线性代数期末复习

第五章 线性变换

倪卫明

复旦大学信息学院

2013, June

第5章 线性变换

1. 线性变换概念

线性变换的定义: 数域 P 上 n 维线性空间 V 的一个变换 T, $\forall \alpha, \beta \in V$, 若满足下列条件:

(1). $T(\alpha+\beta) = T(\alpha) + T(\beta)$.

(2). $T(k\alpha) = kT(\alpha)$, $\quad \forall k \in P$.

则称变换 T 为**线性变换**.

线性变换性质:

(a) $T(\mathbf{0}) = \mathbf{0}$, $\quad T(-\alpha) = -T(\alpha)$

(b) 线性变换保持线性组合与线性关系式不变, 即

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m$$
$$T(\beta) = k_1 T(\alpha_1) + k_2 T(\alpha_2) + \cdots + k_m T(\alpha_m)$$

(c) 线性相关的向量组经线性变换后仍线性相关, 即

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m = \mathbf{0}$$
$$k_1 T(\alpha_1) + k_2 T(\alpha_2) + \cdots + k_m T(\alpha_m) = \mathbf{0}$$

第5章 线性变换

1. 线性变换概念

线性变换的运算: 设 V 是数域 P 上的线性空间, $L(V)$ 表示 V 上全体线性变换的集合, 在 $L(V)$ 上引入加法、数量乘法、乘法和逆变换运算.

(a) 加法: $\forall T_1, T_2 \in L(V), \forall \alpha \in V$

$$(T_1 + T_2)(\alpha) = T_1(\alpha) + T_2(\alpha)$$

(b) 数量乘法: $\forall T \in L(V), \forall \alpha \in V, \forall k \in P$

$$(kT)(\alpha) = k[T(\alpha)]$$

(c) 乘法: $\forall T_1, T_2 \in L(V), \forall \alpha \in V$

$$(T_1 T_2)(\alpha) = T_1[T(\alpha)]$$

(d) 逆变换: 设 $T \in L(V)$, 若存在变换 S, 使得

$$ST = TS = T_E$$

则称变换 S 为 T 的**逆变换**.

(e) 方幂: 设 $T \in L(V)$, 约定

$$T^k = T \cdot T \cdots T$$
$$T^0 = T_E$$

线性变换的矩阵表示: 设 V 是数域 P 上 n 维线性空间, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 是 V 的一组基, $T \in L(V)$, $\forall \xi \in V$, 有

$$\xi = \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix} \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$$
$$T(\xi) = \begin{bmatrix} T(\varepsilon_1) & T(\varepsilon_2) & \cdots & T(\varepsilon_n) \end{bmatrix} \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$$

说明若知道 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 在变换 T 下的像, 则任意一个向量在变换 T 下的像就可知了.

第5章 线性变换

1. 线性变换概念

用 $T(\varepsilon_1), T(\varepsilon_2), \cdots, T(\varepsilon_n)$ 表示基 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 在变换 T 下的像, 则有

$$
\begin{aligned}
& T\begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix} \\
= & \begin{bmatrix} T(\varepsilon_1) & T(\varepsilon_2) & \cdots & T(\varepsilon_n) \end{bmatrix} \\
= & \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \\
= & \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix} \mathbf{A}
\end{aligned}
\qquad (1)
$$

称式(1)中的矩阵 $\mathbf{A}$ 为 T 在基 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 下的矩阵.

反之, 对于 $\forall \alpha_1, \alpha_2, \ldots, \alpha_n \in V$, 在基 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 下, 总唯一存在 $T \in L(V)$, 使

$$T(\varepsilon_i) = \alpha_i, \quad (i = 1, 2, \ldots, n)$$

由此说明线性变换 T 与变换矩阵之间一一对应.

第5章 线性变换

1. 线性变换概念

设 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 是 n 维线性空间 V 的一组基，线性变换 $T_1, T_2 \in L(V)$ 在这组基下对应的矩阵分别为 $\mathbf{A}, \mathbf{B}$，则

(a) 线性变换的和对应于矩阵的和，即 $T_1 + T_2 \leftrightarrow \mathbf{A} + \mathbf{B}$.

(b) 线性变换的数乘对应于矩阵的数量乘积，即 $kT_1 \leftrightarrow k\mathbf{A}$.

(c) 线性变换的乘积对应于矩阵的乘积，即 $T_1 T_2 \leftrightarrow \mathbf{AB}$.

(d) 线性变换的可逆变换对应于可逆矩阵，且逆变换对应于逆矩阵.

设线性变换 T 在基 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 下的矩阵为 $\mathbf{A}$，向量 ξ 在这组基下的坐标为 $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$，$T(\xi)$ 在这组基下的坐标为 $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}^T$，则

$$
\begin{aligned}
\xi &= \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \ldots & \varepsilon_n \end{bmatrix} \mathbf{x} \\
T(\xi) &= \begin{bmatrix} T(\varepsilon_1) & T(\varepsilon_2) & \ldots & T(\varepsilon_n) \end{bmatrix} \mathbf{x} \\
\begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \ldots & \varepsilon_n \end{bmatrix} \mathbf{y} &= \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \ldots & \varepsilon_n \end{bmatrix} \mathbf{A}\mathbf{x}
\end{aligned}
$$

因此，有: $\mathbf{y} = \mathbf{A}\mathbf{x}$.

线性变换的矩阵与所取的基有关，不同的基下，同一线性变换可能有不同的矩阵，设 n 维线性空间 V，线性变换 T 在基 I: $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 和基 II: $\eta_1, \eta_2, \ldots, \eta_n$ 下对应的矩阵分别为 $\mathbf{A}$ 和 $\mathbf{B}$，从基I到基 II的过渡矩阵为 $\mathbf{M}$，即

$$\begin{aligned}
T(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) &= \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix} \mathbf{A} \\
T(\eta_1, \eta_2, \cdots, \eta_n) &= \begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_n \end{bmatrix} \mathbf{B} \\
\begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_n \end{bmatrix} &= \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix} \mathbf{M}
\end{aligned}$$

于是

$$\begin{aligned}
T(\eta_1, \eta_2, \cdots, \eta_n) &= T\left(\begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix} \mathbf{M}\right) \\
&= T(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) \mathbf{M} \\
&= \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix} \mathbf{AM} \\
\begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_n \end{bmatrix} \mathbf{B} &= \begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_n \end{bmatrix} \mathbf{M}^{-1}\mathbf{AM}
\end{aligned}$$

因而有: $\mathbf{B} = \mathbf{M}^{-1}\mathbf{AM}$.

对于 n 阶矩阵 $\mathbf{A}$ 和 $\mathbf{B}$, 若存在 n 满秩阵 $\mathbf{P}$ 使 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\mathbf{B}$ 称矩阵 $\mathbf{A}$ **相似**于 $\mathbf{B}$, 记为 $\mathbf{A}\sim\mathbf{B}$. 显然, 同一线性变换在 不同基下的矩阵之间相似. 反之, 两个相似的矩阵可以看作同一线性变换在不同基下对应的矩阵.

特征值与特征向量

设 T 是数域 P 上线性空间 V 中的一个线性变换, 对于数域 P 上一个数 λ_0, 若存在一个非零向量 ξ, 使得

$$T(\xi)=\lambda_0\xi$$

则称 λ_0 为 T 的**特征值**, 称 ξ 为 T 的属于特征值 λ_0 的一个**特征 向量**. 若令 $V_{\lambda_0}=\left\{\xi\middle|T(\xi)=\lambda_0\xi,\xi\in V\right\}$, 可验证 V_{λ_0} 构成 V 的线性子空间, 称为线性变换 T 的属于特征值 λ_0 的**特征子空间**.

特征值与特征向量的求法

设 V 是数域 P 上的 n 维线性空间, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 是 V 的一组基, 线性变换 T 在这组基下的矩阵是 $\mathbf{A} = [a_{ij}]_{n\times n}$, 设 λ_0 是 T 的一个 特征值, 属于 λ_0 的特征向量 ξ 在这组基下的坐标为:
$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$, 则 $T(\xi)$ 的坐标就是 $\mathbf{Ax}$, 因而有:

$$\mathbf{Ax} = \lambda_0 \mathbf{x} \Rightarrow (\lambda_0 \mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}, \quad \mathbf{x} \neq \mathbf{0}$$

因要求 $\mathbf{x} \neq \mathbf{0}$, 根据方程组解理论 $Rank(\lambda_0 \mathbf{E} - \mathbf{A})$ 必须小于 n, 即 $\det(\lambda_0 \mathbf{E} - \mathbf{A}) = 0$, 为了求得特征值 λ_0, 引入

$$f(\lambda) = \det(\lambda \mathbf{E} - \mathbf{A})$$

其中 λ 是未知量, 称 $f(\lambda)$ 为 $\mathbf{A}$ 的**特征多项式**. $f(\lambda) = 0$ 的根称为 $\mathbf{A}$ 的特征根(或$\mathbf{A}$的特征值). 再将求得的特征值代入 $(\lambda \mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$, 称该方程组为 $\mathbf{A}$ 的**特征方程组**. 它的解即 λ 对应的特征向量.

特征多项式、特征向量性质:($f_{\mathbf{A}}(\lambda)=|\lambda\mathbf{E}-\mathbf{A}|$)

(a) 特征多项式 $f_{\mathbf{A}}(\lambda)$ 是首系数为1的 n 次多项式.

(b) $f_{\mathbf{A}}(\lambda)$ 的 λ^{n-1} 项的系数为 $-tr(\mathbf{A})$ (这里 $tr(\mathbf{A})$ 表示矩阵 $\mathbf{A}$ 的迹).

(c) $f_{\mathbf{A}}(\lambda)$ 的常数项为 $(-1)^n|\mathbf{A}|$.

(d) 相似矩阵有相同的特征多项式. 设矩阵 $\mathbf{A}\sim\mathbf{B}$, 即 $\exists\mathbf{P}$ 可逆, 使得: $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\mathbf{B}$, 则

$$\begin{aligned} f_{\mathbf{B}}(\lambda) &= |\lambda\mathbf{E}-\mathbf{B}|=\left|\lambda\mathbf{P}^{-1}\mathbf{P}-\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right| \\ &= \left|\mathbf{P}^{-1}\right||\lambda\mathbf{E}-\mathbf{A}||\mathbf{P}|=f_{\mathbf{A}}(\lambda) \end{aligned}$$

(e) (Hamilton-Cayley) **定理**: $f_{\mathbf{A}}(\mathbf{A})=\mathbf{0}$.

(f) 属于不同特征值的特征向量线性独立.

设 n 阶矩阵 $\mathbf{A}$ 的特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_s$, 则

$$f_{\mathbf{A}}(\lambda) = \prod_{i=1}^{s} (\lambda - \lambda_i)^{n_i}$$

其中正整数 n_i 称为特征值 λ_i 的**代数重数**, 它所对应的特征方程为 $(\lambda_i \mathbf{E} - \mathbf{A}) \mathbf{x} = \mathbf{0}$, 称 $n - Rank(\lambda_i \mathbf{E} - \mathbf{A})$ 为 λ_i 的**几何重数**. 显然, 特征方程的解空间构成 λ_i 的特征子空间 V_{λ_i}, 其维数就是 λ_i 的几何重数. 关于特征值有如下性质:

(a) 矩阵 $\mathbf{A}$ 所有特征值的积等于矩阵 $\mathbf{A}$ 的行列式, 即

$$\prod_{i=1}^{s} (\lambda_i)^{n_i} = |\mathbf{A}|.$$

(b) 矩阵 $\mathbf{A}$ 的所有特征值之和等于矩阵的迹, 即

$$\sum_{i=1}^{s} n_i \lambda_i = tr(\mathbf{A})$$

(c) $\mathbf{A}$ 可逆 $\iff$ 0 不是 $\mathbf{A}$ 的特征值.

(d) $\mathbf{A}^{-1}$ 的特征多项式为

$$f_{\mathbf{A}^{-1}}(\lambda) = \left|\lambda \mathbf{E} - \mathbf{A}^{-1}\right| = \prod_{i=1}^{s} \left(\lambda - \lambda_i^{-1}\right)^{n_i}$$

且若 α 是 $\mathbf{A}$ 属于特征值 λ_i 的特征向量, 则 α 也是 $\mathbf{A}^{-1}$ 的属于特征值 λ_i^{-1} 的特征向量.

(e) 特征值的几何重数不超过其代数重数.

矩阵可对角化指矩阵是否可相似于一对角矩阵.

(a) n 阶矩阵 $\mathbf{A}$ 可对角化 $\Longleftrightarrow$ $\mathbf{A}$ 有 n 个线性无关的特征向量.

(b) n 阶矩阵 $\mathbf{A}$ 可对角化 $\Longleftrightarrow$ $\mathbf{A}$ 的每个特征值的代数重数与几何重数相等. 特别的, $\mathbf{A}$ 有 n 个互不相同的特征值, 则 $\mathbf{A}$ 可对角化.

矩阵对角化: 设 n 阶矩阵 $\mathbf{A}$ 的特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_n$, 相应的特征向量 为 $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, 并令 $\mathbf{\Lambda} = diag\,(\lambda_1, \lambda_2, \ldots, \lambda_n)$ 表示由特征值作为对角元形成的对角矩阵, $\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}$ 表示由特征(列)向量构成的矩阵, 则

$$\mathbf{AX} = \begin{bmatrix} \lambda_1\mathbf{x}_1 & \lambda_2\mathbf{x}_2 & \cdots & \lambda_n\mathbf{x}_n \end{bmatrix} = \mathbf{X\Lambda}$$

因 n 个特征向量线性无关, 所以 $\mathbf{X}$ 可逆, 有:$\mathbf{X}^{-1}\mathbf{AX} = \mathbf{\Lambda}$

例:将矩阵 $\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & -1 \\ 1 & 0 & 1 \end{bmatrix}$对角化.

解:先求 $\mathbf{A}$ 的特征值和特征向量, 得$\mathbf{A}$ 的特征值为 $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = 2$, 再求 $\mathbf{A}$ 相应于 $\lambda_1, \lambda_2, \lambda_3$ 的特征向量为:
$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, 有 $\mathbf{X}^{-1}\mathbf{AX} =$

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

若矩阵 $\mathbf{A}$ 是实对称阵，则 $\mathbf{A}$ 的特征值都是实数，且属于 $\mathbf{A}$ 的不同特征值的 特征向量一定正交. 而且，是对称矩阵总能找到一个 n 阶正交矩阵 $\mathbf{P}$，使 $\mathbf{P}^{-1}\mathbf{AP}$ 是对角阵. 即 $\mathbf{P}^{-1}\mathbf{AP} = diag\,(\lambda_1, \lambda_2, \cdots, \lambda_n)$. 即任何 一个实对称矩阵总存在正交矩阵，使它化为对角阵. 找这样的正交矩阵的步骤：

(a) 求出实对称矩阵 $\mathbf{A}$ 的特征多项式的全部根 $\lambda_1, \lambda_2, \ldots, \lambda_s (\lambda_i \neq \lambda_j)$;

(b) 对每个 $\lambda_i (i = 1, 2, \ldots, s)$，解特征方程 $(\lambda_i \mathbf{E} - \mathbf{A})\,\mathbf{x} = \mathbf{0}$，得到基础解系：$\mathbf{x}_{i1}, \mathbf{x}_{i2}, \ldots, \mathbf{x}_{in_i}$，其中 n_i 是 λ_i 的代数重数，它们是 λ_i 对应的特征向量.

(c) 通过 Gram-Schmidt 法，将 $\mathbf{x}_{i1}, \mathbf{x}_{i2}, \ldots, \mathbf{x}_{in_i}$ 化为 一组标准正交向量;

(d) 最终将所有的标准正交特征向量按照特征值排序得矩阵 $\mathbf{P}$，这就是所求的正交矩阵 $\mathbf{P}$. 它使得 $\mathbf{P}^T\mathbf{AP} = \mathbf{\Lambda}$，其中 $\mathbf{\Lambda}$ 是所有特征值构成的对角阵.

设 $g(x)$ 是关于 x 的多项式，求 $g(\mathbf{A})$，有下列方法：

一：利用Hamilton-Cayley定义，设 $\mathbf{A}$ 的特征多项式为 $f_{\mathbf{A}}(x)$，则 $f_{\mathbf{A}}(\mathbf{A})=\mathbf{0}$. 将多项式作分解：$g(x)=f_{\mathbf{A}}(x)q(x)+r(x)$，其中 $q(x)$ 为商多项式，$r(x)$为余多项式，可知 $r(x)$ 中关于 x 的最高次小于 $f_{\mathbf{A}}(x)$ 中 x 的最高次，而 $g(\mathbf{A})=r(\mathbf{A})$.

二：若 A 可对角化，即 $\mathbf{X}^{-1}\mathbf{A}\mathbf{X}=\mathbf{\Lambda}\Rightarrow\mathbf{A}=\mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$，则关于 $\mathbf{A}$ 的 n 次方,有 $\mathbf{A}^n=\mathbf{X}\mathbf{\Lambda}^n\mathbf{X}^{-1}$，其中 $\mathbf{\Lambda}^n$ 是对角矩阵 容易求得.

三：先采用方法一求得 $r(\mathbf{A})$，再利用方法二计算 $\mathbf{A}$ 的幂次方.

1. 线性变换概念

设 T 是 n 维欧式空间 V 中的一个线性变换, 若 $\forall \alpha, \beta \in V$, 都有

$$(T(\alpha), T(\beta)) = (\alpha, \beta)$$

则称 T 是一个**正交变换**.

T 是 n 维欧式空间的线性变换, 则
T 是正交变换 $\Longleftrightarrow$ $\forall \alpha \in V$, 有 $\|T(\alpha)\| = \|\alpha\|$. $\Longleftrightarrow$
$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 是 V 的一个标准正交基, 则
$T(\varepsilon_1), T(\varepsilon_2), \ldots, T(\varepsilon_n)$ 也是 V的一组标准正交基.

线性变换与矩阵一一对应,则

(a) 正交变换与正交矩阵一一对应;
(b) 正交变换的乘积还是正交变换;
(c) 正交变换是可逆变换, 且其逆变换也是正交变换;
(d) 正交变换对应矩阵的行列式等于 $+1$ 或 -1; 行列式等于 $+1$ 的正交变换称为第一类正交变换, 行列式等于 -1 的正交变换称为第二类正交变换.