

## 等距映射,酉矩阵和矩阵分解

矩阵的分解是算法中的重要的一环.我们将开始介绍矩阵分解的相关内容.

### 1.等距映射

保持范数的映射十分重要.我们为其专门命名.

#### 1.1 定义:等距映射

线性映射  $S \in \mathcal{L}(V, W)$  被称为等距映射,如果  $\|Sv\| = \|v\|$  对于任意  $v \in V$  成立.

从等距映射的定义可以看出每个等距映射都是单射.欧几里得空间中的旋转映射就是等距映射的一种.

#### 1.2 等距映射的性质

设  $S \in \mathcal{L}(V, W)$ ,  $e_1, \dots, e_n$  为  $V$  的规范正交基,  $f_1, \dots, f_m$  为  $W$  的规范正交基,那么下列命题是等价的.

- (a)  $S$  是等距映射.
- (b)  $S^*S = I$ .
- (c)  $\langle Su, Sv \rangle = \langle u, v \rangle$  对于任意  $u, v \in V$  成立.
- (d)  $Se_1, \dots, Se_n$  为  $W$  中的规范正交组.
- (e)  $\mathcal{M}(S, (e_1, \dots, e_n), (f_1, \dots, f_m))$  的列形成  $\mathbb{F}^m$  中关于欧几里得内积的规范正交组.

#### Proof.

我们将对上述命题逐一推理.

(a) $\Rightarrow$ (b):因为  $S$  是等距映射,于是对于任意  $v \in V$  有

$$\langle (I - S^*S)v, v \rangle = \langle v, v \rangle - \langle S^*Sv, v \rangle = \|v\|^2 - \langle Sv, Sv \rangle = \|v\|^2 - \|Sv\|^2 = 0$$

于是  $I - S^*S = \mathbf{0}$ , 即  $S^*S = I$ .

(b) $\Rightarrow$ (c):对于任意  $u, v \in V$ , 有

$$\langle Su, Sv \rangle = \langle S^*Su, v \rangle = \langle Iu, v \rangle = \langle u, v \rangle$$

于是 (b) 蕴含 (c).

(c) $\Rightarrow$ (d):对于任意  $j, k \in \{1, \dots, n\}$  且  $j \neq k$  有  $\langle Se_j, Se_k \rangle = \langle e_j, e_k \rangle = 0$ .

对于任意  $k \in \{1, \dots, n\}$  又有  $\langle Se_k, Se_k \rangle = \langle e_k, e_k \rangle = 1$ . 于是  $Se_1, \dots, Se_n$  是  $W$  中的规范正交组.

(d) $\Rightarrow$ (e):令  $A = \mathcal{M}(S, (e_1, \dots, e_n), (f_1, \dots, f_m))$ . 对于  $k, r \in \{1, \dots, n\}$ , 我们有

$$\langle A_{\cdot, k}, A_{\cdot, r} \rangle = \sum_{j=1}^m A_{j, k} \overline{A_{j, r}} = \left\langle \sum_{j=1}^m A_{j, k} f_j, \sum_{j=1}^m A_{j, r} f_j \right\rangle = \langle Se_k, Se_r \rangle = \begin{cases} 1, & k = r \\ 0, & k \neq r \end{cases}$$

于是 $A$ 的列是 $\mathbb{F}^m$ 中的规范正交组.

(e) $\Rightarrow$ (a):上式表明 $Se_1, \dots, Se_n$ 是 $W$ 中的规范正交组.对于任意 $v := a_1e_1 + \dots + a_ne_n \in V$ 有

$$Sv = a_1Se_1 + \dots + a_nSe_n$$

于是

$$\|Sv\|^2 = a_1^2 + \dots + a_n^2$$

$$\|v\|^2 = a_1^2 + \dots + a_n^2$$

于是 $\|Sv\| = \|v\|$ ,因而 $S$ 是等距映射.

当等距的算子同时具有可逆性时,就引出了我们要讨论的第二种算子.

## 2.酉算子(么正算子)

### 2.1 定义:酉算子

算子 $T \in \mathcal{L}(V)$ 被称为酉的(么正的),如果 $T$ 是可逆等距映射.

你似乎会觉得奇怪,因为有限维内积空间上的酉算子就是等距算子.然而我们总应该记得,算子将向量映射至原本的向量空间,而一般地线性映射则是两个空间之间的.此外,在遇到无限维向量空间时,我们的定义中仍应当保留可逆这一条件,毕竟无限维向量空间上的单射不一定可逆.

酉算子具有如下的性质.请注意与1.2中等距映射的性质作比较.

### 2.2 酉算子的性质

设 $S \in \mathcal{L}(V)$ ,  $e_1, \dots, e_n$ 为 $V$ 的规范正交基,那么下列命题是等价的.

- (a)  $S$ 是酉算子.
- (b)  $S^*S = SS^* = I$ .
- (c)  $S$ 可逆且 $S^{-1} = S^*$ .
- (d)  $Se_1, \dots, Se_n$ 为 $V$ 的规范正交基.
- (e)  $\mathcal{M}(S, (e_1, \dots, e_n))$ 的行形成 $\mathbb{F}^n$ 中关于欧几里得内积的规范正交组.
- (f)  $S^*$ 是酉算子.

### Proof.

我们将对上述命题逐一推理.

(a) $\Rightarrow$ (b):根据1.2可知 $S^*S = I$ .等式两端右乘 $S^{-1}$ 有 $SS^* = SS^{-1} = I$ ,于是 $S^*S = SS^* = I$ .

(b) $\Rightarrow$ (c):根据可逆性和逆的定义即可知(b)蕴含(c).

(c) $\Rightarrow$ (d):由(c)可知 $S^{-1} = S^*$ ,于是 $SS^* = I$ .根据1.2中(b)和(d)的等价关系可知 $Se_1, \dots, Se_n$ 是 $V$ 中的规范正交组.又因为 $\dim V = n$ ,于是 $Se_1, \dots, Se_n$ 是 $V$ 的规范正交基.

(d) $\Rightarrow$ (e):由1.2中(a)和(d)的等价关系可知 $S$ 是酉算子.于是

$$(S^*)^* S^* = SS^* = I$$

由1.2中(a)和(b)的等价关系可知 $S^*$ 是等距映射,于是 $\mathcal{M}(S^*, (e_1, \dots, e_n))$ 的列形成 $\mathbb{F}^n$ 的规范正交基.

又因为 $\mathcal{M}(S, (e_1, \dots, e_n))$ 是 $\mathcal{M}(S^*, (e_1, \dots, e_n))$ 的共轭转置,于是 $\mathcal{M}(S)$ 的行形成 $\mathbb{F}^n$ 的规范正交基.

(e) $\Rightarrow$ (f): $\mathcal{M}(S^*, (e_1, \dots, e_n))$ 的列形成 $\mathbb{F}^n$ 的规范正交基,于是 $S^*$ 是等距映射,因而是酉算子.

(f) $\Rightarrow$ (a):由于(a)蕴含(f),于是 $(S^*)^*$ 是酉算子,即 $S$ 是酉算子.

回顾一下我们所做的 $\mathbb{C}$ 与 $\mathcal{L}(V)$ 的类比.在这一类比下,复数 $z$ 对应于算子 $S \in \mathcal{L}(V)$ , $\bar{z}$ 对应 $S^*$ .实数( $z = \bar{z}$ )对应于自伴算子( $S = S^*$ ),而非负数对应于正算子.

$\mathbb{C}$ 的另一个重要子集为单位圆,由所有满足 $|z| = 1$ (即 $z\bar{z} = 1$ )的复数组成.在我们的类比中,这一条件即 $SS^* = I$ ,表明 $S$ 是酉算子.接下来的结果表明这体现在酉算子的特征值上.

### 2.3 酉算子的特征值

设 $\lambda$ 是酉算子的特征值,那么 $|\lambda| = 1$ .

**Proof.**

设 $S \in \mathcal{L}(V)$ 是酉算子且 $\lambda$ 是 $S$ 的特征值.不妨令 $v \in V$ 使得 $v \neq \mathbf{0}$ 且 $Sv = \lambda v$ ,于是

$$|\lambda| \|v\| = \|\lambda v\| = \|Sv\| = \|v\|$$

于是 $|\lambda| = 1$ ,命题得证.

接下来这一命题结合了复谱定理刻画了复内积空间上的酉算子的性质.

### 2.4 复内积空间上的酉算子

设 $\mathbb{F} = \mathbb{C}$ 且 $S \in \mathcal{L}(V)$ ,那么下列命题是等价的.

(a)  $S$ 是酉算子.

(b) 存在 $S$ 的特征向量构成的 $V$ 的规范正交基,使得其对应的特征值的绝对值均为1.

**Proof.**

(a) $\Rightarrow$ (b):由 $S$ 是酉算子表明 $S$ 是正规的,于是由复谱定理可得存在 $S$ 的特征向量构成的 $V$ 的规范正交基.

而 $S$ 的每个特征值均为1,于是(a)蕴含(b).

(b) $\Rightarrow$ (a):令 $e_1, \dots, e_n$ 是 $S$ 的特征向量构成的 $V$ 的规范正交基,其对应特征值为 $\lambda_1, \dots, \lambda_n$ .于是

$$\langle Se_j, Se_k \rangle = \langle \lambda_j e_j, \lambda_k e_k \rangle = \lambda_j \overline{\lambda_k} \langle e_j, e_k \rangle = \begin{cases} 1, j = k \\ 0, j \neq k \end{cases}$$

于是 $Se_1, \dots, Se_n$ 也是 $V$ 的规范正交基.根据2.2中(a)和(d)的等价关系可知 $S$ 是酉算子.

### 3.矩阵分解

在本小节中,我们将研究的重心转向矩阵.除非另行说明,我们约定 $n \times n$ 矩阵对应的算子都在具有欧几里得内积的 $\mathbb{F}^n$ 上,并且选定的基为 $\mathbb{F}^n$ 的标准基.容易看出, $\mathbb{F}^n$ 的标准基同时也是规范正交基.我们先做酉矩阵的定义.

#### 3.1 定义:酉矩阵

$n \times n$ 矩阵被称为酉矩阵,如果它的列形成 $\mathbb{F}^n$ 的规范正交基.

考察1.2中(a)和(e)的等价关系,可知酉矩阵对应的算子就是酉算子.又根据2.2中(a)和(e)的等价关系,这定义中的列也可以替换为行.下面给出了一些酉矩阵的性质.

#### 3.2 酉矩阵的性质

设 $Q$ 是 $n \times n$ 矩阵,那么下列命题是等价的.

- (a)  $Q$ 是酉矩阵.
- (b)  $Q$ 的行形成 $\mathbb{F}^n$ 中的规范正交组.
- (c)  $\|Qv\| = \|v\|$ 对任意 $v \in \mathbb{F}^n$ 都成立.
- (d)  $QQ^* = Q^*Q = I$ ,  $I$ 是对角线元素均为1的对角矩阵.

上面几个命题的等价关系的证明是不难的,参照酉算子的性质即可.

利用酉矩阵,我们可以对矩阵进行分解.

#### 3.3 矩阵的QR分解

设 $A$ 是各列线性无关的 $n \times n$ 方阵,那么存在唯一一对 $n \times n$ 方阵 $Q$ 和 $R$ ,其中 $Q$ 是酉矩阵而 $R$ 是对角线上仅有正数的上三角矩阵,使得 $A = QR$ .

#### Proof.

令 $v_1, \dots, v_n$ 表示 $A$ 的列.将其视为 $\mathbb{F}^n$ 的元素并应用Gram-Schmidt过程得到 $\mathbb{F}^n$ 的规范正交基 $e_1, \dots, e_n$ .

那么 $\text{span}(e_1, \dots, e_k) = \text{span}(v_1, \dots, v_k)$ 对任意 $k \in \{1, \dots, n\}$ 成立.令 $R$ 是 $n \times n$ 矩阵,定义为

$$R_{j,k} = \langle v_k, e_j \rangle$$

当 $j > k$ 时,  $e_j \in (\text{span}(e_1, \dots, e_k))^\perp$ , 从而 $e_j$ 与 $v_k$ 正交, 因而 $R_{j,k} = \langle e_j, v_k \rangle = 0$ . 这表明 $R$ 是上三角矩阵. 令 $Q$ 是各列为 $e_1, \dots, e_n$ 的正交矩阵. 那么对于任意 $k \in \{1, \dots, n\}$ 有

$$(QR)_{\cdot,k} = \sum_{j=1}^n R_{j,k} Q_{\cdot,j} = \sum_{j=1}^n \langle v_k, e_j \rangle e_j = v_k$$

于是 $A = QR$ .

Gram-Schmidt过程表明每个 $v_k$ 都是 $e_k$ 的正数倍加上 $e_1, \dots, e_{k-1}$ 的线性组合得到, 于是 $\langle v_k, e_k \rangle > 0$ , 即 $R$ 的对角线上的元素都是正数. 这一点也得证.

最后, 假定还有么正的 $\hat{Q}$ 和对角线上仅有正数的上三角的 $\hat{R}$ 满足 $A = \hat{Q}\hat{R}$ .

令 $q_1, \dots, q_n$ 表示 $\hat{Q}$ 的列. 将矩阵乘法视作列的线性组合, 可知对于任意 $k \in \{1, \dots, n\}$ 有

$$\text{span}(v_1, \dots, v_k) = \text{span}(q_1, \dots, q_k)$$

且 $\langle v_k, q_k \rangle > 0$ . 我们已经证明了满足上述条件的规范正交组是唯一的, 这表示 $q_k = e_k$ 对任意 $k \in \{1, \dots, n\}$ 成立.

于是 $\hat{Q} = Q$ , 这表明 $\hat{R} = R$ , 于是QR分解具有唯一性.

矩阵的QR分解可以用于解线性方程组. 具体来说, 设 $A$ 是各列线性无关的 $n \times n$ 方阵, 而 $b \in \mathbb{F}^n$ . 求解方程 $Ax = b$ , 可以先对 $A$ 做QR分解, 得到 $QRx = b$ . 方程两端左乘 $Q^*$  (即 $Q$ 的共轭转置), 得到 $Rx = Q^*b$ . 由于 $R$ 是上三角矩阵, 于是求解是十分方便的, 只需按逆序求出各个位置上的元素即可.

#### 4. 科列斯基分解

我们从内积的角度刻画可逆正算子作为本小节的开始.

##### 4.1 可逆正算子

自伴算子 $T \in \mathcal{L}(V)$ 是可逆正算子, 当且仅当 $\langle Tv, v \rangle > 0$ 对任意非零的 $v \in V$ 都成立.

##### Proof.

首先假设 $T$ 是可逆正算子. 对于非零的 $v \in V$ , 由于 $T$ 可逆则有 $Tv \neq \mathbf{0}$ , 于是 $\langle Tv, v \rangle \neq 0$ , 于是 $\langle Tv, v \rangle > 0$ . 现在假设 $\langle Tv, v \rangle > 0$ 对任意非零的 $v \in V$ 都成立, 那么 $\text{null } T = \{\mathbf{0}\}$ , 于是 $T$ 是单射, 进而 $T$ 可逆.

接下来, 我们将上述结果转化为矩阵语言.

##### 4.2 定义: 正定矩阵

矩阵 $B \in \mathbb{F}^{n,n}$ 被称为**正定的**, 如果 $B = B^*$ 且 $\langle Bx, x \rangle > 0$ 对任意非零的 $x \in \mathbb{F}^n$ 都成立.

正定矩阵可以被精妙地分解为下三角矩阵和其共轭转置的乘积.

### 4.3 科列斯基分解

设 $B$ 是正定矩阵,那么存在唯一的对角线上仅含正数的上三角矩阵 $R$ 使得 $B = R^* R$ .

#### Proof.

由于 $B$ 是正定矩阵,于是存在可逆矩阵 $A$ 使得 $B = A^* A$ .

令 $A = QR$ 是 $A$ 的QR分解,那么 $A = Q^* R^*$ .于是我们有

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

为证明该分解是唯一的,再令 $S$ 满足 $B = S^* S$ ,其中 $S$ 是对角线上仅含正数的上三角矩阵.于是 $S$ 是可逆的.

在此式两端右乘 $S^{-1}$ 可得 $BS^{-1} = S^*$ .令 $A$ 是前面提到的使得 $B = A^* A$ 的矩阵,于是

$$\begin{aligned}(AS^{-1})^* (AS^{-1}) &= (S^*)^{-1} A^* AS^{-1} \\ &= (S^*)^{-1} BS^{-1} \\ &= (S^*)^{-1} S^* \\ &= I\end{aligned}$$

于是 $AS^{-1}$ 是酉算子.因此, $A = (AS^{-1}) S$ 这一分解是QR分解.根据QR分解的唯一性,可知 $S = R$ ,这就证明了科列斯基分解的唯一性.