資工四 402410012 許恒誠

1.

想要找 maximum-likelihood estimate of $\boldsymbol{\vartheta}$

必須要 maximizes$P(\boldsymbol{X}|\boldsymbol{\theta}) = \prod_{i=1}^{d} \theta_i^{x_i} (1-\theta_i)^{1-x_i}$

取 logarithm

$$l(\boldsymbol{\theta}) = lnP(\boldsymbol{X}|\boldsymbol{\theta})$$

$$\widehat{\boldsymbol{\theta}} = arg\, maxl(\boldsymbol{\theta})$$

$$l(\boldsymbol{\theta}) = \sum_{i=1}^{d} ln\,(\theta_i^{x_i}(1-\theta_i)^{1-x_i})$$

整理一下

$$= \sum_{i=1}^{d} x_i ln\,\theta_i + (1-x_i)ln(1-\theta_i)$$

$$= \sum_{i=1}^{d} x_i ln\,\theta_i + ln(1-\theta_i) - x_i ln(1-\theta_i)$$

所以

$$\boldsymbol{\nabla_\theta} l(\boldsymbol{\theta}) = \sum_{i=1}^{d} \boldsymbol{\nabla_\theta} ln\,(\theta_i^{x_i}(1-\theta_i)^{1-x_i})$$

$$= \sum_{i=1}^{d} (\frac{x_i}{\theta_i} + \frac{-1}{1-\theta_i} - \frac{-x_i}{1-\theta_i})$$

$$= \sum_{i=1}^{d} (\frac{x_i}{\theta_i} + \frac{x_i - 1}{1-\theta_i})$$

帶入

$$\boldsymbol{\nabla_\theta} l(\boldsymbol{\theta}) = 0$$

$$0 = \sum_{i=1}^{d} (\frac{x_i}{\theta_i} + \frac{x_i - 1}{1-\theta_i})$$

$$\frac{x_i}{\theta_i} + \frac{x_i - 1}{1-\theta_i} = 0$$

$$\theta_i = x_i$$

當$\theta_i = x_i$ 時，$P(\boldsymbol{X}|\boldsymbol{\theta}) = \prod_{i=1}^{d} \theta_i^{x_i} (1-\theta_i)^{1-x_i}$最大。

2.

$$p(\boldsymbol{\theta}|\boldsymbol{D}^0) = p(\boldsymbol{\theta}) = \begin{cases} e^{-\theta} & ,\theta \geq 0 \\ 0 & ,\text{otherwise} \end{cases}$$

$$p(\boldsymbol{D}^n|\boldsymbol{\theta}) = p(\boldsymbol{X}_n|\boldsymbol{\theta})p(|\boldsymbol{D}^{n-1}|\boldsymbol{\theta})$$

第一筆 data x_1

$$p(x_1|\theta) = \begin{cases} e^{\theta - x_1} & ,x_1 \geq \theta \\ 0 & ,\text{otherwise} \end{cases}$$

$$p(\theta|\boldsymbol{D}^1) \propto p(x_1|\theta)\, p(\boldsymbol{\theta}|\boldsymbol{D}^0) = \begin{cases} e^{-x_1} & ,x_1 \geq \theta \geq 0 \\ 0 & ,\text{otherwise} \end{cases}$$

第二筆 data x_2

$$p(x_2|\theta) = \begin{cases} e^{\theta - x_2} & ,x_2 \geq \theta \\ 0 & ,\text{otherwise} \end{cases}$$

$$p(\theta|\boldsymbol{D}^2) \propto p(x_2|\theta)\, p(\boldsymbol{\theta}|\boldsymbol{D}^1) = \begin{cases} e^{\theta - x_1 - x_2} & ,x_1, x_2 \geq \theta \geq 0 \\ 0 & ,\text{otherwise} \end{cases}$$

…持續至第 n 筆 data x_n

$$p(x_n|\theta) = \begin{cases} e^{\theta - x_n} & ,x_n \geq \theta \\ 0 & ,\text{otherwise} \end{cases}$$

$$p(\theta|\boldsymbol{D}^n) \propto p(x_n|\theta)\, p(\boldsymbol{\theta}|\boldsymbol{D}^{n-1}) = \begin{cases} e^{(n-1)\theta - (x_1 + x_2 + \cdots + x_n)} & ,x_1, x_2, \ldots, x_n \geq \theta \geq 0 \\ 0 & ,\text{otherwise} \end{cases}$$

3.

(a)

假設X是以n個隨機變數（其中的每個隨機變數是也是一個向量，當然是一個行向量）組成的列向量，

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$$

並且μ_i是其第i個元素的期望值，即，$\mu_i = \mathrm{E}(X_i)$,其中X_i是列向量中的一個元素。共變異數矩陣的第i，j項（第i，j項是一個共變異數）被定義為如下形式：

$$\Sigma_{ij} = \mathrm{cov}(X_i, X_j) = \mathrm{E}\left[(X_i - \mu_i)(X_j - \mu_j)^\top \right]$$

而共變異數矩陣為：

$$\Sigma = \mathrm{E}\left[(\mathbf{X} - \mathrm{E}[\mathbf{X}]) (\mathbf{X} - \mathrm{E}[\mathbf{X}])^\top \right]$$

$$= \begin{bmatrix} \mathrm{E}[(X_1 - \mu_1)(X_1 - \mu_1)] & \mathrm{E}[(X_1 - \mu_1)(X_2 - \mu_2)] & \cdots & \mathrm{E}[(X_1 - \mu_1)(X_n - \mu_n)] \\ \mathrm{E}[(X_2 - \mu_2)(X_1 - \mu_1)] & \mathrm{E}[(X_2 - \mu_2)(X_2 - \mu_2)] & \cdots & \mathrm{E}[(X_2 - \mu_2)(X_n - \mu_n)] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}[(X_n - \mu_n)(X_1 - \mu_1)] & \mathrm{E}[(X_n - \mu_n)(X_2 - \mu_2)] & \cdots & \mathrm{E}[(X_n - \mu_n)(X_n - \mu_n)] \end{bmatrix}$$

又

- An intuitive **unbiased** estimator for $\mathbf{\Sigma}$ is given by

$$\mathbf{C} = \frac{1}{n-1} \sum_{k=1}^{n} (\mathbf{x}_k - \hat{\boldsymbol{\mu}})(\mathbf{x}_k - \hat{\boldsymbol{\mu}})^t \tag{21}$$

  where $\mathbf{C}$ is called **sample covariance matrix**.

以及

- The solution to this problem involves the so-called **scatter matrix** $\mathbf{S}$ defined by

$$\mathbf{S} = \sum_{k=1}^{n} (\mathbf{x}_k - \mathbf{m})(\mathbf{x}_k - \mathbf{m})^t \tag{54}$$

所以先算出 covariance matrix，再乘上 n-1(先乘 I 矩陣)，就可以得到 scatter matrix。

```
S =

   13.5236   10.3097    4.6601
   10.3097   55.2695   13.0659
    4.6601   13.0659   70.8436
```

(b)

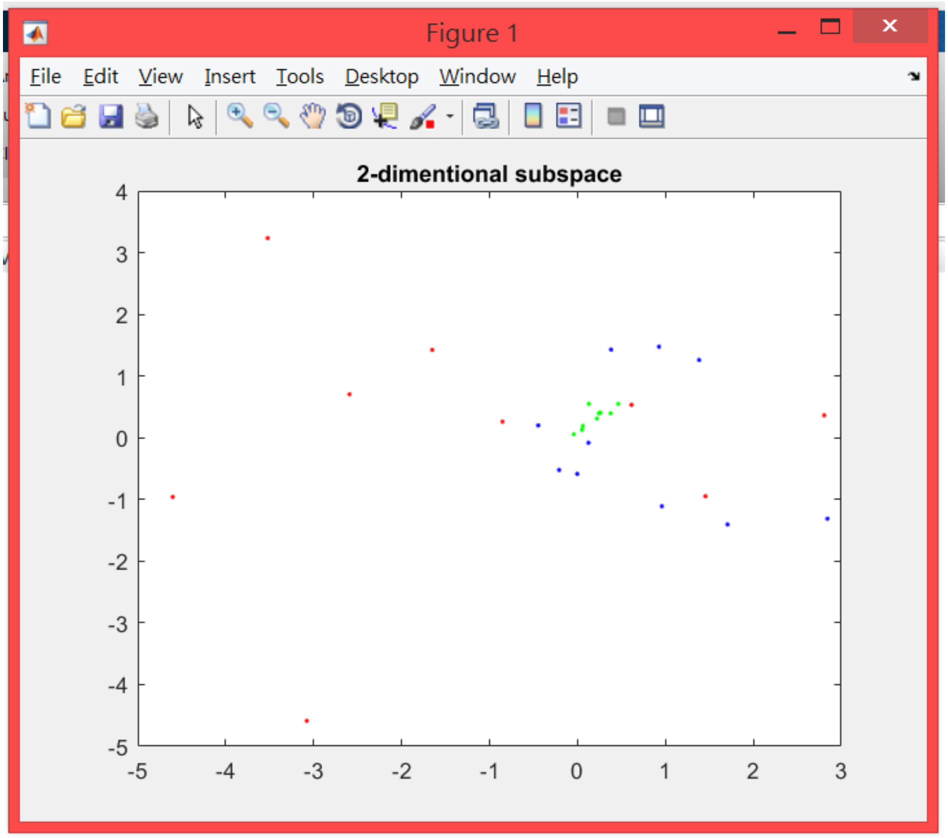
```
eigen_value_1 =          eigen_value_2 =          eigen_value_3 =

   11.0722                  49.0038                  79.5607
```

```
eigen_vector =

    0.9753    0.1710    0.1400
   -0.2192    0.8290    0.5145
   -0.0281   -0.5324    0.8460
```

(c)



(d)

```
MissPercentage =

    0.1667
```

第 4 題沒有寫，所以沒有附 code。