

# 理解高维高斯分布的概率密度函数

$$x \sim (\mu, \Sigma) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left( -\frac{1}{2} (x-\mu)^T \Sigma^{-1} (x-\mu) \right)$$

$\mu$ : 期望       $\Sigma$ : 协方差矩阵  
 $x$ : 自变量       $\Sigma^{-1}$ : 二次型?

$x \in \mathbb{R}^p$  r.v (随机变量)  
 $p$ : 维

考虑矩阵的特征值。  
 若所有特征值均不小于零 ( $\geq$ )，则称为半正定。

若所有特征值均大于零 ( $>$ )，则称为正定。

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} \quad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{pmatrix}_{p \times p}$$

$\Sigma$ : 此节假设正定  
 (一般是半正定)

$1 \times p \quad p \times p \quad p \times 1 \Rightarrow$  一个数      两向量间  
 $(x-\mu)^T \Sigma^{-1} (x-\mu)$  : 马氏距离 (x与 $\mu$ 之间)

$\Sigma = I$ , 马氏距离 = 欧氏距离

特征值分解?

$$\begin{cases} \Sigma = U \Lambda U^T, & U U^T = U^T U = I, & \Lambda = \text{diag}(\lambda_i) \quad i=1,2,\dots,p \\ U = (\mu_1, \mu_2, \dots, \mu_p)_{p \times p} \end{cases}$$

$$\Sigma = U \Lambda U^T$$

$$= (\mu_1, \mu_2, \dots, \mu_p) \begin{pmatrix} \lambda_1 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \end{pmatrix} \begin{pmatrix} \mu_1^T \\ \mu_2^T \\ \vdots \\ \mu_p^T \end{pmatrix}$$

$$= (\mu_1 \lambda_1 \quad \mu_2 \lambda_2 \quad \dots \quad \mu_p \lambda_p) \begin{pmatrix} \mu_1^T \\ \mu_2^T \\ \vdots \\ \mu_p^T \end{pmatrix}$$

$$= \sum_{i=1}^p \mu_i \lambda_i \mu_i^T$$

代  $\Sigma^{-1} = (U \Lambda U^T)^{-1} = (U^T)^{-1} \Lambda^{-1} U^{-1} = U \Lambda^{-1} U^T$

eg:  $z_1 = \begin{pmatrix} z_{11} \\ z_{12} \end{pmatrix} \quad z_2 = \begin{pmatrix} z_{21} \\ z_{22} \end{pmatrix}$

$$(z_1 - z_2)^T \Sigma^{-1} (z_1 - z_2)$$

两向量协方差

$$= (z_{11} - z_{21} \quad z_{12} - z_{22}) \Sigma^{-1} \begin{pmatrix} z_{11} - z_{21} \\ z_{12} - z_{22} \end{pmatrix}$$

当  $\Sigma = I$

$$\frac{(z_{11} - z_{21})^2 + (z_{12} - z_{22})^2}{\text{欧氏距离}}$$

$$= \sum_{i=1}^P u_i \frac{1}{\lambda_i} u_i^T$$

$$\rightarrow \text{diag}(\frac{1}{\lambda_i}) \quad i=1,2,\dots,P$$



$$\boxed{(x-\mu)^T \Sigma^{-1} (x-\mu)} = \Delta$$

$$= (x-\mu)^T \sum_{i=1}^P u_i \frac{1}{\lambda_i} u_i^T (x-\mu)$$

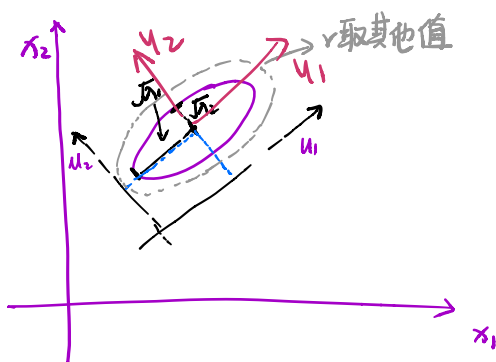
$$= \sum_{i=1}^P \underbrace{(x-\mu)^T u_i}_{\text{与 } i \text{ 有关}} \frac{1}{\lambda_i} \underbrace{u_i^T (x-\mu)}_{\text{与 } i \text{ 有关}} \quad \text{令 } y_i = (x-\mu)^T u_i$$

将  $x$  平移到  $\mu$  后投影到  $u_i$

$$= \sum_{i=1}^P y_i \frac{1}{\lambda_i} y_i^T = \boxed{\sum_{i=1}^P \frac{y_i^2}{\lambda_i}}$$

令  $P=2$ ,  $\Delta = \frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2} = r$  ( $\lambda_1 > \lambda_2$ )

若  $r=1$   $\frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2} = 1 \Rightarrow$  椭圆.



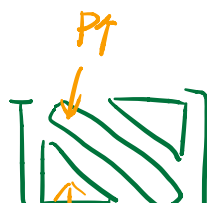
= 二维高斯分布在平面的  
展示  $\Rightarrow$  山丘状的横切面  
 $\Rightarrow$  椭圆曲线.  $r$  不同, 椭圆不同.

若  $\lambda_1=1 \Rightarrow$  圆形曲线.

局限性

$\Delta >$  参数过多的

对称矩阵

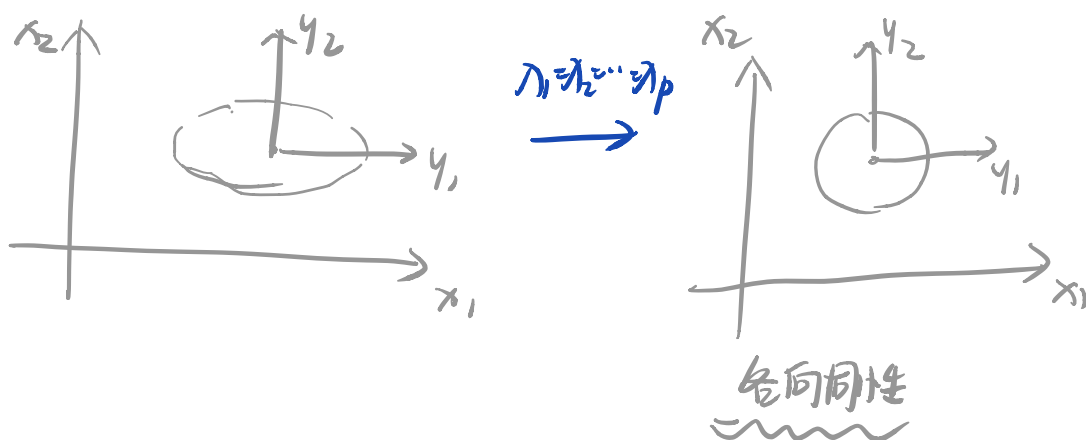


$$\Sigma_{p \times p} \Rightarrow \text{参数: } \frac{p^2 - p}{2} + p \quad \frac{p^2 - p}{2}$$

$$= \frac{p(p+1)}{2} = O(p^2)$$

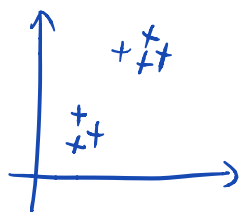
简化:  $\Sigma \xrightarrow{\text{变换}} \text{对角矩阵 } \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_p \end{pmatrix} \Rightarrow y_i = (x - \mu)^T \underbrace{u_i}_{\downarrow}$

此时图像:



c2> 自身局限

eg: 数据分布:



则不能只用一个高斯分布来描述.

使用 GMM  $\Rightarrow$  2个高斯混合  $\checkmark$ .