线性映射

颜文斌

清华大学

内容提要

- **线性映射和矩阵**
- 线性映射的性质
- 换基
- 若当标准型

映射（mapping）

- **映射**：S，S'是两个集合，如果对于任意S中的元素，都有一个S'中的元素和它对应，这个对应就叫一个映射
- 如果我们把一个映射记做$f: S \to S'$。u是S中的一个元素，$f(u)$叫做u在映射f下的**像**（image）。W是S中的一个子集，$f(W)$叫做W在映射f下的**像**
- 映射的**复合**（composition）：$f: U \to V, g: V \to W$。
$$g \circ f: U \to W, \forall\, x \in U, \qquad g \circ f(x) = g(f(x))$$
- 映射的复合满足**结合律**：$g \circ f \circ h = (g \circ f) \circ h = g \circ (f \circ h)$

映射

- **单射**（injection）：映射$F: S \to S'$是单射，如果$\forall\, x, y \in S, x \neq y, \mathrm{s.t.}\, F(x) \neq F(y)$
- **满射**（surjection）：映射$F: S \to S'$是满射，如果$F(S) = S'$
- **双射**（bijection）：既是单射也是满射的映射
- **恒等映射**（identity map）：$I_S: S \to S, \mathrm{s.t.}\, \forall\, x \in S, I_S(x) = x$
- **逆映射**（inverse map）：对于映射$F: S \to S'$如果存在一个映射$G: S' \to S$使得$G \circ F = I_S$和$F \circ G = I_{S'}$，则称映射$F: S \to S'$**可逆**，G被称作F的**逆**

可逆映射的性质

- **定理**：映射$f\colon S \to S'$可逆当且仅当f是双射
- 证明：
  - 首先假设$x, y \in S$，$g\colon S' \to S$是 f的逆
  - 如果$f(x) = f(y)$，则$x = g\big(f(x)\big) = g\big(f(y)\big) = y$，所以$f$是单射
  - $\forall\, z \in S'$，我们有$f\big(g(z)\big) = z$，所以$\exists\, x = g(z), \mathrm{s.t.}\, f(x) = z$，所以$f$是满射。这样我们就从左边推出了右边
  - 接下来假设$f\colon S \to S'$是双射，那么因为f是满射$\forall\, z \in S'$，我们有$x \in S, \mathrm{s.t.}\, f(x) = z$。又因为$f$是单射，所以$x$是唯一的。所以可以定义$g(z) = x$，则$g$是$f$的逆映射。这样我们就从右边推出了左边

线性映射（linear mapping）

- **定义（线性映射）**：V，W是两个（实）线性空间，映射$T: V \to W$是（实）线性映射，如果T满足以下两个条件

1. 对于任意的$\boldsymbol{u}, \boldsymbol{v} \in V$，有$T(\boldsymbol{u}+\boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$
2. 对于任意的实数c，有$T(c\boldsymbol{u}) = cT(\boldsymbol{u})$

- 线性映射也被称为**线性变换**（linear transformation）
- **推论**：$T: V \to W$是线性映射，则$T(\mathbf{0}) = \mathbf{0}$

线性映射例

- 例1：$\boldsymbol{a} = (1, 2, 3), \boldsymbol{v} = (x, y, z)$。映射$T(v) = \boldsymbol{a} \cdot \boldsymbol{v}$是从$\mathbb{R}^3$到$\mathbb{R}$的线性映射
- 例2：$T(\boldsymbol{v}) = \boldsymbol{a} \cdot \boldsymbol{v} + b$不是线性的，因为0没有映到0
- 例3：$T(\boldsymbol{v}) = \|\boldsymbol{v}\|$不是线性的，因为$\|\boldsymbol{v} + \boldsymbol{w}\| \le \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$，而且$\|-\boldsymbol{v}\| = \|\boldsymbol{v}\|$
- 例4：假设$T(\boldsymbol{v})$是线性映射，$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_r\boldsymbol{v}_r = \boldsymbol{0}$，则$T(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_r\boldsymbol{v}_r) = c_1T(\boldsymbol{v}_1) + \cdots + c_rT(\boldsymbol{v}_r) = \boldsymbol{0}$

线性映射例

- 例5：考虑所有可导函数$f(x)$构成的线性空间，映射$T(f)=\frac{df}{dx}$是线性的
  - $\frac{d}{dx}\big(af(x)+bg(x)\big)=a\frac{df(x)}{dx}+b\frac{dg(x)}{dx}$
- 例6（恒等映射）：$\mathrm{id}: V\to V,\ \mathrm{id}(\boldsymbol{u})=\boldsymbol{u}$
- 例7（零映射）：$\forall\,\boldsymbol{u}\in V, T(\boldsymbol{u})=0$
- 例8：考虑mxn的矩阵A，定义线性映射$L_A: \mathbb{R}^n\to\mathbb{R}^m$，使得$\forall\,\boldsymbol{x}\in\mathbb{R}^n$，$L_A(\boldsymbol{x})=A\boldsymbol{x}$

怎么决定线性映射？

- **定理**：V和W是线性空间，$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$是$V$中的一组基，$\{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n\}$是$W$中的任意n个元素。则存在唯一的线性映射$T: V \to W$使得
$$T(\boldsymbol{v}_1) = \boldsymbol{w}_1, \cdots, T(\boldsymbol{v}_n) = \boldsymbol{w}_n$$
- **证明（存在性）**：
  - $\forall\, \boldsymbol{v} \in V$，$\boldsymbol{v}$可以唯一写成$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$的线性组合$\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n$
  - 定义映射$T(\boldsymbol{v}) = c_1\boldsymbol{w}_1 + \cdots + c_n\boldsymbol{w}_n$，下面证明$T(\boldsymbol{v})$是线性映射
  - $T(c\boldsymbol{v}) = T(cc_1\boldsymbol{v}_1 + \cdots + cc_n\boldsymbol{v}_n) = cc_1\boldsymbol{w}_1 + \cdots + cc_n\boldsymbol{w}_n = cT(\boldsymbol{v})$
  - 假设$\boldsymbol{u} = x_1\boldsymbol{v}_1 + \cdots + x_n\boldsymbol{v}_n$，$T(\boldsymbol{v} + \boldsymbol{u}) = T((c_1+x_1)\boldsymbol{v}_1 + \cdots + (c_n+x_n)\boldsymbol{v}_n) = (c_1+x_1)\boldsymbol{w}_1 + \cdots + (c_n+x_n)\boldsymbol{w}_n) = T(\boldsymbol{v}) + T(\boldsymbol{u})$

怎么决定线性映射？

- **定理**：V和W是线性空间，$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$是$V$中的一组基，$\{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n\}$是$W$中的任意n个元素。则存在唯一的线性映射$T: V \to W$使得
$$T(\boldsymbol{v}_1) = \boldsymbol{w}_1, \cdots, T(\boldsymbol{v}_n) = \boldsymbol{w}_n$$
- **证明（唯一性）**：
  - 假设存在另一个线性映射$F: V \to W$使得$F(\boldsymbol{v}_1) = \boldsymbol{w}_1, \cdots, F(\boldsymbol{v}_n) = \boldsymbol{w}_n$
  - $\forall \boldsymbol{v} \in V$，，$\boldsymbol{v}$可以唯一写成$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$的线性组合$\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n$
  - $F(\boldsymbol{v}) = c_1F(\boldsymbol{v}_1) + \cdots + c_nF(\boldsymbol{v}_n) = c_1\boldsymbol{w}_1 + \cdots + c_n\boldsymbol{w}_n = c_1T(\boldsymbol{v}_1) + \cdots + c_nT(\boldsymbol{v}_n) = T(\boldsymbol{v})$

怎么决定线性映射？

- **定理**：V和W是线性空间，$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$是$V$中的一组基，$\{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n\}$是$W$中的任意n个元素。则存在唯一的线性映射$T: V \to W$使得
$$T(\boldsymbol{v}_1) = \boldsymbol{w}_1, \cdots, T(\boldsymbol{v}_n) = \boldsymbol{w}_n$$
- 只要知道一个线性映射在基上的值，就唯一决定了整个线性映射

矩阵和线性映射

- 考虑mxn的矩阵，定义线性映射$L_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$，使得$\forall\, \boldsymbol{x} \in \mathbb{R}^n$, $L_A(\boldsymbol{x}) = A\boldsymbol{x}$
- 定理：mxn矩阵A、B，如果$L_A = L_B$，则$A = B$
- 证明：
  - 根据L_A, L_B定义可知$\forall\, \boldsymbol{x} \in \mathbb{R}^n, A\boldsymbol{x} = B\boldsymbol{x}$，所以$A_i \cdot \boldsymbol{x} = B_i \cdot \boldsymbol{x}$，其中$A_i$（$B_i$）是$A$（$B$）的第i行。
  - $(A_i - B_i) \cdot \boldsymbol{x} = 0$，也就是说向量$A_i - B_i$和$\mathbb{R}^n$中的所有向量都正交，所以$A_i - B_i$只能是零向量
  - 上面的论证对任意行都成立，所以$A = B$

线性映射和矩阵（标准基版本）

- **定理**：设$L: \mathbb{R}^n \to \mathbb{R}^m$是线性映射，则存在唯一的矩阵$A$使得$L = L_A$
- 证明：
  - 假设$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$是$\mathbb{R}^n$中的标准基，$\{\boldsymbol{f}_1, \cdots, \boldsymbol{f}_m\}$是$\mathbb{R}^m$中的标准基，任何$\mathbb{R}^n$中的向量$\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$
  - $L(\boldsymbol{x}) = x_1 L(\boldsymbol{e}_1) + \cdots + x_n L(\boldsymbol{e}_n)$
  - 因为$L(\boldsymbol{e}_i)$是$\mathbb{R}^m$中的向量，所以可以写成$\{\boldsymbol{f}_1, \cdots, \boldsymbol{f}_m\}$的线性组合，也就是说$L(\boldsymbol{e}_i) = a_{1i}\boldsymbol{f}_1 + \cdots + a_{mi}\boldsymbol{f}_{\boldsymbol{m}}$

线性映射和矩阵（标准基版本）

- **定理**：设$L: \mathbb{R}^n \to \mathbb{R}^m$是线性映射，则存在唯一的矩阵$A$使得$L = L_A$
- 证明（续）：
  - 因为$L(\boldsymbol{e}_i)$是$\mathbb{R}^m$中的向量，所以可以写成$\{\boldsymbol{f}_1, \cdots, \boldsymbol{f}_m\}$的线性组合，也就是说$L(\boldsymbol{e}_i) = a_{1i}\boldsymbol{f}_1 + \cdots + a_{mi}\boldsymbol{f}_m$
  - $L(\boldsymbol{x}) = x_1 L(\boldsymbol{e}_1) + \cdots + x_n L(\boldsymbol{e}_n) = x_1(a_{11}\boldsymbol{f}_1 + \cdots + a_{m1}\boldsymbol{f}_m) + \cdots + x_n(a_{1n}\boldsymbol{f}_1 + \cdots + a_{mn}\boldsymbol{f}_m) = (a_{11}x_1 + \cdots a_{1n}x_n)\boldsymbol{f}_1 + \cdots + (a_{m1}x_1 + \cdots a_{mn}x_n)\boldsymbol{f}_m = \begin{bmatrix} a_{11}x_1 + \cdots a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots a_{mn}x_n \end{bmatrix}$
  - $L(\boldsymbol{x}) = A\boldsymbol{x} = L_A(\boldsymbol{x})$。我们证明了存在性，唯一性由之前定理保证

线性映射和矩阵（标准基版本）

- **定理**：设$L: \mathbb{R}^n \to \mathbb{R}^m$是线性映射，则存在唯一的矩阵$A$使得$L = L_A$
- 说明：$L\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = A\boldsymbol{x} = \begin{bmatrix} a_{11}x_1 + \cdots a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots a_{mn}x_n \end{bmatrix} \in \mathbb{R}^m$，给出了矩阵和向量乘法的自然定义

例

- 例1:$F(x,y,z)=(x,y)$，对应的矩阵为$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}$
- 例2:$\mathrm{id}: \mathbb{R}^n \to \mathbb{R}^n$，对应的矩阵为单位矩阵$\begin{bmatrix}1 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & 1\end{bmatrix}$

内容提要

- 线性映射和矩阵
- **线性映射的性质**
- 换基
- 若当标准型

线性映射的加法

- $T: V \to W$，$F: V \to W$为两个线性映射，我们定义这两个线性映射的**和**为一个新的线性映射$H = T + F: V \to W$，满足$\forall\, \boldsymbol{u} \in V$，$(T + F)(\boldsymbol{u}) = T(\boldsymbol{u}) + F(\boldsymbol{u})$
- $T + F$也是一个线性映射
- 例：$L_A(\boldsymbol{x}) = A\boldsymbol{x},\ L_B(\boldsymbol{x}) = B\boldsymbol{x}$
  - $(L_A + L_B)(\boldsymbol{x}) = A\boldsymbol{x} + B\boldsymbol{x} = (A + B)\boldsymbol{x} = L_{A+B}(\boldsymbol{x})$
- 利用线性映射和矩阵的对应，线性映射的加法等价于矩阵的加法
- 加法的零元：零映射$\forall\, \boldsymbol{u} \in V, T(\boldsymbol{u}) = 0$
- 加法的逆：$(-T)(\boldsymbol{u}) = -T(\boldsymbol{u})$，$(-T) + T$是零映射

线性映射的数乘

- $T: V \to W$ 为线性映射，我们定义这线性映射和c的**数乘**为一个新的线性映射$F = cT: V \to W$，满足$\forall\, \boldsymbol{u} \in V$，$(cT)(\boldsymbol{u}) = c(T(\boldsymbol{u}))$
- cT也是一个线性映射
- 例：$L_A(\boldsymbol{x}) = A\boldsymbol{x}$
  - $(cL_A)(\boldsymbol{x}) = cA\boldsymbol{x} = L_{cA}(\boldsymbol{x})$
- 利用线性映射和矩阵的对应，线性映射的数乘等价于矩阵的数乘
- $1T = T$，$-T = (-1)T$

线性映射构成线性空间

- **性质**：所有从$V \rightarrow W$的线性映射集合$\{T\}$构成一个线性空间
  - 加法和数乘按照之前的定义
  - 加法零元是零映射
  - 满足线性空间的所有8条公理（交换律、结合律、分配律等等）
- 和dimVxdimW的矩阵构成的线性空间有一一对应

线性映射的核（Kernel）

- **定义**：线性映射$F: V \to W$的核$\mathrm{Ker}\ F$是所有满足$F(\boldsymbol{v}) = \mathbf{0}$的向量$\boldsymbol{v}$的集合
- **性质**：$\mathrm{Ker}\ F$是V中的一个线性子空间
- 说明：如果$V = \mathbb{R}^n$，$W = \mathbb{R}^m$，矩阵A为F对应的矩阵，则$\mathrm{Ker}\ F = N(A)$

核和单射

- **性质**：$F: V \to W$是线性映射，以下两个条件是等价的
  1. $\mathrm{Ker}\, F$只有零向量
  2. F是单射，换句话说，如果V中的元素$\boldsymbol{v}, \boldsymbol{w}$满足$F(\boldsymbol{v}) = F(\boldsymbol{w})$，则$\boldsymbol{v} = \boldsymbol{w}$
- 证明：
  - （矩阵版本）对应的矩阵零空间是0， $A\boldsymbol{v} = \boldsymbol{b}$如果有解则必有唯一解。
  - （抽象版本）假设V中的元素$\boldsymbol{v}, \boldsymbol{w}$满足$F(\boldsymbol{v}) = F(\boldsymbol{w})$，则$F(\boldsymbol{v} - \boldsymbol{w}) = F(\boldsymbol{v}) - F(\boldsymbol{w}) = \boldsymbol{0}$
  - 因为假设，$\mathrm{Ker} F$只有零向量，则$\boldsymbol{v} - \boldsymbol{w} = \boldsymbol{0}$，所以$\boldsymbol{v} = \boldsymbol{w}$

核的性质

- **定理**：$F: V \to W$是线性映射且$\mathrm{Ker}\, F = \{\mathbf{0}\}$。如果$\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$是$V$中的线性无关向量，则$F(\boldsymbol{v}_1), \cdots, F(\boldsymbol{v}_n)$是$W$中线性无关的向量
- 证明：
  - （矩阵版本）对应矩阵的零空间为$\{\mathbf{0}\}$，所以列满秩，列之间线性无关
  - （抽象版本）假设$x_1 F(\boldsymbol{v}_1) + \cdots + x_n F(\boldsymbol{v}_n) = 0$
  - 根据线性性，$F(x_1 v_1 + \cdots + x_n v_n) = 0$
  - 又因为$\mathrm{Ker} F = \{\mathbf{0}\}$，则$x_1 v_1 + \cdots + x_n v_n = 0$
  - $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$是线性无关的，则$x_1 = x_2 = \cdots = x_n = 0$，所以$F(\boldsymbol{v}_1), \cdots, F(\boldsymbol{v}_n)$线性无关

线性映射的像（image）

- **定义**：线性映射$F: V \to W$的像$\mathrm{Im}F$是W中所有在V中有原像的向量的集合（$\forall\, \boldsymbol{w} \in \mathrm{Im}F, \exists\, \boldsymbol{v} \in V, s.t. F(\boldsymbol{v}) = \boldsymbol{w}$）
- 性质： $\mathrm{Im}F$是W中的线性子空间
- 说明：如果$V = \mathbb{R}^n$， $W = \mathbb{R}^m$，矩阵A为F对应的矩阵，则$\mathrm{Im}F = C(A)$

像和核的关系

- **定理**：V是一个线性空间，$L: V \to W$是一个线性映射，则
$$\dim V = \dim \operatorname{Ker} L + \dim \operatorname{Im} L$$
- 证明：
  - （矩阵版本）$\dim V = \dim N(A) + \dim C(A^T) = \dim N(A) + \dim C(A)$
  - （抽象版本）如果$\operatorname{Im} L = \{0\}$，则$\operatorname{Ker} L = V$，则定理成立
  - 否则假设$\{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_s\}$是$\operatorname{Im} L$的一组基。设$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_s\}$分别是$\{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_s\}$的原像，也就是说$L(\boldsymbol{v}_i) = \boldsymbol{w}_i$，$i = 1, \cdots, s$
  - 另，如果$\operatorname{Ker} L \neq \{0\}$，假设$\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_q\}$是$\operatorname{Ker} L$的一组基
  - 如果我们能证明$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_s, \boldsymbol{u}_1, \cdots, \boldsymbol{u}_q\}$是中的一组基，则定理得证

像和核的关系

- **定理**：V是一个线性空间，$L: V \to W$是一个线性映射，则
$$\dim V = \dim \operatorname{Ker} L + \dim \operatorname{Im} L$$
- 证明：
  - 如果我们能证明$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_s, \boldsymbol{u}_1, \cdots, \boldsymbol{u}_q\}$是中的一组基，则定理得证
  - 首先要证明任何一个V中的向量可以写成$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_s, \boldsymbol{u}_1, \cdots, \boldsymbol{u}_q\}$的线性组合
  - $\forall\, \boldsymbol{v} \in V, L(\boldsymbol{v}) \in \operatorname{Im} L$，所以$L(\boldsymbol{v}) = x_1\boldsymbol{w}_1 + \cdots + x_s\boldsymbol{w}_s$
  - 因为$L(\boldsymbol{v}_i) = \boldsymbol{w}_i$，所以$L(x_1\boldsymbol{v}_1 + \cdots + x_s\boldsymbol{v}_s) = x_1\boldsymbol{w}_1 + \cdots + x_s\boldsymbol{w}_s = L(\boldsymbol{v})$
  - 所以$L(\boldsymbol{v}) - L(x_1\boldsymbol{v}_1 + \cdots + x_s\boldsymbol{v}_s) = L(\boldsymbol{v} - (x_1\boldsymbol{v}_1 + \cdots + x_s\boldsymbol{v}_s)) = 0$
  - $\boldsymbol{v} - (x_1\boldsymbol{v}_1 + \cdots + x_s\boldsymbol{v}_s) \in \operatorname{Ker} L$，可以写成$\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_q\}$的线性组合
  - 所以$\boldsymbol{v} - (x_1\boldsymbol{v}_1 + \cdots + x_s\boldsymbol{v}_s) = y_1u_1 + \cdots + y_qu_q$

像和核的关系

- **定理**：V是一个线性空间，$L: V \to W$是一个线性映射，则
$$\dim V = \dim \operatorname{Ker} L + \dim \operatorname{Im} L$$
- 证明：
  - 如果我们能证明$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_s, \boldsymbol{u}_1, \cdots, \boldsymbol{u}_q\}$是中的一组基，则定理得证
  - 接下来要证明 $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_s, \boldsymbol{u}_1, \cdots, \boldsymbol{u}_q\}$线性无关
  - 假设 $(x_1\boldsymbol{v}_1 + \cdots + x_s\boldsymbol{v}_s) + \left(y_1\boldsymbol{u}_1 + \cdots + y_q\boldsymbol{u}_q\right) = \boldsymbol{0}$
  - $\boldsymbol{0} = L\left((x_1\boldsymbol{v}_1 + \cdots + x_s\boldsymbol{v}_s) + \left(y_1\boldsymbol{u}_1 + \cdots + y_q\boldsymbol{u}_q\right)\right) = L(x_1\boldsymbol{v}_1 + \cdots + x_s\boldsymbol{v}_s) = x_1\boldsymbol{w}_1 + \cdots + x_s\boldsymbol{w}_s$，$\{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_s\}$线性无关，所以$x_i = 0$
  - 所以$y_1\boldsymbol{u}_1 + \cdots + y_q\boldsymbol{u}_q = \boldsymbol{0}$，又因为$\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_q\}$线性无关，所以$y_i = 0$

例

- 线性映射$L: \mathbb{R}^3 \rightarrow \mathbb{R}$：$L(x, y, z) = 3x - 2y + z$
- 对应的矩阵：$A = (3 \quad -2 \quad 1)$
- 像：$\mathbb{R}$，一维
- 核：$C\left(\begin{bmatrix} 2/3 & -1/3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$，二维

例

- 线性映射$\frac{d}{dx}: P^n \to P^{n-1}$：$P^n$是最高次数不超过n的多项式构成的线性空间，$\dim P^n = n + 1$
- 像：P^{n-1}，维度为n
- 核：常数P^0，维度为1

双射的像和核

- **定理**： $L: V \to W$是一个线性映射，假设$\dim V = \dim W$。如果$\operatorname{Ker} L = \{\mathbf{0}\}$，或者$\operatorname{Im} L = W$，则$L$是双射
- 证明：
  - 假设$\operatorname{Im} L = W$，根据定义， L是满射
  - 又由前面的定理可知$\dim \operatorname{Ker} L = \dim V - \dim \operatorname{Im} L = \dim W - \dim \operatorname{Im} L = 0$，所以$\operatorname{Ker} L = \{\mathbf{0}\}$。由此可知$L$是单射
  - 所以L是双射
  - $\operatorname{Ker} L = \{\mathbf{0}\}$的证明类似

内容提要

- 线性映射和矩阵
- 线性映射的性质
- **换基**
- 若当标准型

线性空间、基、坐标向量

- 假设$\mathcal{B} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$是线性空间$V$上的一组基
- 对于V中的任意向量$\boldsymbol{v}$，我们有$\boldsymbol{v} = x_1\boldsymbol{v}_1 + \cdots + x_n\boldsymbol{v}_n$
- **坐标向量**：$x_{\mathcal{B}}(v) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$被称作向量$\boldsymbol{v}$在基$\mathcal{B}$下的坐标向量
- 通过坐标向量和基$\mathcal{B}$，我们有V和$\mathbb{R}^n$的一一映射
  - $\boldsymbol{v} \to \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$
  - 这个映射和基的选取有关

线性变换和矩阵

- $L: V \to W$是一个线性映射， $\mathcal{B} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$是$V$上的一组基，$\mathcal{B}' = \{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m\}$是$W$上的一组基
- 对于V中的任意向量$\boldsymbol{v}$，我们有$\boldsymbol{v} = x_1\boldsymbol{v}_1 + \cdots + x_n\boldsymbol{v}_n$
  - $L(\boldsymbol{v}) = x_1 L(\boldsymbol{v}_1) + \cdots + x_n L(\boldsymbol{v}_n)$
  - $L(\boldsymbol{v}_i)$是W中的向量，所以$L(\boldsymbol{v}_i) = a_{1i}\boldsymbol{w}_1 + \cdots + a_{mi}\boldsymbol{w}_m$
  - $L(\boldsymbol{v}) = x_1(a_{11}\boldsymbol{w}_1 + \cdots + a_{m1}\boldsymbol{w}_m) + \cdots + x_n(a_{1n}\boldsymbol{w}_1 + \cdots + a_{mn}\boldsymbol{w}_m) = (a_{11}x_1 + \cdots a_{1n}x_n)\boldsymbol{w}_1 + (a_{m1}x_1 + \cdots + a_{mn}x_n)\boldsymbol{w}_m$
- $L(\boldsymbol{v})$在$\mathcal{B}'$上的坐标向量是$(a_{11}x_1 + \cdots a_{1n}x_n, \cdots, a_{m1}x_1 + \cdots + a_{mn}x_n)^T$

线性变换和矩阵

- $L: V \to W$是一个线性映射， $\mathcal{B} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$是$V$上的一组基，$\mathcal{B}' = \{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m\}$是$W$上的一组基
- **定理**：存在唯一的mxn矩阵$M_{\mathcal{B}'}^{\mathcal{B}}(L)$，使得$\forall\, \boldsymbol{v} \in V$，$x_{\mathcal{B}'}\big(L(\boldsymbol{v})\big) = M_{\mathcal{B}'}^{\mathcal{B}}(L) x_{\mathcal{B}}(\boldsymbol{v})$。$x_{\mathcal{B}}(\boldsymbol{v})$是 $\boldsymbol{v}$ 在 $\mathcal{B}$ 上的坐标向量， $x_{\mathcal{B}'}\big(L(\boldsymbol{v})\big)$是$L(\boldsymbol{v})$在$\mathcal{B}'$上的坐标向量
- 构造：$M_{\mathcal{B}'}^{\mathcal{B}}(L)$的第j列就是$L\big(\boldsymbol{v}_j\big)$在$\mathcal{B}'$上的坐标向量，换句话说$L\big(\boldsymbol{v}_j\big) = \sum_i \boldsymbol{w}_i \big(M_{\mathcal{B}'}^{\mathcal{B}}(L)\big)_{ij}$
- 注：线性映射对应的矩阵跟基的选取有关，之前的结果是两组基都取标准基的特殊情形

例

- 例1：$L\colon \mathbb{R}^n \to \mathbb{R}^m$，我们之前学习了怎么构造$M_{e_m}^{e_n}(L)$
- 例2：$L\colon V \to W$，$\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$，$\mathcal{B}' = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3\}$
  - $L(\boldsymbol{v}_1) = 3\boldsymbol{w}_1 - \boldsymbol{w}_2 + 17\boldsymbol{w}_3$，$L(\boldsymbol{v}_2) = \boldsymbol{w}_1 + \boldsymbol{w}_2 - \boldsymbol{w}_3$
  - $M_{\mathcal{B}'}^{\mathcal{B}}(L) = \begin{bmatrix} 3 & 1 \\ -1 & 1 \\ 17 & -1 \end{bmatrix}$

例

- 例3

Example 3 The input basis of $\boldsymbol{v}$'s is $1, x, x^2, x^3$. The output basis of $\boldsymbol{w}$'s is $1, x, x^2$. Then $\boldsymbol{T}$ **takes the derivative**: $\boldsymbol{T(v)} = \dfrac{\boldsymbol{dv}}{\boldsymbol{dx}}$ and A = "derivative matrix".

If $\boldsymbol{v} = c_1 + c_2x + c_3x^2 + c_4x^3$

then $\dfrac{d\boldsymbol{v}}{dx} = \mathbf{1}c_2 + \mathbf{2}c_3x + \mathbf{3}c_4x^2$

$$A\boldsymbol{c} = \begin{bmatrix} 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{2} & 0 \\ 0 & 0 & 0 & \mathbf{3} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} c_2 \\ 2c_3 \\ 3c_4 \end{bmatrix}$$

$M_{\mathcal{B}'}^{\mathcal{B}}$的性质

- $M_{\mathcal{B}'}^{\mathcal{B}}$是所有 $L: V \rightarrow W$ 线性变换到$\dim W \times \dim V$矩阵的映射
- $M_{\mathcal{B}'}^{\mathcal{B}}$是一个线性映射
  - $M_{\mathcal{B}'}^{\mathcal{B}}(f + g) = M_{\mathcal{B}'}^{\mathcal{B}}(f) + M_{\mathcal{B}'}^{\mathcal{B}}(g)$
  - $M_{\mathcal{B}'}^{\mathcal{B}}(cf) = cM_{\mathcal{B}'}^{\mathcal{B}}(f)$
- $M_{\mathcal{B}'}^{\mathcal{B}}$是一个**双射**

线性变换的复合与矩阵乘法

- 考虑两个线性映射$L_1: U \to V$， $L_2: V \to W$ 。$\mathcal{B} = \{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_l\}$是$V$上的一组基，$\mathcal{B}' = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$是$V$上的一组基，$\mathcal{B}'' = \{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_m\}$是$W$上的一组基
- $M_{\mathcal{B}''}^{\mathcal{B}}(L_2 \circ L_1) = M_{\mathcal{B}''}^{\mathcal{B}'}(L_2) M_{\mathcal{B}'}^{\mathcal{B}'}(L_1)$
- 线性映射的**复合**等价于对应矩阵的**乘法**
  - 同一个线性空间V: 前一个矩阵的列数=后一个矩阵的行数
  - 可以自然得到矩阵乘法的规则

同一个线性空间，不同的基

- 线性空间V上有两组基， $\mathcal{B}=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$， $\mathcal{B}'=\{\boldsymbol{w}_1,\cdots,\boldsymbol{w}_n\}$。同一个向量$\boldsymbol{v}$在不同的基上有不同的坐标向量$x_{\mathcal{B}}(\boldsymbol{v})$， $x_{\mathcal{B}'}(\boldsymbol{v})$
- **推论**：$x_{\mathcal{B}'}(\boldsymbol{v})=M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id})x_{\mathcal{B}}(\boldsymbol{v})$，或者说$\boldsymbol{v}_j=\sum_i \boldsymbol{w}_i\left(M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id})\right)_{ij}$
- 注意：一般$M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id})$不是单位矩阵，但是$M_{\mathcal{B}}^{\mathcal{B}}(\mathrm{id})=I$
- 利用线性变换复合和矩阵乘法的对应，我们有
- **定理**：$M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id})M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id})=I=M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id})M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id})$，也就是说$M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id})$**可逆**，而且逆是$M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id})$

换基

- **定理**：线性映射$L: V \to W$，$\mathcal{B}, \mathcal{B}'$是$V$上两组基，$\mathcal{C}, \mathcal{C}'$是$W$上的两组基。则
$$M_{\mathcal{C}'}^{\mathcal{B}'}(L) = M_{\mathcal{C}'}^{\mathcal{C}}(\mathrm{id}) M_{\mathcal{C}}^{\mathcal{B}}(L) M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}) = \left(M_{\mathcal{C}}^{\mathcal{C}'}(\mathrm{id})\right)^{-1} M_{\mathcal{C}}^{\mathcal{B}}(L) M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id})$$
- 证明：用线性映射的复合和矩阵乘法的对应
- **推论**：线性映射$L: V \to V$，$\mathcal{B}, \mathcal{B}'$是$V$上两组基，则
$$M_{\mathcal{B}'}^{\mathcal{B}'}(L) = M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}) M_{\mathcal{B}}^{\mathcal{B}}(L) M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}) = \left(M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id})\right)^{-1} M_{\mathcal{B}}^{\mathcal{B}}(L) M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id})$$
- 所以矩阵相似变换就是换基，矩阵对角化就是找到描述同一个线性变换的最好的基

矩阵分解和换基

- LDU分解：$PA = LDU$，$L^{-1}PAU^{-1} = D$， D相抵标准型
- QR分解：$A = QR$，$Q^{-1}A = R$，正交变换变成上三角
- 对角化：$\Lambda = Q^{-1}AQ$， Λ相似标准型
  - 对称矩阵可以正交对角化：正交变换
- 奇异值分解：$A = U\Sigma V$，$\Sigma = U^{T}AV^{T}$

内容提要

- 线性映射和矩阵
- 线性映射的性质
- 换基
- **若当标准型**

矩阵对角化

- 不是所有方阵都可以对角化
- nxn矩阵A可以对角化
  - 有n个线性独立的特征向量
  - 所有特征值的几何重数=代数重数
- 如果nxn矩阵A只有s<n个线性独立的特征向量，怎么把A变成最接近对角矩阵的形式？

若当标准型（Jordan normal form）

- **定理**：nxn矩阵A有s个线性独立的特征向量，则存在可逆矩阵B，使得$B^{-1}AB = \begin{bmatrix} J_1 & 0 & 0 & 0 \\ 0 & J_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & J_s \end{bmatrix}$，其中$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \lambda_i & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \lambda_i \end{bmatrix}$，其中$\lambda_i$是第i个线性独立的特征向量的特征值