314652021 羅瑋翔

Assignment 5.

1. Given $f(x) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$, where $x, \mu \in \mathbb{R}^k$, Σ is a k-by-k positive definite matrix and $|\Sigma|$ is its determinant.
Show that $\int_{\mathbb{R}^k} f(x)\,dx = 1$

Let $u = \Sigma^{-\frac{1}{2}}(x-\mu)$, $x = \Sigma^{\frac{1}{2}}u - \mu \Rightarrow dx = (|\Sigma|)^{\frac{1}{2}} du$.

$\int_{\mathbb{R}^k} f(x)\,dx = \int_{\mathbb{R}^k} \frac{1}{\sqrt{(2\pi)^k|\Sigma|}} e^{-\frac{(x-\mu)^T\Sigma^{-1}(x-\mu)}{2}} dx = \int_{\mathbb{R}^k} \frac{1}{\sqrt{(2\pi)^k}} e^{(-\frac{u^Tu}{2})} du$

$= \frac{1}{(2\pi)^{\frac{k}{2}}} \cdot \int_{\mathbb{R}^k} e^{(-\frac{1}{2}\sum_{i=1}^{k} u_i^2)} du = \frac{1}{(2\pi)^{\frac{k}{2}}} \prod_{i=1}^{k} \int_{\mathbb{R}} e^{(-\frac{1}{2}u_i^2)} du = \frac{1}{(2\pi)^{\frac{k}{2}}} \cdot \prod_{i=1}^{k} \sqrt{2\pi} = 1$ #

2. Let A, B be n-by-n matrices and x be a n-by-1 vector.
(a) Show that $\frac{\partial}{\partial A} tr(AB) = B^T$
(b) Show that $x^TAx = tr(xx^TA)$.
(c) Derive the maximum likelihood estimators for a multivariate Gaussian.

(a) $tr(AB) = \sum_{i,j} A_{ij}B_{ji} \Rightarrow \frac{\partial}{\partial A_{ij}} tr(AB) = B_{ji} \Rightarrow \frac{\partial}{\partial A} tr(AB) = B^T$

(b) $xx^T = \begin{bmatrix} x_1^2 & x_1x_2 & \cdots & x_1x_n \\ x_2x_1 & x_2^2 & & \\ \vdots & & \ddots & \\ x_nx_1 & \cdots & & x_n^2 \end{bmatrix}$, $(xx^TA)_{ii} = \sum_{j=1}^{n} (xx)^T_{ij} A_{ji} = \sum_{j=1}^{n} x_ix_j A_{ji}$.

$\Rightarrow tr(xx^TA) = \sum_{i=1}^{n} (xx^T)A_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_ix_jA_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{n} x_jA_{ji}x_i = x^TAx$. #.

(c) Let $x_1, x_2, \dots, x_N \sim N(\mu, \Sigma)$, and its likelihood is $L(\mu, \Sigma) = \prod_{i=1}^{N} f(x_i)$.

取 log 後, $\ell(\mu, \Sigma) = -\frac{N}{2}\log((2\pi)^k) - \frac{N}{2}\log|\Sigma| - \frac{1}{2}\sum_{i=1}^{N}(x_i-\mu)^T\Sigma^{-1}(x_i-\mu)$

$= -\frac{Nk}{2}\log 2\pi - \frac{N}{2}\log|\Sigma| - \frac{1}{2}\sum_{i=1}^{N} tr((x_i-\mu)^T\Sigma^{-1}(x_i-\mu))$.

① 对 μ 求偏導. $\frac{\partial}{\partial\mu}\ell(\mu,\Sigma) = -\frac{1}{2}\sum_{i=1}^{N}\frac{\partial}{\partial\mu} tr((x_i-\mu)^T\Sigma^{-1}(x_i-\mu))$. (*)

Since Σ^{-1} is symmetric (Σ is positive definite). we have $\frac{\partial}{\partial x}x^TAx = (A+A^T)x. = 2Ax$.

$\Rightarrow \frac{\partial}{\partial\mu}(x_i-\mu)^T\Sigma^{-1}(x_i-\mu) = -2\Sigma^{-1}(x_i-\mu)$. (*) becomes $\Sigma^{-1}\sum_{i=1}^{N}(x_i-\mu)$. (**).

Set (**) to 0, we have $\Sigma^{-1}\sum_{i=1}^{N}(x_i-\mu) = 0 \Rightarrow \sum_{i=1}^{N}(x_i-\mu) = 0$. Thus $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{N} x_i = \bar{x}$.

$\Rightarrow \ell(\mu,\Sigma) = -\frac{Nk}{2}\log(2\pi) - \frac{N}{2}\log|\Sigma| - \frac{1}{2}\sum_{i=1}^{N} tr((x_i-\bar{x})^T\Sigma^{-1}(x_i-\bar{x}))$.

② 对 Σ 求偏導. $\frac{\partial}{\partial \Sigma} l(\mu, \Sigma) = \frac{\partial}{\partial \Sigma^{-1}} \left(-\frac{Nk}{2} \log 2\pi - \frac{N}{2} \log |\Sigma| - \frac{1}{2} \sum_{i=1}^{N} tr\left((x_i - \bar{x})^T \Sigma^{-1} (x_i - \bar{x}) \right) \right)$.

$$= \frac{\partial}{\partial \Sigma^{-1}} \left(-\frac{N}{2} \log |\Sigma| - \frac{1}{2} \sum_{i=1}^{N} tr\left((x_i - \bar{x})^T \Sigma^{-1} (x_i - \bar{x}) \right) \right)$$

$$= -\frac{N}{2} \cdot \Sigma - \frac{1}{2} \sum_{i=1}^{N} \frac{\partial}{\partial \Sigma^{-1}} \left(tr\left(\Sigma^{-1} (x_i - \bar{x})(x_i - \bar{x})^T \right) \right).$$

$$= -\frac{N}{2} \cdot \Sigma + \frac{1}{2} \sum_{i=1}^{N} (x_i - \bar{x})^T (x_i - \bar{x}) \quad — (*)$$

Set $(*)$ to 0, $-\frac{N}{2} \Sigma + \frac{1}{2} \sum_{i=1}^{N} (x_i - \bar{x})^T (x_i - \bar{x}) = 0.$

$$\Rightarrow \Sigma = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^T (x_i - \bar{x}) \ \#.$$

3. Unanswered Questions.

上課中有提到 Assignment 4 中，沒有辦法很好的找出 decision boundary. 因為我們是使用線性函數，那麼，我們該怎麼找出一個可以將 valid 和 invalid points 区分的很好的 decision boundary ?

連在平面上只是一個 valid 的網格点也可以和周遭 invalid 区分的那種.