假设数据集为 $\mathcal{D}=\{x_i\}_{i=1}^n$,其中单样本 $x_i\in\mathbb{R}^d, d>1$ 且服从高维高斯分布:

$$p(x_i; \mu, \Sigma) = (2\pi)^{-rac{d}{2}} det(\Sigma)^{-rac{1}{2}} exp \Bigg\{ -rac{1}{2} (x_i - \mu)^ op \Sigma^{-1} (x_i - \mu) \Bigg\}$$

试利用MLE求解 μ, Σ .

注:可能会用到的公式

$$egin{aligned} rac{\partial \ log \ det(\Sigma)}{\partial \Sigma} &= (\Sigma^{-1})^{ op} \ & \ rac{\partial \ (x^{ op}A^{-1}x)}{\partial A} &= -(A^{-1})^{ op}xx^{ op}(A^{-1})^{ op} \end{aligned}$$

更多公式可参考这里

Ans:

在整个数据集上,有

$$egin{aligned} \log P(D;\mu,\Sigma) &= \log \Big(\prod_{i=1}^n p(x_i;\mu,\Sigma)\Big) \ &= \sum_{i=1}^n \log p(x_i;\mu,\Sigma) \ &= \sum_{i=1}^n -rac{1}{2} (x_i-\mu)^ op \Sigma^{-1} (x_i-\mu) -rac{1}{2} \log det(\Sigma) + const \end{aligned}$$

(1) 关于 μ 求导,并令其为0:

$$egin{align} rac{\partial \log P(D;\mu,\Sigma)}{\partial \mu} &= \sum_{i=1}^n \left[\Sigma^{-1}(x_i - \mu)
ight] \doteq 0 \ \Longrightarrow \mu &= rac{1}{n} \sum_{i=1}^n x_i \end{aligned}$$

(2) 关于 Σ 求导,并令其为0:

$$egin{aligned} rac{\partial \log P(D;\mu,\Sigma)}{\partial \Sigma} &= \sum_{i=1}^n \left[rac{1}{2}\Sigma^{-1}(x_i-\mu)(x_i-\mu)^ op \Sigma^{-1} - rac{1}{2}\Sigma^{-1}
ight] \doteq 0 \ \implies \Sigma &= rac{1}{n}\sum_{i=1}^n (x_i-\mu)(x_i-\mu)^ op \end{aligned}$$