为 φ 的**奇异值**. 通过下面的命题, 我们将看到奇异值反映了映射的一些信息.

命题 6.2. $\varphi : V \rightarrow W$ 是线性映射, 那么

- (1) φ 是单射当且仅当 φ 的奇异值都非零;
- (2) φ 的正奇异值的个数等于 $\dim \operatorname{im} \varphi$, 所以 φ 是满射当且仅当 φ 的正奇异值的个数等于 $\dim W$.

SECTION 6 奇异值分解和极分解

Proof. (1) 根据 [命题 6.1](#), $\ker \varphi = \ker \varphi^* \varphi$, 所以 φ 是单射当且仅当 $\varphi^* \varphi$ 是单射, 而 $\varphi^* \varphi$ 是 V 上的线性变换, 所以 $\varphi^* \varphi$ 是单射当且仅当 $\varphi^* \varphi$ 可逆, 当且仅当 $\varphi^* \varphi$ 的特征值全不为零, 当且仅当 φ 的奇异值全不为零.

(2) 根据谱定理, $\varphi^* \varphi$ 在某组标准正交基下的表示矩阵为对角阵, 对角线为所有特征值, 再根据 [命题 6.1](#), 所以 $\dim \operatorname{im} \varphi = \dim \operatorname{im} \varphi^* \varphi$ 为 $\varphi^* \varphi$ 的非零特征值的个数(计重数), 即 φ 的正奇异值的个数. \square

定理 6.3 (线性映射的奇异值分解). 令 V 是 n 维内积空间, W 是 m 维内积空间, $\varphi : V \rightarrow W$ 是线性映射, 设 φ 的所有正奇异值为 s_1, \dots, s_k , 那么存在 V 的一个正交向量组 e_1, \dots, e_k 和 W 的一个正交向量组 f_1, \dots, f_k 使得

$$\varphi(v) = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_k \langle v, e_k \rangle f_k,$$

其中 $v \in V$.

Proof. 设 φ 的所有奇异值为 s_1, \dots, s_n , 后面 $n - k$ 个奇异值为零. 由于 $\varphi^* \varphi$ 是半正定算子, 根据谱定理, 存在 V 的标准正交基 e_1, \dots, e_n 使得 $\varphi^* \varphi$ 在这组基下的表示矩阵为对角阵

$$\operatorname{diag}(s_1^2, \dots, s_n^2).$$

对于 $1 \leq i \leq k$, 令

$$f_i = \frac{1}{s_i} \varphi(e_i).$$

那么

$$\langle f_i, f_j \rangle = \frac{1}{s_i s_j} \langle \varphi(e_i), \varphi(e_j) \rangle = \frac{1}{s_i s_j} \langle e_i, \varphi^* \varphi(e_j) \rangle = \frac{1}{s_i s_j} \langle e_i, s_j^2 e_j \rangle = \frac{s_j}{s_i} \langle e_i, e_j \rangle = \delta_{ij},$$

所以 f_1, \dots, f_k 两两正交.

任取 $v \in V$, 那么

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n,$$

所以

$$\begin{aligned} \varphi(v) &= \langle v, e_1 \rangle \varphi(e_1) + \cdots + \langle v, e_n \rangle \varphi(e_n) \\ &= s_1 \langle v, e_1 \rangle f_1 + \cdots + s_k \langle v, e_k \rangle f_k \\ &\quad + \langle v, e_{k+1} \rangle \varphi(e_{k+1}) + \cdots + \langle v, e_n \rangle \varphi(e_n), \end{aligned}$$

注意到 $k + 1 \leq i \leq n$ 的时候有 $\varphi^* \varphi(e_i) = 0$, 所以 $e_i \in \ker \varphi^* \varphi = \ker \varphi$, 即 $\varphi(e_i) = 0$, 所以

$$\varphi(v) = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_k \langle v, e_k \rangle f_k. \quad \square$$

SECTION 6 奇异值分解和极分解

定理 6.3 会带来一些有意思的结果. 我们将 e_1, \dots, e_k 扩充为 V 的标准正交基 e_1, \dots, e_n , 将 f_1, \dots, f_k 扩充为 W 的标准正交基 f_1, \dots, f_m , 那么 $1 \leq i \leq k$ 时,

$$\varphi(e_i) = s_i f_i,$$

$k+1 \leq i \leq n$ 时,

$$\varphi(e_i) = 0.$$

这表明 φ 在这两组基下的表示矩阵为一个 $m \times n$ 的“对角阵”, 对角线依次为 φ 的非零奇异值. 特别地, 如果 $W = V$, 那么这表明对于任意算子 φ , 总能选取两组不同的标准正交基, 使得 φ 在这两组正交基下的表示矩阵是对角阵.

我们定义 $m \times n$ 阶对角阵为只有 a_{ii} ($1 \leq i \leq \min\{m, n\}$) 不为零的矩阵.

定理 6.4 (矩阵的奇异值分解). 设 A 是 \mathbb{F} 上的 $m \times n$ 矩阵, 那么存在 m 阶酉(正交)矩阵 P , n 阶酉(正交)矩阵 Q 以及 $m \times n$ 的对角矩阵 D , 使得

$$A = PDQ^*,$$

其中 D 的对角线非零元素依次为 A 的奇异值.

Proof. 令 $\varphi : \mathbb{F}^n \rightarrow \mathbb{F}^m$ 为矩阵 A 相对于标准基表示的线性映射. 设 φ 的奇异值为 s_1, \dots, s_n , 前 k 个不为零. 根据 **定理 6.3**, 存在 \mathbb{F}^n 的正交向量组 e_1, \dots, e_k 和 \mathbb{F}^m 的正交向量组 f_1, \dots, f_k 使得

$$\varphi(v) = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_k \langle v, e_k \rangle f_k.$$

将 e_1, \dots, e_k 扩充为 \mathbb{F}^n 的标准正交基 e_1, \dots, e_n , f_1, \dots, f_k 扩充为 \mathbb{F}^m 的标准正交基 f_1, \dots, f_m . 令 Q 为 e_1, \dots, e_n 按列拼成的矩阵, P 为 f_1, \dots, f_m 拼成的矩阵. 此时 P, Q 为酉(正交)矩阵. 设 u_1, \dots, u_n 为 \mathbb{F}^n 的标准基, 那么 $1 \leq i \leq k$ 时有

$$(AQ - PD)u_i = Ae_i - Ps_i u_i = s_i f_i - s_i P u_i = s_i f_i - s_i f_i = 0,$$

当 $k+1 \leq i \leq n$ 时有

$$(AQ - PD)u_i = Ae_i - 0 = 0 - 0 = 0,$$

所以 $AQ = PD$, 故 $A = PDQ^*$. \square

定理 6.5 (线性映射的极分解). 对于内积空间 V 上的算子 φ , 存在酉(正交)算子 σ 使得

$$\varphi = \sigma \sqrt{\varphi^* \varphi},$$

显然 φ 可逆当且仅当 σ 是唯一确定的.

SECTION 6 奇异值分解和极分解

Proof. 设 s_1, \dots, s_k 是 φ 的正奇异值, 根据奇异值分解, 存在 V 的正交向量组 e_1, \dots, e_k 和 f_1, \dots, f_k 使得

$$\varphi(v) = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_k \langle v, e_k \rangle f_k \quad \forall v \in V,$$

此时

$$\begin{aligned} \langle v, \varphi^*(u) \rangle &= \langle \varphi(v), u \rangle = s_1 \langle v, e_1 \rangle \langle f_1, u \rangle + \cdots + s_k \langle v, e_k \rangle \langle f_k, u \rangle \\ &= \langle v, s_1 \langle u, f_1 \rangle e_1 + \cdots + s_k \langle u, f_k \rangle e_k \rangle, \end{aligned}$$

所以

$$\varphi^*(u) = s_1 \langle u, f_1 \rangle e_1 + \cdots + s_k \langle u, f_k \rangle e_k \quad \forall u \in V.$$

于是对于任意的 $v \in V$, 有

$$\begin{aligned} \varphi^* \varphi(v) &= s_1 \langle v, e_1 \rangle \varphi^*(f_1) + \cdots + s_k \langle v, e_k \rangle \varphi^*(f_k) \\ &= s_1^2 \langle v, e_1 \rangle e_1 + \cdots + s_k^2 \langle v, e_k \rangle e_k, \end{aligned}$$

此时

$$\sqrt{\varphi^* \varphi}(v) = s_1 \langle v, e_1 \rangle e_1 + \cdots + s_k \langle v, e_k \rangle e_k,$$

直接验证可知这确实是 $\varphi^* \varphi$ 的半正定平方根. 这表明此时 $\sqrt{\varphi^* \varphi}$ 将基 e_i 送到 $s_i e_i$, 所以自然就会构造 σ 将 e_i 再送到 f_i .

将 e_1, \dots, e_k 和 f_1, \dots, f_k 扩充为标准正交基 e_1, \dots, e_n 和 f_1, \dots, f_n . 定义算子 σ 为

$$\sigma(v) = \langle v, e_1 \rangle f_1 + \cdots + \langle v, e_n \rangle f_n,$$

那么

$$\|\sigma(v)\|^2 = \langle v, e_1 \rangle^2 + \cdots + \langle v, e_n \rangle^2 = \|v\|^2,$$

所以 σ 是酉 (正交) 算子.

于是

$$\sigma \sqrt{\varphi^* \varphi}(v) = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_k \langle v, e_k \rangle f_k = \varphi(v),$$

所以 $\varphi = \sigma \sqrt{\varphi^* \varphi}$ 即为所求. \square

定理 6.6 (矩阵的极分解). 设 A 是 \mathbb{F} 上的 n 阶矩阵, 那么存在酉 (正交) 矩阵 S 使得

$$A = S \sqrt{A^* A},$$

A 可逆当且仅当这样的 P 是唯一的. 此外, 在实际计算过程中, 我们一般不会直接计算 $\sqrt{A^* A}$, 而是先计算 A 的奇异值分解, 得到 $A = PDQ^*$, 其中 P, Q 为酉 (正交) 矩阵, 然后得到 $A = (PQ^*)(QDQ^*)$, 此时可以验证奇异值分解得到的 $QDQ^* = \sqrt{A^* A}$.

SECTION 6 奇异值分解和极分解

Proof. 令 $\varphi : \mathbb{F}^n \rightarrow \mathbb{F}^m$ 为矩阵 A 相对于标准基表示的线性映射. 根据线性映射的极分解, 存在酉(正交) 算子 σ 使得

$$\varphi = \sigma \sqrt{\varphi^* \varphi},$$

比对两端在标准基下的表示矩阵即得 $A = S \sqrt{A^* A}$.

设 φ 的奇异值为 s_1, \dots, s_n , 前 k 个不为零. 根据 定理 6.3, 存在 \mathbb{F}^n 的正交向量组 e_1, \dots, e_k 和 \mathbb{F}^m 的正交向量组 f_1, \dots, f_k 使得

$$\varphi(v) = s_1 \langle v, e_1 \rangle f_1 + \dots + s_k \langle v, e_k \rangle f_k.$$

将 e_1, \dots, e_k 扩充为 \mathbb{F}^n 的标准正交基 e_1, \dots, e_n , f_1, \dots, f_k 扩充为 \mathbb{F}^m 的标准正交基 f_1, \dots, f_m . 令 Q 为 e_1, \dots, e_n 按列拼成的矩阵, P 为 f_1, \dots, f_m 拼成的矩阵. 根据 定理 6.4 的证明, 此时 $A = PDQ^*$ 就是 A 的奇异值分解. 根据 定理 6.5 的证明, 我们知道 $\sqrt{\varphi^* \varphi}$ 将 e_i 送到 $s_i e_i$. 所以 $A = PDQ^* = (PQ^*)(QDQ^*)$ 时, 有

$$QDQ^*e_i = QDu_i = s_i Qu_i = s_i e_i,$$

其中 u_i 表示 \mathbb{F}^n 的标准基. 所以 $\sqrt{\varphi^* \varphi}$ 在标准基下的表示矩阵就是 $QDQ^* = \sqrt{A^* A}$. \square

例 6.7. 计算实矩阵

$$A = \begin{pmatrix} 1 & 2 & 5 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

的奇异值分解和极分解.

Solution. 计算得

$$A^T A = \begin{pmatrix} 2 & 2 & 6 \\ 2 & 5 & 12 \\ 6 & 12 & 30 \end{pmatrix},$$

故 $A^T A$ 的特征多项式为 $-x(x - 1)(x - 36)$, 所以 A 的奇异值为 $6, 1, 0$. 仿照 例 4.7, 可得

$$Q = \begin{pmatrix} \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{30}} & -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{6}} \\ \frac{5}{\sqrt{30}} & 0 & -\frac{1}{\sqrt{6}} \end{pmatrix}, \quad Q^{-1}(A^T A)Q = \begin{pmatrix} 36 & & \\ & 1 & \\ & & 0 \end{pmatrix}.$$

记 Q 的三个列向量为 e_1, e_2, e_3 . 根据 定理 6.3 的证明, 我们计算

$$\begin{aligned} f_1 &= \frac{1}{s_1} Ae_1 = \frac{1}{\sqrt{30}}(5, 1, 2)^T, \\ f_2 &= \frac{1}{s_2} Ae_2 = \frac{1}{\sqrt{5}}(0, 2, -1)^T, \end{aligned}$$

SECTION 7 MOORE-PENROSE 广义逆和最小二乘

然后将 f_1, f_2 扩充为 \mathbb{R}^3 的一组标准正交基, 可添加 $f_3 = \frac{1}{\sqrt{6}}(-1, 1, 2)^T$, 故得到 A 的奇异值分解

$$P = \begin{pmatrix} \frac{5}{\sqrt{30}} & 0 & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{30}} & -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{6}} \end{pmatrix}, \quad A = P \begin{pmatrix} 6 & & \\ & 1 & \\ & & 0 \end{pmatrix} Q^T.$$

对于极分解, 只需要计算

$$PQ^T = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad QDQ^T = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 5 \end{pmatrix},$$

故 A 的极分解为

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 5 \end{pmatrix}.$$

此时可以验证

$$QDQ^T = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 5 \end{pmatrix} = \sqrt{A^T A}. \quad \blacksquare$$

7 Moore-Penrose 广义逆和最小二乘

设 $\varphi : V \rightarrow W$ 是内积空间之间的线性映射, $w \in W$, 考虑方程

$$\varphi(v) = w,$$

只有 $w \in \text{im } \varphi$ 时才会有解, 此时解集为陪集 $v + \ker \varphi$. 那么在无解或者解不唯一的时候, 能否找到 $v \in V$ 使得 $\|\varphi(v) - w\|$ 尽可能小? (本节的范数均指代内积诱导的范数) 这便是线性方程组的最小二乘问题. 而 Moore-Penrose 广义逆提供了解决此问题的一个优美的视角.

根据同态基本定理, 我们知道有同构

$$V/\ker \varphi \simeq \text{im } \varphi,$$

注意到

$$\dim(\ker \varphi)^\perp = \dim V - \dim \ker \varphi = \dim V/\ker \varphi,$$

所以 $(\ker \varphi)^\perp \simeq V/\ker \varphi$, 同构映射显然为 $v \mapsto v + \ker \varphi$, 实际上就是商同态 $V \rightarrow V/\ker \varphi$ 在 $(\ker \varphi)^\perp$ 上的限制. 于是我们有同构

$$(\ker \varphi)^\perp \simeq V/\ker \varphi \simeq \text{im } \varphi,$$

SECTION 7 MOORE-PENROSE 广义逆和最小二乘

我们把这个同构记为 $\varphi|_{(\ker \varphi)^\perp} : (\ker \varphi)^\perp \rightarrow \text{im } \varphi$, 这是可逆的. 自然我们就会考虑 $\varphi|_{(\ker \varphi)^\perp}^{-1} : \text{im } \varphi \rightarrow (\ker \varphi)^\perp \rightarrow V$, 如果能将定义域扩大到整个 W , 直觉上便有理由将其称为 φ 的“广义逆”.

对于 $w \in W$, 由于 $W = \text{im } \varphi \oplus (\text{im } \varphi)^\perp$, 所以 $w = w_1 + w_2$, 其中 $w_1 \in \text{im } \varphi, w_2 \in (\text{im } \varphi)^\perp$, 于是我们只需要考虑正交投影 $P_{\text{im } \varphi} : W \rightarrow W$ 即可将 w 送到 $P(w) = w_1 \in \text{im } \varphi$. 于是我们定义 $\varphi^\dagger = \varphi|_{(\ker \varphi)^\perp}^{-1} \circ P_{\text{im } \varphi}$, 其满足

$$W \rightarrow \text{im } \varphi \rightarrow (\ker \varphi)^\perp \rightarrow V, \quad w \mapsto w_1 \mapsto \varphi|_{(\ker \varphi)^\perp}^{-1}(w_1).$$

这个映射 φ^\dagger 被称为 φ 的 **Moore-Penrose 广义逆**, 也被称为**伪逆**. 显然

$$\ker \varphi^\dagger = (\text{im } \varphi)^\perp, \quad \text{im } \varphi^\dagger = (\ker \varphi)^\perp.$$

回顾 **命题 1.6**, 这表明 φ^\dagger 和 φ^* 的核和像都是相同的.

命题 7.1. 令 $\varphi : V \rightarrow W$ 是线性映射, 那么

- (1) $\varphi \varphi^\dagger = P_{\text{im } \varphi}, P_{\text{im } \varphi}$ 表示 W 到 $\text{im } \varphi$ 的正交投影;
- (2) $\varphi^\dagger \varphi = P_{(\ker \varphi)^\perp}, P_{(\ker \varphi)^\perp}$ 表示 V 到 $(\ker \varphi)^\perp$ 的正交投影;
- (3) 若 φ 可逆, 那么 $\varphi^\dagger = \varphi^{-1}$;
- (4) $(\varphi^\dagger)^\dagger = \varphi$.

Proof. (1) 根据定义显然.

(2) 任取 $v \in V$, 那么

$$\varphi^\dagger \varphi(v) = \varphi|_{(\ker \varphi)^\perp}^{-1} P_{\text{im } \varphi} \varphi(v) = \varphi|_{(\ker \varphi)^\perp}^{-1} \varphi(v) = P_{(\ker \varphi)^\perp}(v),$$

所以 $\varphi^\dagger \varphi = P_{(\ker \varphi)^\perp}$.

(3) φ 可逆表明 $(\ker \varphi)^\perp = V$ 以及 $\text{im } \varphi = W$, 由 (1) 和 (2) 即得

$$\varphi \varphi^\dagger = 1_W, \varphi^\dagger \varphi = 1_V,$$

即 $\varphi^\dagger = \varphi^{-1}$.

(4) 注意到 $V = (\ker \varphi) \oplus (\ker \varphi)^\perp$. 任取 $v \in \ker \varphi$, 那么

$$(\varphi^\dagger)^\dagger(v) = (\varphi^\dagger)|_{\text{im } \varphi}^{-1} P_{(\ker \varphi)^\perp}(v) = 0 = \varphi(v).$$

任取 $u \in (\ker \varphi)^\perp$, 那么

$$(\varphi^\dagger)^\dagger(u) = (\varphi^\dagger)|_{\text{im } \varphi}^{-1} P_{(\ker \varphi)^\perp}(u) = (\varphi^\dagger)|_{\text{im } \varphi}^{-1}(u),$$

而 $(\varphi^\dagger)^{-1}(u) = \varphi(u) + \ker \varphi^\dagger = \varphi(u) + (\text{im } \varphi)^\perp$, 又因为 $\varphi(u) \in \text{im } \varphi$, 所以

$$(\varphi^\dagger)^\dagger(u) = (\varphi^\dagger)|_{\text{im } \varphi}^{-1}(u) = \varphi(u).$$

故 $(\varphi^\dagger)^\dagger = \varphi$. \square

定理 7.2 (广义逆的刻画). 令 $\varphi : V \rightarrow W$ 是线性映射, 那么广义逆 φ^\dagger 由满足下列四条性质, 并且由这四条性质唯一确定.

- (1) $\varphi\varphi^\dagger\varphi = \varphi$;
- (2) $\varphi^\dagger\varphi\varphi^\dagger = \varphi^\dagger$;
- (3) $\varphi\varphi^\dagger$ 是自伴随算子;
- (4) $\varphi^\dagger\varphi$ 是自伴随算子.

Proof. 首先我们说明广义逆满足这四条性质. (1) 和 (2) 根据 命题 7.1 即得. 对于 (3), 根据 命题 7.1, 任取 $w \in W$, 有

$$\varphi\varphi^\dagger(w) = P_{\text{im } \varphi}(w),$$

若 $w \in \text{im } \varphi$, 那么

$$\begin{aligned} \langle (\varphi\varphi^\dagger)^*(w), u \rangle &= \langle P_{\text{im } \varphi}^*(w), u \rangle = \langle w, P_{\text{im } \varphi}(u) \rangle = \langle P_{\text{im } \varphi}(w), u + P_{\text{im } \varphi}(u) - u \rangle \\ &= \langle P_{\text{im } \varphi}(w), u \rangle + \langle P_{\text{im } \varphi}(w), P_{\text{im } \varphi}(u) - u \rangle = \langle P_{\text{im } \varphi}(w), u \rangle, \end{aligned}$$

最后一个等号是因为 $P_{\text{im } \varphi}(u) - u \in (\text{im } \varphi)^\perp$. 这表明

$$(\varphi\varphi^\dagger)^*(w) = P_{\text{im } \varphi}(w) = (\varphi\varphi^\dagger)(w).$$

若 $w \in (\text{im } \varphi)^\perp$, 由于

$$\ker P_{\text{im } \varphi}^* = (\text{im } P_{\text{im } \varphi})^\perp = (\text{im } \varphi)^\perp,$$

所以

$$P_{\text{im } \varphi}^*(w) = 0 = P_{\text{im } \varphi}(w),$$

故此时 $(\varphi\varphi^\dagger)^*(w) = (\varphi\varphi^\dagger)(w) = 0$. 由于 $W = (\text{im } \varphi) \oplus (\text{im } \varphi)^\perp$, 所以 $(\varphi\varphi^\dagger)^* = \varphi\varphi^\dagger$. (4) 同理.

下面我们说明唯一性. 假设 σ 和 τ 都满足上面的四条性质, 那么

$$\begin{aligned} \sigma &= \sigma\varphi\sigma = (\sigma\varphi)^*\sigma = \varphi^*\sigma^*\sigma = (\varphi\tau\varphi)^*\sigma^*\sigma = \varphi^*\tau^*\varphi^*\sigma^*\sigma \\ &= (\tau\varphi)^*(\sigma\varphi)^*\sigma = \tau\varphi\sigma\varphi\sigma = \tau\varphi\sigma, \end{aligned}$$

另一方面, 有

$$\begin{aligned} \tau &= \tau\varphi\tau = \tau(\varphi\tau)^* = \tau\tau^*\varphi^* = \tau\tau^*(\varphi\sigma\varphi)^* = \tau\tau^*\varphi^*\sigma^*\varphi^* \\ &= \tau(\varphi\tau)^*(\varphi\sigma)^* = \tau\varphi\tau\varphi\sigma = \tau\varphi\sigma, \end{aligned}$$

所以 $\sigma = \tau$. □

定理 7.3 (最小二乘解). $\varphi : V \rightarrow W$ 是线性映射, 给定 $w \in W$, 那么

(1) 任取 $v \in V$, 有

$$\|\varphi(\varphi^\dagger(w)) - w\| \leq \|\varphi(v) - w\|,$$

等号成立当且仅当 $v \in \varphi^\dagger(w) + \ker \varphi$.

(2) 如果 $v \in \varphi^\dagger(w) + \ker \varphi$, 那么

$$\|\varphi^\dagger(w)\| \leq \|v\|,$$

等号成立当且仅当 $v = \varphi^\dagger(w)$.

这些性质表明 $\varphi^\dagger(w)$ 就是方程 $\varphi(v) = w$ 的最小二乘解.

Proof. (1) 我们有

$$\varphi(v) - w = (\varphi(v) - \varphi\varphi^\dagger(w)) + (\varphi\varphi^\dagger(w) - w),$$

第一项 $\varphi(v) - \varphi\varphi^\dagger(w) \in \text{im } \varphi$. 第二项

$$\varphi\varphi^\dagger(w) - w = P_{\text{im } \varphi}(w) - w \in (\text{im } \varphi)^\perp,$$

根据勾股定理, 就得到了

$$\|\varphi(v) - w\|^2 = \|\varphi(v) - \varphi\varphi^\dagger(w)\|^2 + \|\varphi\varphi^\dagger(w) - w\|^2 \geq \|\varphi\varphi^\dagger(w) - w\|^2.$$

等号成立当且仅当 $\varphi(v) = \varphi\varphi^\dagger(w)$, 当且仅当 $v - \varphi^\dagger(w) \in \ker \varphi$.

(2) 由于

$$v = (v - \varphi^\dagger(w)) + \varphi^\dagger(w),$$

其中 $v - \varphi^\dagger(w) \in \ker \varphi$, $\varphi^\dagger(w) \in (\ker \varphi)^\perp$, 所以

$$\|v\|^2 = \|v - \varphi^\dagger(w)\|^2 + \|\varphi^\dagger(w)\|^2 \geq \|\varphi^\dagger(w)\|^2.$$

等号成立当且仅当 $v = \varphi^\dagger(w)$. □

下面我们研究广义逆的具体表达形式.

定理 7.4. 令 V 是 n 维内积空间, W 是 m 维内积空间, $\varphi : V \rightarrow W$ 是线性映射, 设 φ 的所有正奇异值为 s_1, \dots, s_k , 根据奇异值分解 6.3, 存在 V 的一个正交向量组 e_1, \dots, e_k 和 W 的一个正交向量组 f_1, \dots, f_k 使得

$$\varphi(v) = s_1 \langle v, e_1 \rangle f_1 + \dots + s_k \langle v, e_k \rangle f_k,$$

其中 $v \in V$. 那么

$$\varphi^\dagger(w) = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \dots + \frac{\langle w, f_k \rangle}{s_k} e_k,$$

其中 $w \in W$.

SECTION 7 MOORE-PENROSE 广义逆和最小二乘

Proof. 奇异值分解告诉我们 $\text{im } \varphi = \text{span}(f_1, \dots, f_k)$ 以及 $\ker \varphi = \text{span}(e_1, \dots, e_k)^\perp$. 故

$$P_{\text{im } \varphi}(w) = \langle w, f_1 \rangle f_1 + \dots + \langle w, f_k \rangle f_k,$$

又因为 $\varphi(e_i) = s_i f_i$, 所以

$$\begin{aligned} P_{\text{im } \varphi}(w) &= \frac{\langle w, f_1 \rangle}{s_1} \varphi(e_1) + \dots + \frac{\langle w, f_k \rangle}{s_k} \varphi(e_k) \\ &= \varphi \left(\frac{\langle w, f_1 \rangle}{s_1} e_1 + \dots + \frac{\langle w, f_k \rangle}{s_k} e_k \right), \end{aligned}$$

故

$$\varphi^{-1}(P_{\text{im } \varphi}(v)) = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \dots + \frac{\langle w, f_k \rangle}{s_k} e_k + \ker \varphi \in (\ker \varphi)^\perp + \ker \varphi,$$

所以

$$\varphi^\dagger(w) = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \dots + \frac{\langle w, f_k \rangle}{s_k} e_k. \quad \square$$

定理 7.5 (矩阵的广义逆). 矩阵 $A \in M_{m,n}(\mathbb{F})$, 设 A 的奇异值分解为 $A = PDQ^*$, 其中 $D \in M_{m,n}(\mathbb{F})$ 且对角线依次为 A 的非零奇异值 s_1, \dots, s_k . 那么

$$A^\dagger = QD^\dagger P^* \in M_{n,m}(\mathbb{F}),$$

其中 $D^\dagger \in M_{n,m}(\mathbb{F})$ 是 D 广义逆, 容易发现其对角线依次为 $1/s_1, \dots, 1/s_k$.

Proof. 直接验证可知其满足 [定理 7.2](#) 的四条性质, 根据唯一性, 所以这就是 A 的广义逆. 我们也可以从 [定理 7.4](#) 直接构造 A^\dagger .

令 $\varphi : \mathbb{F}^n \rightarrow \mathbb{F}^m$ 为矩阵 A 相对于标准基表示的线性映射. 设 φ 的奇异值为 s_1, \dots, s_n , 前 k 个不为零. 根据 [定理 7.4](#), 存在 \mathbb{F}^n 的正交向量组 e_1, \dots, e_k 和 \mathbb{F}^m 的正交向量组 f_1, \dots, f_k 使得

$$\varphi^\dagger(w) = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \dots + \frac{\langle w, f_k \rangle}{s_k} e_k,$$

将 e_1, \dots, e_k 扩充为 \mathbb{F}^n 的标准正交基 e_1, \dots, e_n , f_1, \dots, f_k 扩充为 \mathbb{F}^m 的标准正交基 f_1, \dots, f_m . 令 Q 为 e_1, \dots, e_n 按列拼成的矩阵, P 为 f_1, \dots, f_m 拼成的矩阵. 设 u_1, \dots, u_m 为 \mathbb{F}^m 的标准基, 那么 $1 \leq i \leq k$ 时有

$$(A^\dagger P - QD^\dagger)(u_i) = A^\dagger f_i - Qu_i/s_i = e_i/s_i - e_i/s_i = 0,$$

$k+1 \leq i \leq m$ 时有

$$(A^\dagger P - QD^\dagger)(u_i) = A^\dagger f_i - 0 = 0 - 0 = 0,$$

所以 $A^\dagger = QD^\dagger P^*$. \square

最后我们介绍广义逆的另一种重要表达形式.

引理 7.6. 对于任意 n 阶 Hermite (实对称) 矩阵 A , 有

$$P_{\text{im } A} = \lim_{\epsilon \rightarrow 0} (A + \epsilon I_n)^{-1} A = \lim_{\epsilon \rightarrow 0} A (A + \epsilon I_n)^{-1}.$$

Proof. 我们知道 A 的特征值都是实数, 设 A 绝对值最小的非零特征值为 λ , 那么 $0 < |\epsilon| < |\lambda|$ 的时候, $A + \epsilon I_n$ 的特征值非零, 所以此时 $A + \epsilon I_n$ 可逆. 任取 $x \in \mathbb{F}^n$, 设 $x = x_1 + x_2$, 其中 $x_1 \in \text{im } A$, $x_2 \in (\text{im } A)^\perp$. A 正规表明 $\ker A = (\text{im } A)^\perp$, 所以 $x_2 \in \ker A$, 所以 $Ax = Ax_1$. 设 $x_1 = Ay$, 那么

$$(A + \epsilon I_n)^{-1} Ax = (A + \epsilon I_n)^{-1} A^2 y,$$

根据谱定理, 存在正交矩阵 P 使得 $A = PDP^{-1}$, 其中 $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ 是 A 的全部特征值. 所以

$$(A + \epsilon I_n)^{-1} Ax = (A + \epsilon I_n)^{-1} A^2 y = P(D + \epsilon I)^{-1} D^2 P^{-1} y,$$

因此

$$\lim_{\epsilon \rightarrow 0} (A + \epsilon I_n)^{-1} Ax = \lim_{\epsilon \rightarrow 0} P(D + \epsilon I_n)^{-1} D^2 P^{-1} y = PDP^{-1} y = Ay = x_1 = P_{\text{im } A} x,$$

这就表明

$$P_{\text{im } A} = \lim_{\epsilon \rightarrow 0} (A + \epsilon I_n)^{-1} A.$$

另一边同理. \square

定理 7.7. 对于 $A \in M_{m,n}(\mathbb{F})$, 有

$$A^\dagger = \lim_{\epsilon \rightarrow 0} (A^* A + \epsilon^2 I_m)^{-1} A^* = \lim_{\epsilon \rightarrow 0} A^* (A A^* + \epsilon^2 I_m)^{-1}.$$

Proof. 由于

$$A^* (A A^* + \epsilon^2 I_m) = A^* A A^* + \epsilon^2 A^* = (A^* A + \epsilon^2 I_n) A^*,$$

并且 $A A^*$ 和 $A^* A$ 都是半正定矩阵, 所以 $A A^* + \epsilon^2 I_m$ 和 $A^* A + \epsilon^2 I_n$ 一定可逆, 所以

$$(A^* A + \epsilon^2 I_n)^{-1} A^* = A^* (A A^* + \epsilon^2 I_m)^{-1},$$

故若极限存在则二者一定相等.

任取 $x \in \mathbb{F}^m$, 那么 $x = x_1 + x_2$, 其中 $x_1 \in \text{im } A$, $x_2 \in (\text{im } A)^\perp = \ker A^*$, 所以 $A^* x = A^* x_1$, 设 $x_1 = Ay$, 所以

$$(A^* A + \epsilon^2 I_n)^{-1} A^* x = (A^* A + \epsilon^2 I_n)^{-1} A^* A y,$$

根据引理 7.6,

$$P_{\text{im } A^* A} = \lim_{\epsilon \rightarrow 0} (A^* A + \epsilon^2 I_n)^{-1} A^* A,$$

又因为 $\text{im } A^* A = \text{im } A^* = (\ker A)^\perp$, 所以

$$P_{(\ker A)^\perp} y = \lim_{\epsilon \rightarrow 0} (A^* A + \epsilon^2 I_n)^{-1} A^* x \in (\ker A)^\perp,$$

所以

$$AA^\dagger x = P_{\text{im } A} x = x_1 = Ay = AP_{(\ker A)^\perp} y,$$

这就表明

$$A^\dagger x = P_{(\ker A)^\perp} y = \lim_{\epsilon \rightarrow 0} (A^* A + \epsilon^2 I_n)^{-1} A^* x,$$

即

$$A^\dagger = \lim_{\epsilon \rightarrow 0} (A^* A + \epsilon^2 I_n)^{-1} A^*. \quad \square$$

推论 7.8. 给定矩阵 $A \in M_{m,n}(\mathbb{F})$.

(1) 若 A 列满秩, 即 $\text{rank } A = n \leq m$, 那么

$$A^\dagger = (A^* A)^{-1} A^*.$$

(2) 若 A 行满秩, 即 $\text{rank } A = m \leq n$, 那么

$$A^\dagger = A^* (A A^*)^{-1}.$$

Proof. (1) A 列满秩表明 A 作为 $\mathbb{F}^n \rightarrow \mathbb{F}^m$ 的线性映射是单射, 故 $\ker A^* A = \ker A = 0$, 所以 $A^* A$ 可逆, 所以

$$A^\dagger = \lim_{\epsilon \rightarrow 0} (A^* A + \epsilon^2 I_n)^{-1} A^* = (A^* A)^{-1} A^*.$$

当然, 也可以直接利用 定理 7.2 来证明 $(A^* A)^{-1} A^*$ 是 A 的广义逆.

(2) 同理. \square

推论 7.9. 给定矩阵 $A \in M_{m,n}(\mathbb{F})$, 那么

$$(A^*)^\dagger = (A^\dagger)^*.$$

Proof. 注意到 $AA^* + \epsilon^2 I_m$ 是 Hermite 矩阵, 所以

$$(A^*)^\dagger = \lim_{\epsilon \rightarrow 0} (AA^* + \epsilon^2 I_m)^{-1} A = \lim_{\epsilon \rightarrow 0} (A^* (AA^* + \epsilon^2 I_m)^{-1})^* = (A^\dagger)^*. \quad \square$$

8 QR 分解和 Cholesky 分解

我们首先回顾 Gram-Schmidt 正交化方法.

定理 8.1 (Gram-Schmidt 正交化). 设 e_1, \dots, e_n 是内积空间 V 的一组基, 那么存在两两正交的非零向量 f_1, \dots, f_n 使得对于 $1 \leq k \leq n$, 有

$$\text{span}(f_1, \dots, f_k) = \text{span}(e_1, \dots, e_k).$$

Proof. 取 $f_1 = e_1$. 假设已经选好 f_1, \dots, f_{k-1} , 由于我们希望 $f_k \in \text{span}(e_1, \dots, e_k)$ 并且 $f_k \notin \text{span}(e_1, \dots, e_{k-1})$, 所以待到系数

$$f_k = e_k + c_{k-1}f_{k-1} + \dots + c_1f_1,$$

对于 $1 \leq i \leq k-1$, 令 $\langle f_k, f_i \rangle = 0$, 即

$$\langle e_k, f_i \rangle + c_i \langle f_i, f_i \rangle = 0,$$

解得

$$c_i = -\frac{\langle e_k, f_i \rangle}{\|f_i\|^2}.$$

所以

$$f_k = e_k - \frac{\langle e_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1} - \dots - \frac{\langle e_k, f_1 \rangle}{\|f_1\|^2} f_1$$

即满足条件. \square

定理 8.2 (QR 分解). 设 A 是 \mathbb{F} 上的可逆矩阵, 那么存在唯一的酉(正交)矩阵 Q 和对角线为正数的上三角矩阵 R 使得

$$A = QR.$$

Proof. 令 $v_1, \dots, v_n \in \mathbb{F}^n$ 为 A 的列向量组, A 可逆表明 v_1, \dots, v_n 组成 \mathbb{F}^n 的一组基. 采用 Gram-Schmidt 正交化, 得到 \mathbb{F}^n 的一组正交基 f_1, \dots, f_n 使得

$$\text{span}(f_1, \dots, f_k) = \text{span}(v_1, \dots, v_k) \quad \forall 1 \leq k \leq n.$$

其中

$$f_k = v_k - \frac{\langle v_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1} - \dots - \frac{\langle v_k, f_1 \rangle}{\|f_1\|^2} f_1,$$

即

$$v_k = f_k + \frac{\langle v_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1} + \dots + \frac{\langle v_k, f_1 \rangle}{\|f_1\|^2} f_1.$$

写成矩阵的形式, 即

$$(v_1 \ v_2 \ \cdots \ v_n) = \left(\frac{f_1}{\|f_1\|} \ \frac{f_2}{\|f_2\|} \ \cdots \ \frac{f_n}{\|f_n\|} \right) \begin{pmatrix} \|f_1\| & \frac{\langle v_2, f_1 \rangle}{\|f_1\|} & \cdots & \frac{\langle v_n, f_1 \rangle}{\|f_1\|} \\ & \|f_2\| & \cdots & \frac{\langle v_n, f_2 \rangle}{\|f_2\|} \\ & & \ddots & \vdots \\ & & & \|f_n\| \end{pmatrix},$$

记 $e_i = f_i / \|f_i\|$, 那么 e_1, \dots, e_n 是 V 的标准正交基, 此时 $\|f_i\| = \langle v_i, e_i \rangle$, 所以

$$(v_1 \ v_2 \ \cdots \ v_n) = (e_1 \ e_2 \ \cdots \ e_n) \begin{pmatrix} \langle v_1, e_1 \rangle & \langle v_2, e_1 \rangle & \cdots & \langle v_n, e_1 \rangle \\ & \langle v_2, e_2 \rangle & \cdots & \langle v_n, e_2 \rangle \\ & & \ddots & \vdots \\ & & & \langle v_n, e_n \rangle \end{pmatrix},$$

上式右端即为所求的 Q 和 R .

假设另一对酉 (正交) 矩阵 Q' 和对角线为正数的上三角矩阵 R' 也满足 $A = Q'R'$. 那么 $Q'R' = QR$, 由于 R 可逆, 所以 $R'R^{-1} = Q'^{-1}Q$, 此时 $R'R^{-1}$ 仍然是对角线为正数的上三角矩阵, $Q'^{-1}Q$ 仍然是酉 (正交) 矩阵. 对于一个对角线为正数的上三角矩阵

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix},$$

其同时为正交矩阵表明行向量和列向量的模长均为 1, 故 $a_{11} = 1$, 进而 $a_{12} = \cdots = a_{1n} = 0$, 进而 $a_{22} = 1$, 进而 $a_{23} = \cdots = a_{2n} = 0$, 以此类推, 所以这样的矩阵必为单位阵. 故 $R'R^{-1} = I_n$, 即 $R' = R$, 进而 $Q' = Q$. \square

定理 8.3 (Cholesky 分解). 设 A 是 \mathbb{F} 上的正定矩阵, 那么存在唯一的对角线全为正数的上三角矩阵 R 使得

$$A = R^*R.$$

Proof. 根据 [定理 5.1](#), 存在矩阵 B 使得 $A = B^*B$, A 可逆表明 B 也可逆. 根据 QR 分解, 存在酉 (正交) 矩阵 Q 和对角线为正数的上三角矩阵 R 使得 $B = QR$, 所以

$$A = B^*B = R^*Q^*QR = R^*R,$$

这就是我们想要的.

假设对角线为正数的上三角矩阵 S 也使得 $A = S^*S$, 那么 $R^*R = S^*S$, 即 $(S^{-1})^*R^*RS^{-1} = I_n$, 即 $(RS^{-1})^*(RS^{-1}) = I_n$, 所以 RS^{-1} 是酉 (正交) 矩阵. 但 RS^{-1} 仍为对角线为正数的上三角矩阵, 所以 $RS^{-1} = I_n$, 即 $R = S$. \square

参考文献

- [1] Axler S. Linear Algebra Done Right. Springer Nature; 2024.
- [2] Albert A. Regression and the Moore-Penrose Pseudoinverse.
- [3] Roman S, Axler S, Gehring FW. Advanced Linear Algebra. New York: Springer; 2005 Mar 22.
- [4] 李炯生, 查建国. 线性代数. 中国科学技术大学出版社; 1989.
- [5] 姚慕生, 吴泉水, 谢启鸿. 高等代数学. 复旦大学出版社; 2014.