§5 线性变换 II

5.1 恒同变换与基变换

设$\sigma : V \to V$ 为n维向量空间V上的恒同变换, 则 $\sigma(\mathbf{v}) = \mathbf{v},\ \forall \mathbf{v} \in V.$
若取V的一组基$\alpha_1, \cdots, \alpha_n$, 则$\sigma(\alpha_j) = \alpha_j,\ j = 1, \cdots, n.$
于是, 在这同一组基下, $\sigma(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n) I_n,$
恒同变换σ的矩阵是I_n.

若取V的两组不同的基$\alpha_1, \cdots, \alpha_n;\ \beta_1, \cdots, \beta_n$. 则有

$$\sigma(\alpha_j) = \alpha_j = \sum_{i=1}^{n} p_{ij}\beta_i$$

5.1 恒同变换与基变换

即

$$(\alpha_1 \cdots \alpha_n) = \sigma(\alpha_1 \cdots \alpha_n) = (\beta_1 \cdots \beta_n) \underbrace{\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}_{n\times n}}_{P \quad \text{可逆}}$$

则恒同变换σ在两组基下的矩阵表示P为V的这两组基之间的基变换矩阵.

5.1 恒同变换与基变换

例(1):

设$\mathbb{R}^3$中有输入基$\varepsilon_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \varepsilon_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \varepsilon_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$，输出基

$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$ 求基变换矩阵.

解:

$$\sigma(\varepsilon_1\, \varepsilon_2\, \varepsilon_3) = (\varepsilon_1\, \varepsilon_2\, \varepsilon_3) = (\mathbf{e}_1\, \mathbf{e}_2\, \mathbf{e}_3) \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}}_{P}$$

所以这两组基的基变换矩阵是矩阵P.

5.1 恒同变换与基变换

记号如例(1).

例(2):

设$\mathbb{R}^3$中有输入基$\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, 输出基$\varepsilon_1, \varepsilon_2, \varepsilon_3$, 求基变换矩阵.

解: $$\sigma(\mathbf{e}_1\ \mathbf{e}_2\ \mathbf{e}_3) = (\mathbf{e}_1\ \mathbf{e}_2\ \mathbf{e}_3) = (\varepsilon_1\ \varepsilon_2\ \varepsilon_3)P^{-1}$$

所以这两组基的基变换矩阵是矩阵P^{-1}.

5.2 图像压缩——基变换的应用

一张256×256像素的灰度图像 $\longleftrightarrow \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \in \mathbb{C}^N, N = 256^2.$

$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_N\mathbf{e}_N$, 其中$\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_N$是$\mathbb{C}^N$的标准基.

5.2 图像压缩——基变换的应用

换基底(以$\mathbb{C}^4$为例):

小波基: $\mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{w}_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \mathbf{w}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{w}_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$

Fourier基: $\xi_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \xi_2 = \begin{pmatrix} 1 \\ i \\ i^2 \\ i^3 \end{pmatrix}, \xi_3 = \begin{pmatrix} 1 \\ i^2 \\ i^4 \\ i^6 \end{pmatrix}, \xi_4 = \begin{pmatrix} 1 \\ i^3 \\ i^6 \\ i^9 \end{pmatrix}$

5.2 图像压缩——基变换的应用

常用的图像压缩方法JPEG(Joint Photographic Expert Group)就是作基变换, 使用Fourier基.

输入信号$\mathbf{x}$

无损压缩 $\downarrow$ 换基

$\mathbf{x}$在新基$\mathbf{W}$下的坐标$\mathbf{c} = W^{-1}\mathbf{x}$

有损压缩 $\downarrow$

$\hat{\mathbf{c}}$

重建信号 $\downarrow$

$$\mathbf{x} = \sum_{i=1}^{N} \hat{c}_i \mathbf{w}_i = W\hat{\mathbf{c}}$$

$$\begin{aligned}
\mathbf{x} &= c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_N\mathbf{w}_N \\
&= \begin{pmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_N \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} \\
&= W\mathbf{c}
\end{aligned}$$

关键: 取好基底, 作基变换.

5.3 线性变换在不同基下的矩阵

设V是n维向量空间,$\alpha_1, \cdots, \alpha_n$和$\beta_1, \cdots, \beta_n$是V的两组基, 且

$$(\beta_1 \ \cdots \ \beta_n) = (\alpha_1 \ \cdots \ \alpha_n)P$$

其中P可逆. 设$\sigma : V \to V$是V上的线性变换,σ在这两组基下的矩阵分别是A, B, 即

$$\sigma(\alpha_1 \ \cdots \ \alpha_n) = (\alpha_1 \ \cdots \ \alpha_n)A$$

$$\sigma(\beta_1 \ \cdots \ \beta_n) = (\beta_1 \ \cdots \ \beta_n)B$$

5.3 线性变换在不同基下的矩阵

则

$$\sigma(\beta_1 \cdots \beta_n) = \sigma((\alpha_1 \cdots \alpha_n)P) = \sigma(\alpha_1 \cdots \alpha_n)P = (\alpha_1 \cdots \alpha_n)AP,$$

又

$$\sigma(\beta_1 \cdots \beta_n) = (\beta_1 \cdots \beta_n)B = (\alpha_1 \cdots \alpha_n)PB.$$

由于$\alpha_1, \cdots, \alpha_n$线性无关, 故有AP = PB, 即

$$B = P^{-1}AP.$$

5.3 线性变换在不同基下的矩阵

定理: n维向量空间V上的线性变换σ在V的不同基下的矩阵是相似矩阵.

从线性变换的复合的角度再看:

$$
\begin{tikzcd}
\{\beta_1, \cdots, \beta_n\} & V \arrow[r, "\sigma"] \arrow[d, "I_1"'] & V & \{\beta_1, \cdots, \beta_n\} \\
\{\alpha_1, \cdots, \alpha_n\} & V \arrow[r, "\sigma"] & V \arrow[u, "I_2"'] & \{\alpha_1, \cdots, \alpha_n\}
\end{tikzcd}
$$

5.3 线性变换在不同基下的矩阵

注: 称n维向量空间V上的线性变换σ在V的一组基下的矩阵A的特征多项式, 特征值, 迹, 行列式分别为线性变换σ的特征多项式, 特征值, 迹, 行列式.

5.4 矩阵分解与基变换

对给定线性变换, 选取适当基, 使其矩阵尽可能简单.
事实上, 矩阵分解都可理解为线性变换在基变换下矩阵间的关系.

设线性变换$\sigma : \mathbb{R}^n \to \mathbb{R}^m$ 在$\mathbb{R}^n$的标准基$\mathbf{e}_1, \cdots, \mathbf{e}_n$和$\mathbb{R}^m$的标准基$\widetilde{\mathbf{e}}_1, \cdots, \widetilde{\mathbf{e}}_m$ 下的矩阵为A. 则 $\sigma : \mathbb{R}^n \to \mathbb{R}^m$ 可表示为

$$\boxed{\sigma(\mathbf{v}) = A\mathbf{v}, \ \forall \mathbf{v} \in \mathbb{R}^n.}$$

5.4 矩阵分解与基变换

- 若改变$\mathbb{R}^n$的基$(\mathbf{v}_1\cdots\mathbf{v}_n)=(\mathbf{e}_1\cdots\mathbf{e}_n)P,$
  则σ在**v**-基和$\widetilde{\mathbf{e}}$-基下的矩阵为AP.

- 若改变$\mathbb{R}^m$的基$(\mathbf{w}_1\cdots\mathbf{w}_m)=(\widetilde{\mathbf{e}}_1\cdots\widetilde{\mathbf{e}}_m)Q,$
  则σ在**e**-基和**w**-基下的矩阵为 $Q^{-1}A.$

- 若$V=W=\mathbb{R}^n,\mathbf{e}=\widetilde{\mathbf{e}},$ 作基变换$(\mathbf{v}_1\cdots\mathbf{v}_n)=(\mathbf{e}_1\cdots\mathbf{e}_n)P,$
  则σ在**v**-基下的矩阵为 $P^{-1}AP.$

5.4 矩阵分解与基变换

(1) $A = S\Lambda S^{-1}$

设线性变换$\sigma : \mathbb{R}^n \to \mathbb{R}^n$ 在$\mathbb{R}^n$的标准基下的矩阵为A, 则σ可表示为

$$\sigma(\mathbf{v}) = A\mathbf{v},\ \forall \mathbf{v} \in \mathbb{R}^n.$$

A有n个线性无关的特征向量$\mathbf{x}_1, \cdots, \mathbf{x}_n$, 则可构成$\mathbb{R}^n$的一组新基.$\sigma$在这组新基下的矩阵为对角阵$\Lambda$:

$$\sigma(\mathbf{x}_1 \cdots \mathbf{x}_n) = (A\mathbf{x}_1 \cdots A\mathbf{x}_n) = \underbrace{(\mathbf{x}_1 \cdots \mathbf{x}_n)}_{S} \underbrace{\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}}_{\Lambda}$$

5.4 矩阵分解与基变换

$$\begin{tikzcd}[column sep=large] \{\mathbf{e}\} & & & \{\mathbf{e}\} \\ & \mathbb{R}^n \arrow[r, "\sigma", "A"'] & \mathbb{R}^n \arrow[d, "S^{-1}"', "id_2"] & \\ & \mathbb{R}^n \arrow[u, "id_1", "S"'] \arrow[r, "\sigma", "\Lambda"'] & \mathbb{R}^n & \\ \{\mathbf{x}\} & & & \{\mathbf{x}\} \end{tikzcd}$$

$$(\mathbf{x}_1 \cdots \mathbf{x}_n) = (\mathbf{e}_1 \cdots \mathbf{e}_n)S$$

$$\sigma(\mathbf{e}_1 \cdots \mathbf{e}_n) = (\mathbf{e}_1 \cdots \mathbf{e}_n)A$$

$$(\mathbf{e}_1 \cdots \mathbf{e}_n) = (\mathbf{x}_1 \cdots \mathbf{x}_n)S^{-1}$$

$$\sigma(\mathbf{x}_1 \cdots \mathbf{x}_n) = (\mathbf{x}_1 \cdots \mathbf{x}_n)\Lambda$$

故 $\Lambda = S^{-1}AS, A = S\Lambda S^{-1}$.

5.4 矩阵分解与基变换

(2) $A = U\Sigma V^T$

$$\begin{tikzcd} \{\mathbf{e}\} \quad \mathbb{R}^n \arrow[r, "\sigma", "A"'] \arrow[d, "id_{\mathbb{R}^n}"', "V^T"] & \mathbb{R}^n \quad \{\widetilde{\mathbf{e}}\} \\ \{\mathbf{v}\} \quad \mathbb{R}^n \arrow[r, "\sigma", "\Sigma"'] & \mathbb{R}^n \quad \{\mathbf{u}\} \arrow[u, "U", "id_{\mathbb{R}^m}"'] \end{tikzcd}$$

A和Σ 为同一线性变换$\sigma:\mathbb{R}^n \to \mathbb{R}^m$在不同基下的矩阵表示.

A为σ 在$\mathbb{R}^n$ 和 $\mathbb{R}^m$ 的标准基下的矩阵:

$$\sigma(\mathbf{e}_1 \cdots \mathbf{e}_n) = (\widetilde{\mathbf{e}}_1 \cdots \widetilde{\mathbf{e}}_m)A,$$

Σ为σ 在$\mathbb{R}^n$的**v**-基和$\mathbb{R}^m$的**u**-基下的矩阵:

$$\sigma(\mathbf{v}_1 \cdots \mathbf{v}_n) = (\mathbf{u}_1 \cdots \mathbf{u}_m)\Sigma.$$

$\mathbf{v}_1 \cdots \mathbf{v}_n : A^TA$的单位正交特征向量基.

$\mathbf{w}_1 \cdots \mathbf{w}_n : AA^T$ 的单位正交特征向量基.

$$id_{\mathbb{R}^n}(\mathbf{e}_1 \cdots \mathbf{e}_n) = (\mathbf{v}_1 \cdots \mathbf{v}_n)V^T,$$

$$id_{\mathbb{R}^m}(\mathbf{u}_1 \cdots \mathbf{u}_m) = (\widetilde{\mathbf{e}}_1 \cdots \widetilde{\mathbf{e}}_m)U,$$

$\sigma = id_{\mathbb{R}^m} \circ \sigma \circ id_{\mathbb{R}^n}$, 故$A = U\Sigma V^T$.

5.5 线性变换的核与像

有两个与线性变换$\sigma : V \to W$密切相关的集合:

$$\boxed{\ker \sigma := \{\mathbf{v} \in V \mid \sigma(\mathbf{v}) = \mathbf{0}\} \qquad (\text{核 kernel})}$$

$$\boxed{Im\sigma := \{\sigma(\mathbf{v}) \mid \mathbf{v} \in V\} \qquad (\text{像 image, 或值域})}$$

命题: $\ker \sigma$是向量空间V的子空间, $Im\sigma$ 是向量空间W的子空间.

5.5 线性变换的核与像

定义: 称 $\dim\ker\sigma$ 为线性变换σ的零度(nullity), 称$\dim Im\sigma$ 为线性变换σ 的秩(rank).

例: 设A为 $m\times n$ 实矩阵, 则线性变换$\sigma:\mathbb{R}^n\to\mathbb{R}^m, \mathbf{x}\mapsto A\mathbf{x}$的核 $\ker\sigma=N(A)$, σ的零度$=n-r(A)$, (故此数也成为矩阵A的零度), σ的像$Im\sigma=C(A)$, σ的秩$=r(A)$.

例: 求导变换 $\sigma=\frac{d}{dx}:P_n\to P_n$ 是一个线性变换, 则

$Im\sigma=P_{n-1}$, $\ker\sigma=\mathbb{R}$, σ 的秩为n, 零度为1.

5.5 线性变换的核与像

定理: 设$\sigma \in \mathcal{L}(V,V)$, $\mathbf{v}_1,\cdots,\mathbf{v}_n$是V的一组基,$\sigma$在这组基下的矩阵是A.则

(1) $Im\sigma = L(\sigma(\mathbf{v}_1),\cdots,\sigma(\mathbf{v}_n))$;

(2) σ的秩$= r(A)$.

证明:

(1) $\forall \mathbf{v} \in V, \mathbf{v} = \sum_{i=1}^{n} c_i\mathbf{v}_i$, 则$\sigma(\mathbf{v}) = \sum_{i=1}^{n} c_i\sigma(\mathbf{v}_i)$,

故 $\sigma(\mathbf{v}) \in L(\sigma(\mathbf{v}_1),\cdots,\sigma(\mathbf{v}_n))$, $Im\sigma \subseteq L(\sigma(\mathbf{v}_1),\cdots,\sigma(\mathbf{v}_n))$. 显然又有 $L(\sigma(\mathbf{v}_1),\cdots,\sigma(\mathbf{v}_n)) \subseteq Im\sigma$. 故

$$Im\sigma = L(\sigma(\mathbf{v}_1),\cdots,\sigma(\mathbf{v}_n)).$$

5.5 线性变换的核与像

(2) 由(1)知, σ 的秩等于向量组$\sigma(\mathbf{v}_1),\cdots,\sigma(\mathbf{v}_n)$的秩. 又A的列向量分别是$\sigma(\mathbf{v}_1),\cdots,\sigma(\mathbf{v}_n)$在基$\mathbf{v}_1,\cdots,\mathbf{v}_n$下的坐标. 而给定向量空间V的一组基, 有线性同构映射$V\to\mathbb{F}^n$, 把V中任一向量与它在这组基下的坐标一一对应起来. 并且线性同构保持向量组的一切线性关系.

因此向量组$\sigma(\mathbf{v}_1),\cdots,\sigma(\mathbf{v}_n)$ 的秩等于它的坐标向量组, 也即A的列向量组的秩, 也即A的秩. 于是σ 的秩$=r(A)$.

5.5 线性变换的核与像

定理:设$\sigma \in \mathcal{L}(V,V), \dim V = n$, 则 $\dim\ker\sigma + \dim Im\sigma = \dim V$.

证明:

设$\dim\ker\sigma = q$, 在$\ker\sigma$中取一组基 $\mathbf{e}_1, \cdots, \mathbf{e}_q$, 将其扩充为V的一组基$\mathbf{e}_1, \cdots, \mathbf{e}_q, \mathbf{e}_{q+1}, \cdots, \mathbf{e}_n$. 由上面定理知,

$$Im\sigma = L(\sigma(\mathbf{e}_1), \cdots, \sigma(\mathbf{e}_q), \sigma(\mathbf{e}_{q+1}), \cdots, \sigma(\mathbf{e}_n)).$$

而$\sigma(\mathbf{e}_1) = \cdots = \sigma(\mathbf{e}_q) = 0$, 于是, $Im\sigma = L(\sigma(\mathbf{e}_{q+1}), \cdots, \sigma(\mathbf{e}_n))$. 故只要证明$\sigma(\mathbf{e}_{q+1}), \cdots, \sigma(\mathbf{e}_n)$线性无关即可.

5.5 线性变换的核与像

设$k_{q+1}\sigma(\mathbf{e}_{q+1})+\cdots+k_n\sigma(\mathbf{e}_n)=0$,

则$\sigma(k_{q+1}\mathbf{e}_{q+1}+\cdots+k_n\mathbf{e}_n)=0$, 即$k_{q+1}\mathbf{e}_{q+1}+\cdots+k_n\mathbf{e}_n\in\ker\sigma$.

于是$k_{q+1}\mathbf{e}_{q+1}+\cdots+k_n\mathbf{e}_n$可由$\ker\sigma$的基线性表出.

设 $k_{q+1}\mathbf{e}_{q+1}+\cdots+k_n\mathbf{e}_n=k_1\mathbf{e}_1+\cdots+k_q\mathbf{e}_q$.

由$\mathbf{e}_1,\cdots,\mathbf{e}_n$线性无关得$k_1=\cdots=k_n=0$. 因此$\sigma(\mathbf{e}_{q+1}),\cdots,\sigma(\mathbf{e}_n)$

线性无关, 构成$Im\sigma$的一组基, 从而$\dim Im\sigma=n-q$. 于是,

$$\dim\ker\sigma+\dim Im\sigma=q+(n-q)=n=\dim V.$$

5.5 线性变换的核与像

注： 尽管$\dim\ker\sigma+\dim Im\sigma=\dim V$，但未必有$\ker\sigma+Im\sigma=V$.

如求导变换 $\sigma=\frac{d}{dx}:P_n\to P_n,\ker\sigma=\mathbb{R},Im\sigma=P_{n-1},\mathbb{R}+P_{n-1}\neq P_n.$

5.5 线性变换的核与像

定理：设V是有限维向量空间, $\sigma \in \mathcal{L}(V,V)$, 则

$$\boxed{\sigma\text{是单射} \Leftrightarrow \sigma\text{是满射} \Leftrightarrow \sigma\text{可逆}}$$

证明:

$$\sigma\text{是单射} \Leftrightarrow \ker\sigma = \{\mathbf{0}\} \Leftrightarrow \dim Im\sigma = \dim V \Leftrightarrow Im\sigma = V.$$

σ可逆$\Rightarrow \exists\tau \in \mathcal{L}(V,V), s.t. \sigma\tau = I_V$. 则$\forall \mathbf{x} \in V, \exists \mathbf{y} = \tau(\mathbf{x}) \in V$, 有

$\sigma(\mathbf{y}) = \sigma(\tau(\mathbf{x})) = (\sigma\tau)(\mathbf{x}) = I_V(\mathbf{x}) = \mathbf{x}$, 即$\sigma$是满射.

5.5 线性变换的核与像

σ是单射$\Rightarrow$ 对V中任一组基$\mathbf{e}_1, \cdots, \mathbf{e}_n$, 有$\sigma(\mathbf{e}_1), \cdots, \sigma(\mathbf{e}_n)$也是$V$的一组基.

事实上, 考虑$k_1\sigma(\mathbf{e}_1) + \cdots + k_n\sigma(\mathbf{e}_n) = \mathbf{0}$, 则

$$\sigma(k_1\mathbf{e}_1 + \cdots + k_n\mathbf{e}_n) = \mathbf{0},$$

故$k_1\mathbf{e}_1 + \cdots + k_n\mathbf{e}_n \subseteq \ker\sigma = \{\mathbf{0}\}$, 即$k_1\mathbf{e}_1 + \cdots + k_n\mathbf{e}_n = \mathbf{0}$.

由$\mathbf{e}_1, \cdots, \mathbf{e}_n$线性无关得$k_1 = \cdots = k_n = 0$. 故$\sigma\mathbf{e}_1, \cdots, \sigma\mathbf{e}_n$线性无关, 也构成$n$维向量空间$V$的一组基.

5.5 线性变换的核与像

于是, 可构造线性变换$\tau \in \mathcal{L}(V, V)$, 使得$\sigma(\mathbf{e}_j) \mapsto \mathbf{e}_j$.

因此

$$\mathbf{e}_j \xmapsto{\sigma} \sigma(\mathbf{e}_j) \xmapsto{\tau} \mathbf{e}_j, \qquad \tau\sigma = I_V;$$

$$\sigma(\mathbf{e}_j) \xmapsto{\tau} \mathbf{e}_j \xmapsto{\sigma} \sigma(\mathbf{e}_j), \qquad \sigma\tau = I_V;$$

故σ可逆, $\sigma^{-1} = \tau$.

注: 这是有限维向量空间V到自身线性变换的特性.

5.5 线性变换的核与像

例: 设A为n阶实方阵,$\mathbf{b}$为n维实向量. 则$\sigma:\mathbf{x}\mapsto A\mathbf{x}$定义了$\mathbb{R}^n\to\mathbb{R}^n$的线性变换. 则方程组$A\mathbf{x}=\mathbf{b}$对任意$\mathbf{b}$都有解$\Longleftrightarrow Im\sigma=\mathbb{R}^n$ $\Longleftrightarrow \sigma$是满射$\Longleftrightarrow \sigma$是单射$\Longleftrightarrow \{0\}=\ker\sigma=N(A)\Longleftrightarrow A\mathbf{x}=\mathbf{0}$只有零解.

5.6 不变子空间

设$\sigma \in \mathcal{L}(V,V)$, 对$Im\sigma, \forall \alpha \in Im\sigma, \sigma(\alpha) \in Im\sigma$;

$$对\ker\sigma, \forall\alpha \in \ker\sigma, \sigma(\alpha) = 0 \in \ker\sigma.$$

定义: 设$\sigma \in \mathcal{L}(V,V)$, W是V的子空间. 若$\forall\alpha \in W, \sigma(\alpha) \in W$, 则称$W$是线性变换$\sigma$的不变子空间.

将σ的作用限制在其不变子空间W上, 记为$\sigma|_W$, 称为σ在W上的限制.
由于W是σ的不变子空间, $\sigma|_W : W \to W$是W上的线性变换.

5.6 不变子空间

对线性变换$\sigma: V \to V$的非平凡不变子空间W, 若存在σ的另一不变子空间U, 且U恰为W的补空间. 于是

$$V = W \oplus U.$$

取V的基$\mathbf{v}_1, \cdots, \mathbf{v}_n$, 使$\mathbf{v}_1, \cdots, \mathbf{v}_k$为$W$的基, $\mathbf{v}_{k+1}, \cdots, \mathbf{v}_n$为$U$的基,则

$$\sigma(\mathbf{v}_1 \cdots \mathbf{v}_k \mathbf{v}_{k+1} \cdots \mathbf{v}_n) = (\mathbf{v}_1 \cdots \mathbf{v}_k \mathbf{v}_{k+1} \cdots \mathbf{v}_n) \begin{pmatrix} A_1 & \mathbf{0} \\ \mathbf{0} & A_2 \end{pmatrix}.$$

其中A_1是$\sigma|_W$在W的基$\mathbf{v}_1, \cdots, \mathbf{v}_k$下的矩阵;

A_2是$\sigma|_U$在U的基$\mathbf{v}_{k+1}, \cdots, \mathbf{v}_n$下的矩阵.

一般地, 不变子空间W没有不变补空间U时, 可取V的基使σ的矩阵表示形如

$$\begin{pmatrix} A_1 & B \\ \mathbf{0} & A_2 \end{pmatrix}.$$

其中A_1是$\sigma|_W$的矩阵表示.

5.7 幂零变换

定义: 若存在自然数m,使得$\sigma^m = 0$(相应地$A^m = 0$),则称线性变换σ(矩阵A)是幂零的.具有上述性质的最小的数m,称为线性变换(矩阵)的幂零次数.

命题: 幂零变换的特征值都是零.

推论: 非零的幂零变换不可能对角化.

5.7 幂零变换

例：设$\sigma : V^n \to V^n$为线性变换, V中有一向量$\mathbf{e} \neq \mathbf{0}$,使得$\mathbf{e}, \sigma\mathbf{e}, \cdots, \sigma^{n-1}\mathbf{e}$构成$V$的一组基,且$\sigma^n\mathbf{e} = 0$. 在基底$e_1 = \sigma^{n-1}\mathbf{e}, \cdots, \mathbf{e}_{n-1} = \sigma\mathbf{e}, \mathbf{e}_n = \mathrm{e}$下,这个线性变换的矩阵是

$$\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{pmatrix} \qquad (*)$$

称这个线性变换是“循环线性变换”.

5.7 幂零变换

注: 对任意向量$\mathbf{x}$, $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, 及$m \leq n$,
有$\sigma^m\mathbf{x} = x_{m+1}\mathbf{e}_1 + \cdots x_n\mathbf{e}_{n-m}$, 特别有$\sigma^n\mathbf{x} = \mathbf{0}$.
所以,循环线性变换是幂零的,且幂零次数是n.

5.7 幂零变换

定理: 对任意幂零线性变换$\sigma: V \to V$,空间V必定可分解为σ的不变线性子空间的直和$V = W_1 \oplus \cdots \oplus W_s$, 使得线性变换$\sigma$在每个子空间$W_i$上诱导的线性变换$\sigma|_{W_i}$是循环线性变换.

5.7 幂零变换

证明: 设$\sigma : V \to V$为任一幂零线性变换,其幂零次数为m. 令

$$V_i := \mathrm{Im}\sigma^i, \qquad i = 0, 1, \cdots, m.$$

由于$\sigma^{i+1} = \sigma^i \sigma$, 故

$$\{0\} = V_m \subset V_{m-1} \subset \cdots \subset V_{i+1} \subset V_i \subset \cdots V_1 \subset V_0 = V.$$

这里规定$\sigma^0 = Id$.

由以上构造知$\sigma(V_i) = V_{i+1}\ (0 \leq i < m)$, 特别$V_{m-1} \subset \ker \sigma$.

5.7 幂零变换

记$p_{m-1}=\dim V_{m-1}$. 对V_{m-1}的任一组基

$$\mathbf{e}_1^{(m-1)},\cdots,\mathbf{e}_{p_{m-1}}^{(m-1)},\qquad (1)$$

有 $\sigma\mathbf{e}_1^{(m-1)}=\cdots=\sigma\mathbf{e}_{p_{m-1}}^{(m-1)}=0.$

因为$\sigma(V_{m-2})=V_{m-1}$,所以在V_{m-2}中有向量 $\mathbf{e}_1^{(m-2)},\cdots,\mathbf{e}_{p_{m-1}}^{(m-2)}$, 使得

$$\sigma\mathbf{e}_1^{(m-2)}=\mathbf{e}_1^{(m-1)},\cdots,\sigma\mathbf{e}_{p_{m-1}}^{(m-2)}=\mathbf{e}_{p_{m-1}}^{(m-1)}.\qquad (2)$$

5.7 幂零变换

断言: 在V_{m-2}中, $\mathbf{e}_1^{(m-1)}, \cdots, \mathbf{e}_{p_{m-1}}^{(m-1)}, \mathbf{e}_1^{(m-2)}, \cdots, \mathbf{e}_{p_{m-1}}^{(m-2)}$ (3)
线性无关.

事实上, 设

$$k_1\mathbf{e}_1^{(m-1)} + \cdots + k_{p_{m-1}}\mathbf{e}_{p_{m-1}}^{(m-1)} + l_1\mathbf{e}_1^{(m-2)} + \cdots + l_{p_{m-1}}\mathbf{e}_{p_{m-1}}^{(m-2)} = 0. \quad (4)$$

把σ作用在上式两边, 得 $l_1\mathbf{e}_1^{(m-1)} + \cdots + l_{p_{m-1}}\mathbf{e}_{p_{m-1}}^{(m-1)} = 0$,
由于$\mathbf{e}_1^{(m-1)}, \cdots, \mathbf{e}_{p_{m-1}}^{(m-1)}$为$V_{m-1}$的一组基，得$l_1 = \cdots = l_{p_{m-1}} = 0$.
代回(4),得到$k_1 = \cdots = k_{p_{m-1}} = 0$. 断言得证.

5.7 幂零变换

于是可将向量组(3)扩充为V_{m-2}的一组基:

$$\mathbf{e}_1^{(m-1)},\cdots,\mathbf{e}_{p_{m-1}}^{(m-1)},\mathbf{e}_1^{(m-2)},\cdots,\mathbf{e}_{p_{m-1}}^{(m-2)},\mathbf{e}_{p_{m-1}+1}^{(m-2)},\cdots,\mathbf{e}_{p_{m-2}}^{(m-2)}. \quad (5)$$

其中$p_{m-2}=\dim V_{m-2}-\dim V_{m-1}$.

断言: 补充的向量$\mathbf{e}_{p_{m-1}+1}^{(m-2)},\cdots,\mathbf{e}_{p_{m-2}}^{(m-2)}$可取自$\ker\sigma$.

5.7 幂零变换

这是因为$V_{m-1}=\sigma(V_{m-2})$,可设

$$\sigma(e_{p_{m-1}+i}^{(m-2)}) = x_{i1}e_1^{(m-1)} + \cdots x_{ip_{m-1}}e_{p_{m-1}}^{(m-1)}.$$

令 $\tilde{e}_{p_{m-1}+i}^{(m-2)} = e_{p_{m-1}+i}^{(m-2)} - x_{i1}e_1^{(m-2)} - \cdots - x_{ip_{m-1}}e_{p_{m-1}}^{(m-2)}.$

于是$\sigma(\tilde{e}_{p_{m-1}+i}^{(m-2)})=0$.下面仍记之为$e_{p_{m-1}+i}^{(m-2)}$.

因为$V_{m-2}=\sigma(V_{m-3})$,在V_{m-3}中有向量

$$\mathbf{e}_1^{(m-3)}, \cdots, \mathbf{e}_{p_{m-2}}^{(m-3)}, \qquad (6)$$

使得 $\sigma(\mathbf{e}_i^{(m-3)}) = \mathbf{e}_i^{(m-2)}$ $(i=1,\cdots,p_{m-2})$.

5.7 幂零变换

运用同样的方法, 可证向量组(5)和(6)构成的向量组线性无关.

于是可以扩充得到V_{m-3}的基:

$$\begin{array}{lllllll}
e_1^{(m-1)}, & \cdots, & e_{p_{m-1}}^{(m-1)}, & & & & \\
e_1^{(m-2)}, & \cdots, & e_{p_{m-1}}^{(m-2)}, & \cdots, & e_{p_{m-2}}^{(m-2)}, & & \\
e_1^{(m-3)}, & \cdots, & e_{p_{m-1}}^{(m-2)}, & \cdots, & e_{p_{m-2}}^{(m-3)}, & \cdots, & e_{p_{m-3}}^{(m-3)}.
\end{array}$$

同理可以证明补充的向量$\mathbf{e}_{p_{m-2}+1}^{(m-3)}, \cdots, \mathbf{e}_{p_{m-3}}^{(m-3)}$可以取自$\ker\sigma$,

即 $\sigma\mathbf{e}_{p_{m-2}+1}^{(m-3)} = \cdots = \sigma\mathbf{e}_{p_{m-3}}^{(m-3)} = 0.$

5.7 幂零变换

一步一步构造，得到$V_0 = V$的一组基：

$$
\begin{array}{cccccccccc}
0 = V_m & & & & & & & & & \\
\cap & & & & & & & & & \\
V_{m-1} & e_1^{(m-1)} & \cdots & e_{p_{m-1}}^{(m-1)}, & & & & & & \\
\cap & \uparrow & & \uparrow & & & & & & \\
V_{m-2} & e_1^{(m-2)}, & \cdots, & e_{p_{m-1}}^{(m-2)}, & \cdots, & e_{p_{m-2}}^{(m-2)}, & & & & \\
\cap & \uparrow & & \uparrow & & \uparrow & & & & \\
V_{m-3} & e_1^{(m-3)}, & \cdots, & e_{p_{m-1}}^{(m-3)}, & \cdots, & e_{p_{m-2}}^{(m-3)}, & \cdots, & e_{p_{m-3}}^{(m-3)}, & & \\
 & \vdots & & \vdots & & \vdots & & \vdots & & \\
\cap & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
V = V_0 & e_1^{(0)}, & \cdots, & e_{p_{m-1}}^{(0)}, & \cdots, & e_{p_{m-2}}^{(0)}, & \cdots, & e_{p_{m-3}}^{(0)}, & \cdots & e_{p_0}^{(0)}. \\
 & W_1 & & W_{p_{m-1}} & & W_{p_{m-2}} & & W_{p_{m-3}} & & W_{p_0}
\end{array}
$$

5.7 幂零变换

在σ的作用下, 每一列向量构成σ的一个不变子空间，σ限制在这些 子空间上是循环线性变换. 每一列中最低的一个向量可作为这个循环线性变换的向量$\mathbf{e}$. V是这些不变子空间的直和,

故定理得证.

注： 由证明得到

$$
\begin{aligned}
&d_m = p_{m-1}\text{个}m\text{阶Jordan块},\\
&d_{m-1} = p_{m-2} - p_{m-1}\text{个}(m-1)\text{阶Jordan块},\\
&d_{m-2} = p_{m-3} - p_{m-2}\text{个}(m-2)\text{阶Jordan块},\\
&\cdots,\\
&d_2 = p_1 - p_2\text{个}2\text{阶Jordan块},\\
&d_1 = p_0 - p_1\text{个}1\text{阶Jordan块},
\end{aligned}
$$

5.7 幂零变换

其中

$$\begin{aligned}
d_m &= \dim V_{m-1}, \\
d_{m-1} &= \dim V_{m-2} - 2\dim V_{m-1}, \\
d_{m-2} &= \dim V_{m-3} - 2\dim V_{m-2} + \dim V_{m-1}, \\
&\cdots \\
d_1 &= \dim V_0 - 2\dim V_1 + \dim V_2.
\end{aligned}$$

5.8 Jordan标准形

回到任意线性变换$\sigma : V \to V$, 特征子空间V_λ是使线性变换$\sigma - \lambda I$在其上作用为零的最大子空间.

类似地可引入

定义: 空间V中使线性变换$\sigma - \lambda I$在其上作用为幂零的最大子空间R_λ称为是属于特征值λ的根子空间. 对非零向量$v \in V$, 若存在整数$m \geq 0$, 使得$(\sigma - \lambda I)^m v = 0$,则称$v$为属于特征值$\lambda$的根向量.

5.8 Jordan标准形

容易证明, 根子空间R_λ是线性变换$\sigma - \lambda I$的不变子空间, 进而也是σ-不变子空间.

命题: 当$\mu \neq \lambda$时, $(\sigma - \mu I)|_{R_\lambda}$是可逆的.

命题: 属于不同特征值的根向量线性无关.

命题: 对于线性变换σ的任何特征值λ, $\dim R_\lambda$等于该特征值的代数重数.

5.8 Jordan标准形

定理: 对任何线性变换$\sigma : V \to V$, 若它的特征值都落在属于$\mathbf{F}$中, 则空间V是该线性变换的根子空间的直和:

$$\boxed{V = R_{\lambda_1} \oplus \cdots \oplus R_{\lambda_m}.}$$

由于$(\sigma - \lambda I)|_{R_\lambda}$为幂零变换. 利用前面关于幂零线性变换的讨论, 每个根子空间R_λ可以分解称若干$\sigma - \lambda I$-不变子空间的W_i直和, $(\sigma - \lambda I)|_{W_i}$是循环线性变换, 则存在一组基使得$(\sigma - \lambda I)|_{W_i}$的矩阵表示是特征值为零的Jordan块(*). 于是,$\sigma|_{W_i}$在这组基下的矩阵表示是一个特征值为λ的Jordan块.

5.8 Jordan标准形

综上所述，我们又证明了

定理： 对任意线性变换$\sigma : V \to V$，若它的特征值都在$\mathbb{F}$中，则存在空间V的一组基，在该基下σ的矩阵为Jordan矩阵，形如

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix}, \quad J_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}.$$

5.8 Jordan标准形

换言之,

定理: 设$\mathbf{F}$为代数闭域, V为$\mathbf{F}$上的有限维线性空间, $\sigma: V \to V$为线性变换.则存在σ的Jordan基$v_1, \cdots, v_n$, 使得σ在旧基下的矩阵A经基底变换,可化为Jordan矩阵J, 即存在非退化矩阵P, 使得$P^{-1}AP = J$.