《线性代数》

14-线性变换 (Linear Transform)

杨启哲

上海师范大学信机学院计算机系

2024 年 5 月 16 日

我们称一个方阵是对称 (Symmetric)的，如果其满足：

$$A = A^{\mathrm{T}}$$

定理 1.

所有实对称矩阵的特征向量是实数，并且每个特征值都有一个对应的实特征向量。

定理 2.

对于一个实对称矩阵 S，如果 λ_1 和 λ_2 是 S 的两个不同的特征值，x_1 和 x_2 是 λ_1 和 λ_2 对应的特征向量，则 x_1 和 x_2 是正交的。

定理 3.

令 S 是一个 $n \times n$ 的实对称矩阵，则 S 可以对角化。更精确的说，存在一个正交矩阵 Q 和对角矩阵 Λ 使得：

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} = Q^T S Q$$

这里 $\lambda_1, \cdots, \lambda_n$ 都是 S 的特征值。显然这是一个当且仅当的关系，因为另一个方向是显然成立的。

从而，给定一个实对称矩阵 S，我们有其谱分解：

$$S = Q\Lambda Q^T = \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T + \cdots + \lambda_n q_n q_n^T$$

上海师范大学 Shanghai Normal University

定义 4.

给定一个 $n \times n$ 的实矩阵A，如果对任意向量 $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$:

1. $x^\mathrm{T}Ax > 0$，则称 A 是正定的。
2. $x^\mathrm{T}Ax \geqslant 0$，则称 A 是半正定的。
3. $x^\mathrm{T}Ax < 0$，则称 A 是负定的。
4. $x^\mathrm{T}Ax \leqslant 0$，则称 A 是半负定的。
5. 若不满足以上任何一种条件，则称矩阵 A 是不定的。

正定矩阵的性质

令 S 是一个对称矩阵，S 是正定的等价于

1. S 的特征值都大于 0.
2. S 的顺序主子式都大于 0。
3. 存在一个矩阵 A，使得 $S = A^\mathrm{T}A$.

一般地，对于 $x_1,\cdots,x_n$ 的一个二次齐次函数：

$$f(x_1,\cdots,x_n)=s_{11}x_1^2+\cdots+s_{nn}x_n^2+2\sum_{1\leqslant i\neq j\leqslant n}s_{ij}x_ix_j$$

我们总可以将其转换成如下的矩阵形式：

$$f(x_1,\cdots,x_n)=x^TSx,\ \text{其中：}\ x=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix},\ S=\begin{bmatrix}s_{11}&s_{12}&\cdots&s_{1n}\\ s_{12}&s_{22}&\cdots&s_{2n}\\ \vdots&\vdots&\ddots&\vdots\\ s_{1n}&s_{2n}&\cdots&s_{nn}\end{bmatrix}$$

特别的，将只含平方项的二次型称为标准二次型，特别的，如果其系数进一步化简为 $1,1,\cdots,1,-1,\cdots,-1$，我们称其为规范二次型。

转换方法

1. 利用对称矩阵的谱分解。
2. 配方法。

主要内容

- 特征空间、代数重数以及几何重数
- 线性变换
- 线性变换的矩阵形式
- 线性变换的像和核
- 对偶性 (Duality)

特征空间、代数重数以及几何重数

给定一个 $n\times n$ 的矩阵 A 和其一个特征值 $\lambda\in\mathbb{C}$，若 x_1,x_2 是其特征向量，则我们有：

$$A(x_1+x_2)=Ax_1+Ax_2=\lambda x_1+\lambda x_2=\lambda(x_1+x_2)$$

$$A(cx_1)=cAx_1=c\lambda x_1=\lambda(cx_1)$$

即其特征向量组成的空间是一个向量空间，我们称为特征空间。

定义 5 **[特征空间 (Eigenspace)].**

给定一个 $n\times n$ 的矩阵 A 和其一个特征值 $\lambda\in\mathbb{C}$，定义：

$$V_\lambda=\{x\in\mathbb{C}^n\mid Ax=\lambda x\}$$

为 A 的特征值 λ 对应的特征空间。

引理 6.

V_λ 是 $\mathbb{C}^n$ 的一个子空间，并且 $\dim(V_\lambda)\geqslant 1$。

考察旋转矩阵:

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

我们有:

$$\det(A - \lambda I) = \lambda^2 + 1 \implies \lambda = \pm i$$

从而我们有:

$$\lambda = i \implies V_i = \{x \in \mathbb{C}^n \mid Ax = ix\} = \text{span}\left\{\begin{bmatrix} 1 \\ i \end{bmatrix}\right\}$$

$$\lambda = -i \implies V_{-i} = \{x \in \mathbb{C}^n \mid Ax = -ix\} = \text{span}\left\{\begin{bmatrix} 1 \\ -i \end{bmatrix}\right\}$$

我们将 V_λ 的维数称为 λ 的几何重数 (Geometric Multiplicity)。

定义 7 [几何重数 (Geometric Multiplicity)].

给定一个 $n \times n$ 的矩阵 A 和其一个特征值 $\lambda \in \mathbb{C}$，λ 的几何重数是其特征空间 V_λ 的维数，即：

$$\begin{aligned}\dim(V_\lambda) &= \dim\left(\{x \in \mathbb{C}^n \mid Ax = \lambda x\}\right) \\ &= \dim\left(\{x \in \mathbb{C}^n \mid (A - \lambda I)x = \mathbf{0}\}\right) \\ &= \dim\left(\mathrm{null}(A - \lambda I)\right) = n - \mathrm{rank}(A - \lambda I)\end{aligned}$$

例 8.

在上述的旋转矩阵 A 中，我们有 $\dim(V_i) = \dim(V_{-i}) = 1$，即 i 和 $-i$ 的几何重数为 1。

考虑 A 的特征多项式:

$$f_A(\lambda) = \det(A - \lambda I) = \prod_{i=1}^{n}(\lambda - \lambda_i) = \prod_{i=1}^{k}(\lambda - \lambda_i)^{m_i}(m_1 + m_2 + \cdots + m_k = n)$$

我们将 m_i 称为 λ_i 的代数重数 (Algebraic Multiplicity)。

定义 9 [代数重数 (Algebraic Multiplicity)].

给定一个 $n \times n$ 的矩阵 A 和其一个特征值 $\lambda \in \mathbb{C}$, λ 的代数重数是其特征多项式中 $(\lambda - \lambda_i)$ 的幂次 m_i，即:

$$m_i = \max\{m \mid (\lambda - \lambda_i)^m 是 f_A(\lambda) 的因子\}$$

例 10.

在上述的旋转矩阵 A 中，我们有 $f_A(\lambda) = (\lambda + i)(\lambda - i)$，即 i 和 $-i$ 的代数重数为 1。

考察如下矩阵:

$$A = \begin{bmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

其特征多项式为:

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} -1-\lambda & 1 & 0 \\ -4 & 3-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{vmatrix} \\ &= (-1-\lambda)\begin{vmatrix} 3-\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} - 1\begin{vmatrix} -4 & 0 \\ 1 & 2-\lambda \end{vmatrix} \\ &= (-1-\lambda)(3-\lambda)(2-\lambda) + 4(2-\lambda) \\ &= (1-\lambda)^2(2-\lambda) \end{aligned}$$

从而其特征值为 $\lambda = 1, 2$, 代数重数分别为 $2, 1$。

$\lambda = 2$ 对应的特征空间

考察 $A - 2I$:

$$A - 2I = \begin{bmatrix} -3 & 1 & 0 \\ -4 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

不难验证 $\mathrm{rank}(A - 2I) = 2$，从而 $\dim(V_2) = n - \mathrm{rank}(A - 2I) = 1$，即 2 的几何重数为 1，更精确的说:

$$V_2 = \{x \in \mathbb{C}^n \mid Ax = 2x\} = \left\{ c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \middle| c \in \mathbb{C} \right\} = \mathrm{span}\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$\lambda = 1$ 对应的特征空间

考察 $A - I$:

$$A - I = \begin{bmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

不难验证 $\mathrm{rank}(A - I) = 2$，从而 $\dim(V_1) = n - \mathrm{rank}(A - I) = 1$，即 1 的几何重数为 1，更精确的说:

$$V_1 = \{x \in \mathbb{C}^n \mid Ax = x\} = \left\{ c \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \middle| c \in \mathbb{C} \right\} = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \right\}$$

注意到1 的代数重数为 2，即代数重数和几何重数并不相等。

几何重数与代数重数的关系

定理 11.

特征值的代数重数大于等于其几何重数。

引理 12.

令 A 和 B 是相似矩阵，即存在可逆矩阵 P 使得 $B = P^{-1}AP$，则我们有：

$$\det(A - \lambda I) = \det(B - \lambda I)$$

即两个相似矩阵不仅有相同的特征值，而且其代数重数也相等。

引理12的证明

上海师范大学 Shanghai Normal University

由于 $B = P^{-1}AP$，我们有:

$$\begin{aligned}
\det(B - \lambda I) &= \det(A - \lambda I) \\
&= \det(P^{-1}AP - P^{-1}(\lambda I)P) \\
&= \det(P^{-1}(A - \lambda I)P) \\
&= \det(P^{-1})\det(A - \lambda I)\det(P) \\
&= \det(A - \lambda I)
\end{aligned}$$

$\square$

$\det(P)\det(P^{-1}) = 1$

注意到，若 P 是可逆的，则我们有:

$$1 = \det(I) = \det(PP^{-1}) = \det(P)\det(P^{-1}) \Longrightarrow \det(P^{-1}) = \frac{1}{\det(P)}$$

定理的直观思路

如果 $\dim(V_\lambda)=i$，意味着可以从中找出 i 个线性无关的特征向量 $x_1,\cdots,x_i$，使得:

$$A\begin{bmatrix}x_1 & \cdots & x_i\end{bmatrix}=\begin{bmatrix}\lambda x_1 & \cdots & \lambda x_i\end{bmatrix}=\begin{bmatrix}x_1 & \cdots & x_i\end{bmatrix}\begin{bmatrix}\lambda & & \\ & \ddots & \\ & & \lambda\end{bmatrix}$$

从而我们可以将 A 转化成如下的形式:

$$A=P\begin{bmatrix}\lambda I_i & B\\ O & C\end{bmatrix}P^{-1}$$

由上述引理可知 λ 的代数重数一定大于等于 i。

定理11的证明 (II)

上海师范大学 Shanghai Normal University

现在令 A 是一个 $n\times n$ 的矩阵，λ_0 是其一个特征值，考虑其特征空间：

$$V_{\lambda_0}=\{x\in\mathbb{C}^n \mid Ax=\lambda_0 x\}$$

令其的一组基为 $v_1,\cdots,v_m$，则 m 就是 λ_0 的几何重数。特别的，我们有：

$$\begin{aligned} A\begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix} &= \begin{bmatrix} \lambda_0 v_1 & \cdots & \lambda_0 v_m \end{bmatrix} \\ &= \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix}\begin{bmatrix} \lambda_0 & & \\ & \ddots & \\ & & \lambda_0 \end{bmatrix} \end{aligned}$$

这里 $\begin{bmatrix} \lambda_0 & & \\ & \ddots & \\ & & \lambda_0 \end{bmatrix}=\lambda_0 I_m$ 是一个 $m\times m$ 的矩阵。

定理11的证明 (III)

我们将 $v_1, \cdots, v_m$ 扩展成 $\mathbb{C}^n$ 的一组基，即:

$$v_1, \cdots, v_m, v_{m+1}, \cdots, v_n$$

从而我们有:

$$\begin{aligned} A\begin{bmatrix} v_1 & \cdots & v_m & v_{m+1} & \cdots & v_n \end{bmatrix} &= \begin{bmatrix} \lambda_0 v_1 & \cdots & \lambda_0 v_m & Av_{m+1} & \cdots & Av_n \end{bmatrix} \\ &= \begin{bmatrix} v_1 & \cdots & v_m & v_{m+1} & \cdots & v_n \end{bmatrix} \begin{bmatrix} \lambda_0 I_m & B \\ O & C \end{bmatrix} \end{aligned}$$

这里 B 是一个 $m \times (n-m)$ 的矩阵，C 是一个 $(n-m) \times (n-m)$ 的矩阵。

定理11的证明 (IV)

上海师范大学 Shanghai Normal University

令 $P = \begin{bmatrix} v_1 & \cdots & v_m & v_{m+1} & \cdots & v_n \end{bmatrix}$，则 P 是可逆矩阵，从而我们有：

$$A = P \begin{bmatrix} \lambda_0 I_m & B \\ O & C \end{bmatrix} P^{-1}$$

从而我们有

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} \lambda_0 I_m & B \\ O & C \end{bmatrix} - \lambda I\right) = \begin{vmatrix} (\lambda_0 - \lambda) I_m & B \\ O & C - \lambda I_{n-m} \end{vmatrix}$$

不难验证：

$$\det(A - \lambda I) = (\lambda_0 - \lambda)^m g(\lambda)$$

这里 $g(\lambda)$ 是一个 $n - m$ 次的多项式，从而 λ_0 的代数重数至少为 m，即：

$$\lambda_0 \text{ 的代数重数} \geqslant \lambda_0 \text{ 的几何重数。}$$

□

回顾矩阵可对角化的条件:

定理 13.

$n \times n$ 的矩阵可对角化当且仅当 A 有n 个线性无关的特征向量。

从而我们可以得到:

定理 14.

$n \times n$ 的矩阵可对角化当且仅当其每个特征值的代数重数等于几何重数。

因此对于实对称矩阵来说:

推论 15.

令 S 是一实对称矩阵，并且 λ 是其一个特征值，则 λ 的几何重数等于其代数重数。

阶段总结

- 特征空间。
- 代数重数与几何重数。
- 代数重数和几何重数之间的关系，及与对角化的联系。

线性变换

考虑函数:

$$y = 3x$$

这是一个线性函数 $f : \mathbb{R} \longrightarrow \mathbb{R}$.

- 几何上来讲，其定义了 $\mathbb{R}^2$ 上或者说是一个二维平面上的一条线。
- 代数上来讲，对任意的 $x, y, c \in \mathbb{R}$ 其满足:

$$f(x+y) = 3(x+y) = 3x + 3y = f(x) + f(y)$$

$$f(cx) = 3cx = cf(x)$$

定义 16 **[Linear Transformation].**

令 V 和 W 实两个向量空间。一个函数 $T : \mathbb{V} \longrightarrow \mathbb{W}$ 被称作是一个线性变换，如果其满足对任意的 $v, w \in \mathbb{V}$ 和实数 $c \in \mathbb{R}$:

$$T(v + w) = T(v) + T(w), \quad T(cv) = cT(v)$$

术语说明!

在很多翻译当中，会将 $V \neq W$ 的情况称为线性映射，而只有 $V = W$ 的情况下称为线性变换。我们的课件则不作这类区分，但当 $V \neq W$ 时，我们会经常显示的表达出 V 和 W。

引理 17.

1. $T(\mathbf{0}) = \mathbf{0}$.
2. $T(cv + dw) = cT(v) + dT(w)$.
3. $T(c_1v_1 + \cdots + c_nv_n) = c_1T(v_1) + \cdots + c_nT(v_n)$

一些例子

1. 内积 (Inner Product)：固定一个向量 $a = (1, 2, 3) \in \mathbb{R}^3$，定义 $T_a(v) = a \cdot v$。则 T_a 是 $\mathbb{R}^3 \to \mathbb{R}$ 上的线性变换。
2. 长度 (Length): 定义 $T(v) = \|v\|$，则 T不是$\mathbb{R}^n \to \mathbb{R}$ 上的线性变换。
3. 旋转 (Rotation)：定义 $\mathbb{V}$ 上的映射:

$$T(v) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} v$$

则 T 是一个 $\mathbb{V}$ 到 $\mathbb{V}$ 上的线性变换，我们也称其为 $\mathbb{V}$ 上的线性变换。

由上述旋转的例子可以看到，我们可以利用矩阵的乘法来定义一个线性变换。

引理 18.

令 A 是一个 $m \times n$ 的矩阵，定义 $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ 为:

$$T_A(x) = Ax$$

则 T_A 是一个 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 上的线性变换。

一个线性变换的前瞻性视角

事实上我们将说明，对于任何一个如下的线性变换:

$$T : \mathbb{V} \longrightarrow \mathbb{W}$$

其中 $\dim(V) = n, \dim(W) = m$，我们都可以将其理解成如下的线性变换:

$$T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

上海师范大学 Shanghai Normal University

我们还是先研究空间的基对线性变换的影响。

引理 19.
令 $\mathbb{V}$ 和 $\mathbb{W}$ 是两个线性空间，并且 $\dim(\mathbb{V}) = \dim(\mathbb{W}) = n$。令 $v_1, \cdots, v_n$ 是 $\mathbb{V}$ 的一组基，则：

1. 令 $T_1, T_2 : \mathbb{V} \longrightarrow \mathbb{W}$ 是两个其上的线性变换，满足对任意的 $i \in [n]$ 都有：

$$T_1(v_i) = T_2(v_i)$$

   则 $T_1 = T_2$。
2. 令 $w_1, \cdots, w_n$ 是 W 中的一组基，则存在唯一的线性变换 $T : \mathbb{V} \longrightarrow \mathbb{W}$ 使得对任意的 $i \in [n]$ 都有：

$$T(v_i) = w_i$$

这说明了，在给定的基下，线性变换是唯一确定的。

引理19的证明

Shanghai Normal University 上海师范大学

令 $v \in \mathbb{V}$。由于 $v_1, \cdots, v_n$ 是 $\mathbb{V}$ 的一组基，从而存在 $c_1, \cdots, c_n$ 使得:

$$v = c_1 v_1 + \cdots + c_n v_n = \sum_{i \in [n]} c_i v_i$$

从而我们有:

$$\begin{aligned} T_1(v) &= T_1\left(\sum_{i \in [n]} c_i v_i\right) = \sum_{i \in [n]} c_i T_1(v_i) \\ &= \sum_{i \in [n]} c_i T_2(v_i) = T_2\left(\sum_{i \in [n]} c_i v_i\right) = T_2(v) \end{aligned}$$

另一方面，注意到 $c_1, \cdots, c_n$ 是唯一的，从而我们有:

$$T(v) = \sum_{i \in [n]} c_i w_i$$

即 T 满足对任意的 $i \in [n]$，有 $T(v_i) = w_i$。 $\square$

系数向量 (一组基下的坐标)

令 $\mathbb{V}$ 是一个 n 维的向量空间，$v_1, \cdots, v_n$ 是 $\mathbb{V}$ 的一组基，我们记作：

$$\overline{v} = v_1, \cdots, v_n$$

注意到，对于任意的 $v \in \mathbb{V}$，存在唯一的$c_1, \cdots, c_n \in \mathbb{R}$ 使得：

$$v = c_1 v_1 + \cdots + c_n v_n$$

定义其系数构成的向量 (coordinate vector) 为：

$$(c_1, \cdots, c_n) \in \mathbb{R}^n$$

其被称作是在基 $\overline{v}$ 下的坐标。

回顾-向量空间 $V = \mathbb{R}^+$

回顾我们之前提过的一个很奇怪的向量空间:

$$\mathsf{V} = \mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\}$$

我们定义了其上的加法 $\oplus$ 和数乘 $\otimes$:

$$x \oplus x' = x \times x', \quad c \otimes x = x^c$$

由于 $\dim(\mathsf{V}) = 1$，我们可以选取如下的一组基:

$$v_1 = 2$$

则对于任意的 $x \in \mathsf{V}$，我们有唯一的 $c \in \mathbb{R}$ 使得:

$$x = c \otimes v_1 = 2^c$$

即: $c = \log_2 x$ 是 x 在基 v_1 下的坐标。

坐标的表示也是线性变换

令 n 维空间 $\mathbb{V}$ 的一组基为 $\overline{v}$，我们定义 $T_{\overline{v}}: \mathbb{V} \longrightarrow \mathbb{R}^n$:

$$T_{\overline{v}}(v) = (c_1, \cdots, c_n)$$

这里 $(c_1, \cdots, c_n)$ 是 v 在基 $\overline{v}$ 下的坐标。

定理 20.

$T_{\overline{v}}$ 是一个 $\mathbb{V}$ 到 $\mathbb{R}^n$ 上的线性变换。并且 $T_{\overline{v}}$ 是一一对应的 (bijiective)。

$V = \mathbb{R}^+$ 上的例子

$$V = \mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\}$$

我们定义了其上的加法 $\oplus$ 和数乘 $\otimes$:

$$x \oplus x' = x \times x', \quad c \otimes x = x^c$$

由于 $\dim(V) = 1$，我们可以选取如下的一组基:

$$v_1 = 2$$

则对于任意的 $x \in V$, $\log_2 x$ 是其在基 v_1 下的坐标。

下列映射:

$$x \mapsto \log_2 x$$

是 $\mathbb{R}^+$ 到 $\mathbb{R}$ 上的一一对应的线性变换。

定理20的证明 (I)

我们只需证明:

1. 对任意的 $u, v \in \mathbb{V}$ 有
$$T_{\overline{v}}(u+v) = T_{\overline{v}}(u) + T_{\overline{v}}(v)$$
2. 对任意的 $v \in \mathbb{V}$ 和 $c \in \mathbb{R}$ 有
$$T_{\overline{v}}(cv) = cT_{\overline{v}}(v)$$
3. $T_{\overline{v}}$ 是一个双射:
   - 满射 (Surjection): 对任意的 $(c_1, \cdots, c_n) \in \mathbb{R}^n$，存在 $v \in \mathbb{V}$ 使得 $T_{\overline{v}}(v) = (c_1, \cdots, c_n)$。
   - 单射 (Injection): 对任意的 $v, u \in \mathbb{V}$，如果 $T_{\overline{v}}(v) = T_{\overline{v}}(u)$，则 $v = u$。

定理20的证明 (II)

令 $v = \sum_{i\in[n]} c_i v_i$，$v' = \sum_{i\in[n]} c_i' v_i$，则我们有:

$$
\begin{aligned}
T_{\overline{v}}(v+v') &= T_{\overline{v}}\left(\sum_{i\in[n]} c_i v_i + \sum_{i\in[n]} c_i' v_i\right) \\
&= T_{\overline{v}}\left(\sum_{i\in[n]} (c_i + c_i') v_i\right) \\
&= (c_1 + c_1', \cdots, c_n + c_n') \\
&= (c_1, \cdots, c_n) + (c_1', \cdots, c_n') \\
&= T_{\overline{v}}\left(\sum_{i\in[n]} c_i v_i\right) + T_{\overline{v}}\left(\sum_{i\in[n]} c_i' v_i\right) = T_{\overline{v}}(v) + T_{\overline{v}}(v')
\end{aligned}
$$

定理20的证明 (III)

令 $v=\sum_{i\in[n]} c_i v_i$ 和 $c\in\mathbb{R}$，则我们有:

$$
\begin{aligned}
T_{\overline{v}}(cv) &= T_{\overline{v}}(c\sum_{i\in[n]} c_i v_i) \\
&= T_{\overline{v}}\left(\sum_{i\in[n]} (cc_i) v_i\right) \\
&= (cc_1,\cdots,cc_n) \\
&= c\,(c_1,\cdots,c_n) \\
&= cT_{\overline{v}}(\sum_{i\in[n]} c_i v_i) = cT_{\overline{v}}(v)
\end{aligned}
$$

最后我们证明 $T_{\overline{V}}$ 是一个双射：

- 令 $v, v' \in \mathbb{V}$ 满足：$T_{\overline{V}}(v) = T_{\overline{V}}(v') = (c_1, \cdots, c_n)$，则我们有：

$$v = \sum_{i \in [n]} c_i v_i = v'$$

  即 $v = v'$。

- 令 $(c_1, \cdots, c_n) \in \mathbb{R}^n$，则定义：

$$u = c_1 v_1 + \cdots + c_n v_n = \sum_{i \in [n]} c_i v_i \in \mathbb{V}$$

  从而我们有：$T_{\overline{V}}(u) = (c_1, \cdots, c_n)$.

$\square$

我们可以看到, $T_{\overline{v}}: \mathbb{V} \longrightarrow \mathbb{R}^n$ 的定义取决于基 $\overline{v}$ 的选择:

$$v_1, \cdots, v_n$$

如果我们选取不同的基 $\overline{v}$ 和 $\overline{v'}$, $T_{\overline{v}}$ 和 $T_{\overline{v'}}$ 有什么关系?

例 21.

考察 $\mathbb{R}^2$, 我们选取两组基:

$$\overline{v_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \overline{v_2} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

则对于 $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ 来说，我们有:

$$T_{\overline{v_1}}(\begin{bmatrix} 2 \\ 3 \end{bmatrix}) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, T_{\overline{v_2}}(\begin{bmatrix} 2 \\ 3 \end{bmatrix}) = \begin{bmatrix} 2.5 \\ -0.5 \end{bmatrix}$$

事实上，我们将证明其中存在一个基变换矩阵M 满足: $T_{\overline{v_1}}(v) = MT_{\overline{v_2}}(v)$.

一些矩阵表示 (I)

上海师范大学 Shanghai Normal University

令 $v_1, \cdots, v_n \in \mathbb{V}$，$A$ 是一个 $n \times m$ 的矩阵。则我们定义：

$$\begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} A$$

为满足：

$$u_i = \sum_{k \in n} A(k, i) v_k$$

的 m 个 $\mathbb{V}$ 中的元素 $u_1, \cdots, u_m \in \mathbb{V}$。

说明

注意，这里的 u_i, v_i 并不是列向量，而是 V 中的一个元素，从而 $\begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ 并不是矩阵，我们使用相同的写法是因为可以证明其有类似的含义。

引理 22.

令 $v_1, \cdots, v_n \in \mathbb{V}$，$A$ 是 $n \times m$ 的矩阵，B 是 $m \times l$ 的矩阵，则我们有：

$$\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} (AB) = (\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} A)B$$

证明. 定义：

$$\begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} A, \quad \begin{bmatrix} w_1 & \cdots & w_l \end{bmatrix} = (\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} A)B$$

则对于任意的 $i \in [l]$，我们有：

$$w_i = \sum_{k \in [m]} B(k,i) u_k = \sum_{k \in [m]} B(k,i) \left(\sum_{j \in [n]} A(j,k) v_j \right) = \sum_{j \in [n]} \left(\sum_{k \in [m]} A(j,k) B(k,i) \right) v_j$$

即：

$$(\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} A)B = \begin{bmatrix} w_1 & \cdots & w_l \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} (AB)$$

$\square$

令 $\overline{v} = v_1, \cdots, v_n$ 是 $\mathbb{V}$ 的一组基，则对于 $v \in \mathbb{V}$ 我们有：

$$v = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} T_{\overline{v}}(v)$$

引理 23.

令 $\overline{v} = v_1, \cdots, v_n$ 和 $\overline{v'} = v_1', \cdots, v_n'$ 是 V 的两组基，则存在一个唯一的 $n \times n$ 矩阵 M 满足：

$$\begin{bmatrix} v_1' & \cdots v_n' \end{bmatrix} = \begin{bmatrix} v_1 & \cdots v_n \end{bmatrix} M$$

定义 24 **[Change of Basis].**

上述 M 被称作基 $\overline{v}$ 到 $\overline{v'}$ 的过渡矩阵。

引理23的证明

由于 $v_1, \cdots, v_n$ 是 $\mathbb{V}$ 的一组基，从而对于任意的 v_i'，存在唯一的 $a_{i1}, \cdots, a_{in}$ 使得:

$$v_i' = a_{i1}v_1 + \cdots + a_{in}v_n = \sum_{k\in[n]} a_{ik}v_k$$

即 $(a_{i1}, \cdots, a_{in}) \in \mathbb{R}^n$ 是 v_i' 在基 $\overline{v}$ 下的坐标。

定义下列矩阵:

$$M = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

则我们有:

$$\begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} M$$

M 的唯一性由 $v_i' = \sum_{k\in[n]} a_{ik}v_k$ 保证。 $\square$

定理 25.

M 是可逆的。

证明. 由引理 23，存在唯一的 $n \times n$ 的矩阵 M' 满足:

$$\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} M'$$

从而:

$$\begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} M = (\begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} M')M = \begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} (M'M)$$

另一方面，我们有: $\begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} = \begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} I$，即 I 也是 $\overline{v'}$ 到 $\overline{v'}$ 的过渡矩阵。根据过渡矩阵的唯一性，我们有:

$$M'M = I$$

□

定理 26.

令 $\overline{v}n$ 和 $\overline{v'}$ 是 $\mathbb{V}$ 的两组基，则对于任意的 $u \in \mathbb{V}$，我们有：

$$T_{\overline{v}}(u) = MT_{\overline{v'}}(u)$$

证明. 注意到：

$$\begin{aligned}
\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} T_{\overline{v}}(u) = u &= \begin{bmatrix} v'_1 & \cdots & v'_n \end{bmatrix} T_{\overline{v'}}(u) \\
&= \left(\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} M \right) T_{\overline{v'}}(u) \\
&= \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} (MT_{\overline{v'}}(u))
\end{aligned}$$

从而由 $v_1, \cdots, v_n$ 是线性无关的，我们有：

$$T_{\overline{v}}(u) = MT_{\overline{v'}}(u)$$

□

坐标变换公式 (II)

给定 $v \in \mathbb{V}$:

- 其在基 $\overline{v}$ 下的坐标为:

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

- 其在基 $\overline{v'}$ 下的坐标为:

$$\begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix}$$

则 $T_{\overline{v}}(v) = MT_{\overline{v'}}(v)$ 告诉我们:

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = M \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix}$$

阶段总结

Shanghai Normal University 上海师范大学

- 线性变换的基本概念。
- 利用坐标来表示向量空间的元素，线性变换 $T_{\overline{V}}$.
- 利用矩阵来表示基变换下的不同坐标。

线性变换的矩阵形式

令 $\mathbb{V}$ 和 $\mathbb{W}$ 是两个向量空间，并且 $\dim(\mathbb{V}) = n$，$\dim(\mathbb{W} = m)$ 特别的：

$$\overline{v} = v_1, \cdots, v_n \text{ 和 } \overline{w} = w_1, \cdots, w_m$$

分别是 $\mathbb{V}$ 和 $\mathbb{W}$ 的一组基。

定义 $T : \mathbb{V} \longrightarrow \mathbb{W}$ 是一个线性变换，这意味着对于 $i \in [n]$, $j \in [m]$，存在 $a_{ij} \in \mathbb{R}$ 使得：

$$
\begin{aligned}
T(v_1) &= a_{11}w_1 + \cdots + a_{1m}w_m \\
T(v_2) &= a_{21}w_1 + \cdots + a_{2m}w_m \\
&\vdots \\
T(v_n) &= a_{m1}w_1 + \cdots + a_{nm}w_m
\end{aligned}
$$

利用矩阵表示线性变换 (II)

给定 $v \in \mathbb{V}$，则存在 $c_1, \cdots, c_n \in \mathbb{R}$:

$$v = c_1 v_1 + \cdots + c_n v_n$$

从而:

$$\begin{aligned}
\mathsf{T}(v) &= \mathsf{T}(c_1 v_1 + \cdots + c_n v_n) \\
&= c_1 \mathsf{T}(v_1) + \cdots + c_n \mathsf{T}(v_n) \\
&= c_1(a_{11} w_1 + \cdots + a_{1m} w_m) + \cdots + c_n(a_{m1} w_1 + \cdots + a_{nm} w_m) \\
&= (a_{11} c_1 + \cdots + a_{m1} c_n) w_1 + \cdots + (a_{1m} c_1 + \cdots + a_{nm} c_n) w_m \\
&= \sum_{j \in [m]} \left(\sum_{i \in [n]} c_i a_{ij} \right) w_j
\end{aligned}$$

即 $\mathsf{T}(v)$ 在基 $\mathbb{W}$ 下的坐标为:

$$\left(\sum_{i \in [n]} c_i a_{i1}, \cdots, \sum_{i \in [n]} c_i a_{im} \right)$$

定义下列矩阵:

$$A_T = \begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1m} & \cdots & a_{nM} \end{bmatrix}$$

则 $T(v)$ 在基 $\overline{w}$ 下的坐标为:

$$A_T \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

定义 27.

A_T 被称作线性变换 T 在基 $\overline{v}$ 和 $\overline{w}$ 下的矩阵。

引理 28.

1.
$$\begin{bmatrix} T(v_1) & \cdots & T(v_n) \end{bmatrix} = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} A_T$$

2. A_T 是唯一的，即如果存在 $m \times n$ 的矩阵 B 使得：
$$\begin{bmatrix} T(v_1) & \cdots & T(v_n) \end{bmatrix} = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} B$$
则我们有 $B = A_T$。

证明.

1. 由 A_T 的定义可直接得到。
2. 由 $T(v_j) = \sum_{i\in[n]} a_{ji} w_i$ 的唯一性可以得到。

□

令 $\mathbb{V}$ 和 $\mathbb{W}$ 是两个向量空间，并且 $\dim(\mathbb{V}) = n$, $\dim(\mathbb{W} = m)$ 特别的：

$$\overline{v} = v_1, \cdots, v_n \text{ 和 } \overline{w} = w_1, \cdots, w_m$$

分别是 $\mathbb{V}$ 和 $\mathbb{W}$ 的一组基。我们有：

定理 29.

令 $v \in \mathbb{V}$ 和 $w \in \mathbb{W}$。若 $x \in \mathbb{R}^n$ 是 v 在基 $\overline{v}$ 下的坐标，$y \in \mathbb{R}^m$ 是 $T(v)$ 在基 $\overline{w}$ 下的坐标，则我们有：

$$T(v) = w \quad \Longleftrightarrow \quad y = A_T x$$

这说明，在给定的基下，线性变换可以用一个大小确定的矩阵来表示。

不妨令 $x = (x_1, \cdots, x_n)$, $y = (y_1, \cdots, y_n)$，从而:

$$v = \sum_{i \in [n]} x_i v_i, \quad w = \sum_{j \in [m]} y_j w_j$$

若 $T(v) = w$，则我们有:

$$\begin{aligned}
T(v) = T\left(\sum_{i \in [n]} x_i v_i\right) &= \sum_{i \in [n]} x_i T(v_i) \\
&= \begin{bmatrix} T(v_1) & \cdots & T(v_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} T(v_1) & \cdots & T(v_n) \end{bmatrix} x \\
&= \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} A_T x = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} (A_T x)
\end{aligned}$$

另一方面，注意到 $w_1, \cdots, w_n$ 是线性无关的:

$$w = \sum_{j \in [m]} y_j w_j = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} y = (\begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} A_T) x = \begin{bmatrix} T(v_1) & \cdots & T(v_n) \end{bmatrix} x$$

$A_T A_S$ 与线性变换的复合

令 $\mathbb{U}, \mathbb{V}, \mathbb{W}$ 是三个向量空间，其基分别为 $\overline{u}, \overline{v}, \overline{w}$。

引理 30.

令 $S : \mathbb{U} \longrightarrow \mathbb{V}$ 和 $T : \mathbb{V} \longrightarrow \mathbb{W}$ 是两个线性变换。

- 定义 $TS : \mathbb{U} \longrightarrow \mathbb{W}$ 为：

$$TS(u) = T(S(u))$$

  则 TS 是一个线性变换。

- 若 A_S 和 A_T 分别是 S 和 T 在基 $\overline{u}, \overline{v}$ 和 $\overline{v}, \overline{w}$ 下的矩阵，则 A_{TS} 是 TS 在基 $\overline{u}, \overline{w}$ 下的矩阵，则：

$$A_{TS} = A_T A_S$$

  即 $A_T A_S$ 是 TS 在基 $\overline{u}, \overline{w}$ 下的矩阵。

一个例子-旋转的复合 (I)

考察 $\mathbb{R}^2$ 上的旋转变换:

$$\mathsf{T}:\mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

其将向量逆时针旋转 θ 角度，从而我们有:

$$\mathsf{T}(v)=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}v$$

我们使用如下的标准基:

$$\bar{e}=\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}$$

则:

$$A_{\mathsf{T}}=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$$

一个例子-旋转的复合 (II)

现在考察 $\mathbb{R}^2$ 上的两个旋转变换 (即令 $\mathbb{U}=\mathbb{V}=\mathbb{W}=\mathbb{R}^2$)，令 S 将向量逆时针旋转 θ 角度，T 将向量逆时针旋转 δ 角度，即:

$$S(v)=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}v,\quad T(v)=\begin{bmatrix}\cos\delta & -\sin\delta\\ \sin\delta & \cos\delta\end{bmatrix}v$$

从而:

$$\begin{aligned}TS(v)=T(S(v))&=\begin{bmatrix}\cos\delta & -\sin\delta\\ \sin\delta & \cos\delta\end{bmatrix}\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}v\\ &=\begin{bmatrix}\cos(\theta+\delta) & -\sin(\theta+\delta)\\ \sin(\theta+\delta) & \cos(\theta+\delta)\end{bmatrix}v\end{aligned}$$

而在标准基 e 下我们有:

$$A_{TS}=\begin{bmatrix}\cos(\theta+\delta) & -\sin(\theta+\delta)\\ \sin(\theta+\delta) & \cos(\theta+\delta)\end{bmatrix}=\begin{bmatrix}\cos\delta & -\sin\delta\\ \sin\delta & \cos\delta\end{bmatrix}\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}=A_TA_S$$

引理30的证明

我们先证明 TS 是一个线性变换。给定 $u, u' \in \mathbb{U}$ 和 $c \in \mathbb{R}$:

$$\mathsf{TS}(u+u') = \mathsf{T}(\mathsf{S}(u+u')) = \mathsf{T}(\mathsf{S}(u)+\mathsf{S}(u')) = \mathsf{T}(\mathsf{S}(u)) + \mathsf{T}(\mathsf{S}(u')) = \mathsf{TS}(u) + \mathsf{TS}(u')$$

$$\mathsf{TS}(cu) = \mathsf{T}(\mathsf{S}(cu)) = \mathsf{T}(c\mathsf{S}(u)) = c\mathsf{T}(\mathsf{S}(u)) = c\mathsf{TS}(u)$$

从而 TS 是一个线性变换。

另一方面，令 $\mathsf{TS}(u)$ 在 $\overline{w}$ 下的坐标为 z，$\mathsf{S}(u)$ 在 $\overline{v}$ 下的坐标为 y，u 在 $\overline{u}$ 下的坐标为 x，则我们有:

$$\mathsf{TS}(u) = \mathsf{T}(\mathsf{S}(u)) \quad \Longrightarrow \quad z = A_{\mathsf{T}} y$$

$$\mathsf{S}(u) = \mathsf{S}(u) \quad \Longrightarrow \quad y = A_{\mathsf{S}} x$$

从而我们有:

$$z = A_{\mathsf{T}} A_{\mathsf{S}} x$$

即:

$$A_{\mathsf{TS}} = A_{\mathsf{T}} A_{\mathsf{S}}$$

□

目前为止对线性变换和矩阵的认知

令 $\mathbb{V}$ 和 $\mathbb{W}$ 是两个向量空间，并且 $\dim(\mathbb{V}) = n$，$\dim(\mathbb{W} = m)$，特别的令其基分别为 $\mathbb{v}$ 和 $\mathbb{w}$。

- 我们定义了线性变换，特别的，$\mathbb{V}$ 到 $\mathbb{W}$ 上所有的线性变换为如下的集合：

$$\mathsf{T}(\mathbb{V}, \mathbb{W}) := \{\mathsf{T} \mid \mathsf{T} : \mathbb{V} \longrightarrow \mathbb{W} \text{ 是一个线性变换。}\}$$

  通过合适的加法运算和数乘运算，$\mathsf{T}(\mathbb{V}, \mathbb{W})$ 是一个向量空间。

- 考虑所有的 $m \times n$ 的实矩阵：

$$\mathbb{M}_{m\times n}(\mathbb{R}) = \{A \mid A\text{是一个 } m \times n \text{ 的实矩阵。}\}$$

- 令 $\mathsf{T} \in \mathsf{T}(\mathbb{V}, \mathbb{W})$，$A_{\mathsf{T}} \in \mathbb{M}_{m\times n}(\mathbb{R})$ 存在且唯一，满足：

$$\mathsf{T}(v) = w \quad \Longleftrightarrow \quad y = A_{\mathsf{T}} x$$

这其实说明了，在给定的基下，线性变换和矩阵是一一对应的。

定理 31.

下述映射：

$$T \mapsto A_T$$

是 $T(\mathbb{V}, \mathbb{W})$ 到 $M_{m\times n}(\mathbb{R})$ 的一一映射。

证明. 令 $T_1, T_2 \in T(\mathbb{V}, \mathbb{W})$，并且假设：

$$A_{T_1} = \begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1m} & \cdots & a_{nm} \end{bmatrix}, \quad A_{T_2} = \begin{bmatrix} b_{11} & \cdots & b_{n1} \\ \vdots & \ddots & \vdots \\ b_{1m} & \cdots & b_{nm} \end{bmatrix}$$

即：

$$\begin{array}{c} T_1(v_1) = a_{11}w_1 + \cdots + a_{1m}w_m \\ \vdots \\ T_1(v_n) = a_{n1}w_1 + \cdots + a_{nm}w_m \end{array}, \quad \begin{array}{c} T_2(v_1) = b_{11}w_1 + \cdots + b_{1m}w_m \\ \vdots \\ T_2(v_n) = b_{n1}w_1 + \cdots + b_{nm}w_m \end{array}$$

定理31的证明续. 若 $T_1 \neq T_2$，则存在 $i \in [n]$ 使得

$$T_1(v_i) \neq T_2(v_i)$$

从而 $(a_{i1}, \cdots, a_{im}) \neq (b_{i1}, \cdots, b_{im})$，即 $A_{T_1} \neq A_{T_2}$。从而映射 $T \mapsto A_T$ 是单射。

另一方面，给定 $A \in \mathbb{M}_{m\times n}(\mathbb{R})$，定义：

$$\begin{aligned} u_i &= a_{11}w_1 + \cdots + a_{1m}w_m \\ &\vdots \\ u_n &= a_{n1}w_1 + \cdots + a_{nm}w_m \end{aligned}$$

定义 $T : \mathbb{V} \longrightarrow \mathbb{W}$ 满足对任意的 $i \in [n]$ 有 $T(v_i) = u_i$。则 $T \in T(\mathbb{V}, \mathbb{W})$ 且 $A_T = A$。即映射 $T \mapsto A_T$ 是满射。 $\square$

可以看到 A_T 是依赖选择的基的。如果选择不同的基会发生什么?

定理 32.

令 $\mathbb{V}$ 是一个 n 维的向量空间, $\overline{v}$ 和 $\overline{v'}$ 是 $\mathbb{V}$ 的两组基。考虑 $\mathbb{V} \longrightarrow \mathbb{V}$ 的一个线性变换 T,令:

- M 是 $\overline{v}$ 到 $\overline{v'}$ 的过渡矩阵。
- A 是在基 $\overline{v}$ 下 T 对应的矩阵。
- B 是在基 $\overline{v'}$ 下 T 对应的矩阵。

则我们有:

$$B = M^{-1}AM$$

相似矩阵!

上述定理说明了,在同一个向量空间上不同基下的线性变换对应的矩阵是相似的。

定理32的证明 (I)

M 是 $\overline{v}$ 到 $\overline{v'}$ 的过渡矩阵，从而：

$$\begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} M$$

M 是可逆的，因此：

$$\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} M^{-1}$$

由 A, B 的定义我们有：

$$\begin{bmatrix} T(v_1) & \cdots & T(v_n) \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} A$$
$$\begin{bmatrix} T(v_1') & \cdots & T(v_n') \end{bmatrix} = \begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} B$$

我们将先证明：

$$\begin{bmatrix} T(v_1') & \cdots & T(v_n') \end{bmatrix} = \begin{bmatrix} T(v_1) & \cdots & T(v_n) \end{bmatrix} M$$

定理32的证明 (II)

令:

$$M = \begin{bmatrix} m_{11} & \cdots & m_{n1} \\ \vdots & \ddots & \vdots \\ m_{1n} & \cdots & m_{nn} \end{bmatrix}$$

从而对任意的 $i \in [n]$，我们有:

$$v_i' = \sum_{j \in [n]} m_{ij} v_j$$

从而:

$$T(v_i') = T\left(\sum_{j \in [n]} m_{ij} v_j\right) = \sum_{j \in [n]} m_{ij} T(v_j)$$

即:

$$\begin{bmatrix} T(v_1') & \cdots & T(v_n') \end{bmatrix} = \begin{bmatrix} T(v_1) & \cdots & T(v_n) \end{bmatrix} \begin{bmatrix} m_{11} & \cdots & m_{n1} \\ \vdots & \ddots & \vdots \\ m_{1n} & \cdots & m_{nn} \end{bmatrix} = \begin{bmatrix} T(v_1) & \cdots & T(v_n) \end{bmatrix} M$$

定理32的证明 (III)

从而我们有:

$$\begin{aligned}
\begin{bmatrix} T(v_1') & \cdots & T(v_n') \end{bmatrix} &= \begin{bmatrix} T(v_1) & \cdots & T(v_n) \end{bmatrix} M \\
&= \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} AM \\
&= \begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} M^{-1}AM \\
&= \begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} (M^{-1}AM)
\end{aligned}$$

注意到:

$$\begin{bmatrix} T(v_1') & \cdots & T(v_n') \end{bmatrix} = \begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} B$$

由线性变换矩阵的唯一性可知:

$$B = M^{-1}AM$$

□

选择最好的基

上海师范大学 Shanghai Normal University

我们知道基的选择会影响一个线性变换对应的矩阵。考察 n 维空间 $\mathbb{V}$ 上的一个线性变换，令其一组基为 $\bar{v} = v_1, \cdots, v_n$. 若对于每一个 $i \in [n]$，都存在 $\lambda_i \in \mathbb{R}$ 使得:

$$T(v_i) = \lambda_i v_i$$

则我们有:

$$A_T = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

与矩阵相类似，我们也称满足 $T(v_i) = \lambda_i v_i$ 的 λ_i 和 v_i 为 T 上的特征值和特征向量。

若尔当标准型 (Jordan Form)

根据之前的学习，我们知道并不是所有矩阵都可以对角化。同样的，对于线性变换，我们也不是所有的线性变换都可以找到一组特征向量作为基。但我们可以寻找尽可能接近对角化的形式，这就是若尔当标准型 (Jordan Form)，但我们在这里不多叙述。。

线性变换的矩阵形式

1. 在给定的基下，线性变换可以用一个唯一的矩阵来表示。
2. 线性变换的复合就是矩阵的乘积。
3. 在给定的基下，线性变换和矩阵是一一对应的。
4. 不同基下的线性变换对应的矩阵是相似的。

线性变换的像和核

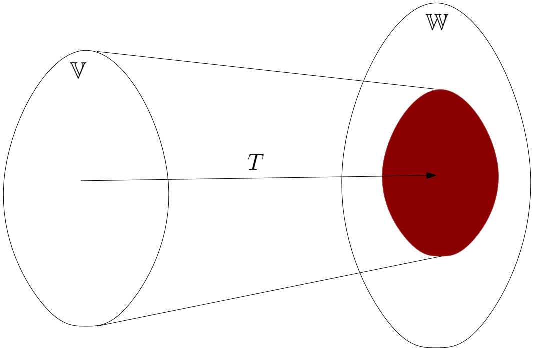
上海师范大学 Shanghai Normal University

本节我们固定两个向量空间 $\mathbb{V}$ 和 $\mathbb{W}$，并且考虑 $\mathbb{V}$ 到 $\mathbb{W}$ 上的一个线性变换 T。

定义 33 **[像 (Image)].**

T 的像 (Image) 是:

$$\mathrm{Im}(\mathrm{T}) := \{\mathrm{T}(\mathrm{v}) \mid \mathrm{v} \in \mathbb{V}\}$$

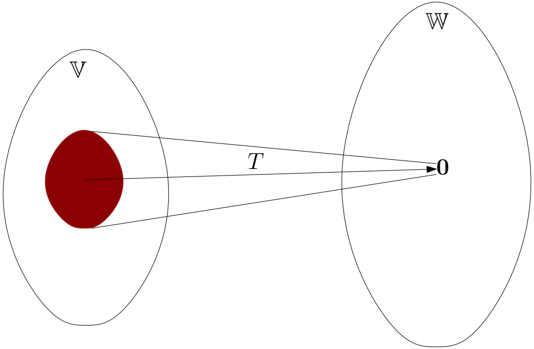


核的定义

定义 34 **[核 (Kernel)].**

T 的核 (Kernel) 是:

$$\mathrm{Ker}(T) := \{v \in \mathbb{V} \mid T(v) = \mathbf{0}\}$$



引理 35.

令 $\mathbb{V}$ 和 $\mathbb{W}$ 是两个向量空间，$\mathsf{T}:\mathbb{V}\longrightarrow\mathbb{W}$ 是一个线性变换。则:

1. $\mathrm{Im}(\mathsf{T})$ 是 $\mathbb{W}$ 的一个子空间。
2. $\mathrm{Ker}(\mathsf{T})$ 是 $\mathbb{V}$ 的一个子空间。

矩阵 A 和线性变换 T_A

考虑一个 $m \times n$ 的矩阵 A:

$$T_A(x) = Ax$$

是一个 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 上的线性变换。我们有:

$$\mathrm{Im}(T_A) = \{T_A(x) \mid x \in \mathbb{R}^n)\} = \{Ax \mid x \in \mathbb{R}^n\} = C(A)$$

$$\mathrm{Ker}(T_A) = \{x \mid T_A(x) = \mathbf{0}\} = \{x \in \mathbb{R}^n \mid Ax = \mathbf{0}\} = N(A)$$

推论 36.

1. 线性变换 T 是单射当且仅当:

$$\mathrm{Ker}(T) = \{\mathbf{0}\}$$

2. 假设 $\mathbb{W}$ 是有限维的，线性变换 T 是满射当且仅当:

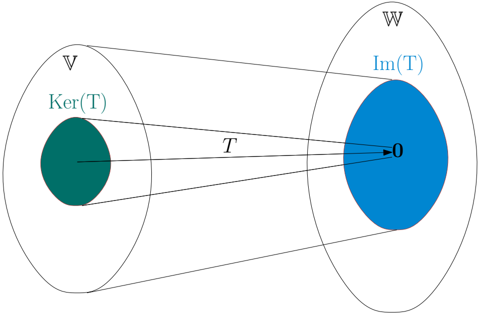
$$\dim(\mathrm{Im}(T)) = \dim(\mathbb{W})$$

Rank-Nullity 定理

定理 37 **[Rank-Nulity 定理].**

令 $\mathbb{V}$ 和 $\mathbb{W}$ 是两个有限维的向量空间，$\mathsf{T}:\mathbb{V}\longrightarrow\mathbb{W}$ 是一个线性变换。则：

$$\dim(\mathbb{V})=\dim(\mathsf{Im}(\mathsf{T}))+\dim(\mathsf{Ker}(\mathsf{T}))$$



推论 38.

令 A 是 $m\times n$ 的矩阵，则：

$$n=\dim(C(A))+\dim(N(A))=\mathsf{rank}(A)+\dim(N(A))$$

假设 $\mathbb{V}$ 存在一组基 $\overline{v} = v_1, \cdots, v_n$，特别的 $\dim(\mathbb{V}) = n$。注意到:

$$\mathrm{Im}(T) = \mathrm{span}\{T(v_1), \cdots, T(v_n)\}$$

从而 $\{T(v_1), \cdots, T(v_n)\}$ 中的一组极大的线性无关的子集构成了 $\mathrm{Im}(T)$ 的一组基。令其为:

$$\overline{w} = w_1, \cdots, w_m$$

从而 $\dim(\mathrm{Im}(T)) = m$，不妨令对于所有的 $i \in [m]$:

$$T(v_i) = w_i$$

否则我们重新排列 v_i 的顺序。

Rank-Nullity 定理的证明 (II)

Ker(T) 是 $\mathbb{V}$ 的子空间，从而 Ker(T) 存在一组基:

$$u_1, \cdots, u_p$$

其中 $p \leqslant n$。

我们下面证明:

$$v_1, \cdots, v_m, u_1, \cdots, u_p \text{是 } \mathbb{V} \text{ 中的一组基。}$$

若上述命题成立，则立马可以得到:

$$\dim(\mathbb{V}) = \dim(\mathrm{Im}(T)) + \dim(\mathrm{Ker}(T))$$

引理 39.

$v_1, \cdots, v_m, u_1, \cdots, u_p$ 是 $\mathbb{V}$ 中的一组基。

证明. 令 $v \in \mathbb{V}$，则 $\mathrm{T}(v) \in \mathrm{Im}(\mathrm{T})$，从而存在 $x_1, \cdots, x_n \in \mathbb{R}$ 使得:

$$\mathrm{T}(v) = \sum_{i=1}^{m} x_i w_i = \sum_{i=1}^{m} x_i \mathrm{T}(v_i) = \mathrm{T}\left(\sum_{i=1}^{m} x_i v_i\right)$$

从而:

$$\mathrm{T}\left(v - \sum_{i=1}^{m} x_i v_i\right) = \mathbf{0} \implies v - \sum_{i=1}^{m} x_i v_i \in \mathrm{Ker}(\mathrm{T})$$

即:

$$v - \sum_{i=1}^{m} x_i v_i \in \mathrm{span}\{u_1, \cdots, u_p\} \implies v = \sum_{i=1}^{m} x_i v_i + \sum_{i=1}^{p} y_i u_i$$

从而我们有:

$$v \in \mathrm{span}\{v_1, \cdots, v_m, u_1, \cdots, u_p\} \implies \mathbb{V} = \mathrm{span}\{v_1, \cdots, v_m, u_1, \cdots, u_p\}$$

上海师范大学 Shanghai Normal University

引理的证明续. 我们还需要证明$v_1, \cdots, v_m, u_1, \cdots, u_p$ 是线性无关的。假设存在 $x_1, \cdots, x_m, y_1, \cdots, y_p \in \mathbb{R}$ 使得:

$$\sum_{i=1}^{m} x_i v_i + \sum_{i=1}^{p} y_i u_i = \mathbf{0}$$

则:

$$\begin{aligned}
\mathbf{0} &= T\left(\sum_{i=1}^{m} x_i v_i + \sum_{i=1}^{p} y_i u_i\right) \\
&= \sum_{i=1}^{m} x_i T(v_i) + \sum_{i=1}^{p} y_i T(u_i) = \sum_{i=1}^{m} x_i T(v_i) = \sum_{i=1}^{m} x_i w_i
\end{aligned}$$

由 $w_1, \cdots, w_m$ 是线性无关的，从而对任意的 $i \in [m]$ 都有 $x_i = 0$，因此我们有:

$$\sum_{i=1}^{p} y_i u_i = \mathbf{0}$$

由 $u_1, \cdots, u_j$ 是线性无关的，我们也可以得到 $y_1 = \cdots = y_p = 0$。从而 $v_1, \cdots, v_m, u_1, \cdots, u_p$ 是线性无关的。 $\square$

一个应用

我们来展现一个 Rank-Nullity 定理的应用。

令 A 是一个 $m \times n$ 的矩阵，B 是一个 $n \times l$ 的矩阵。我们有：

$$\operatorname{rank}(AB) \leqslant \min\{\operatorname{rank}(A), \operatorname{rank}(B)\}$$

我们进一步将证明：

$$\operatorname{rank}(A) + \operatorname{rank}(B) \leqslant \operatorname{rank}(AB) + n$$

一个简短但神秘的证明

$$\begin{aligned}
\operatorname{rank}(AB) + n &= \operatorname{rank}\left(\begin{bmatrix} AB & O \\ O & I \end{bmatrix}\right) = \operatorname{rank}\left(\begin{bmatrix} AB & O \\ B & I \end{bmatrix}\right) \\
&= \operatorname{rank}\left(\begin{bmatrix} O & -A \\ B & I \end{bmatrix}\right) = \operatorname{rank}\left(\begin{bmatrix} B & I \\ O & -A \end{bmatrix}\right) \\
&= \operatorname{rank}\left(\begin{bmatrix} B & -I \\ O & A \end{bmatrix}\right) \geqslant \operatorname{rank}(A) + \operatorname{rank}(B)
\end{aligned}$$

从 Rank-Nullity 定理角度的证明 (I)

由 Rank-Nullity 定理我们有:

$$\mathrm{rank}(A) + \mathrm{rank}(B) \leqslant \mathrm{rank}(AB) + n$$

等价于:

$$(n - \dim(N(A))) + (l - \dim(N(B))) \leqslant (l - \dim(N(AB))) + n$$

即要证:

$$\dim(N(A)) + \dim(N(B)) \geqslant \dim(N(AB))$$

更加清楚的直观

其实从上述式子可以看到，这个不等式是比较显然的，因为 $N(AB)$ 的维数显然不可能超过 $N(A)$ 与 $N(B)$ 的维数之和。

令 $x \in \mathbb{R}^l$，则我们有:

$$x \in N(AB) \iff ABx = \mathbf{0} \iff Bx \in N(A)$$

定义:

$$W = \{Bx \mid x \in \mathbb{R}^l,\ Bx \in N(A)\} \implies W = C(B) \cap N(A) \subseteq \mathbb{R}^n$$

从而 W 是 $N(A)$ 的一个子空间。定义线性变换 $T : N(AB) \implies \mathbb{R}^n$:

$$T(x) = Bx$$

从而我们有:

$$\mathrm{Im}(T) = W$$

$$\mathrm{Ker}(T) = \{x \in N(AB) \mid Bx = 0\} = N(B)$$

从而由 Rank-Nullity 定理我们有:

$$\dim(N(AB)) = \dim(\mathrm{Im}(T)) + \dim(\mathrm{Ker}(T)) = \dim(W) + \dim(N(B)) \leqslant \dim(N(A)) + \dim(N(B))$$

$\square$

阶段总结

- 线性变换的核与像，矩阵的零空间与列空间。
- Rank-Nullity 定理。

对偶性 (Duality)

线性泛函 (Linear Functional)

我们固定一个向量空间 $\mathbb{V}$。我们考察其到 $\mathbb{R}$ 上的线性映射

定义 40 **[线性泛函].**

一个 $\mathbb{V}$ 上的泛函 L 是一个线性映射 $f:\mathbb{V}\longrightarrow\mathbb{R}$，即：$L\in T(\mathbb{V},\mathbb{R})$。

例 41.

1. $F:\mathbb{R}^3\Longrightarrow\mathbb{R}$:
$$F(x,y,z)=4x-5y+2z$$
是一个线性泛函。
2. 固定 $(c_1,\cdots,c_n)\in\mathbb{R}^n$，定义 $F:\mathbb{R}^n\Longrightarrow\mathbb{R}$:
$$F(x)=\sum_{i=1}^{n}c_ix_i$$
是一个线性泛函。

定义 42 **[对偶空间 (Dual Space)].**

$\mathbb{V}$ 的对偶空间$\mathbb{V}'$是 $\mathbb{V}$ 上所有线性泛函的集合。即:

$$\mathbb{V}' = \mathrm{T}(\mathbb{V}, \mathbb{R})$$

引理 43.

$\mathbb{V}'$ 是一个向量空间。

我们先给出对偶基 (Dual Basis) 的概念。

定义 44 **[对偶基 (Dual Basis)].**

令

$$v_1, \cdots, v_n$$

是 V 中的一组基。则 $v_1, \cdots, v_n$ 的对偶基是如下的系列:

$$L_1, \cdots, L_n$$

其中，L_i 是 V 上的线性泛函，满足:

$$L_i(v_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

对偶基是对偶空间的一组基 (I)

我们下面证明，我们给出的对偶基的概念，恰恰是对偶空间的一组基。

定理 45.

假设 $\mathbb{V}$ 是有限维的，则 $\mathbb{V}'$ 中一组基的对偶基是 $\mathbb{V}'$ 的一组基。

推论 46.

$\dim(\mathbb{V}) = \dim(\mathbb{V}')$。

证明. 令 $v_1, \cdots, v_n$ 是 $\mathbb{V}$ 的一组基，$L_1, \cdots, L_n$ 是其对偶基。假设存在 $c_1, \cdots, c_n \in \mathbb{R}$ 使得:

$$T = c_1 L_1 + \cdots + c_n L_n = 0$$

则对任意的 $i \in [n]$:

$$0 = T(v_i) = (c_1 L_1 + \cdots + c_n L_n)(v_i) = c_i L_i(v_i) = c_i$$

从而 $c_1 = \cdots = c_n = 0$，即 $L_1, \cdots, L_n$ 是线性无关的。

证明续. 另一方面，考虑 $T \in \mathbb{V}' = T(\mathbb{V}, \mathbb{R})$，即 T 是一个 $\mathbb{V}$ 到 $\mathbb{R}$ 的线性映射。

对任意的 $i \in [n]$，令：

$$c_i = T(v_i)$$

则定义：

$$T' = c_1 L_1 + \cdots + c_n L_n$$

显然 $T' \in T(\mathbb{V}, \mathbb{R}) = \mathbb{V}'$，并且对任意的 $i \in [n]$：

$$T'(v_i) = c_i L_i(v_i) = c_i = T(v_i)$$

即 $T' = T$，从而：

$$\mathbb{V}' = \mathrm{span}\{L_1, \cdots, L_n\}$$

$\square$

令 $\mathbb{V}$ 和 $\mathbb{W}$ 是两个向量空间。

定义 47 [对偶变换 (Dual Transformation)].

对于一个线性变换 $T \in \mathcal{T}(\mathbb{V}, \mathbb{W})$，定义其对偶变换 $T' \in (\mathbb{W}', \mathbb{V}')$:

$$T'(L) = LT, \text{ 即对任意的 } v \in \mathbb{V} \text{ 有 } T'(L)(v) = L(T(v))$$

我们要说明两件事:

1. 对任意的 $L \in \mathbb{W}'$，$T'(L)$ 是 $\mathbb{V}'$ 上的线性泛函，即是 $\mathbb{V}$ 到 $\mathbb{R}$ 上的线性变换。
2. T' 是一个 $\mathbb{W}'$ 到 $\mathbb{V}'$ 上的线性变换。

注意到:

$$T'(L) = LT,\ \text{即对任意的 } v \in \mathbb{V} \text{ 有 } T'(L)(v) = L(T(v))$$

我们只需验证 $T'(L)$ 是 $\mathbb{V}'$ 到 $\mathbb{R}$ 上的线性变换即可。

- 向量加法: 令 $v_1, v_2 \in \mathbb{V}$, 则:

$$\begin{aligned} T'(L)(v_1 + v_2) &= L(T(v_1 + v_2)) = L(T(v_1) + T(v_2)) \\ &= L(T(v_1)) + L(T(v_2)) = T'(L)(v_1) + T'(L)(v_2) \end{aligned}$$

- 数乘: 令 $v \in \mathbb{V}$, $c \in \mathbb{R}$, 则:

$$T'(L)(cv) = L(T(cv)) = L(cT(v)) = cL(T(v)) = cT'(L)(v)$$

注意到:

$$T'(L) = LT, \text{ 即对任意的 } v \in \mathbb{V} \text{ 有 } T'(L)(v) = L(T(v))$$

- 向量加法：令 $L_1, L_2 \in \mathbb{W}' = T(\mathbb{W}, \mathbb{R})$，则对于任意的 $v \in \mathbb{V}$:

$$\begin{aligned} T'(L_1 + L_2)(v) &= (L_1 + L_2)(T(v)) = L_1(T(v)) + L_2(T(v)) \\ &= T'(L_1)(v) + T'(L_2)(v) \end{aligned}$$

即 $T'(L_1 + L_2) = T'(L_1) + T'(L_2)$.

- 数乘：令 $L \in \mathbb{W}' = T(\mathbb{W}, \mathbb{R})$，$c \in \mathbb{R}$，则对于任意的 $v \in \mathbb{V}$:

$$T'(cL)(v) = (cL)(T(v)) = cL(T(v)) = cT'(L)(v)$$

即 $T'(cL) = cT'(L)$.

给定两个向量空间 $\mathbb{V}$ 和 $\mathbb{W}$:

- 令 $\overline{v} = v_1 \cdots v_n$ 是 $\mathbb{V}$ 的一组基，$L_1^{\overline{v}}, \cdots, L_n^{\overline{v}}$ 是其对偶基，也是 $\mathbb{V}'$ 的一组基。
- 令 $\overline{w} = w_1 \cdots w_m$ 是 $\mathbb{W}$ 的一组基，$L_1^{\overline{w}}, \cdots, L_m^{\overline{w}}$ 是其对偶基，也是 $\mathbb{W}'$ 的一组基。

定理 48.

令 $T \in \mathcal{T}(\mathbb{V}, \mathbb{W})$ 并且 A_T 是 T 在基 $\overline{v}$ 和 $\overline{w}$ 下的矩阵。令 $A_{T'}$ 是其对偶变换 $T' \in \mathcal{T}(\mathbb{W}', \mathbb{V}')$ 在基 $L_1^{\overline{w}}, \cdots, L_m^{\overline{w}}$ 和 $L_1^{\overline{v}}, \cdots, L_n^{\overline{v}}$ 下的矩阵。则:

$$A_{T'} = A_T^{\top}$$

记

$$A_T = \begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1m} & \cdots & a_{nm} \end{bmatrix}, \qquad A_{T'} = \begin{bmatrix} b_{11} & \cdots & b_{m1} \\ \vdots & \ddots & \vdots \\ b_{1n} & \cdots & b_{mn} \end{bmatrix}$$

即:

$$\begin{aligned} T(v_1) &= a_{11}w_1 + \cdots + a_{1m}w_m \\ T(v_2) &= a_{21}w_1 + \cdots + a_{2m}w_m \\ &\vdots \\ T(v_n) &= a_{n1}w_1 + \cdots + a_{nm}w_m \end{aligned}, \qquad \begin{aligned} T'(L_1^{\overline{W}}) &= b_{11}L_1^{\overline{V}} + \cdots + b_{1n}L_n^{\overline{V}} \\ T'(L_2^{\overline{W}}) &= b_{21}L_1^{\overline{V}} + \cdots + b_{2n}L_n^{\overline{V}} \\ &\vdots \\ T'(L_m^{\overline{W}}) &= b_{m1}L_1^{\overline{V}} + \cdots + b_{mn}L_n^{\overline{V}} \end{aligned}$$

定理48的证明 (II)

注意到对偶基的定义:

$$L_i^V(v_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

从而对于任意的 $i \in [m]$, $j \in [n]$ 有:

$$T'(L_i^W)(v_j) = (b_{i1}L_1^V + \cdots + b_{in}L_n^V)(v_j) = b_{ij}L_j^V(v_j) = b_{ij}$$

另一方面，由定义可得:

$$\begin{aligned} T'(L_i^W)(v_j) = L_i^W(T(v_j)) &= L_i^W(a_{j1}w_1 + \cdots + a_{jm}w_m) \\ &= a_{j1}L_i^W(w_1) + \cdots + a_{jm}L_i^W(w_m) \\ &= a_{ji}L_i^W(w_i) = a_{ji} \end{aligned}$$

从而:

$$b_{ij} = T'(L_i^W)(v_j) = a_{ji}$$

$\square$

现在我们来考虑一下对偶变换中的像与核。

定义 49 **[零化子 (Annihilator)].**

令 $\mathbb{V}$ 是一个向量空间，$\mathbb{U} \subseteq \mathbb{V}$ 是其子空间。$\mathbb{U}$ 的零化子，记作 $\mathbb{U}^0$，定义如下：

$$\mathbb{U}^0 = \{\mathrm{L} \in \mathbb{V}' \mid \mathrm{L}(\mathrm{u}) = 0,\ \forall \mathrm{u} \in \mathbb{U}\}$$

引理 50.

$\mathbb{U}^0$ 是 $\mathbb{V}'$ 的一个子空间。

$\mathbb{U}^0$ 是子空间的证明

- 显然零函数 **0** $\in \mathbb{U}^0$。
- 零 L_1，$L_2 \in \mathbb{U}^0$, 则对任意的 $u \in \mathbb{U}$，有:

$$(L_1 + L_2)(u) = L_1(u) + L_2(u) = 0 + 0 = 0$$

即 $L_1 + L_2 \in \mathbb{U}^0$。

- $L \in \mathbb{U}^0$，$c \in \mathbb{R}$，则对任意的 $u \in \mathbb{U}$，有:

$$(cL)(u) = cL(u) = c \cdot 0 = 0$$

即 $cL \in \mathbb{U}^0$。

$\square$

定理 51.

令 $\mathbb{V}$ 是一个有限维的向量空间，$\mathbb{U} \subseteq \mathbb{V}$ 是其子空间。则：

$$\dim(\mathbb{U}) + \dim(\mathbb{U}^0) = \dim(\mathbb{V})$$

证明. 我们使用如下的包含映射 (Inclusion Linear Transformation)：$\mathsf{T}_{\text{incl}} : \mathbb{U} \longrightarrow \mathbb{V}$：

$$\text{对任意的 } u \in \mathbb{U} \text{ 有} \mathsf{T}_{\text{incl}}(u) = u$$

则其对偶映射为 $\mathsf{T}'_{\text{incl}} \in \mathsf{T}(\mathbb{V}', \mathbb{U}')$，则由 Rank-Nullity 定理有：

$$\dim(\mathbb{V}') = \dim(\mathsf{Im}(\mathsf{T}'_{\text{incl}})) + \dim(\mathsf{Ker}(\mathsf{T}'_{\text{incl}}))$$

我们将证明：

- $\mathsf{Im}(\mathsf{T}'_{\text{incl}}) = \mathbb{U}'$。
- $\mathsf{Ker}(\mathsf{T}'_{\text{incl}}) = \mathbb{U}^0$。

从而：

$$\dim(\mathbb{V}) = \dim(\mathbb{V}') = \dim(\mathbb{U}') + \dim(\mathbb{U}^0) = \dim(\mathbb{U}) + \dim(\mathbb{U}^0)$$

令 $L \in \mathbb{U}' = T(\mathbb{U}, \mathbb{R})$。注意到 $\mathbb{V}$ 是有限维的，不妨令其一组基为；

$$v_1, \cdots, v_m, v_{m+1}, \cdots, v_n$$

注意到 $\mathbb{U}$ 是 $\mathbb{V}$ 的子空间，不妨进一步假设 $v_1, \cdots, v_m$ 是 $\mathbb{U}$ 的一组基。我们定义 $L_V \in T(\mathbb{V}, \mathbb{R})$：

$$L_V(v_i) = \begin{cases} L(v_i) & i \in [m] \\ 0 & i \in [m+1, n] \end{cases}$$

则对任意的 $u \in \mathbb{U}$ 有：

$$T'_{\mathrm{incl}}(L_V)(u) = L_V(T_{\mathrm{incl}}(u)) = L_V(u) = L(u)$$

这里最后一个等号是因为 $u \in \mathbb{U}$ 可以写成：

$$u = u_1 v_1 + \cdots + u_m v_m + 0 v_{m+1} + \cdots + 0 v_n$$

的形式，从而：

$$\mathrm{Im}(T'_{\mathrm{incl}}) = \mathbb{U}'$$

$\mathrm{Ker}(T'_{\mathrm{incl}}) = \mathbb{U}^0$

上海师范大学 Shanghai Normal University

对于 $\mathrm{Ker}(T'_{\mathrm{incl}})$ 有:

$$\begin{aligned}
\mathrm{Ker}(T'_{\mathrm{incl}}) &= \{L \in \mathbb{V}' \mid \mathrm{Ker}(T'_{\mathrm{incl}})(L) = \mathbf{0}\} \\
&= \{L \in \mathbb{V}' \mid \text{对任意的 } u \in \mathbb{U} \text{ 有 } T'_{\mathrm{incl}}(L)(u) = 0\} \\
&= \{L \in \mathbb{V}' \mid \text{对任意的 } u \in \mathbb{U} \text{ 有 } L(T_{\mathrm{incl}}(u)) = 0\} \\
&= \{L \in \mathbb{V}' \mid \text{对任意的 } u \in \mathbb{U} \text{ 有 } L(u) = 0\} \\
&= \mathbb{U}^0
\end{aligned}$$

$\square$

定理 52.

假设 $\mathbb{V}$ 和 $\mathbb{W}$ 是有限维的向量空间，令 $\mathsf{T} \in \mathsf{T}(\mathbb{V}, \mathbb{W})$，则:

1. $\mathsf{Ker}(\mathsf{T}') = (\mathsf{Im}(\mathsf{T}))^0$。
2. $\dim(\mathsf{Ker}(\mathsf{T}')) = \dim(\mathsf{Ker}(\mathsf{T})) + \dim(\mathbb{W}) - \dim(\mathbb{V})$。

证明. 我们先利用第一个结论来证明第二个，注意到:

$$\begin{aligned}
\dim(\mathsf{Ker}(\mathsf{T}')) &= \dim((\mathsf{Im}(\mathsf{T}))^0) \\
&= \dim(\mathbb{W}) - \dim(\mathsf{Im}(\mathsf{T})) \\
&= \dim(\mathbb{W}) - (\dim(\mathbb{V}) - \dim(\mathsf{Ker}(\mathsf{T}))) \\
&= \dim(\mathsf{Ker}(\mathsf{T})) + \dim(\mathbb{W}) - \dim(\mathbb{V})
\end{aligned}$$

- 令 $L \in \mathrm{Ker}(T')$，即 $T'(L) = \mathbf{0} \in \mathbb{V}' = T(\mathbb{V}, \mathbb{R})$，也就是对任意的 $v \in \mathbb{V}$ 有

$$T'(L)(v) = L(T(v)) = 0$$

从而对于任意的 $w \in \mathrm{Im}(T)$，我们都有 $L(w) = 0$，即 $L \in (\mathrm{Im}(T))^0$，即:

$$\mathrm{Ker}(T') \subseteq (\mathrm{Im}(T))^0$$

- 另一方面，考察 $L \in (\mathrm{Im}(T))^0$，这意味着对任意的 $w \in \mathrm{Im}(T)$ 都有

$$L(w) = 0$$

从而对任意的 $u \in \mathbb{V}$:

$$T'(L)(u) = L(T(u)) = 0$$

即: $L \in \mathrm{Ker}(T)'$，也就是$(\mathrm{Im}(T))^0 \subseteq \mathrm{Ker}(T')$。

$\square$

定理 53.

假设 $\mathbb{V}$ 和 $\mathbb{W}$ 是有限维的向量空间，令 $T \in \mathcal{T}(\mathbb{V}, \mathbb{W})$，则：

1. $\dim(\operatorname{Im}(T')) = \dim(\operatorname{Im}(T))$.
2. $\operatorname{Im}(T') = (\operatorname{Ker}(T))^0$。

证明.

1. $\dim(\operatorname{Im}(T')) = \dim(\mathbb{W}') - \dim(\operatorname{Ker}(T')) = \dim(\mathbb{W}) - \dim((\operatorname{Im}(T))^0) = \dim(\operatorname{Im}(T))$.
2. 令 $L \in \operatorname{Im}(T') \subseteq \mathbb{V}'$，则存在 $L_w \in \mathbb{W}'$ 使得：$T'(L_w) = L$

   注意到对任意的 $u \in \operatorname{Ker}(T)$ 我们有：

   $$L(u) = T'(L_w)(u) = L_w(T(u)) = L_w(\mathbf{0}) = 0$$

   从而$L \in (\operatorname{Ker}(T))^0$，即 $\operatorname{Im}(T)' \subseteq (\operatorname{Ker}(T))^0$，另一方面，我们有：

   $$\dim(\operatorname{Im}(T)') = \dim(\operatorname{Im}(T)) = \dim(\mathbb{V}) - \dim(\operatorname{Ker}(T)) = \dim((\operatorname{Ker}(T))^0)$$

   从而$\operatorname{Im}(T)' = (\operatorname{Ker}(T))^0$。

定理 54.

令 A 是一个 $m \times n$ 的矩阵，则：row-rank(A) = column-rank(A).

证明. 定义 $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$:

$$T(v) = Av$$

特别的，在标准基 $e_1, \cdots, e_n$ 下的 T 的矩阵为 A。从而其对偶变换 T' 在对应的对偶基下的矩阵为

$$A^T$$

因此:

$$\begin{aligned}\text{column-rank}(A) &= \dim(C(A)) = \dim(\mathrm{Im}(T)) \\ &= \dim(\mathrm{Im}(T')) = \dim(C(A^T)) = \text{column-rank}(A^T) = \text{row-rank}(A)\end{aligned}$$

□

阶段总结

- 线性泛函的概念。
- 对偶空间、对偶基。
- 对偶变换，和转置矩阵的关系。
- 对偶变换的核与像的性质。