理解高维高斯分布的概率密度函数

x：自变量

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

μ：期望　Σ：协方差矩阵

2次型 ?

$x \in R^p$　r.v（随机变量）

p维

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

Σ：此节假设正定

（一般是半正定）

$$\underset{1\times p}{(x-\mu)^T}\ \underset{p\times p}{\Sigma^{-1}}\ \underset{p\times 1}{(x-\mu)} \Rightarrow \text{一个数}$$

：马氏距离（x与μ之间）　两向量间

$\Sigma = I$，马氏距离 = 欧氏距离

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 $\Sigma = I$

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

特征值分解？

$$\begin{cases} \Sigma = U \Lambda U^T, \quad UU^T = U^T U = I, \quad \Lambda = diag(\lambda_i) \quad i = 1, 2 \cdots p \\ U = (u_1, u_2 \cdots u_p)_{p \times p} \end{cases}$$

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

$$= (u_1, u_2, \cdots u_p) \begin{pmatrix} \lambda_1 & \cdots & 0 \\ 0 & \lambda_2 \cdots & 0 \\ \vdots & \vdots & \\ 0 & 0 & \lambda_p \end{pmatrix} \begin{pmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_p^T \end{pmatrix}$$

$$= (u_1 \lambda_1 \quad u_2 \lambda_2 \cdots u_p \lambda_p) \begin{pmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_p^T \end{pmatrix}$$

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

代入

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

$$= \boxed{\sum_{i=1}^{p} u_i \frac{1}{\lambda_i} u_i^T} \qquad \longrightarrow \operatorname{diag}\left(\frac{1}{\lambda_i}\right) \quad i=1,2\cdots p$$

$$\Downarrow$$

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

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

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

令 $y_i = (x-\mu)^T u_i$

将x平移μ后投影到u_i

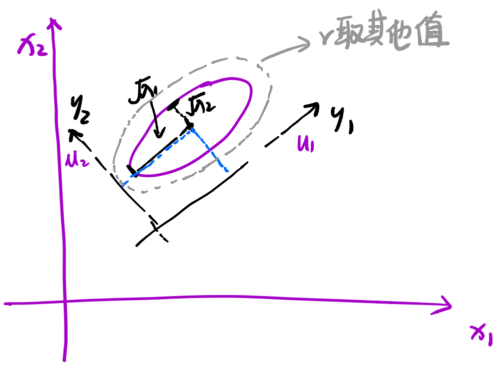
与i无关

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

马氏距离

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 ⇒ 山丘状的横切面

⇒ 椭圆曲线. r不同, 椭圆不同.

若 $\lambda_i = 1$ ⇒ 圆形曲线.