HW2 Zhao Nailin 1003003

2020年2月26日 12:32

1.1 $Ent(p) = -\sum_{i=1}^{n} p_i \log p_i$, $\sum_{i=1}^{n} p_i = 1$

$\mathcal{L}(p, \lambda) = -\sum_{i=1}^{n} p_i \log p_i + \lambda(\sum_{i=1}^{n} p_i - 1)$

$\frac{\partial \mathcal{L}}{\partial p_i} = -[1 + \log p_i] + \lambda$, $\forall i \in \{1, 2, \dots, n\}$

In order to find maximum, $\frac{\partial \mathcal{L}}{\partial p_i} = 0$

$\therefore \frac{\partial \mathcal{L}}{\partial p_i} = 0$

$-[1 + \log p_i] + \lambda = 0$

$\lambda = 1 + \log p_i$

$\log p_i = \lambda - 1$

$e^{\log p_i} = e^{\lambda - 1}$

$p_i = e^{\lambda - 1}$, $\forall i \in \{1, 2, \dots, n\}$

$\because \sum_{i=1}^{n} p_i = 1$

$\therefore n p_i = 1$

$p_i = \frac{1}{n}$, $\forall i \in \{1, 2, \dots, n\}$

$\therefore$ When this matrix is at $(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$, it achieves at maximum entropy.

1.2. Since entropy is always greater than or equal to zero (given)

$\Rightarrow$ lowest bound of $Ent(p) = 0$

If we take $p_i = 1$, $p_j = 0$

$\log p_i$ increases w en p_i cecreases

$\therefore Ent(p) = -\sum_{i=1}^{n} p_i \cdot \log(p_i)$

$= (-1) \cdot \log(1) + (n-1) \cdot \lim_{x \to 0} \frac{\log x}{\frac{1}{x}}$

$= (-1) \cdot 0 + (n-1) \cdot \lim_{x \to 0} \frac{\log x}{x^{-1}}$

$= (n-1) \cdot \lim_{x \to 0} \log x \cdot (x)$

when $x \to 0$, $Ent(p) \to 0$

$\therefore p_i = 1$ & $p_i = 0$ minimizes entropy.

2. A: $n \times n$ matrix

B: $n \times p$ matrix

C: $p \times n$ matrix

D: $p \times p$ matrix

M: $(A - BD^{-1}C)^{-1}$

$$\begin{bmatrix} M & -MBD^{-1} \\ -D^{-1}CM & D^{-1}+D^{-1}CMBD^{-1} \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

$$= \begin{bmatrix} AM - CMBD^{-1} & BM - MB \\ -AD^{-1}CM + CD^{-1} + CD^{-1}CMBD^{-1} & -BD^{-1}CM + I + CMBD^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} M(A - BD^{-1}C) & 0 \\ CD^{-1}M(-A + BD^{-1}C) + CD^{-1} & I \end{bmatrix}$$

$$= \begin{bmatrix} MM^{-1} & 0 \\ CD^{-1}M(-M^{-1}) + CD^{-1} & I \end{bmatrix}$$

$$= \begin{bmatrix} I & 0 \\ -CD^{-1} + CD^{-1} & I \end{bmatrix}$$

$$= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

$\because A^{-1}A = I$

$\therefore \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is the inverse of RHS

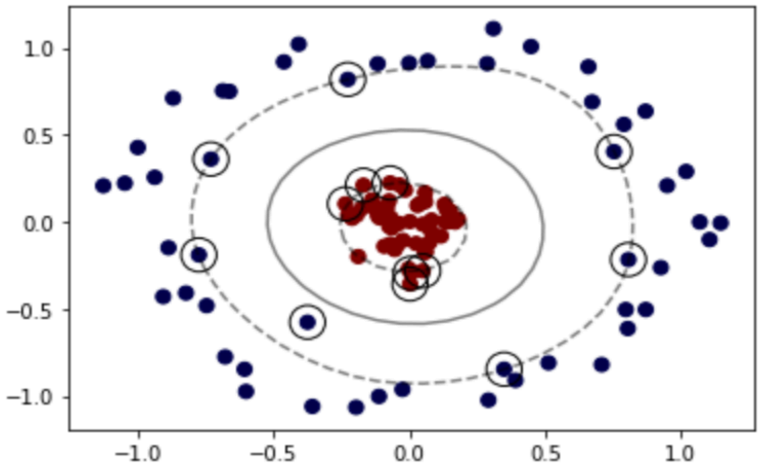
$\therefore \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} M & -MBD^{-1} \\ -D^{-1}CM & D^{-1}+D^{-1}CMBD^{-1} \end{bmatrix}$ (shown)

Qn3.

```
Test set accuracy:  88.31 %
```

Qn4.

```
default kernel: RBF, default C:1
```



```
kernel: RBF, C:3
kernel: RBF, C:10
```