线性代数第二次自救

定义4.3.1(可对角化) 若方阵A相似于一个对角矩阵，则称A可对角化.

问题3：若 $P^{-1}AP=D$(对角阵)，P和D有什么特点？(从形式$AP=PD$考虑)

A_n可对角化 $\Leftrightarrow$ A_n有n个无关特征向量

矩阵能否对角化与是否满秩（特征值是否非零）、特征根是否有重根之间没有必然的联系。这些是矩阵的不同性质，彼此独立。具体说明如下：

1. 可对角化与满秩（特征值非零）

- **可对角化但不满秩**：例如零矩阵 $\mathbf{0}$，它可对角化（本身就是对角矩阵），但秩为0，特征值全为零。
- **满秩但不可对角化**：例如二阶Jordan块 $A=\begin{pmatrix}1&1\\0&1\end{pmatrix}$，它满秩（可逆），特征值均为1（非零），但不可对角化（缺乏足够的线性无关特征向量）。

2. 可对角化与特征根有重根

- **有重根但可对角化**：例如单位矩阵 I，特征值全为1（重根），但它可对角化（本身就是对角矩阵）。
- **有重根且不可对角化**：如上文的Jordan块 A，特征值1是二重的，但不可对角化。

3. 满秩与特征根有重根

- **满秩且有重根**：例如单位矩阵 I，满秩且特征值全为1（重根）。
- **满秩且无重根**：例如对角矩阵 $\mathrm{diag}(1,2)$，满秩且特征值互异。

4. 特征值非零与可对角化

- 特征值非零（即满秩）不能保证可对角化（如Jordan块），特征值有零也不妨碍可对角化（如零矩阵）。

定理4.3.1 n阶矩阵可对角化的充要条件是有n个线性无关的特征向量；且对角矩阵的主对角线由特征值(可按任意次序)构成，相似变换矩阵由属于相应特征值的特征向量构成.

定理4.3.2 属于不同特征值的特征向量线性无关.

推论4.3.3 若n阶矩阵有n个互不相同的特征值，则矩阵可对角化.

注 由哈密顿-凯莱定理，任意方阵A都满足$f(A)=O$，其中$f(\lambda)=|\lambda E-A|$.

重特征值条件下的对角化

定理4.3.4 若$\lambda_1, \lambda_2, \dots, \lambda_m$是矩阵$A$的不同特征值，而$A$的属于$\lambda_i$的线性无关的特征向量为$\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{is_i}$ (i=1,2, … ,m)，则向量组

$$\alpha_{11}, \dots, \alpha_{1s_1}, \alpha_{21}, \dots, \alpha_{2s_2}, \dots, \alpha_{m1}, \dots, \alpha_{ms_m},$$

线性无关.

定理4.3.5 设λ_0是n阶方阵A的k重特征值，则A的属于特征值λ_0的线性无关的特征向量个数不超过k.

说明：

$$\underbrace{\xi_1, \xi_2, \dots, \xi_s}_{\lambda_0}, \eta_{s+1}, \dots, \eta_n \text{ 线性无关}$$

- 将向量 $e_1, e_2, \dots, e_n$依次添加到向量组$\xi_1, \xi_2, \dots, \xi_s$中
- 当添加的向量使得添加后的向量组线性相关时，删去该添加的向量
- 最后得扩充的线性无关组$\xi_1, \xi_2, \dots, \xi_s, \eta_{s+1}, \dots, \eta_n$.

$$A(\xi_1, \dots, \xi_s, \eta_{s+1}, \dots, \eta_n)=(\lambda_0\xi_1, \dots, \lambda_0\xi_s, \varphi_{s+1}, \dots, \varphi_n)$$

$$= \underbrace{(\xi_1, \dots, \xi_s, \eta_{s+1}, \dots, \eta_n)}_{P} \underbrace{\left(\begin{array}{ccc|cc} \lambda_0 & & & * & * \\ & \ddots & & * & * \\ & & \lambda_0 & * & * \\ \hline & & & * & * \\ & & & * & * \end{array}\right)}_{B}$$

结论：$P^{-1}AP=B$

$$|\lambda E-A|=|\lambda E-B|=(\lambda-\lambda_0)^s g(\lambda)=(\lambda-\lambda_0)^k f(\lambda)\text{，其中 } f(\lambda_0)\neq 0$$

* A的特征值λ_0的线性无关特征向量个数(λ_0几何重数)$\leq\lambda_0$的重数(λ_0代数重数)

定理4.3.6 n阶方阵A可对角化的充要条件是每个k_i重特征值λ_i对应的特征矩阵$\lambda_i E-A$的秩为$n-k_i$.

将矩阵A对角化的具体步骤：

(1) 解特征方程 $|\lambda E-A|=0$ 得特征值 $\lambda=\lambda_1(s_1$重$), \dots, \lambda_m(s_m$重$)$；

(2) 对每个特征值 λ_i, i=1,2,…,m，解齐次方程组

$$(\lambda_i E-A)x=\theta,$$

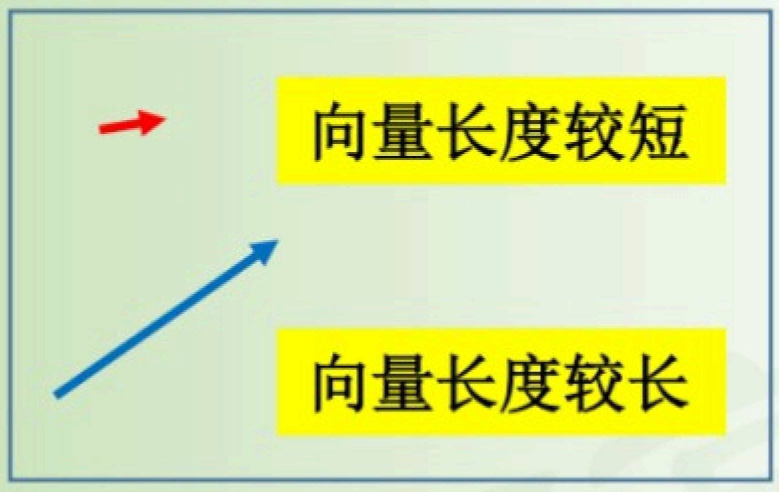
得一个基础解系(即无关特征向量) $\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{ir_i}$.
若有某个 i 使得 $r_i<s_i$ (λ_i几何重数<代数重数)，则矩阵 A 不可对角化；

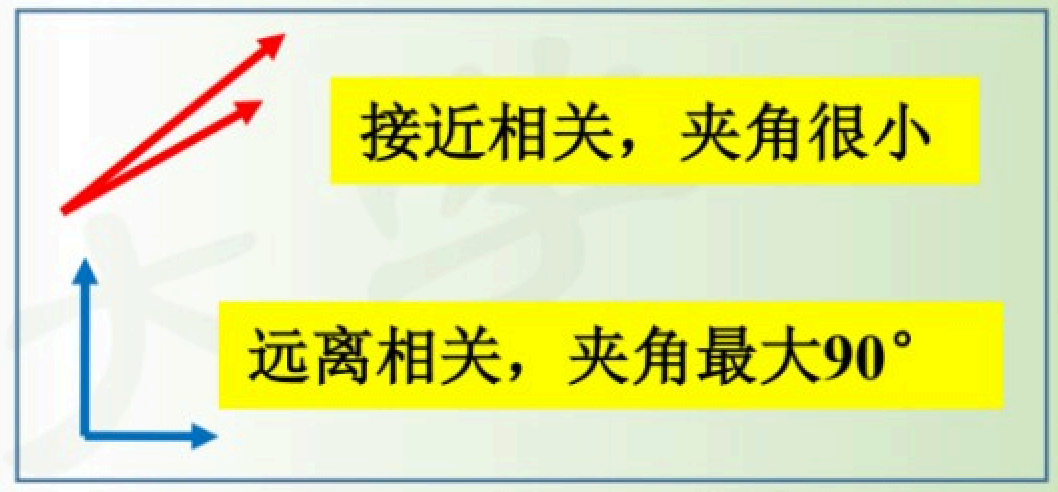
(3) 当所有的 $r_i=s_i$, i=1,2, … ,m，则令

$$P=(\alpha_{11}, \dots, \alpha_{1s_1}, \alpha_{21}, \dots, \alpha_{2s_2}, \dots, \alpha_{m1}, \dots, \alpha_{ms_m}),$$

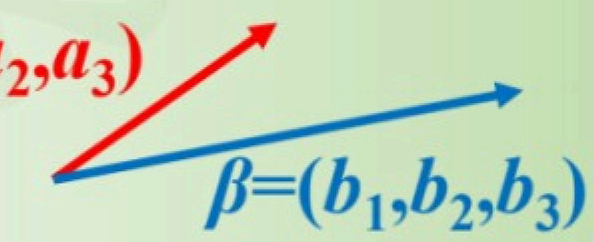
即得 $P^{-1}AP=\text{diag}(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \dots, \lambda_m, \dots, \lambda_m)$.

两个向量，除了线性相关、线性无关，还需要其它更多的信息，如向量长度，它们相关或无关的程度(接近相关还是远离相关).





从二维、三维向量的长度、夹角关系我们知道可以用向量的内积来表示向量的长度和两个向量之间的夹角.



右图两个向量内积为：$(\alpha,\beta)=a_1b_1+a_2b_2+a_3b_3$.

长度：$|\alpha|=\sqrt{(\alpha,\alpha)}$ 夹角：α,β夹角=arccos($(\alpha,\beta)/(|\alpha|\cdot|\beta|)$)

定义4.4.1 (向量内积) 设α,β为n维向量，用列向量表示为$\alpha=(a_1,a_2,\dots,a_n)^T$, $\beta=(b_1,b_2,\dots,b_n)^T$. 若$\alpha,\beta$为实向量，则称$a_1b_1+a_2b_2+\dots+a_nb_n$为$\alpha,\beta$的**实内积**；若$\alpha,\beta$为复向量，则称 $a_1\bar{b}_1+a_2\bar{b}_2+\cdots+a_n\bar{b}_n$ 为α,β的**复内积**；统称为向量的**内积**，记为(α,β)，并称$\|\alpha\|=\sqrt{(\alpha,\alpha)}$为向量的**长度**或**模**；称模为1的向量为**单位向量**.

共轭转置

显然，实向量α,β的内积$(\alpha,\beta)=\alpha^T\beta$；复向量$\alpha,\beta$的内积$(\alpha,\beta)=\alpha^T\bar{\beta}=\bar{\beta}^T\alpha=\beta^H\alpha$.

实内积的基本性质：

(1) $(\alpha,\beta)=(\beta,\alpha)$;

(2) $(k\alpha,\beta)=k(\alpha,\beta)$;

(3) $(\alpha+\gamma,\beta)=(\alpha,\beta)+(\gamma,\beta)$;

(4) $(\alpha,\alpha)\geq 0$；$(\alpha,\alpha)=0$ 当且仅当$\alpha=\theta$.

说明 利用定义 $(\alpha,\beta)=a_1b_1+a_2b_2+\dots+a_nb_n$易得结论

定义4.4.2 (向量夹角、向量正交) 若$(\alpha,\beta)=0$，则称α和β**正交**或**垂直**. 若α,β均为非零实向量，则称 $\arccos\left(\frac{(\alpha,\beta)}{\|\alpha\|\cdot\|\beta\|}\right)$ 为向量α和β的**夹角**.

定义4.4.3 (正交向量组、法正交组) 若一个不含零向量的向量组中的向量两两正交，则称该向量组为**正交向量组**；若一个正交向量组中的 向量均为单位向量，则该向量组称为**标准正交向量组**，简称**法正交组**.

定理4.4.2 (施密特(Schmidt)正交化) 由线性无关向量组$\alpha_1,\alpha_2,\dots,\alpha_n$可构造出与之等价的正交向量组$\xi_1,\xi_2,\dots,\xi_n$. 并且$\xi_i$可以表示成$\alpha_1,\alpha_2,\dots,\alpha_i$, $i=1,2,\dots,n$的线性组合.

施密特(Schmidt)正交化说明 用构造法.

第一步，取$\xi_1=\alpha_1\neq\theta$.

第二步，取$\xi_2=\alpha_2-k_{21}\xi_1=\alpha_2-k_{21}\alpha_1$，要满足$\xi_1, \xi_2$ 构成正交组，即满足：$\xi_2\neq\theta$；$(\xi_2, \xi_1)=0$

(1) 若$\xi_2=\theta$，则$\xi_2=\alpha_2-k_{21}\alpha_1=\theta$，即$\alpha_1, \alpha_2$线性相关，矛盾.

(2) $(\xi_2, \xi_1)=(\alpha_2, \xi_1)-k_{21}(\xi_1, \xi_1)=0$，只要取$k_{21}=(\alpha_2, \xi_1)/(\xi_1, \xi_1)$.

依次下去到第 i 步之前，则我们已经由$\alpha_1, \dots, \alpha_{i-1}$构造出等价正交组$\xi_1, \dots, \xi_{i-1}$且$\xi_k$可表示成$\alpha_1, \dots, \alpha_k$, $k=1,2,\dots,i-1$的线性组合.

第i步，取 $\xi_i=\alpha_i-k_{i1}\xi_1-k_{i2}\xi_2-\dots-k_{i,i-1}\xi_{i-1}$，

要求满足$\xi_1, \dots, \xi_{i-1}, \xi_i$构成正交组，需 $\xi_i\neq\theta$；$(\xi_i, \xi_k)=0, k=1,\dots,i-1$

(1) 若$\xi_i=\theta$，因为ξ_k可以表示成$\alpha_1, \dots, \alpha_k$ 的线性组合，则

$\xi_i=\alpha_i-k_{i1}\xi_1-\dots-k_{i,i-1}\xi_{i-1}=\alpha_i-t_{i1}\alpha_1-\dots-t_{i,i-1}\alpha_{i-1}=\theta$, 即$\alpha_1, \dots,\alpha_i$ 线性相关,矛盾.

(2) $(\xi_i, \xi_k)=(\alpha_i, \xi_k)-k_{i1}(\xi_1, \xi_k)-\dots-k_{ik}(\xi_k, \xi_k)-\dots-k_{i,i-1}(\xi_{i-1}, \xi_k)$

$=(\alpha_i, \xi_k)-0-0-\dots-k_{ik}(\xi_k, \xi_k)-0-\dots-0=(\alpha_i, \xi_k)-k_{ik}(\xi_k, \xi_k)=0,$

只要取$k_{ik}=(\alpha_i, \xi_k)/(\xi_k, \xi_k)$, $k=1,2,\dots,i-1$.

一直下去, 最后构成正交组$\xi_1,\dots,\xi_n$.等价性由$\alpha_i=k_{i1}\xi_1+\dots+k_{i,i-1}\xi_{i-1}+\xi_i$可知.

注 由无关组$\alpha_1, \dots,\alpha_n$可构造出等价法正交组$\beta_1,\dots,\beta_n$，且β_i可由$\alpha_1,\dots, \alpha_i$表示.

例4.4.4 将3个线性无关4维向量组 $\{\alpha_1, \alpha_2, \alpha_3\}$ 标准正交化，其中

$$\alpha_1=\begin{pmatrix}1\\0\\0\\1\end{pmatrix}, \alpha_2=\begin{pmatrix}1\\1\\0\\1\end{pmatrix}, \alpha_3=\begin{pmatrix}1\\1\\1\\0\end{pmatrix}.$$

解 先正交化，令

$$\xi_1=\alpha_1=\begin{pmatrix}1\\0\\0\\1\end{pmatrix},\qquad \xi_2=\alpha_2-\frac{(\alpha_2,\xi_1)}{\|\xi_1\|^2}\xi_1=\begin{pmatrix}1\\1\\0\\1\end{pmatrix}-\frac{2}{2}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},$$

$$\xi_3=\alpha_3-\frac{(\alpha_3,\xi_1)}{\|\xi_1\|^2}\xi_1-\frac{(\alpha_3,\xi_2)}{\|\xi_2\|^2}\xi_2=\begin{pmatrix}1\\1\\1\\0\end{pmatrix}-\frac{1}{2}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}-\frac{1}{1}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}=\begin{pmatrix}1/2\\0\\1\\-1/2\end{pmatrix}.$$

再单位化，

$$\beta_1=\frac{1}{\|\xi_1\|}\xi_1=\frac{1}{\sqrt{2}}\xi_1=\begin{pmatrix}1/\sqrt{2}\\0\\0\\1/\sqrt{2}\end{pmatrix},\qquad \beta_2=\frac{1}{\|\xi_2\|}\xi_2=\xi_2=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\qquad \beta_3=\frac{1}{\|\xi_3\|}\xi_3=\sqrt{\frac{2}{3}}\xi_3=\begin{pmatrix}1/\sqrt{6}\\0\\2/\sqrt{6}\\-1/\sqrt{6}\end{pmatrix}.$$

得标准正交向量组为 $\{\beta_1, \beta_2, \beta_3\}$.

施密特(Schmidt)正交化的矩阵形式(正交分解)

若 $A\in\mathbf{R}^{m\times n}$，$r(A)=n$，则有 分解 $A=QR$，
其中$Q\in\mathbf{R}^{m\times n}$, $R\in\mathbf{R}^{n\times n}$，且$Q^{\mathrm{T}}Q=E$，$R$为上三角可逆矩阵.

说明：将A按列分块，$A=(\alpha_1,\alpha_2,\dots,\alpha_n)$，利用施密特正交化定理得

$\xi_1=\alpha_1$, $\xi_2=\alpha_2-k_{21}\xi_1$, … , $\xi_n=\alpha_n-k_{n1}\xi_1-\dots-k_{n,n-1}\xi_{n-1}$，

即 $\alpha_1=\xi_1$, $\alpha_2=k_{21}\xi_1+\xi_2,\dots,\alpha_n=k_{n1}\xi_1+\dots+k_{n,n-1}\xi_{n-1}+\xi_n$，

即 $(\alpha_1,\dots,\alpha_n)=(\xi_1,\dots,\xi_n)R_1$，其中$R_1=\begin{pmatrix}1 & k_{21} & \cdots & k_{n1}\\ 0 & 1 & \cdots & k_{n2}\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{pmatrix}$，

令 $\beta_i=\|\xi_i\|^{-1}\xi_i$, $i=1,2,\dots,n$, $Q=(\beta_1,\dots,\beta_n)$, $D=\mathrm{diag}(\|\xi_1\|,\|\xi_2\|,\dots,\|\xi_n\|)$.

则有 $A=(\xi_1,\dots,\xi_n)R_1=QDR_1=QR$，则 $R=DR_1=\begin{pmatrix}\|\xi_1\| & * & \cdots & *\\ & \|\xi_2\| & \cdots & *\\ & & \ddots & \vdots\\ & & & \|\xi_n\|\end{pmatrix}$

由$\beta_i^{\mathrm{T}}\beta_j=\delta_{ij}$知$Q^{\mathrm{T}}Q=(\beta_i^{\mathrm{T}}\beta_j)=E$. 由 $|R|=\|\xi_1\|\times\dots\times\|\xi_n\|\neq0$知 R是上三角可逆阵.

* 推广此处的矩阵Q，则完整的标准正交向量组 $\beta_1,\beta_2,\dots,\beta_n$ 构成的矩阵 $A=(\beta_1,\beta_2,\dots,\beta_n)$ 有关系 $A^{\mathrm{T}}A=(\beta_i^{\mathrm{T}}\beta_j)=E$ ，A称为正交矩阵.

定义4.4.4 (正交矩阵) 若实方阵A满足 $A^{\mathrm{T}}A=E$，则称A为正交矩阵.

注 若A, B是同阶的正交矩阵，则AB也是正交矩阵 .

$(AB)^{\mathrm{T}}AB=B^{\mathrm{T}}A^{\mathrm{T}}AB$
$=B^{\mathrm{T}}EB=B^{\mathrm{T}}B=E$

正交矩阵的性质

定理4.4.3 对于方阵A，下列条件互为等价：

(1) A为正交矩阵（$A^{\mathrm{T}}A=E$）；

(2) $A^{\mathrm{T}}=A^{-1}$；

(3) $AA^{\mathrm{T}}=E$；

(4) A的列向量构成标准正交列向量组；

(5) A的行向量构成标准正交行向量组.

定理4.4.4 设A为n阶正交矩阵，λ为A的特征值，α为n维列向量，则有

(1) $|A|^2=1$；

(2) $\overline{(A\alpha)}^{\mathrm{T}}(A\alpha)=\overline{\alpha}^{\mathrm{T}}\alpha$;

(3) $|\lambda|=1$.

实对称矩阵对角化特点：

- 实对称矩阵一定可以对角化
- 与实对称矩阵相似的对角矩阵是实对角矩阵（对角元为实数）
- 存在正交的相似变换矩阵，使得实对称矩阵对角化

后两条隐含了如下含义：

•实对称矩阵的特征值都是实数

•实对称矩阵属于不同特征值的特征向量正交

第一条表示的是：

•对任意特征值λ，有 $r(\lambda E-A)+\lambda$的重数$=n$ (矩阵阶数) 或几何重数=代数重数

实对称矩阵特征值与特征向量

定理4.5.1 实对称矩阵的特征值均为实数.

证明思路 $A\xi=\lambda\xi$，

考虑 $A\xi$，凑成对称形式 $\bar{\xi}^{T}A\xi$，且 $\bar{\xi}^{T}\xi=\|\xi\|^2$，于是

$\bar{\xi}^{T}A\xi=\lambda\bar{\xi}^{T}\xi=\lambda\|\xi\|^2$，共轭转置$\bar{\xi}^{T}A\xi=(\xi^{T}\bar{A}\bar{\xi})^{T}=(\overline{\bar{\xi}^{T}A\xi})^{T}=\bar{\lambda}\|\xi\|^2$

则 $\lambda\|\xi\|^2=\bar{\lambda}\|\xi\|^2$，即 $\lambda=\bar{\lambda}$.

注意 当矩阵为实矩阵，特征值也是实数时，我们只考虑实的特征向量.

定理4.5.2 实对称矩阵的属于不同特征值的特征向量相互正交.

证明思路 $A\xi_1=\lambda_1\xi_1$， $A\xi_2=\lambda_2\xi_2$，

将 $A\xi_1$, $A\xi_2$ 凑成对称形式 $\xi_1^{T}A\xi_2=\xi_2^{T}A\xi_1$ $(\because \xi_1^{T}A\xi_2=(\xi_1^{T}A\xi_2)^{T}=\xi_2^{T}A^{T}\xi_1)$

于是 $\lambda_2\xi_1^{T}\xi_2=\lambda_1\xi_2^{T}\xi_1=\lambda_1\xi_1^{T}\xi_2$，则 $(\lambda_2-\lambda_1)\xi_1^{T}\xi_2=0$，即 $\xi_1^{T}\xi_2=0$.

引理4.5.3 设有实n维单位列向量β，则必能找到n-1个向量与β一起构成由n个向量组成的标准正交向量组.

证明思路 用 $\beta, e_1, e_2, \dots, e_n$ 构成n个无关向量 $\beta, \alpha_2, \alpha_3, \dots, \alpha_n$,

利用施密特正交化方法标准正交化：$\beta_1, \beta_2, \beta_3, \dots, \beta_n$, $(\beta_1=\beta)$.

定理4.5.4 若A是实对称矩阵，则存在同阶的正交矩阵P使得$P^{T}AP$是实对角矩阵，从而实对称矩阵可对角化.

证明思路 数学归纳法：m+1阶矩阵A，$Aq_1=\lambda_1 q_1$, $Q_1=(q_1,q_2,\dots,q_{m+1})$正交阵

$$Q_1^{T}AQ_1=\begin{pmatrix} q_1^{T}\\ q_2^{T}\\ \vdots\\ q_{m+1}^{T}\end{pmatrix}A(q_1,q_2,\cdots,q_{m+1})=\begin{pmatrix} q_1^{T}\\ q_2^{T}\\ \vdots\\ q_{m+1}^{T}\end{pmatrix}(\lambda_1 q_1,Aq_2,\cdots,Aq_{m+1})=\begin{pmatrix}\lambda_1 & * & \cdots & *\\ 0 & * & \cdots & *\\ \vdots & \vdots & & \vdots\\ 0 & * & \cdots & *\end{pmatrix}=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & * & \cdots & *\\ \vdots & \vdots & B & \vdots\\ 0 & * & \cdots & *\end{pmatrix}$$

矩阵对称

B对称

进一步利用归纳假设：$Q_2^{T}BQ_2=\Lambda$，B为上述右下块，则

$$P^{T}AP=\begin{pmatrix}1 & 0\\ 0 & Q_2\end{pmatrix}^{T}Q_1^{T}AQ_1\begin{pmatrix}1 & 0\\ 0 & Q_2\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & Q_2^{T}\end{pmatrix}\begin{pmatrix}\lambda_1 & 0\\ 0 & B\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & Q_2\end{pmatrix}=\begin{pmatrix}\lambda_1 & 0\\ 0 & \Lambda\end{pmatrix},\text{其中}P^{T}P=E$$

* 对方阵A，存在正交矩阵P使得$P^{\mathrm{T}}AP$是上三角矩阵.

定义5.1.1 (二次型) 含有n个变量的在某个数域上的二次齐次多项式

$$f(x_1, x_2, \dots, x_n) = a_{11}x_1^2+2a_{12}x_1x_2+2a_{13}x_1x_3+\dots+2a_{1n}x_1x_n$$
$$a_{22}x_2^2+2a_{23}x_2x_3+\dots+2a_{2n}x_2x_n \qquad (5.2)$$
$$+\dots\dots\dots$$
$$+a_{nn}x_n^2$$

称为二次型； 若全部 $a_{ij}\in\mathrm{R}$，则称式(5.2)中的f为实二次型；若有元素 $a_{ij}\in\mathrm{C}$，则称式(5.2)中的f为复二次型 .

(5.2)式的f可改写为

$$\begin{aligned} f(x_1, x_2, \dots, x_n) = {} & a_{11}x_1^2 + a_{12}x_1x_2+a_{13}x_1x_3+\dots+a_{1n}x_1x_n \\ & +a_{12}x_2x_1 + a_{22}x_2^2 +a_{23}x_2x_3+\dots+a_{2n}x_2x_n \\ & +\dots\dots\dots \\ & +a_{1n}x_nx_1+a_{2n}x_nx_2+a_{3n}x_nx_3+\dots+a_{nn}x_n^2 . \end{aligned}$$

用矩阵表示为: $f(x_1, \dots, x_n) =x^{\mathrm{T}}Ax$,

称为二次型f的矩阵表示，其中 $x=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}$.

二次型 $f(x_1, x_2, \dots, x_n) = x^{\mathrm{T}}Ax$ ，则

f与对称矩阵A 一一对应.

定义5.1.2 (二次型的矩阵) 称二次型

$$f(x_1, x_2, \dots, x_n) = x^{\mathrm{T}}Ax$$

中的对称矩阵A为二次型f的矩阵，A的秩称为二次型f的秩.

定义5.1.3 (线性变换、非退化线性变换) 称如下的变换

$$\begin{cases} x_1 = c_{11}y_1 + c_{12}y_2 + \cdots + c_{1n}y_n, \\ x_2 = c_{21}y_1 + c_{22}y_2 + \cdots + c_{2n}y_n, \\ \cdots\cdots\cdots\cdots \\ x_n = c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nn}y_n, \end{cases}$$

为由 $x_1, x_2, \dots, x_n$ 到 $y_1, y_2, \dots, y_n$ 的一个线性变换.

若线性变换的系数行列式

$$|P| = \begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{vmatrix} \neq 0,$$

则称该线性变换为非异线性变换或非退化线性变换.

若 $|P|=0$，则称该线性变换为奇异线性变换或退化线性变换.

若 P 为正交矩阵，则称该线性变换为正交变换.

定义5.1.4 (合同、合同变换) 设A和B是两个同阶方阵，若存在一个可逆矩阵P，使得有$B=P^{T}AP$，则称A合同于B. 称B为A的合同矩阵，而称P为A到B的合同变换矩阵.

矩阵的合同关系是一个等价关系，满足：(1) 自反性(2) 对称性(3) 传递性.

定义5.1.5 (二次型的标准形) 二次型 $f(x_1, x_2, \dots, x_n)$ 经过非退化线性变换后得到一个只包含变量平方项的二次型 $d_1y_1^2+d_2y_2^2+\dots+d_ny_n^2$，称为原二次型的标准形 .

定理5.1.1 存在非退化的线性变换将实二次型化为标准形，且平方项系数可按任意次序排列；存在可逆矩阵将实对称矩阵合同变换为实对角矩阵，且对角元素可以任意次序排列.

正交变换法化实二次型 $f(x)=x^{T}Ax$ 为标准形的步骤：

(1) 求解矩阵A的特征方程 $|\lambda E-A|=0$ ，解得特征值 $\lambda=\lambda_1, \lambda_2, \dots, \lambda_r$ ；

(2) 对每一个特征值 $\lambda=\lambda_i$ (s_i重)，求出齐次线性方程组

$$(\lambda_i E-A)\,x=\theta$$

的基础解系(即特征向量的极大无关组) $\xi_{i1}, \xi_{i2}, \dots, \xi_{is_i}$ ，并标准正交化为 $\eta_{i1}, \eta_{i2}, \dots, \eta_{is_i}$ ；

(3) 将标准正交化的特征向量作为列构成正交矩阵

$$P=(\eta_1, \eta_2, \dots, \eta_n),$$

则非退化线性变换 $x=Py$ 将实二次型 $f(x)$ 化为标准形

$$f(x)=g(y)=y^{T}\Lambda y,$$

其中 $\Lambda=\mathrm{diag}(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \dots, \lambda_r, \dots, \lambda_r)$ 。

配方法化实二次型 $f(x)=x^{T}Ax$ 为标准形的步骤：

反复对可能出现的以下两种情况进行处理：

情况1 式中有非零平方项，例如若非零平方项为 $a_{11}x_1^2$，则将式中所有含 x_1 的项配成一个平方项 $a_{11}\left(x_1+\frac{a_{12}}{a_{11}}x_2+\cdots+\frac{a_{1n}}{a_{11}}x_n\right)^2$，并令非退化线性变换为

$$\begin{cases} y_1=x_1+\frac{a_{12}}{a_{11}}x_2+\cdots+\frac{a_{1n}}{a_{11}}x_n, \\ y_2=x_2, \\ \cdots\cdots \\ y_n=x_n, \end{cases}$$

则可将原式化为不含 y_1 的交叉项的式子。

情况2 式中无非零平方项有非零交叉项，这时我们可以用一个线性变换配出平方项。例如，若有非零交叉项为 $2a_{12}x_1x_2$ ，则作如右边的非退化线性变换就可将原式化为含有y_1,y_2的平方项的式子，再按情况1进行处理。

$$\begin{cases} x_1=y_1+y_2, \\ x_2=y_1-y_2, \\ x_3=y_3, \\ \cdots\cdots \\ x_n=y_n, \end{cases}$$

合同变换法化实二次型 $f(x)=x^{\mathrm{T}}Ax$ 为标准形：

对矩阵 $B=\begin{pmatrix} A \\ E \end{pmatrix}$

对k=1,2,…,n，做一系列初等列变换，消去k行的所有非对角元素。再对称地消去k列的所有非对角元素，最后得到矩阵

$$\begin{pmatrix} \Lambda \\ P \end{pmatrix}=\begin{pmatrix} P^{\mathrm{T}} & \\ & E \end{pmatrix}\begin{pmatrix} A \\ E \end{pmatrix}P=\begin{pmatrix} P^{\mathrm{T}}AP \\ P \end{pmatrix}$$

其中 Λ 为对角矩阵，而 P 为非退化线性变换矩阵，变换为 $x=Py$.
原二次型化为标准形：

$$f(x)=g(y)=y^{\mathrm{T}}\Lambda y .$$

标准形中找最简单且唯一的形式——规范形

实二次型：$f(x_1, x_2, \dots, x_n)=x^{\mathrm{T}}Ax$的标准形

$$g(y_1,y_2,\dots,y_n)=d_1y_1^2+d_2y_2^2+\dots+d_ny_n^2$$

最简形式为：$\pm y_1^2\pm y_2^2\pm\dots\pm y_r^2$　　其中：r为二次型的秩

若是复二次型：最简形式可最终化为　$y_1^2+y_2^2+\dots+y_r^2$
若是实二次型：最简形式只能化为　$y_1^2+\dots+y_p^2-y_{p+1}^2-\dots-y_r^2$

定义5.1.6 (实(复)二次型的规范形) 实二次型$f(x_1, x_2, \dots, x_n)$经过非退化的实线性变换得到如下形式的二次型

$$z_1^2+\dots+z_p^2-z_{p+1}^2-\dots-z_r^2\ ,\quad r\le n,$$

称为原二次型的**实规范形**，r称为该二次型的秩；
复二次型经过非退化的复线性变换得到如下形式的二次型

$$z_1^2+z_2^2+\dots+z_r^2\ ,\quad r\le n,$$

称为原二次型的**复规范形**，r称为该二次型的秩 .

定理5.1.2 存在非退化的复线性变换将复二次型化为复规范形

$$z_1^2+z_2^2+\dots+z_r^2\ ;$$

存在复可逆矩阵将复对称矩阵合同变换为 $\mathrm{diag}(E_r, O_{n-r})$，其中 r为二次型矩阵的秩.

定理5.1.3 (惯性定理) 存在非退化的实线性变换将实二次型化为实规范形

$$z_1^2+\dots+z_p^2-z_{p+1}^2\dots-z_r^2\ ;$$

存在实可逆矩阵将实对称矩阵合同变换为 $\mathrm{diag}(E_p, -E_{r-p}, O_{n-r})$，其中 r为二次型矩阵的秩，p是唯一确定的 .

定义5.1.7 (正惯性指数、负惯性指数) 若实二次型的实规范形为

$$z_1^2+\dots+z_p^2-z_{p+1}^2-\dots-z_r^2\ ,\quad r\le n,$$

则称p为原二次型的**正惯性指数**；称$r-p$为原二次型的**负惯性指数** .

推论5.1.4 若实二次型矩阵 A 合同于对角矩阵 $B=\mathrm{diag}(b_{11},\ldots,b_{nn})$，则正对角元个数为实二次型的正惯性指数，负对角元个数为实二次型的负惯性指数，非零对角元个数为二次型的秩．

推论5.1.5 实二次型矩阵 A 的正特征值个数为正惯性指数，负特征值个数为负惯性指数，非零特征值个数为二次型的秩．

补充例5A (习题五4)证明反对称矩阵均合同于形式为 $\begin{pmatrix} & & & & & & 1 \\ & & & & & \cdot^{\cdot^{\cdot}} & \\ & & & & 1 & & \\ & & & 0 & & & \\ & & \cdot^{\cdot^{\cdot}} & & & & \\ & 0 & & & & & \\ -1 & & & & & & \\ \cdot^{\cdot^{\cdot}} & & & & & & \\ -1 & & & & & & \end{pmatrix}$ 的矩阵.

证 用数学归纳法，设 $A=\begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ -a_{12} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & \cdots & 0 \end{pmatrix}$.

当$n=1$时$A=O$；

$n=2$时，$A=O$，或 $E(2(1/a_{12}))^{\mathrm{T}}AE(2(1/a_{12}))=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$，成立.

假设$n<m$时结论成立．

当 $n=m>2$时，若$A=O$结论成立. 否则 $A=\begin{pmatrix} 0 & a_{12} & \cdots & a_{1m} \\ -a_{12} & 0 & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1m} & -a_{2m} & \cdots & 0 \end{pmatrix}$ 中有$a_{ij}\neq 0, i<j$,

令 $P_1=\begin{cases} E(m(1/a_{1m})), & a_{1m}\neq 0 \\ E(j,m)E(1,i)E(m(1/a_{ij})), & a_{1m}=0 \end{cases}$ 则有 $P_1^{\mathrm{T}}AP_1=A_2=\begin{pmatrix} 0 & c_{12} & \cdots & 1 \\ -c_{12} & 0 & \cdots & c_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -c_{2m} & \cdots & 0 \end{pmatrix}$

再令 $P_2=\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & -c_{12} & -c_{13} & \cdots & 1 \end{pmatrix}\begin{pmatrix} 1 & -c_{2m} & -c_{3m} & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$

则 $P_2^{\mathrm{T}}A_2P_2=A_3=\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & d_{23} & \cdots & d_{2,m-1} & 0 \\ 0 & -d_{23} & 0 & \cdots & d_{3,m-1} & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & -d_{2,m-1} & -d_{3,m-1} & \cdots & 0 & 0 \\ -1 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}=\begin{pmatrix} 0 & 0 & 1 \\ 0 & B & 0 \\ -1 & 0 & 0 \end{pmatrix}, B^{\mathrm{T}}=-B$

由归纳假设，存在可逆矩阵Q，使得 $Q^{\mathrm{T}}BQ=\begin{pmatrix} 0 & \cdots & 1 \\ \vdots & \cdot^{\cdot^{\cdot}} & \vdots \\ -1 & \cdots & 0 \end{pmatrix}$.

令 $P=P_1P_2\begin{pmatrix} 1 & 0 & 0 \\ 0 & Q & 0 \\ 0 & 0 & 1 \end{pmatrix}$， 则有 $P^{\mathrm{T}}AP=\begin{pmatrix} 0 & \cdots & 1 \\ \vdots & \cdot^{\cdot^{\cdot}} & \vdots \\ -1 & \cdots & 0 \end{pmatrix}$.

5.2 正定二次型

矛盾方程组 $Bx=b$，我们可以求最小二乘解，即满足 $\|Bx-b\|$ 最小的 x，而 $\|Bx-b\|$ 的最小值对应于

$$\|Bx-b\|^2=(Bx-b)^{\mathrm{T}}(Bx-b)=x^{\mathrm{T}}B^{\mathrm{T}}Bx-2b^{\mathrm{T}}Bx+b^{\mathrm{T}}b$$

的最小值，这一类的问题我们看成是求二次函数

$$f(x)=\frac{1}{2}x^{\mathrm{T}}Ax+b^{\mathrm{T}}x+c$$

的最小值的问题，其中 $x=(x_1, x_2,\ldots, x_n)^{\mathrm{T}}$，$A^{\mathrm{T}}=A$.

该函数 $f(x)$，当 A 有负特征值如 $\lambda_0<0$ 时，设 λ_0 的单位特征向量是 ξ_0，取$x=k\xi_0$，则有

$$f(x)=\frac{1}{2}\lambda_0k^2+(b^{\mathrm{T}}\xi_0)k+c,$$

由 $\frac{1}{2}\lambda_0<0$ 知 $f(x)$ 没有最小值.

当 A 有0特征值如 $\lambda_1=0$ 时，设 λ_1 的单位特征向量是 ξ_1，取 $x=k\xi_1$，则有

$$f(x)=(b^{\mathrm{T}}\xi_1)k+c,$$

当 $b^{\mathrm{T}}\xi_1\neq 0$ 时，可知 $f(x)$ 没有最小值.

故函数 $f(x)$ 只有当 A 的特征值都是正数时才能谈求 $f(x)$ 最小值的问题.

当 A 的特征值都是正数时，易知$f(x)$的二次项部分 $0.5x^{\mathrm{T}}Ax$可以化成标准形 $0.5\lambda_1y_1^2+\ldots+0.5\lambda_ny_n^2$，平方项全是正项.

再看二次型：

$$f(x,y,z)=(x,y,z)\begin{pmatrix}8&-8&8\\-8&11&-11\\8&-11&13\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}.$$

可以化成标准形

$$f(x,y,z)=(x',y',z')P^{\mathrm{T}}\begin{pmatrix}8&-8&8\\-8&11&-11\\8&-11&13\end{pmatrix}P\begin{pmatrix}x'\\y'\\z'\end{pmatrix}=(x',y',z')\begin{pmatrix}8&0&0\\0&3&0\\0&0&2\end{pmatrix}\begin{pmatrix}x'\\y'\\z'\end{pmatrix}$$

$$=8x'^2+3y'^2+2z'^2.$$

非退化线性变换为

$$\begin{pmatrix}x\\y\\z\end{pmatrix}=P\begin{pmatrix}x'\\y'\\z'\end{pmatrix},\text{ 其中}P=\begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix}.$$

称这类标准形是全正项平方项的二次型为正定二次型.

称二次型矩阵

特征值全正的实对称矩阵A， $\begin{pmatrix}8&-8&8\\-8&11&-11\\8&-11&13\end{pmatrix}$

为正定矩阵.

定义5.2.1 (正定二次型、正定矩阵) 设 $f(x)=x^{\mathrm{T}}Ax$ 为实二次型，若当实向量 $x\neq\theta$ 时都有 $x^{\mathrm{T}}Ax>0$，则称 f 为正定二次型，称 A 为正定矩阵；
当 $x\neq\theta$ 时都有 $x^{\mathrm{T}}Ax<0$，则称 f 为负定二次型，称 A 为负定矩阵；
当 $x\neq\theta$ 时都有 $x^{\mathrm{T}}Ax\geq 0$，则称 f 为半正定二次型，称 A 为半正定矩阵；
当 $x\neq\theta$ 时都有 $x^{\mathrm{T}}Ax\leq 0$，则称 f 为半负定二次型，称 A 为半负定矩阵 .

注 有时为了强调正定矩阵的对称性，也称对称正定矩阵 .

定义5.2.2 (矩阵的顺序主子式和主子式)

矩阵

$$A=(a_{ij})_{n\times n}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{nn} \end{pmatrix}$$

的左上角 i 行 i 列

$(1\leq i\leq n)$ 构成的行列式

$$\begin{vmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{ii} \end{vmatrix}$$

称为矩阵 A 的 i 阶顺序主子式.

矩阵 A 的 $i_1,i_2,\ldots,i_k$ 行和 $i_1,i_2,\ldots,i_k$ 列 $(1\leq i_1<i_2<\ldots<i_k\leq n)$ 的元素构成的行列式

$$\begin{vmatrix} a_{i_1i_1} & a_{i_1i_2} & \cdots & a_{i_1i_k} \\ a_{i_2i_1} & a_{i_2i_2} & \cdots & a_{i_2i_k} \\ \vdots & \vdots & & \vdots \\ a_{i_ki_1} & a_{i_ki_2} & \cdots & a_{i_ki_k} \end{vmatrix}$$

称为矩阵 A 的 k 阶主子式 .

正定矩阵的判定

定理5.2.1 若 A 为 n 阶的实对称矩阵，则下列条件互为等价：
(1) A 为正定矩阵；(即 $x^{\mathrm{T}}Ax>0$, 当 $x\neq\theta$)
(2) A 的特征值均为正；(即 $\lambda_i(A)>0, i=1,2,\ldots,n$)
(3) A 的正惯性指数为 n；(即 $P^{\mathrm{T}}AP=E$)
(4) A 的各阶顺序主子式均为正 .

证明思路：
先证明(1)、(2)、(3)等价

$P^{\mathrm{T}}AP=E$, $A=P^{-\mathrm{T}}P^{-1}$
$x^{\mathrm{T}}Ax=x^{\mathrm{T}}P^{-\mathrm{T}}P^{-1}x=y^{\mathrm{T}}y>0$

(1) A 为正定矩阵

$A\xi=\lambda\xi$
$0<\xi^{\mathrm{T}}A\xi=\lambda\xi^{\mathrm{T}}\xi$

(2) A 的特征值均为正 $\longrightarrow$ (3) A 的正惯性指数为 n

推论5.1.5

半正定矩阵的判定

定理5.2.2 若A为n阶的实对称矩阵，则下列条件互为等价：

(1) A为半正定矩阵；(即$x^{\mathrm{T}}Ax\geq 0$, 当$x\neq\theta$)

(2) A 的特征值大于等于零； (即$\lambda_i(A)\geq 0, i=1,2,\ldots,n$)

(3) A的正惯性指数为 $\mathrm{r}(A)$；(即 $P^{\mathrm{T}}AP=\mathrm{diag}(E_r,O)$)

(4) A的各阶主子式非负 .

证明思路:

先证明(1)、(2)、(3)等价

(1) A为半正定矩阵

$0\leq \xi^{\mathrm{T}}A\xi=\lambda\xi^{\mathrm{T}}\xi$

(2) A 的特征值大于等于零 $\longrightarrow$ (3) A的正惯性指数为 $\mathrm{r}(A)$

推论5.1.5

$P^{\mathrm{T}}AP=\mathrm{diag}(E_r,O)$,

$A=P^{-\mathrm{T}}\mathrm{diag}(E_r,O)P^{-1}=P^{-\mathrm{T}}\Lambda P^{-1}$

$x^{\mathrm{T}}Ax=x^{\mathrm{T}}P^{-\mathrm{T}}\Lambda P^{-1}x=y^{\mathrm{T}}\Lambda y\geq 0$

例5.2.6 t取何值时，二次型

$$f(x_1,x_2,x_3)=2x_1^2+2x_1x_2+4x_1x_3+tx_2^2-2x_2x_3+3x_3^2$$ 是正定二次型？

解 二次型矩阵为 $A=\begin{pmatrix}2&1&2\\1&t&-1\\2&-1&3\end{pmatrix}$.

用顺序主子式判别法：$\det(2)=2>0, \begin{vmatrix}2&1\\1&t\end{vmatrix}=2t-1>0, \begin{vmatrix}2&1&2\\1&t&-1\\2&-1&3\end{vmatrix}=\begin{vmatrix}2&0&0\\1&t-0.5&-2\\2&-2&1\end{vmatrix}=2t-9>0.$

解两个关于t的不等式得到解集：$t>4.5$，故当$t\in(4.5,+\infty)$时，该二次型为正定二次型 .

解法二 二次型矩阵 $A=\begin{pmatrix}2&1&2\\1&t&-1\\2&-1&3\end{pmatrix}\xrightarrow{\text{合同变换}}\begin{pmatrix}2&0&0\\0&t-0.5&-2\\0&-2&1\end{pmatrix}\xrightarrow{\text{合同变换}}\begin{pmatrix}2&0&0\\0&t-4.5&0\\0&0&1\end{pmatrix}$.

易知当$t>4.5$时，该二次型为正定二次型 .

解法三 $f(x_1,x_2,x_3)=2x_1^2+2x_1x_2+4x_1x_3+tx_2^2-2x_2x_3+3x_3^2$ 是

$$=2(x_1+0.5x_2+x_3)^2+(t-4.5)x_2^2+(2x_2-x_3)^2=2y_1^2+(t-4.5)y_2^2+y_3^2,$$

易知当$t>4.5$时，正惯性指数为变量个数，故二次型正定 .

定义6.2.2 (维数) 设V是数域K上的线性空间，

(1)如果在V中可以找到任意多个线性无关的向量,则称V是**无限维线性空间**；

(2)如果存在有限多个向量$\alpha_1, \alpha_2, \ldots, \alpha_n \in V$，$n \geq 1$满足：

1) $\alpha_1, \alpha_2, \ldots, \alpha_n$**线性无关**

2) V中任一向量都可由$\alpha_1, \alpha_2, \ldots, \alpha_n$**线性表出**.

则称V是**有限维线性空间**，称$\alpha_1, \alpha_2, \ldots, \alpha_n$是$V$的一组**基**(或**基底**)，$\alpha_i$叫第$i$个**基向量**，基向量的个数$n$称为线性空间$V$的**维数**，记为$\dim(V)=n$，并称$V$是$n$维线性空间.

用基 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 表示基向量 $\beta_1, \beta_2, \ldots, \beta_n$

$$\beta_1 = (\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \beta_2 = (\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} a_{12} \\ \vdots \\ a_{n2} \end{pmatrix}, \cdots, \beta_n = (\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix},$$

两组基 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 与 $\beta_1, \beta_2, \ldots, \beta_n$ 的关系可用矩阵表示：

$$(\beta_1, \beta_2, \cdots, \beta_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = (\alpha_1, \alpha_2, \cdots, \alpha_n)P$$

$\Rightarrow$ $A = (\alpha_1, \alpha_2, \cdots, \alpha_n)$
$B = (\beta_1, \beta_2, \cdots, \beta_n)$
$(A|B) \rightarrow (E|P)$. 初等行变换

P称为从基$\alpha_1, \ldots, \alpha_n$到基$\beta_1, \ldots, \beta_n$的**过渡矩阵**.

例6.2.7实际上给出了从一个已知基构造另外基的方法：
新的基

$$(\beta_1, \ldots, \beta_n) = (\alpha_1, \ldots, \alpha_n) P, \quad (P\text{可逆})$$

例6.2.8 在线性空间 $P_n[x]$ 中，多项式

$$p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$$

在基 $1, x, x^2, \ldots, x^n$ 下的坐标是 $(a_0, a_1, \ldots, a_n)$，若取另一组基

$$1, (x-a), (x-a)^2, \ldots, (x-a)^n,$$

则多项式$p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$ 按泰勒公式展开为

$$p(x) = p(a) + p'(a)(x-a) + \frac{p''(a)}{2!}(x-a)^2 + \cdots + \frac{p^{(n)}(a)}{n!}(x-a)^n.$$

因此，$p(x)$ 在新的一组基 $1, (x-a), (x-a)^2, \ldots, (x-a)^n$ 下的坐标是

$$(p(a), p'(a), \frac{p''(a)}{2!}, \cdots, \frac{p^{(n)}(a)}{n!}).$$

6.2.2 基变换与坐标变换

若有两组基 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 与 $\beta_1, \beta_2, \ldots, \beta_n$，有关系：

$$(\beta_1, \ldots, \beta_n) = (\alpha_1, \ldots, \alpha_n) P,$$

向量α用两组基表示为：

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \ldots + x_n\alpha_n, \quad \alpha = y_1\beta_1 + y_2\beta_2 + \ldots + y_n\beta_n,$$

则有关系：

$$\alpha = (\beta_1, \cdots, \beta_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (\alpha_1, \alpha_2, \cdots, \alpha_n)P\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

于是不同基下的坐标的关系为：$P\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$，或者$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = P^{-1}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

重要式子：

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)P\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$$

即：$\beta = \alpha \cdot P$,

$Y = P^{-1}X$

β的基时坐标、 α的基时坐标

定理6.2.2 设 $\alpha_1,\alpha_2,\dots,\alpha_n$ 与 $\beta_1,\beta_2,\dots,\beta_n$ 是n维线性空间V的两组基，并且

$$(\beta_1,\dots,\beta_n)=(\alpha_1,\dots,\alpha_n)P,$$

基底变换公式

若V中任意元素α在这两组基下的坐标分别是 $(x_1,x_2,\dots,x_n)$ 和 $(y_1,y_2,\dots,y_n)$，则

$$\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=P\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix},\text{或者}\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}=P^{-1}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}.$$

坐标变换公式

基底变换公式：$(\beta_1,\dots,\beta_n)=(\alpha_1,\dots,\alpha_n)P$，$(\alpha_1,\dots,\alpha_n)=(\beta_1,\dots,\beta_n)P^{-1}$

从基底$\alpha_1,\dots,\alpha_n$ 到基底 $\beta_1,\dots,\beta_n$的过渡矩阵：P

坐标变换公式：$y=P^{-1}x$， $x=Py$

定义7.3.1 (线性空间的同构) 数域K上的两个线性空间V和W称为**同构**，如果由V到W有一个一一映射 f 满足以下两个条件：

(1) $f(\alpha+\beta)=f(\alpha)+f(\beta)$； 表示 $\alpha+\beta\leftrightarrow\alpha'+\beta'$

(2) $f(k\alpha)=kf(\alpha)$． 表示 $k\alpha\leftrightarrow k\alpha'$

这里$\alpha,\beta\in V$, $k\in K$，这样的映射 f 称为**同构映射**．

定理7.3.1 设V与W都是数域K上的有限维线性空间，则它们同构的充要条件是它们的维数相同．

定义6.3.1 (子空间) 设W是数域K上的线性空间V的一个非空子集，若W关于V上的加法和数乘也构成数域K上的一个线性空间，则称W是V的一个**线性子空间**，简称子空间，记为$W\subseteq V$，若$W\neq V$，记为$W\subset V$．

定理6.3.1 线性空间V的一个非空子集W是V的子空间的充要条件是

(1) 对任意的 $\alpha,\beta\in W$，有$\alpha+\beta\in W$，即对加法封闭；

(2) 对任意的$\alpha\in W$，$\lambda\in K$，有 $\lambda\alpha\in W$，即对数乘封闭．

定理6.3.2 线性空间V的一个非空子集W是V的子空间的充要条件是对任意的 $\alpha,\beta\in W$，$\lambda,\mu\in K$，有$\lambda\alpha+\mu\beta\in W$．

定义6.3.2 设V是数域K上的线性空间，向量组$\alpha_1,\alpha_2,\dots,\alpha_s\in V$，由这组向量所有可能的线性组合构成的集合

$$W(\alpha_1,\alpha_2,\dots,\alpha_s)=\{\alpha\mid\alpha=\sum k_i\alpha_i,\ k_i\in K,\ i=1,2,\dots,s\}$$

是非空集合，且构成V的子空间，称为由向量组$\alpha_1,\alpha_2,\dots,\alpha_s$**生成的子空间**，记作 $\mathrm{span}\{\alpha_1,\alpha_2,\dots,\alpha_s\}$ 或 $L\{\alpha_1,\alpha_2,\dots,\alpha_s\}$

特别地，零子空间是由零向量生成的子空间 $\mathrm{span}\{0\}$．

定义6.3.3 (子空间的交与和) 设W_1与W_2是数域K上线性空间V的两个子空间，定义W_1与W_2的交为

$$W_1\cap W_2=\{\,\alpha \mid \alpha\in W_1,\ \alpha\in W_2\,\},$$

W_1与W_2的和为

$$W_1+W_2=\{\,\gamma \mid \gamma=\alpha+\beta，对所有\alpha\in W_1，\beta\in W_2\,\}.$$

定理6.3.3 数域K上线性空间V的两个子空间W_1与W_2的交与和仍是V的子空间.

子空间的交与和满足交换律和结合律：子空间$W_1, W_2, W_3\subseteq V$

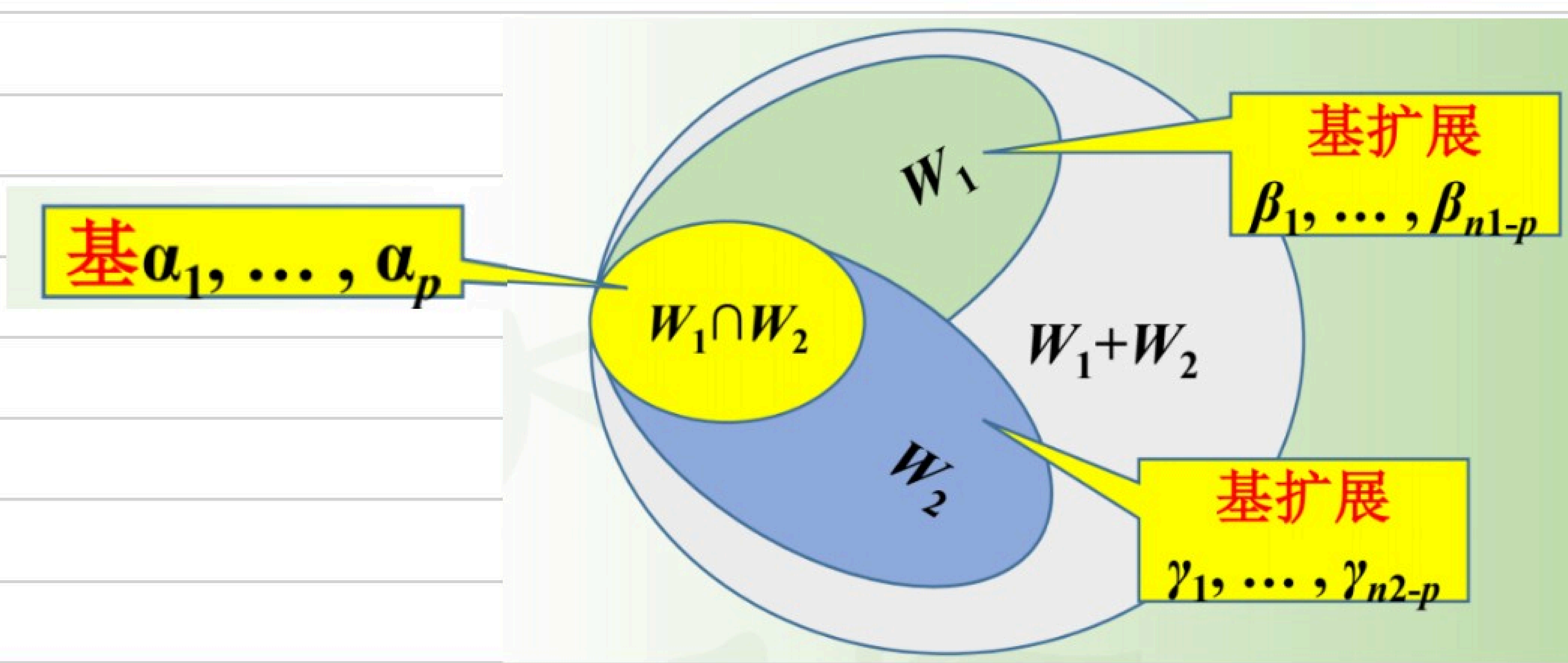
$$W_1+W_2=W_2+W_1,$$
$$(W_1+W_2)+W_3=W_1+(W_2+W_3),$$
$$W_1\cap W_2=W_2\cap W_1,$$
$$(W_1\cap W_2)\cap W_3=W_1\cap(W_2\cap W_3).$$

多个子空间的交与和：子空间$W_1, W_2, \ldots, W_m\subseteq V$

$$W_1\cap W_2\cap\ldots\cap W_m=(W_1\cap W_2\cap\ldots\cap W_{m-1})\cap W_m,$$
$$W_1+W_2+\ldots+W_m=(W_1+W_2+\ldots+W_{m-1})+W_m.$$

定理6.3.4 若W_1, W_2是线性空间V的两个有限维子空间，则

$$\dim W_1+\dim W_2=\dim(W_1+W_2)+\dim(W_1\cap W_2).$$



定义6.3.4 (直和) 若W_1+W_2中任一向量只能唯一地表示为子空间W_1的一个向量与子空间W_2的一个向量的和，则称W_1+W_2是直和(或直接和)，记为$W_1\oplus W_2$或$W_1\dot{+}W_2$. 若$W=W_1\oplus W_2$，则称在W内W_1是W_2的补空间，或W_2是W_1的补空间.

定理6.3.5 W_1+W_2是直和的充要条件是 $W_1\cap W_2=\{0\}$.

推论6.3.6 W_1+W_2是直和的充要条件是 $\dim(W_1+W_2)=\dim W_1+\dim W_2$.

W_1+W_2是直和表示子空间W_1和W_2没有非零重叠部分.

定义6.3.5 设$W_1, W_2, \dots, W_m$是线性空间V的子空间，若

(1) $W_1+W_2+\dots+W_m=V$；

(2) $W_1\cap W_2=\{0\}$，$(W_1+W_2)\cap W_3=\{0\}$，…，$(W_1+W_2+\dots+W_{m-1})\cap W_m=\{0\}$，

则称V是$W_1, W_2, \dots, W_m$的**直和**，记作

$$V=W_1\oplus W_2\oplus\dots\oplus W_m .$$

下面我们讨论如何求出子空间W_1、W_2的和空间W_1+W_2、交空间$W_1\cap W_2$.

利用线性空间中向量及关系一一对应坐标及关系

线性空间	V	$\Leftrightarrow$	K^n
	$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$	$\Leftrightarrow$	$e_1, e_2, \dots, e_n$
向量	α	$\Leftrightarrow$	x
关系1	$k_1\alpha_1+k_2\alpha_2+\dots+k_r\alpha_r=0$	$\Leftrightarrow$	$k_1\beta_1+k_2\beta_2+\dots+k_r\beta_r=\theta$
关系2	$\alpha=k_1\alpha_1+k_2\alpha_2+\dots+k_r\alpha_r$	$\Leftrightarrow$	$x=k_1\beta_1+k_2\beta_2+\dots+k_r\beta_r$
关系3	$\alpha_1, \alpha_2, \dots, \alpha_r$ 线性无关	$\Leftrightarrow$	$\beta_1, \beta_2, \dots, \beta_r$ 线性无关
不同基与坐标	$(\eta_1,\dots,\eta_n)=(\varepsilon_1,\dots,\varepsilon_n)P$	$\Leftrightarrow$	$y=P^{-1}x$ (坐标关系)

- 子空间W_1的基$\xi_1,\dots,\xi_m$，W_2的基$\eta_1,\dots,\eta_n$，对应坐标$x_1,\dots,x_m$和$y_1,\dots,y_n$.
- $W_1+W_2=\text{span}\{\xi_1,\dots,\xi_m, \eta_1,\dots,\eta_n\}$对应坐标空间 $\text{span}\{x_1,\dots,x_m, y_1,\dots,y_n\}$，求$\{x_1,\dots,x_m, y_1,\dots,y_n\}$极大无关组对应的$\{\xi_1,\dots,\xi_m, \eta_1,\dots,\eta_n\}$即为$W_1+W_2$的基.
- $W_1\cap W_2=\{\gamma|\gamma=k_1\xi_1+\dots+k_m\xi_m=t_1\eta_1+\dots+t_n\eta_n\}$对应坐标空间满足$k_1x_1+\dots+k_mx_m=t_1y_1+\dots+t_ny_n$，解方程组$k_1x_1+\dots+k_mx_m-t_1y_1-\dots-t_ny_n=\theta$ 得解$(k_1,\dots, k_m, -t_1,\dots, -t_n)$，即$W_1\cap W_2=\{\gamma|\gamma=k_1\xi_1+\dots+k_m\xi_m\}=\{\gamma|\gamma=t_1\eta_1+\dots+t_n\eta_n\}$.

定义6.4.1 (线性变换) 设V_1, V_2都是数域K上线性空间，根据某一规则T，对V_1中的任一元素α，有V_2中的唯一元素α'与之对应，即 $T\alpha=\alpha'$，则称T为V_1到V_2的**映射**.

如果V_1到V_2的映射T还满足

$$T(\alpha+\beta)=T\alpha+T\beta, \quad T(\lambda\alpha)=\lambda T\alpha$$

其中$\alpha, \beta\in V_1, \lambda\in K$，则称$T$为$V_1$到$V_2$的**线性映射**. 在 $V_1=V_2=V$ 时，称这个T为V上的**线性变换**.

例6.4.2 设 $A=(a_{ij})_{m\times n}\in \mathbf{R}^{m\times n}$，对 $x=(x_1, x_2,\dots, x_n)^{\mathrm{T}}\in \mathbf{R}^n$，令

$$Ax=y,$$

其中 $y=(y_1, y_2,\dots, y_m)^{\mathrm{T}}\in \mathbf{R}^m$, $y_i=\sum_{j=1}^{n} a_{ij}x_j, (i=1,2,\cdots,m)$,

则容易验证，$A(\lambda x+\mu y)=\lambda(Ax)+\mu(Ay)$，故$A$是$\mathbf{R}^n$到$\mathbf{R}^m$的线性映射.

** 线性映射T中V_2不一定是T的值域，如例6.4.2，这样可考虑多个映射.

线性映射T的基本性质：

(1) $T0=0$；

(2) 对任意的$\alpha_1, \alpha_2, \dots, \alpha_m \in V_1$，有

$$T(k_1\alpha_1+k_2\alpha_2+\dots+k_m\alpha_m)=k_1T\alpha_1+k_2T\alpha_2+\dots+k_mT\alpha_m;$$

(3) 若$\alpha_1, \alpha_2, \dots, \alpha_m$线性相关，则$T\alpha_1, T\alpha_2, \dots, T\alpha_m$也线性相关；反之不成立.

说明： (1) 用$T(\lambda\alpha)=\lambda T\alpha$，取$\lambda=0$

(2) 反复用$T(k_1\alpha_1+k_2\alpha_2)=k_1T\alpha_1+k_2T\alpha_2$，

$$T(k_1\alpha_1+k_2\alpha_2+k_3\alpha_3)=k_1T\alpha_1+T(k_2\alpha_2+k_3\alpha_3)=k_1T\alpha_1+k_2T\alpha_2+k_3T\alpha_3, \dots$$

(3) 用(2)的结论，反之设$T\alpha_1=T\alpha_2=\dots=T\alpha_m=0$

定义6.4.2 (像空间与核空间) (1) 若V_1, V_2都是数域K上线性空间，设线性映射T：$V_1 \to V_2$，则称V_1中所有元素的像的集合$\{T\alpha \mid \alpha \in V_1\}$为**像空间**，记作 R ($T$) 或 Im($T$)，简记为 R .

(2) 对于V_1到V_2的线性映射T，称集合 $N=N(T)=\{\alpha \mid T\alpha=0', \alpha \in V_1\}$为$T$的**核空间**，也记作 ker($T$)，其中 0'是$V_2$的零元.

定理6.4.1 线性映射的像空间与核空间是线性空间.

♥ 线性映射T: $V_1 \to V_2$，$\alpha_1, \dots, \alpha_n$为$V_1$的基，则$\mathrm{Im}(T)=\mathrm{span}\{T\alpha_1, \dots, T\alpha_n\}$.

定理6.4.2 设对V的两个线性变换T_1和T_2，有 ($\varepsilon_1, \varepsilon_2, \cdots \varepsilon_n$ 为V的一组基)

$$T_1\varepsilon_i=T_2\varepsilon_i,\ i=1,2,\dots,n,$$

则 $T_1=T_2$

(这里两个线性变换相等是指它们对V的任一向量的像相等).

该定理告诉我们：线性变换T能够由基的像 $T\varepsilon_1, T\varepsilon_2, \dots, T\varepsilon_n$ 表示

引理6.4.1 设$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$是线性空间$V$的一组基，任给$\alpha_1, \alpha_2, \dots, \alpha_n \in V$，则一定存在唯一的线性变换$T$，使得

$$T\varepsilon_i=\alpha_i,\ i=1,2,\dots,n.$$

构造： $T(x_1\varepsilon_1+x_2\varepsilon_2+\dots+x_n\varepsilon_n)=x_1\alpha_1+x_2\alpha_2+\dots+x_n\alpha_n$

定理6.4.3 在线性空间V的一组基$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$下，$V$上的线性变换$T$与$n$阶方阵$A$一一对应，且它们的对应关系是

$$(T\varepsilon_1, T\varepsilon_2, \dots, T\varepsilon_n)=(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)A.$$

即A的第i个列向量是$T\varepsilon_i$在基$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$下的坐标.

若α和$T\alpha$的坐标为x和y，则有 $y=Ax$.

定理6.4.4 在n维线性空间V中取定两组基底$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$与$\omega_1, \omega_2, \dots, \omega_n$，设由基底$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$到$\omega_1, \omega_2, \dots, \omega_n$的过渡矩阵为$P$ (P可逆)，即$(\omega_1, \omega_2, \dots, \omega_n)=(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)P$，并设$V$上的线性变换$T$在这两组基底下的矩阵分别是$A$和$B$，则

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

即A相似于B（$A \sim B$）.

定理6.4.5 设A, B都是n阶方阵，则$A \sim B$的充要条件是：它们是n维的线性空间V上的某个线性变换T在不同基底下的矩阵.

数域K上的线性空间V与K^n的对应关系

基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下

线性空间 $V \longleftrightarrow K^n$

线性变换 $T \longleftrightarrow A$

向量 $\begin{cases} \xi,\eta \longleftrightarrow x,y \\ \eta=T\xi \longleftrightarrow y=Ax \end{cases}$

关系1 $T\xi=0 \longleftrightarrow Ax=\theta$

关系2 $k_1\alpha_1+\cdots+k_r\alpha_r=0 \longleftrightarrow k_1x_1+\cdots+k_rx_r=\theta$

关系3 $\beta=k_1\alpha_1+\cdots+k_r\alpha_r \longleftrightarrow y=k_1x_1+\cdots+k_rx_r$

关系4 $\alpha_1,\alpha_2,\cdots,\alpha_r$无关 $\longleftrightarrow$ $x_1,x_2,\cdots,x_r$无关

关系5 $\ker(T) \longleftrightarrow$ $Ax=\theta$ 的解集，即解空间

$\ker(T)$的基 $\longleftrightarrow$ $Ax=\theta$ 的基础解系

关系6 $\mathrm{Im}(T)=span\{T\varepsilon_1,\cdots,T\varepsilon_n\} \longleftrightarrow \{Ax \mid x\in \mathrm{R}^n\}$即$A$列所有组合

$\mathrm{Im}(T)$的基

即$\{T\varepsilon_1,\cdots,T\varepsilon_n\}$的极大无关组 $\longleftrightarrow$ A 的列的极大无关组