复习高等代数

一. 线性空间

定义 设F是包含 0 和 1 的数集，若F中任意两个数的和差积商（0 不作除数）仍然在F中，则称F是一个数域。

如有理数集Q，实数集R，复数集C都是数域，但整数集Z不是数域。

定义 设V是一个非空集合，F是一个数域。在V中定义两种运算：加法与数乘。

（1）**加法**：对V中任意两个元素α和β，定义V中一个唯一确定的元素η与之对应，称为α与β的和，记为$\eta=\alpha+\beta$。

（2）**数乘**：对V中任意一个元素α和数域F中的任意一个数k，定义V中一个唯一确定的元素δ与之对应，称为k与α的数量乘积，记为$\delta=k\alpha$。

规定集合V中定义的上述加法与数乘运算满足以下八条规律：（设α,β,γ是集合V中的任意元素，k,l是数域F中的任意数）

（1）加法交换律，$\alpha+\beta=\beta+\alpha$；

（2）加法结合律，$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$；

（3）存在零元θ，$\alpha+\theta=\alpha$；

（4）存在负元α^*，$\alpha+\alpha^*=\theta$；

（5）存在单位元数 1，$1\cdot\alpha=\alpha$；

（6）数乘结合律，$k(l\alpha)=(kl)\alpha$；

（7）分配律，$(k+l)\alpha=k\alpha+l\alpha$；

（8）分配律，$k(\alpha+\beta)=k\alpha+k\beta$。

则称集合V是数域F上的一个线性空间。

全体n维实向量在通常的向量加法与数乘运算下，构成实数域R上的一个线性空间，称为n维向量空间，记为R^n。

全体实随机变量在通常的加法与数乘运算下，也构成实数域R上的一个线性空间，称为随机变量空间。

需要注意的是，n维向量空间只是线性空间的一个例子，还有其他各种各样的线性空间。

二. 内积与长度

向量的数量积：R^n中两个n维向量$\alpha=(a_1,a_2,\cdots,a_n)^T$与$\beta=(b_1,b_2,\cdots,b_n)^T$，数

$$\sum_{i=1}^{n}a_ib_i=a_1b_1+a_2b_2+\cdots+a_nb_n$$

称为向量α与β的数量积或内积，记为$\alpha\bullet\beta$或(α,β)。

注：（1）数量积$\alpha\bullet\beta$中间的点乘号“•”不能省略。

（2）设α与β是n维列向量，数量积$\alpha\bullet\beta$看作矩阵乘积为$\alpha^T\beta$，写作$\alpha\beta$无意义。

一般地，定义线性空间V上内积运算。 尖括号 < >　　内积是一种映射

定义 设V是数域F上的线性空间。对V中任意两个元素α和β，定义F中一个唯一确定的数与之对应，如果满足以下四条规律：（设α,β,γ是线性空间V中的任意元素，k是数域F中的任意数）

（1）交换律，$(\alpha,\beta)=(\beta,\alpha)$； $\langle\alpha,\beta\rangle=\langle\beta,\alpha\rangle$

（2）加法分配率，$(\alpha,\beta+\gamma)=(\alpha,\beta)+(\alpha,\gamma)$；

（3）数乘结合律，$k(\alpha,\beta)=(k\alpha,\beta)$；

（4）自身非负性，$(\alpha,\alpha)\geq 0$，等号成立当且仅当$\alpha=\theta$为零元素。

则称为α与β的内积，记为(α,β)。 $\langle\alpha,\beta\rangle$

显然n维向量的数量积就是内积，两个随机变量的协方差也是内积（将常数看作随机性为零的量）。

定义 设V是数域F上定义有内积运算的线性空间。对V中任意元素α，称$\sqrt{(\alpha,\alpha)}$为α的长度，记为$\|\alpha\|$。

定义 设V是数域F上定义有内积运算的线性空间。对V中任意两个非零元素α和β，称

$$\arccos\frac{(\alpha,\beta)}{\|\alpha\|\cdot\|\beta\|}$$

为α与β的夹角。

定义 设V是数域F上定义有内积运算的线性空间。若$(\alpha,\beta)=0$，则称α与β正交。

注： α与β正交的几何意义是α与β相互垂直，夹角为90°。

高等代数与概率论中相关概念的对应关系：

高等代数	概率论
内积(α,β)	协方差$\mathrm{Cov}(X,Y)$
长度$\|\alpha\|=\sqrt{(\alpha,\alpha)}$	标准差$\sqrt{\mathrm{Var}(X)}$
夹角余弦$\dfrac{(\alpha,\beta)}{\|\alpha\|\cdot\|\beta\|}$	相关系数$\dfrac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}$
正交$(\alpha,\beta)=0$	不相关$\mathrm{Cov}(X,Y)=0$

三． 标准正交基与正交矩阵

将长度等于1的向量称为单位向量。并且以下所讨论的向量均为列向量。

若n维向量空间R^n中的一组非零向量$\alpha_1,\alpha_2,\cdots,\alpha_m$两两正交，称之为正交向量组。若正交向量组中每一个向量都是单位向量，则称之为正交单位向量组。正交向量组必定线性无关。

定义 若n维向量空间R^n中n个单位向量$\alpha_1,\alpha_2,\cdots,\alpha_n$两两正交，则称之为$R^n$中一组标准正交基。

注： 标准正交基的几何意义是直角坐标系，且R^n中任一向量都可由标准正交基唯一线性表示。

设$\alpha_1,\alpha_2,\cdots,\alpha_n$是$R^n$中一组标准正交基，则

$$(\alpha_i,\alpha_j)=\alpha_i^T\alpha_j=\begin{cases}1, & i=j;\\ 0, & i\neq j.\end{cases}$$

定义 设$\alpha_1,\alpha_2,\cdots,\alpha_n$是$R^n$中一组标准正交基，称$n$阶矩阵$C=(\alpha_1,\alpha_2,\cdots,\alpha_n)$为正交矩阵。

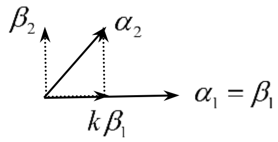
设C为正交矩阵，则

$$C^TC=\begin{pmatrix}\alpha_1^T\\ \alpha_2^T\\ \vdots\\ \alpha_n^T\end{pmatrix}(\alpha_1,\alpha_2,\cdots,\alpha_n)=\begin{pmatrix}\alpha_1^T\alpha_1 & \alpha_1^T\alpha_2 & \cdots & \alpha_1^T\alpha_n\\ \alpha_2^T\alpha_1 & \alpha_2^T\alpha_2 & \cdots & \alpha_2^T\alpha_n\\ \vdots & \vdots & \ddots & \vdots\\ \alpha_n^T\alpha_1 & \alpha_n^T\alpha_2 & \cdots & \alpha_n^T\alpha_n\end{pmatrix}=\begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{pmatrix}=E。$$

因此$C^T=C^{-1}$。

对于R^n中任意的线性无关向量组$\alpha_1,\alpha_2,\cdots,\alpha_m$，可构造与之等价的正交单位向量组$\beta_1^*,\beta_2^*,\cdots,\beta_m^*$。其方法称为施密特标准正交化。

正交化：取 $\beta_1=\alpha_1$；再取 $\beta_2=\alpha_2-k\beta_1$，确定常数 k，使得 β_2 与 β_1 正交。



$$(\beta_1,\beta_2)=(\beta_1,\alpha_2)-k(\beta_1,\beta_1)=0\text{，}$$

可得

$$k=\frac{(\beta_1,\alpha_2)}{(\beta_1,\beta_1)}\text{，}$$

故

$$\beta_2=\alpha_2-\frac{(\beta_1,\alpha_2)}{(\beta_1,\beta_1)}\beta_1\text{。}$$

进一步，取 $\beta_3=\alpha_3-k_1\beta_1-k_2\beta_2$，确定常数 k_1,k_2，使得 β_3 与 β_1,β_2 都正交。

$$(\beta_1,\beta_3)=(\beta_1,\alpha_3)-k_1(\beta_1,\beta_1)-0=0\text{，}\quad(\beta_2,\beta_3)=(\beta_2,\alpha_3)-0-k_2(\beta_2,\beta_2)=0\text{，}$$

可得

$$k_1=\frac{(\beta_1,\alpha_3)}{(\beta_1,\beta_1)}\text{，}\quad k_2=\frac{(\beta_2,\alpha_3)}{(\beta_2,\beta_2)}\text{，}$$

故

$$\beta_3=\alpha_3-\frac{(\beta_1,\alpha_3)}{(\beta_1,\beta_1)}\beta_1-\frac{(\beta_2,\alpha_3)}{(\beta_2,\beta_2)}\beta_2\text{。}$$

以此类推，得到与 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 等价的正交向量组 $\beta_1,\beta_2,\cdots,\beta_m$。

单位化：令

$$\beta_1^*=\frac{\beta_1}{\|\beta_1\|},\beta_2^*=\frac{\beta_2}{\|\beta_2\|},\cdots,\beta_m^*=\frac{\beta_m}{\|\beta_m\|}\text{，}$$

可得 $\beta_1^*,\beta_2^*,\cdots,\beta_m^*$ 是与 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 等价的正交单位向量组。

特别是，当 $m=n$ 时，可得 $\beta_1^*,\beta_2^*,\cdots,\beta_n^*$ 是 R^n 中一组标准正交基。而当 $m<n$ 时，总能在 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 后面再增加 $n-m$ 个 n 维向量，得到 n 个线性无关的 n 维向量 $\alpha_1,\alpha_2,\cdots,\alpha_m,\cdots,\alpha_n$，再正交化单位化，得到 R^n 中一组标准正交基 $\beta_1^*,\beta_2^*,\cdots,\beta_n^*$，也就是能将其扩充为一组标准正交基。再进一步组成正交矩阵

$$C=(\beta_1^*,\beta_2^*,\cdots,\beta_n^*)\text{。}$$

定义　设两组变量 $x_1,x_2,\cdots,x_n$ 与 $y_1,y_2,\cdots,y_n$ 满足关系

$$\begin{cases}x_1=c_{11}y_1+c_{12}y_2+\cdots+c_{1n}y_n\\x_2=c_{21}y_1+c_{22}y_2+\cdots+c_{2n}y_n\\\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\x_n=c_{n1}y_1+c_{n2}y_2+\cdots+c_{nn}y_n\end{cases}\text{，矩阵形式}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\c_{21}&c_{22}&\cdots&c_{2n}\\\cdots&\cdots&\cdots&\cdots\\c_{n1}&c_{n2}&\cdots&c_{nn}\end{pmatrix}\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}\text{，}$$

即 $X=CY$，称为由 $X=(x_1,x_2,\cdots,x_n)^T$ 到 $Y=(y_1,y_2,\cdots,y_n)^T$ 的线性变换，矩阵 C 为变换阵。

若变换阵 C 可逆，即 $|C|\neq 0$，则称 $X=CY$ 是可逆线性变换。若变换阵 C 为正交阵，则称 $X=CY$ 是正交变换。

注： 正交变换保持内积、长度、夹角、正交性不变。

设 C 为正交阵，正交变换 $X_1=CY_1$，$X_2=CY_2$，有

$$(X_1,X_2)=X_1^TX_2=(CY_1)^TCY_2=Y_1^TC^TCY_2=Y_1^TEY_2=Y_1^TY_2=(Y_1,Y_2),$$

$$\|X_1\|=\sqrt{(X_1,X_1)}=\sqrt{(Y_1,Y_1)}=\|Y_1\|。$$

即正交变换前后向量内积、长度不变，进一步可得向量夹角、正交性不变。

四．特征值与特征向量

定义 设 A 为 n 阶方阵，若存在数 λ 和 n 维非零向量 X，使得 $AX=\lambda X$，则称 λ 是 A 的特征值，X 是 A 的对应于特征值 λ 的特征向量。

如 $A=\begin{pmatrix}1&4\\2&3\end{pmatrix}$，取 $X=\begin{pmatrix}1\\1\end{pmatrix}$，有

$$AX=\begin{pmatrix}1&4\\2&3\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}5\\5\end{pmatrix}=5X,$$

故 $\lambda=5$ 是 A 的特征值，$X=\begin{pmatrix}1\\1\end{pmatrix}$ 是对应的特征向量。

又如 $A=\begin{pmatrix}1&1&1\\1&2&3\\1&3&5\end{pmatrix}$，取 $X=\begin{pmatrix}1\\-2\\1\end{pmatrix}$，有

$$AX=\begin{pmatrix}1&1&1\\1&2&3\\1&3&5\end{pmatrix}\begin{pmatrix}1\\-2\\1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}=0X,$$

故 $\lambda=0$ 是 A 的特征值，$X=\begin{pmatrix}1\\-2\\1\end{pmatrix}$ 是对应的特征向量。

一般地，设 λ 是 A 的特征值，X 是对应的特征向量，有 $AX=\lambda X$，得齐次线性方程组 $(\lambda E-A)X=O$，由于特征向量 X 是非零向量，即 $(\lambda E-A)X=O$ 有非零解，则其系数行列式 $|\lambda E-A|=0$。设

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix},$$

则行列式

$$|\lambda E-A|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}=\lambda^n+c_{n-1}\lambda^{n-1}+\cdots+c_1\lambda+c_0$$

为 n 次多项式。

定义 设 A 为 n 阶矩阵，称 $\lambda E-A$ 为矩阵 A 的特征矩阵，其行列式 $|\lambda E-A|$ 为 A 的特征多项式，$|\lambda E-A|=0$ 为 A 的特征方程，特征方程的根称为特征根，也就是特征值。

结论：n 阶方阵必有 n 个特征根（包括复数根，且重根按重数计算）。

求特征值与特征向量的步骤：

（1）求解特征方程 $|\lambda E-A|=0$，得 n 阶矩阵 A 的 n 个特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$。

（2）对每一个特征值 λ_i，求解齐次线性方程组 $(\lambda_i E-A)X=O$，其全部非零解都是 A 的特征向量。

例 设 $A=\begin{pmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0\end{pmatrix}$，求 A 的特征值与特征向量。

解：特征方程

$$|\lambda E-A|=\begin{vmatrix}\lambda & -1 & -1\\ -1 & \lambda & -1\\ -1 & -1 & \lambda\end{vmatrix}=(\lambda-2)(\lambda+1)^2=0,$$

得特征值 $\lambda_1=2,\lambda_2=\lambda_3=-1$。

取 $\lambda_1=2$，求解齐次线性方程组 $(2E-A)X=O$，因

$$2E-A=\begin{pmatrix}2 & -1 & -1\\ -1 & 2 & -1\\ -1 & -1 & 2\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 0 & -1\\ 0 & 1 & -1\\ 0 & 0 & 0\end{pmatrix},$$

得基础解系 $X_1=\begin{pmatrix}1\\1\\1\end{pmatrix}$，故对应于 $\lambda_1=2$ 的全部特征向量为 $k_1X_1=k_1\begin{pmatrix}1\\1\\1\end{pmatrix}$，$(k_1\neq 0)$。

取 $\lambda_2=\lambda_3=-1$，求解 $(-E-A)X=O$，因

$$-E-A=\begin{pmatrix}-1 & -1 & -1\\ -1 & -1 & -1\\ -1 & -1 & -1\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 1 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix},$$

得基础解系 $X_2=\begin{pmatrix}-1\\1\\0\end{pmatrix}$，$X_3=\begin{pmatrix}-1\\0\\1\end{pmatrix}$，对应于 $\lambda_2=\lambda_3=-1$ 的全部特征向量为

相似矩阵 有相同的特征值

$$k_2X_2+k_3X_3=k_2\begin{pmatrix}-1\\1\\0\end{pmatrix}+k_3\begin{pmatrix}-1\\0\\1\end{pmatrix}$$，（k_2,k_3 不全为 0）。

注：（1）对应于不同特征值的特征向量线性无关。

（2）k 重特征值对应的线性无关特征向量可以是一个，也可以是多个，但不超过 k 个。

五． 相似矩阵与矩阵对角化

定义 设 A,B 为 n 阶方阵，若存在可逆矩阵 P，使得 $P^{-1}AP=B$，则称 A 与 B 相似，记为 $A\sim B$，且矩阵 P 称为变换阵。

如 $A=\begin{pmatrix}1&4\\2&3\end{pmatrix}$，取 $P=\begin{pmatrix}-2&1\\1&1\end{pmatrix}$，$|P|=-3\neq 0$，$P$ 可逆，有

$$P^{-1}AP=-\frac{1}{3}\begin{pmatrix}1&-1\\-1&-2\end{pmatrix}\begin{pmatrix}1&4\\2&3\end{pmatrix}\begin{pmatrix}-2&1\\1&1\end{pmatrix}=\begin{pmatrix}-1&0\\0&5\end{pmatrix}=B,$$

故

$$A=\begin{pmatrix}1&4\\2&3\end{pmatrix}\sim B=\begin{pmatrix}-1&0\\0&5\end{pmatrix}。$$

相似矩阵有很多共同的性质，而对角阵是一种简单的矩阵，讨论一般 n 阶方阵能否相似于对角阵。

定理 n 阶方阵 A 相似于对角阵的充分必要条件是 A 有 n 个线性无关的特征向量。

证明：必要性：设 A 相似于对角阵 $\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$，且变换阵 $P=(X_1,X_2,\cdots,X_n)$。因变换阵 P 可逆，则 P 的列向量组 $X_1,X_2,\cdots,X_n$ 线性无关。

由于 $P^{-1}AP=\Lambda$，有 $AP=P\Lambda$，即

$$AP=A(X_1,X_2,\cdots,X_n)=(AX_1,AX_2,\cdots,AX_n)$$

$$=P\Lambda=(X_1,X_2,\cdots,X_n)\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix}=(\lambda_1X_1,\lambda_2X_2,\cdots,\lambda_nX_n),$$

则 $AX_1=\lambda_1X_1,AX_2=\lambda_2X_2,\cdots,AX_n=\lambda_nX_n$，即 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 A 的特征值，$X_1,X_2,\cdots,X_n$ 是对应的特征向量。因 $X_1,X_2,\cdots,X_n$ 线性无关，所以 A 有 n 个线性无关的特征向量。

充分性：设 A 有 n 个线性无关的特征向量 $X_1,X_2,\cdots,X_n$，且 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是相应的特征值，则有

$$AX_1=\lambda_1X_1,AX_2=\lambda_2X_2,\cdots,AX_n=\lambda_nX_n,$$

令对角阵 $\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$，且变换阵 $P=(X_1,X_2,\cdots,X_n)$，有 P 可逆，可得 $AP=P\Lambda$，即 $P^{-1}AP=\Lambda$，所以 A 相似于对角阵 Λ。

相似对角化

此定理表明：若 n 阶方阵 A 有 n 个线性无关的特征向量，则 A 相似于对角阵，且对角阵 Λ 的主对角线上元素就是 A 的特征值，变换阵 P 的列向量组就是 A 的线性无关特征向量。

推论 若 n 阶方阵 A 有 n 个互不相同的特征值，则 A 必相似于对角阵。

推论 若 n 阶方阵 A 有重根，则 A 相似于对角阵的充分必要条件是 A 的每个 k 重特征根恰好对应于 k 个线性无关特征向量。

下面给出矩阵对角化的一个应用。设 $x_1=x_1(t), x_2=x_2(t),\cdots,x_n=x_n(t)$ 是 n 个一元函数，考虑 n 维常系数线性微分方程组

$$\begin{cases}\dfrac{dx_1}{dt}=a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n;\\ \dfrac{dx_2}{dt}=a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n;\\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\\ \dfrac{dx_n}{dt}=a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n.\end{cases}，\text{矩阵形式}\ \frac{d}{dt}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},$$

即 $\dfrac{dX}{dt}=AX$。如果 n 阶方阵 A 相似于对角阵，存在可逆矩阵 P，使得 $P^{-1}AP=\Lambda$ 为对角阵。作可逆线性变换 $X=PY$，有 $P\dfrac{dY}{dt}=APY$，即 $\dfrac{dY}{dt}=P^{-1}APY=\Lambda Y$，得到

$$\begin{cases}\dfrac{dy_1}{dt}=\lambda_1 y_1;\\ \dfrac{dy_2}{dt}=\lambda_2 y_2;\\ \quad\vdots\\ \dfrac{dy_n}{dt}=\lambda_n y_n.\end{cases}$$

解得

$$y_i=C_ie^{\lambda_i t},\quad C_i\text{为任意常数},\quad i=1,2,\cdots,n,$$

再根据 $X=PY$，可得到 $x_i=x_i(t)$ 的解。

六. 实对称阵的正交相似对角化

实对称阵A对应不同特征值的特征向量正交

不加证明的给出以下定理。

定理 实对称阵特征值必为实数。

定理 实对称阵不同特征值的特征向量相互正交。

定理 n 阶实对称阵必有 n 个线性无关特征向量，因此必定相似于对角阵。

n 阶实对称阵有 n 个线性无关特征向量，将其正交化单位化。由于不同特征值的特征向量相互正交，因此正交化只需在同一特征值内线性无关特征向量之间进行，正交化后仍为其特征向量。从而得到 n 阶实对称阵的 n 个线性无关特征向量构成 R^n 的一组标准正交基，再进一步组成正交矩阵。

这样 n 阶实对称阵 A 有 n 个特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$，对应有 n 个线性无关特征向量 $X_1,X_2,\cdots,X_n$，将其正交化单位化，得到 $\eta_1^*,\eta_2^*,\cdots,\eta_n^*$，令对角阵 $\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$，矩阵 $C=(\eta_1^*,\eta_2^*,\cdots,\eta_n^*)$，$C$ 为正交阵，

施密特正交化

则实对称阵 A 可正交相似对角化

$$C^{-1}AC=C^TAC=\Lambda。$$

七. 二次型及其标准型

定义 n 个变量 $x_1, x_2, \cdots, x_n$ 的二次齐次多项式

$$f(x_1, x_2, \cdots, x_n) = a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n + a_{22}x_2^2 + \cdots + 2a_{2n}x_2x_n + \cdots + a_{nn}x_n^2$$

称为一个 n 元二次型。

一般地，先将二次型一般形式写成一般对称形式：

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) = a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n \\ + \cdots \cdots \cdots \cdots \cdots \cdots \cdots \\ + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 \end{aligned},$$

这里，设 $a_{ij} = a_{ji}$，记

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

有

$$f(x_1, x_2, \cdots, x_n) = (x_1 \quad x_2 \quad \cdots \quad x_n) \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{pmatrix}$$

$$= (x_1 \quad x_2 \quad \cdots \quad x_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= X^T AX,$$

故二次型的矩阵形式为 $f = X^T AX$，这里矩阵 A 为对称阵，称为二次型 f 的系数矩阵。

二次型的系数矩阵：平方项系数不变写在主对角线上，乘积项系数分半写在对称位置上。只含平方项的二次型 $f(x_1, x_2, \cdots, x_n) = a_1x_1^2 + a_2x_2^2 + \cdots + a_nx_n^2$，称为标准形，其系数矩阵是对角阵

标准形的系数矩阵是对角阵

$$\Lambda = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}。$$

二次型 $f = X^T AX$，作可逆线性变换 $X = CY$，有

$$f = (CY)^T A(CY) = Y^T C^T ACY 。$$

设 $B = C^T AC$，B 为对称阵，也就是变换后新二次型的系数矩阵。

定义 设 A, B 为 n 阶方阵，若存在可逆矩阵 C，使得 $C^T AC = B$，则称 A 与 B 合同，且矩阵 C 称为变换阵。

正交变换下、合同与相似一致

注：A 与 B 相似，即 $C^{-1}AC = B$；A 与 B 合同，即 $C^T AC = B$。矩阵相似与合同是两个不同的概念。当 C 为正交阵时，相似与合同才是一致的。

前面已经知道，对称阵 A 必正交相似于对角阵 Λ，从而对称阵 A 必合同于对角阵 Λ。这样二次型 $f = X^T AX$ 通过正交变换 $X = CY$ 化为只含平方项的标准形 $f = Y^T \Lambda Y$，这就是正交变换法化二次型为标准形。

正交变换法化二次型为标准形的步骤：

（1）写出二次型 f 的系数矩阵 A，即 $f = X^T AX$。

（2）求解特征方程 $|\lambda E - A| = 0$，得 n 阶矩阵 A 的 n 个特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n$。

（3）对每一个特征值 λ_i，求解齐次线性方程组 $(\lambda_i E - A)X = O$，得到基础解系。所有特征值对应的基础解系构成 A 的 n 个线性无关特征向量 $X_1, X_2, \cdots, X_n$。

（4）将 $X_1, X_2, \cdots, X_n$ 正交化单位化，得到 R^n 的一组标准正交基 $\eta_1^*, \eta_2^*, \cdots, \eta_n^*$。

（5）令对角阵 $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$，矩阵 $C = (\eta_1^*, \eta_2^*, \cdots, \eta_n^*)$，$C$ 为正交阵。使得实对称阵 A 正交相似对角于对角阵，$C^{-1}AC = C^T AC = \Lambda$。

（6）二次型 $f = X^T AX$ 通过正交变换 $X = CY$ 化为只含平方项的标准形 $f = Y^T \Lambda Y$。