矩阵分析及其应用

孔贺

南科大，自动化与智能制造学院

kongh@sustech.edu.cn

课程结构

- 一. 线性代数基本知识回顾（2 个学时）
- 二. 线性空间与线性映射（10 个学时）
- 三. λ-矩阵与 Jordan 标准型（8 个学时）
- 四. 内积空间与矩阵分解（20 个学时）
- 五. 范数及其他（8 个学时）

2.1 线性空间

定义：给定一个非空集合 $\mathbf{V}$ 和数域 $\mathbf{F}$。若

- 1. 存在映射

$$\sigma : \mathbf{V} \times \mathbf{V} \to \mathbf{V},$$

  也记为 $(\alpha,\beta) \mapsto \sigma(\alpha,\beta)$, 即对于 $\mathbf{V}$ 中的任意两个元素 α 和 β，在 $\mathbf{V}$ 中都有唯一的元素 ϵ 与它们对应，称 ϵ 为 α 和 β 的和，记为 $\epsilon = \alpha + \beta$，我们称 σ 为 $\mathbf{V}$ 上的“加法”。

- 2. 在集合 $\mathbf{V}$ 和数域 $\mathbf{F}$ 之间存在映射

$$\gamma : \mathbf{V} \times \mathbf{F} \to \mathbf{V},$$

  也记为 $(v,k) \mapsto \gamma(v,k)$, 即对于 $\mathbf{V}$ 中的任意一个元素 v 和数域 $\mathbf{F}$ 中的任一数 k，在 $\mathbf{V}$ 中都有唯一的元素 η 与它们对应，称之为 k 和 v 的数乘，记为 $\eta = k \cdot v = kv$，我们称 γ 为集合 $\mathbf{V}$ 和数域 $\mathbf{F}$ 之间的“数乘法”。

- 3. 且以上两种映射（运算）满足“通常的运算法则”。

我们则称集合 $\mathbf{V}$ 为数域 $\mathbf{F}$ 上的线性空间。

几点补充说明

- 关于数域 $\mathbf{F}$：包括 0 和 1，对于 $+-\times\div$ 四种运算封闭。最常用的为实数域 $\mathbf{R}$ 和虚数域 $\mathbf{C}$。
- 集合的笛卡尔积（Cartesian product）：$\mathbf{V}\times\mathbf{F}$，$\mathbf{V}\times\mathbf{V}$
- 要习惯用映射的观念去理解运算
- 映射的两种表示方式：(1) $\gamma:\mathbf{X}\to\mathbf{Y}$; (2) $\gamma: x\mapsto y$
- 数乘法的数在左边或者右边的区别
- “加法”和“数乘法”
- 定义中的“隐性要求”

重要说明："通常的运算法则"，8 条

加法

- 1. 交换律
- 2. 结合律
- 3. 零元素
- 4. 负元素

数乘法

- 5. 对集合 $\mathbf{V}$ 中元素加法的分配律
- 6. 对数域 $\mathbf{F}$ 中数加法的分配律
- 7. 与数域 $\mathbf{F}$ 中乘法的关系
- 8. 用数 1 做乘法

线性空间的经典例子

- 1. 数域 $\mathbf{F}$ 上的标准线性空间 $\mathbf{F}^n$。
- 2. 几何空间作为线性空间：$\mathbf{V}$ 为空间中有向线段的集合，数域 $\mathbf{F}$ 为实数域 $\mathbf{R}$；加法为平行四边形法则，数乘法为同向或反向伸缩（因此我们也将线性空间及其元素分别称向量空间及向量）
- 3. 函数空间 $\mathcal{F}(I, \mathbf{R}^n)$，即 $\mathcal{F}$ 为定义在 I 上，取值于 $\mathbf{R}^n$ 的函数的集合；其中 I 为数轴上的一个区间；加法和数乘法的常规定义）。

与矩阵 $A \in \mathbf{R}^{m\times n}$ 有关的四个子空间（第二次见面）

- (1) 列空间：$\mathbf{V}=\{y \in \mathbf{R}^m \mid y = Ax,\ A \in \mathbf{R}^{m\times n}, x \in \mathbf{R}^n\}$，即 A 的像空间（或值域），常用 $im(A)$ (或 $R(A)$，$C(A)$) 表示
- (2) 零空间：$\mathbf{V}=\{x \mid Ax = 0,\ A \in \mathbf{R}^{m\times n}, x \in \mathbf{R}^n\}$，即齐次线性方程组的解 $Ax = 0$ 构成的集合，即 $Ax = 0$ 的解空间，也称为矩阵 A 的核（或零）空间，常用 $ker(A)$ (或 $N(A)$) 表示
- (3) 行空间：$R(A^{\mathrm{T}}) \subset \mathbf{R}^n$
- (4) 左零空间：$N(A^{\mathrm{T}}) \subset \mathbf{R}^m$

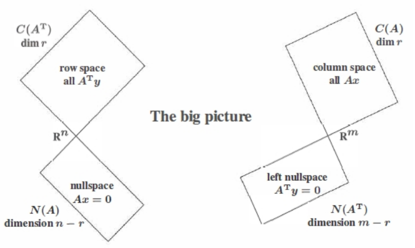

图: Fig. 3.5, Gilbert Strang, Introduction to Linear Algebra (5th Edition)

其他例子

- 1. 维数为 $m \times n$ 的实矩阵的集合 $\mathbf{R}^{m \times n}$。
- 2. 维数为 m 的实向量的集合 $\mathbf{R}^{m}$。
- 3. 次数小于 n，变量为 x 的实系数多项式的集合。

以上集合均为实数域上的线性空间。

“特殊”例子

1. 我们用 $\mathbf{R}$ 和 $\mathbf{R}^{++}$ 来分别表示所有实数和正实数的集合，且分别定义“加法” $\oplus$ 和 “数乘法”$\odot$：

- $a \oplus b := ab$， $\forall a, b \in \mathbf{R}^{++}$
- $k \odot a := a^k$， $\forall a \in \mathbf{R}^{++}, k \in \mathbf{R}$

可以验证，$\mathbf{R}^{++}$ 构成实数域 $\mathbf{R}$ 上的线性空间。

不是线性空间的例子

- 1. 假设 $\mathbf{V}$ 是由系数在实数域 $\mathbf{R}$ 上且次数为 n 的 n 次多项式构成的集合，其加法和数乘法按照常规来定义，则 $\mathbf{V}$ 不是 $\mathbf{R}$ 上的线性空间。
- 2. 假设非齐次线性方程组 $Ax = b$ 有解；但是，其解所组成的集合不构成线性空间。

回忆：上次课讲过的“向量组”与“向量组线性相关性”

5. 向量组的线性无关性

- （1）给定 $v_1, v_2, \cdots, v_n$ 为 $\mathbf{R}^n$ 中的一个向量组，如果存在不全为零的 n 个实数，$k_1, k_2, \cdots, k_n \in \mathbf{R}$，使得

$$v_1k_1 + v_2k_2 + \cdots + v_nk_n = 0,$$

  则称向量组为线性相关的。
- （2）如果向量组不是线性相关的，则称它为线性无关的。

向量组、抽象矩阵及线性相关性

定义：假设集合 $\mathbf{V}$ 为数域 $\mathbf{F}$ 上的线性空间

- 称 $\mathbf{V}$ 中的向量组成的有限序列 $v_1, v_2, \cdots, v_n$ 为一个向量组；
- 向量组按顺序排成的行/列，称为向量组拼成的抽象矩阵，记为 $\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$

定义：向量组的线性相关性

- 1. 如果存在不全为零的 n 个数，$k_1, k_2, \cdots, k_n \in \mathbf{F}$，使得

$$v_1 k_1 + v_2 k_2 + \cdots + v_n k_n = 0,$$

  则称向量组为线性相关的。
- 2. 如果向量组不是线性相关的，则称它为线性无关的。

说明：向量组线性相关性的矩阵表示。

定理：假设集合 $\mathbf{V}$ 为数域 $\mathbf{F}$ 上的线性空间；$v_1, v_2, \cdots, v_n$，β 为 $\mathbf{V}$ 中的向量；且

- 1. $v_1, v_2, \cdots, v_n$ 组成的向量组线性无关。
- 2. $v_1, v_2, \cdots, v_n$，β 组成的向量组线性相关。

则我们有

- 1. β 可由 $v_1, v_2, \cdots, v_n$ 线性表出。
- 2. 且表出是唯一的。

两个向量组之间的线性表示关系

定义：假设 $\mathbf{V}$ 为 $\mathbf{F}$ 上的线性空间，v 为 $\mathbf{V}$ 中的任一向量，$\alpha_1, \alpha_2, \cdots, \alpha_p$ 和 $\beta_1, \beta_2, \cdots, \beta_q$ 为 $\mathbf{V}$ 中的两个向量组。

- 1．如果存在 p 个数，$k_1, k_2, \cdots, k_p \in \mathbf{F}$，使得

$$\alpha_1 k_1 + \alpha_2 k_2 + \cdots + \alpha_p k_p = v, \tag{1}$$

  则称向量 v 可由向量组 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 线性表示。

- 2．如果每个 β_i, $i = 1, 2, \cdots, q$, 都可以由向量组 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 线性表示，则称向量组 $\beta_1, \beta_2, \cdots, \beta_q$ 可以由向量组 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 线性表示。通常记为 $\{\beta_j\} \leq_{lin} \{\alpha_i\}$。

说明：矩阵表达；线性表示的传递性。

向量组的极大线性无关子组

定义：假设 $\mathbf{V}$ 为 $\mathbf{F}$ 上的线性空间，$\alpha_1, \alpha_2, \cdots, \alpha_p$ 为 $\mathbf{V}$ 中的一个向量组。向量组 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 的一个子组 $\beta_1, \beta_2, \cdots, \beta_s$ 为其极大线性无关子组，如果后者满足

- 1. 无关性：$\{\beta_j\}$ 线性无关。
- 2. 极大性：若 $\{\gamma_1, \gamma_2, \cdots, \gamma_k\}$ 也是 $\{\alpha_i\}$ 的子组，同时有 $\{\beta_j\}$ 是 $\{\gamma_t\}$ 的子组且 $s < k$，则 $\{\gamma_1, \gamma_2, \cdots, \gamma_k\}$ 线性相关。

说明：

- 1. “极大性”等同于“生成性”。
- 2. 极大线性无关子组不必是唯一的。
- 3. 但是不同的极大线性无关子组所含向量的个数是唯一的（反证法；扁的齐次方程方程有非零解）；通常称为向量组的秩。

2.2 基、坐标与坐标变换

假设 $\mathbf{V}$ 为 $\mathbf{F}$ 上的线性空间。若存在正整数 n 及 $\mathbf{V}$ 中的向量组 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 使得:

- 1. $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关。
- 2. 任意向量 $v \in \mathbf{V}$ 均可由 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性表示，也即

$$v = \alpha_1 k_1 + \alpha_2 k_2 + \cdots + \alpha_n k_n = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}$$

其中 $k_1, k_2, \cdots, k_n \in \mathbf{F}$

则我们称:

- $\mathbf{V}$ 为 $\mathbf{F}$ 上的有限维线性空间；基向量组中向量的个数 n 称为 $\mathbf{V}$ 的维数，通常记为 $\dim V$；称向量组 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 为 $\mathbf{V}$ 的一个基或者坐标系；称 $\begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix}$ 为基矩阵；
- 列向量 $\begin{bmatrix} k_1 & k_2 & \cdots & k_n \end{bmatrix}^{\mathrm{T}}$ 为向量 v 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标或坐标向量。

说明：无限维空间的例子。

几点性质

- 1 给定向量 v 在给定基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标的唯一性。
- 2. 线性空间维数的唯一性：假设 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 和 $\beta_1, \beta_2, \cdots, \beta_n$ 分别是 $\mathbf{V}$ 的两个基，则 $m = n$。
- 3. 基（坐标系）实现了抽象线性空间到标准线性空间的一一对应。

关于第 3 点性质的证明

- a. 首先证明 $\sigma : \mathbf{V} \to \mathbf{F}^n$ 为映射（$f: \mathbf{A} \to \mathbf{B}$ 的两点要求：（I）$\forall a \in \mathbf{A}$, $\exists b \in \mathbf{B}$ 与之对应；（II）与 a 对应的 b 是唯一的)。
- b. 再证明为一一对应（III）$\mathbf{F}^n$ 中的每一个元素都有原象；（IV）$\mathbf{F}^n$ 中的每一个元素的原象都是唯一的)。

标准线性空间的标准基与一般基

- 向量组

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \ e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \ \cdots, \ e_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

  被称为标准线性空间 $\mathbf{F}^n$ 的标准基，其构成单位矩阵。因此，$\forall g \in \mathbf{F}^n$，有 $g = \begin{bmatrix} e_1 & e_2 & \cdots & e_n \end{bmatrix} g = I_n g$。
- 给定向量在不同基下的坐标变换。
- 向量组 $v_1, v_2, \cdots, v_n$ 构成 $\mathbf{F}^n$ 的一个基的充要条件是该向量组拼成的矩阵 $\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$ 非奇异。任意向量 $v \in \mathbf{F}^n$ 在该基下的坐标是非齐次方程的解：$\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} x = v$。
- 反之，我们可以说，求解具有非奇异系数矩阵的非齐次线性方程的代数问题，其实是将右端向量沿系数矩阵的列向量组构成的坐标系展开这一几何问题。

基变换与坐标变换

非零线性空间的基不是唯一的，所以一个向量在不同基下的坐标也是不同的，那么不同基和坐标之间有什么关系？

- 问题 1. 假设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 和 $\beta_1, \beta_2, \cdots, \beta_n$ 分别是 $\mathbf{V}$ 的两个基，且有：

$$\beta_i = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix} \begin{bmatrix} p_{1i} \\ p_{2i} \\ \vdots \\ p_{ni} \end{bmatrix},$$

那么，我们称 n 阶方阵 P

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

为基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 到基 $\beta_1, \beta_2, \cdots, \beta_n$ 的过渡矩阵，且 P 可逆。

- 问题 2：一个给定向量的坐标在不同基的坐标之间是什么关系？

2.3 子空间

定义：假设 $\mathbf{V}$ 为 $\mathbf{F}$ 上的线性空间。若存在非空集合 $\mathbf{W} \subseteq \mathbf{V}$ 使得

- 1. 对任意的 $\alpha, \beta \in \mathbf{W}$，均有 $\alpha + \beta \in \mathbf{W}$。
- 2. 对任意的 $\alpha \in \mathbf{W}$，$k \in \mathbf{F}$, 均有 $\alpha \cdot k \in \mathbf{W}$。

则称 $\mathbf{W}$ 为 $\mathbf{V}$ 的子空间。

- 容易证明：$\mathbf{W}$ 也构成数域 $\mathbf{F}$ 上的线性空间。
- 平凡子空间，非平凡子空间。
- 举例：二维或者三维空间中，起点为原点，终点在同一过原点的直线上的所有（共线）向量的集合；三维空间中过原点的共面向量集。

与矩阵 $A \in \mathbf{R}^{m\times n}$ 有关的四个子空间（第三次见面）

- (1) 列空间：$\mathbf{V} = \{y \in \mathbf{R}^m \mid y = Ax,\ A \in \mathbf{R}^{m\times n}, x \in \mathbf{R}^n\}$，即 A 的像空间（或值域），常用 $im(A)$ (或 $R(A)$，$C(A)$) 表示
- (2) 零空间：$\mathbf{V} = \{x \mid Ax = 0,\ A \in \mathbf{R}^{m\times n}, x \in \mathbf{R}^n\}$，即齐次线性方程组的解 $Ax = 0$ 构成的集合，即 $Ax = 0$ 的解空间，也称为矩阵 A 的核（或零）空间，常用 $ker(A)$ (或 $N(A)$) 表示
- (3) 行空间：$R(A^{\mathrm{T}}) \subset \mathbf{R}^n$
- (4) 左零空间：$N(A^{\mathrm{T}}) \subset \mathbf{R}^m$

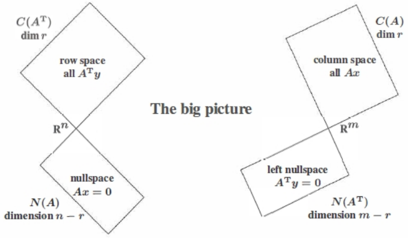

图: Fig. 3.5, Gilbert Strang, Introduction to Linear Algebra (5th Edition)

向量组生成的子空间和子空间的生成组

- 定理：给定向量组 $\alpha_1, \alpha_2, \cdots, \alpha_p \in \mathbf{V}$，则

$$\begin{aligned}\mathbf{W} &= span\{\alpha_1, \alpha_2, \cdots, \alpha_p\} \\ &= \{\alpha_1 c_1 + \alpha_2 c_2 + \cdots + \alpha_p c_p \mid c_i \in \mathbf{F},\ i = 1, 2, \cdots, p\}\end{aligned}$$

  是 $\mathbf{V}$ 的一个子空间（按照定义证明），且我们称 $\mathbf{W}$ 为向量组 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 所张成的子空间。
- 若给定 $\mathbf{V}$ 的一个线性子空间 $\mathbf{H}$，且 $\mathbf{H}$ 可由 $\beta_1, \beta_2, \cdots, \beta_r$ 张成，则称向量组 $\beta_1, \beta_2, \cdots, \beta_r$ 为子空间 $\mathbf{H}$ 的一个向量生成组，简称生成组。
- 类似于线性空间，子空间也可以定义维数和基等概念。
- 生成组的概念提供了子空间的一种表示方法。
- $im(A)$ 用生成组的概念的解读视角。

子空间的交、和

- 定义：假设 $\mathbf{V}$ 为 $\mathbf{F}$ 上的线性空间。假设 $\mathbf{V}_1$、$\mathbf{V}_2$ 是 $\mathbf{V}$ 是两个子空间。则我们分别称

$$\mathbf{V}_1 \cap \mathbf{V}_2 = \{v \mid v \in \mathbf{V}_1, v \in \mathbf{V}_2\}$$
$$\mathbf{V}_1 + \mathbf{V}_2 = \{v = v_1 + v_2 \mid v_1 \in \mathbf{V}_1, v_2 \in \mathbf{V}_2\}$$

  为 $\mathbf{V}_1$ 和 $\mathbf{V}_2$ 的交空间、和空间。
- 可以证明：交空间、和空间均为子空间。而子空间的并（集合意义上）不是子空间（二维平面中的反例）。

子空间的直和、补子空间

- 定义：假设 $\mathbf{V}$ 为 $\mathbf{F}$ 上的线性空间，且 $\mathbf{V}_1$、$\mathbf{V}_2$、$\mathbf{W}$ 是 $\mathbf{V}$ 的子空间。若 $\mathbf{V}_1 \cap \mathbf{V}_2 = \{0\}$，则我们称 $\mathbf{V}_1 + \mathbf{V}_2$ 为 $\mathbf{V}_1$ 和 $\mathbf{V}_2$ 的直和，记为 $\mathbf{V}_1 \oplus \mathbf{V}_2$。
- 若 $\mathbf{V}_1 \oplus \mathbf{V}_2 = \mathbf{W}$，则我们称 $\mathbf{W}$ 是 $\mathbf{V}_1$、$\mathbf{V}_2$ 的直和，或者说 $\mathbf{V}_1$、$\mathbf{V}_2$ 是 $\mathbf{W}$ 的一个直和分解。
- 子空间的直和与子空间的和相比，不同之处在于，前者将和空间中的元素表为两个子空间中元素之和时，表示是唯一的。即，假设 $\mathbf{W}$ 是 $\mathbf{V}_1$、$\mathbf{V}_2$ 的直和，若 $w = v_1 + v_2$，$w = v_1^{'} + v_2^{'}$，其中，$w \in \mathbf{W}$，$v_1, v_1^{'} \in \mathbf{V}_1$，$v_2, v_2^{'} \in \mathbf{V}_2$，则一定有 $v_1 = v_1^{'}$，$v_2 = v_2^{'}$。
- 若 $\mathbf{V}_1 \oplus \mathbf{V}_2 = \mathbf{V}$，则我们称 $\mathbf{V}_1$、$\mathbf{V}_2$ 是线性空间 $\mathbf{V}$ 的一对互补的子空间，或称 $\mathbf{V}_1$ 是 $\mathbf{V}_2$ 的补子空间。

2.4 线性映射

定义：假设 $\mathbf{V}$ 为 $\mathbf{F}$ 上的线性空间。我们称映射 $\sigma : \mathbf{V}_1 \rightarrow \mathbf{V}_2$ 为 $\mathbf{V}_1$ 到 $\mathbf{V}_2$ 的线性映射，如果其保持加法和数乘法，即

- 1. $\forall a, \beta \in \mathbf{V}_1$, $\sigma(a+\beta) = \sigma(a) + \sigma(\beta)$（和的像等于像的和）
- 2. $\forall a \in \mathbf{V}_1$, $k \in \mathbf{F}$, $\sigma(a \cdot k) = \sigma(a) \cdot k$（倍数的像等于像的倍数）

一个线性空间到其本身的线性映射也称为该线性空间上的线性变换。若映射 $\sigma : \mathbf{V}_1 \rightarrow \mathbf{V}_2$ 是可逆的，则称 σ 为线性同构。

- 例子：恒等映射、零映射、由实矩阵构造的映射等。
- 性质：线性相关的原像，经过线性映射之后的像，依然是线性相关的；线性无关的原像，经过线性映射之后的像，则不一定是线性无关的。
- 性质：矩阵与标准线性空间之间线性映射两个事物的等同性。

线性映射的矩阵表示

定义：给定 $\mathbf{F}$ 上的线性空间 $\mathbf{V}$、$\mathbf{W}$ 以及线性映射 $\mathcal{A}: \mathbf{V} \to \mathbf{W}$，其中 $\dim(\mathbf{V}) = n$, $\dim(\mathbf{W}) = m$, 且 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ 为 $\mathbf{V}$ 的一个基（入口基），$\eta_1, \eta_2, \cdots, \eta_m$ 为 $\mathbf{W}$ 的一个基（出口基）。假设 $\epsilon_j \in \mathbf{V}$ 在 $\mathcal{A}$ 下的像 $\mathcal{A}(\epsilon_j) \in \mathbf{W}$ 在出口基下的坐标为 $a_j = \begin{bmatrix} a_{1j} & a_{2j} & \cdots & a_{mj} \end{bmatrix}^T \in \mathbf{F}^m$, 即

$$\mathcal{A}(\epsilon_j) = \begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_m \end{bmatrix} \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix},$$

则由 $a_1, a_2, \cdots, a_n \in \mathbf{F}^m$ 拼成的矩阵

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} = [a_{ij}]_{m \times n}$$

称为 $\mathcal{A}$ 在相应的入口基和出口基下的表示。可记为:

$$\mathcal{A}\left(\begin{bmatrix} \epsilon_1 & \epsilon_2 & \cdots & \epsilon_n \end{bmatrix}\right) = \begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_m \end{bmatrix} A$$

线性映射及其矩阵表示的唯一性

- 线性映射的矩阵表示

$$\mathcal{A}\left(\begin{bmatrix} \text{入口基} \\ \text{矩阵} \end{bmatrix}\right) = \begin{bmatrix} \text{出口基} \\ \text{矩阵} \end{bmatrix} \begin{bmatrix} \text{表示} \\ \text{矩阵} \end{bmatrix}$$

- 给定线性映射、入口基与出口基，那么相应的矩阵表示是唯一的。
- 给定入口基、出口基及相应维数的矩阵，那么存在唯一的线性映射与之对应。

线性映射与坐标之间的联系

定理： 假设线性映射 $\mathcal{A}: \mathbf{V} \to \mathbf{W}$ 在入口基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ 和出口基 $\eta_1, \eta_2, \cdots, \eta_m$ 下的坐标表示为 $A \in \mathbf{F}^{m\times n}$。假设 $\alpha \in \mathbf{V}$ 在入口基下的坐标为 $x \in \mathbf{F}^n$，那么 $\mathcal{A}(\alpha) \in \mathbf{W}$ 在出口基下的坐标为 $Ax \in \mathbf{F}^m$。

证明： 需要证明

$$\mathcal{A}(\alpha) = \begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_m \end{bmatrix} (Ax)$$

由以下计算可得

$$\begin{aligned}
&\begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_m \end{bmatrix} (Ax) \\
&= \left(\begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_m \end{bmatrix} A \right) x \\
&= \mathcal{A} \left(\begin{bmatrix} \epsilon_1 & \epsilon_2 & \cdots & \epsilon_n \end{bmatrix} \right) x \\
&= \mathcal{A} \left(\begin{bmatrix} \epsilon_1 & \epsilon_2 & \cdots & \epsilon_n \end{bmatrix} x \right) \\
&= \mathcal{A}(\alpha)
\end{aligned}$$

例子 1：微分算子的矩阵表示

考虑微分算子 $\mathcal{D}:\mathbf{R}_4[x]\to\mathbf{R}_3[x]$

- 选定线性空间 $\mathbf{R}_4[x]$ 的一组基 $1,x,x^2,x^3$，选定线性空间 $\mathbf{R}_3[x]$ 的一组基 $1,x,x^2$。那么，我们有

$$\mathcal{D}\left(\begin{bmatrix} 1 & x & x^2 & x^3 \end{bmatrix}\right)=\begin{bmatrix} 0 & 1 & 2x & 3x^2 \end{bmatrix}$$
$$=\begin{bmatrix} 1 & x & x^2 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

即，微分算子 $\mathcal{D}:\mathbf{R}_4[x]\to\mathbf{R}_3[x]$ 的矩阵表示为

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

- 矩阵表示的优势：计算 $\mathcal{D}(f(x))$ 的两种方式及其区别。

例子 2：旋转矩阵的表示

- 2 维空间中的旋转矩阵的推导
- 线性映射的视角

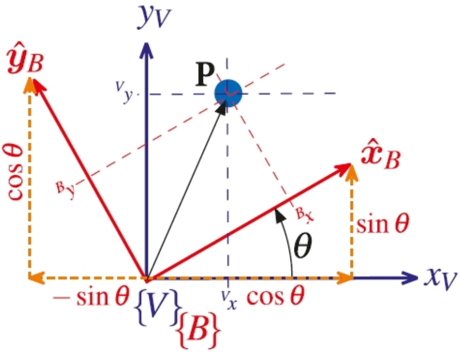

图: Fig. 2.7, Peter Corke, Robotics, Vision and Control (Second Edition)

例子 3：几何空间中的镜面反射

线性映射在不同对基下的矩阵表示之间的关系

- **定理**：给定线性映射 $\mathcal{A}:\mathbf{V}\to\mathbf{W}$。假设：$\alpha_1,\alpha_2,\cdots,\alpha_n$ 和 $\alpha_1^{'},\alpha_2^{'},\cdots,\alpha_n^{'}$ 为 $\mathbf{V}$ 的两组基，P 为 $\{\alpha_i\}$ 到 $\left\{\alpha_i^{'}\right\}$ 的过渡矩阵；$\beta_1,\beta_1,\cdots,\beta_m$ 和 $\beta_1^{'},\beta_2^{'},\cdots,\beta_n^{'}$ 为 $\mathbf{W}$ 的两组基，Q 为 $\{\beta_j\}$ 到 $\left\{\beta_j^{'}\right\}$ 的过渡矩阵。若线性映射 $\mathcal{A}$ 在基 $\{\alpha_i\}$ 与 $\{\beta_i\}$ 下的矩阵表示为 A。则其在基 $\left\{\alpha_i^{'}\right\}$ 与 $\left\{\beta_j^{'}\right\}$ 下的矩阵表示为

$$B=Q^{-1}AP$$

- **定义**：两个矩阵 $A,B\in\mathbf{F}^{m\times n}$ 被称为是等价的，如果存在 n 阶非奇异矩阵 $P\in\mathbf{F}^{n\times n}$ 和 m 阶非奇异矩阵 $Q\in\mathbf{F}^{m\times m}$ 使得

$$B=Q^{-1}AP$$

2.5 矩阵的等价与相似

定义：两个矩阵 $A, B \in \mathbf{F}^{m\times n}$ 被称为是等价的，记为 $A \sim B$，如果存在 n 阶非奇异矩阵 $P \in \mathbf{F}^{n\times n}$ 和 m 阶非奇异矩阵 $Q \in \mathbf{F}^{m\times m}$ 使得

$$AP = QB$$

说明：

- 1. 线性代数中矩阵等价的定义
- 2. 矩阵等价的几何意义（$x \mapsto y,\ y = Ax;\ x = Px',\ y = Qy',\ y' = Bx'$）
- 3. 问题：给定 $A \in \mathbf{F}^{m\times n}$，如何选择非奇异矩阵 $P \in \mathbf{F}^{n\times n}$，$Q \in \mathbf{F}^{m\times m}$，使得 $B = Q^{-1}AP$ “最简单”？

初等行列变换的标准型：对任意矩阵 $A \in \mathbf{F}^{m\times n}$，可以经过初等行列变换将 A 化为标准型，即存在 n 阶非奇异矩阵 $P \in \mathbf{F}^{n\times n}$ 和 m 阶非奇异矩阵 $Q \in \mathbf{F}^{m\times m}$ 使得

$$Q^{-1}AP = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \leftrightarrow AP = Q\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

说明：

- 几何意义
- 输入输出“解耦”的视角

矩阵的相似

定义：两个矩阵 $A, B \in \mathbf{F}^{n\times n}$ 被称为是相似的，记为 $A \simeq B$，如果存在 n 阶非奇异矩阵 $P \in \mathbf{F}^{n\times n}$ 使得

$$AP = PB$$

说明：

- 1. 矩阵相似的几何意义
- 2. “相似最简型”问题：给定 $A \in \mathbf{F}^{n\times n}$，如何选择非奇异矩阵 $P \in \mathbf{F}^{n\times n}$，使得 $B = P^{-1}AP$ “最简单”？

方阵的不变子空间

- 定义：给定矩阵 $A \in \mathbf{F}^{n\times n}$，如果存在 $\mathbf{W} \subseteq \mathbf{F}^n$ 使得 $A(\mathbf{W}) \subseteq \mathbf{W}$，其中
$$A(\mathbf{W}) = \{Ax \mid x \in \mathbf{W}\}$$
则称 $\mathbf{W}$ 为 A 的不变子空间。
- 矩阵不变子空间的例子：矩阵的核、像以及零空间（这里指的是只包含零元素的空间）和全空间。

方阵的不变子空间与方阵的块三角化的等同性

定理：假设 $B = P^{-1}AP$，且矩阵 P, B 相应的分块如下

$$P = \begin{bmatrix} P_1 & P_2 \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

则

- $B_{21} = 0$，当且仅当 $im(P_1)$ 为 A 的不变子空间
- $B_{12} = 0$，当且仅当 $im(P_2)$ 为 A 的不变子空间
- $B_{21} = 0, B_{12} = 0$，当且仅当 $im(P_1)$ 和 $im(P_2)$ 均为 A 的不变子空间（两者实际上是互补的子空间）

基于不变子空间将矩阵块三角化、块对角化的方法

假设 $\mathbf{W} \subseteq \mathbf{F}^n$ 是给定矩阵 $A \in \mathbf{F}^{n\times n}$ 的不变子空间，如何构造 P 使得矩阵 A 上三角化？

- 假设 $\dim \mathbf{W} = t$，任选 $\mathbf{W}$ 的一组基 $p_1, \cdots, p_t$, 将其扩充为

$$p_1, p_2, \cdots, p_t, p_{t+1}, \cdots, p_{n-1}, p_n$$

  并使之成为 $\mathbf{F}^n$ 的一组基，将其拼成的矩阵记为 P，则 $P^{-1}AP$ 即为上三角矩阵。

同样的，如何使矩阵块对角化？

- 假设 $\mathbf{W}, \mathbf{U} \subseteq \mathbf{F}^n$ 为给定矩阵 $A \in \mathbf{F}^{n\times n}$ 的两个互补的非平凡的不变子空间，即 $\mathbf{W} \oplus \mathbf{U} \subseteq \mathbf{F}^n$，则任选 $\mathbf{W}, \mathbf{U}$ 的基，它们组成了 $\mathbf{F}^n$ 的基；用该基拼成的矩阵做相似变换，则可使矩阵 A 块对角化。

n 阶方阵可相似对角化的条件

- 定义：给定 n 阶方阵 $A \in \mathbf{F}^{n\times n}$，以及 $\mathbf{W} = span\{p\} \subseteq \mathbf{F}^n$，其中 $p \neq 0$。称 $\mathbf{W}$ 为矩阵 A 的一维不变子空间，如果

$$Ap \in span\{p\}$$

也即，存在 $\lambda \in \mathbf{F}$ 使得

$$Ap = p\lambda$$

- 定义：给定 n 阶方阵 $A \in \mathbf{F}^{n\times n}$，$\lambda \in \mathbf{F}$，以及非零向量 $p \in \mathbf{F}^n$，$p \neq 0$。若有

$$Ap = p\lambda$$

则称 λ 为矩阵 A 的一个特征值，p 为其对应的特征向量；或者称数和向量对 (λ, p) 为矩阵 A 的一个特征值和特征向量。

- 定理：给定 n 阶方阵 $A \in \mathbf{F}^{n\times n}$，其可被相似对角化等价于存在 n 个互补的一维不变子空间，即 n 个线性无关的特征向量。
- 证明：

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \leftrightarrow AP = P \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

$$\leftrightarrow A \begin{bmatrix} p_1 & p_2 & \cdots & p_n \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \\ \cdots \\ p_n \end{bmatrix}^T \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

$$\leftrightarrow Ap_i = p_i \lambda_i,\ for\ i = 1, 2, \cdots, n$$

n 阶方阵特征值与特征向量的基本性质

- 给定 n 阶方阵 $A \in \mathbf{F}^{n\times n}$，那么不同特征值对应的特征向量线性无关（几何解释）。
- 考虑复数域上的矩阵 $A \in \mathbf{C}^{n\times n}$，及其一个特征值 $\mu \in \mathbf{C}$，即 $|\mu I_n - A| = 0$。
  - （1）A 的特征值 μ 的**代数重数**是指因式 $\lambda - \mu$ 在特征多项式 $|\lambda I_n - A|$ 的质因式分解中出现的次数，也就是 μ 做为多项式 $|\lambda I_n - A|$ 的根的重数。
  - （2）A 的特征值 μ 的**几何重数**指的是

$$
\begin{aligned}
&\mu\text{对应的线性无关的特征向量的个数}\\
&= \text{方程}(\mu I_n - A)x = 0\text{解空间的维数}\\
&= \dim(\ker(\mu I_n - A))\\
&= n - rank(\mu I_n - A)
\end{aligned}
$$

  - （3）代数重数不小于几何重数
- 对称矩阵的特征值为实数。半正定矩阵特征值为非负实数。

矩阵相似对角化的应用

- 例 1：给定初始条件，求解常系数微分方程组的解

$$\begin{cases} \frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ \frac{dx_n}{dt} = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{cases}$$

- 例 2. 给定数域上的矩阵 $A \in \mathbf{F}^{n\times n}$，求解

$$A^{99}$$

- 最后的问题：并不是每一个 n 阶方阵 $A \in \mathbf{F}^{n\times n}$，都可以被相似对角化。若不可被相似对角化，那么最简单可以变为什么形式？