CS 534 Homework 0

1. Yes.

2. institution name, course department, course name, course grade

	institution name	course department	course name	course grade
i	Central South University	Information Science and Engineering	Mathematical Statistics	A
	Central South University	ISE	Probability Theory B	A
ii	Central South University	ISE	Linear Algebra A	A

iii No

iv No

3. $P(\text{Bob} = \text{green}) = P(\text{Bob}=\text{green} \mid \text{Alice} \neq \text{green}) + P(\text{Bob}=\text{green} \mid \text{Alice} = \text{green})$

$$= \frac{7}{12} \times \frac{5}{11} + \frac{5}{12} + \frac{4}{11} = \frac{55}{132}$$

4. To program A: $f(a) = \begin{cases} 1 & 0 \le a \le 1 \\ 0 & a<0 \text{ or } a>1 \end{cases}$, $P(A) = \int_0^a f(a)\,da$.

By using program A:

$$P(X=-1) = \int_0^{0.2} f(a)\,da = 0.2$$

$$P(X=0) = \int_{0.2}^{0.65} f(a)\,da = 0.45$$

$$P(X=1) = \int_{0.65}^{1} f(a)\,da = 0.35$$

in other words, if the number generated by program A is between 0 and 0.2, $X=-1$. If the number is between 0.2 and 0.65, $X=0$. If the number is between 0.65 and 1, $X=1$.

5. See PDF uploaded by computer.

6. let $\vec{x}=\vec{e_1}=[1,0,\cdots,0]\in R^n$, $\vec{w}=[w_1,w_2,\cdots,w_n]$, k scalar.

$\because$ $\langle\vec{e_1},\vec{w}\rangle+\langle\vec{e_1},\vec{e_1}\rangle = w_1+1 = \langle\vec{e_1},\vec{e_1}+\vec{w}\rangle$

and $\langle\vec{e_1},k\vec{w}\rangle = kw_1 = k\langle\vec{e_1},\vec{w}\rangle$

$\therefore$ There at least exists a vector $\vec{x}=\vec{e_1}\in R^n$ such that

$f(\vec{w})=\langle\vec{x},\vec{w}\rangle$, for any $\vec{w}\in R^n$.

7.

i let $\frac{\partial\left(\frac{1}{2}\|\vec{y}-\vec{X}\cdot\vec{w}\|^2\right)}{\partial\vec{w}}=0$,

$$\frac{\partial\left(\frac{1}{2}\|\vec{y}-\vec{X}\cdot\vec{w}\|^2\right)}{\partial\vec{w}} = \frac{\partial\left[\frac{1}{2}(\vec{y}-\vec{X}\cdot\vec{w})^T(\vec{y}-\vec{X}\cdot\vec{w})\right]}{\partial\vec{w}}$$

$$= \frac{1}{2}(-2\vec{X}^T)(\vec{y}-\vec{X}\cdot\vec{w}) = \vec{X}^T\vec{X}\cdot\vec{w}+(-\vec{X}^T\vec{y})$$

$$\frac{\partial(\vec{X}^T\vec{X}\vec{w}-\vec{X}^T\cdot\vec{y})}{\partial\vec{w}} = \vec{X}^T\cdot\vec{X}.$$

$\because$ $\vec{X}^T\vec{X}$ is a positive matrix.

$\therefore$ $(\vec{X}^T\vec{X}\cdot\vec{w}-\vec{X}^T\cdot\vec{y})$ is incremental change.

So if there is a solution to $\vec{X}^T\cdot\vec{X}\cdot\vec{w}-\vec{X}^T\cdot\vec{y}=0$.

The solution will result in a minimum of $(\frac{1}{2}\|\vec{y}-\vec{X}\cdot\vec{w}\|^2$

So. when $\mathrm{rank}(\vec{X})=\mathrm{rank}(\vec{X},\vec{y})=p$, the optimal vector $\vec{w}^*=\vec{X}^+\cdot\vec{y}$.

because $\vec{X}^T\cdot\vec{X}\cdot\vec{w}-\vec{X}^T\cdot\vec{y}=0 \Rightarrow \vec{X}^T(\vec{X}\cdot\vec{w}-\vec{y})=0. \Rightarrow \vec{X}\cdot\vec{w}=\vec{y}$

$\Rightarrow \vec{w}=(\vec{X}^T\vec{X})^{-1}\vec{X}^T\vec{y} \Rightarrow \vec{w}=\vec{X}^+\cdot\vec{y}$, Then $\Rightarrow \vec{w}^*=\vec{X}^+\cdot\vec{y}$

ii If $\vec{X}$ is not full rank, it does not work because it means there will be many solutions in $\vec{X}\cdot\vec{w}=\vec{y}$.

In this case, $\vec{X}^T\cdot\vec{X}\cdot\vec{w}-\vec{X}^T\vec{y}=0$, where $\vec{X}^T\cdot\vec{X}$ is $p\times p$ matrix, $\vec{X}^T\cdot\vec{y}\in R^n$.

$\Rightarrow (\vec{X}^T\cdot\vec{X})\vec{w}=\vec{X}^T\cdot\vec{y}$, $\mathrm{rank}(\vec{X}^T\cdot\vec{X})<p \Rightarrow \vec{w}=(\vec{X}^T\cdot\vec{X})^+(\vec{X}^T\cdot\vec{y})$. where $(\vec{X}^T\vec{X})^+$ can be constructed from the singular value $(\vec{X}^T\vec{X})=UDV^T$, by $(\vec{X}^T\vec{X})^+=V_r D_r^{-1} U_r^T$ where U_r, D_r, V_r are matrices with the degenerated rows and columns deleted.

iii let $\frac{\partial(\frac{1}{2}\|\vec{y}-\vec{X}\cdot\vec{w}\|^2+\frac{c}{2}\|\vec{w}\|^2)}{\partial\vec{w}}=0$.

$$\frac{\partial(\frac{1}{2}\|\vec{y}-\vec{X}\cdot\vec{w}\|^2+\frac{c}{2}\|\vec{w}\|^2)}{\partial\vec{w}}=\vec{X}^T\vec{X}\cdot\vec{w}-\vec{X}^T\cdot\vec{y}+c\vec{w}$$

$$=(\vec{X}^T\vec{X}+c\vec{I}_p)\vec{w}-\vec{X}^T\cdot\vec{y}$$

$$\frac{\partial[(\vec{X}^T\cdot\vec{X}\cdot+c\vec{I}_p)\vec{w}-\vec{X}^T\cdot\vec{y}]}{\partial\vec{w}}=\vec{X}^T\cdot\vec{X}+c\vec{I}_p \Rightarrow \text{positive}$$

$\therefore$ the solution to $[(\vec{X}^T\vec{X}+c\vec{I}_p)\vec{w}-\vec{X}^T\cdot\vec{y}]$ will result in a minimum of $(\frac{1}{2}\|\vec{y}-\vec{X}\vec{w}\|^2+\frac{c}{2}\|\vec{w}\|^2)$, i.e. $\vec{w}^*$

According to the analysis of i and ii, we ~~get~~ can get,

$$\vec{w}^*=(\vec{X}^T\vec{X}+c\vec{I}_p)^{+}(\vec{X}^T\cdot\vec{y}).$$

8. $p(\vec{x};\vec{\mu},\vec{\Sigma}) = \dfrac{1}{\sqrt{(2\pi)^n|\vec{\Sigma}|}} e^{\left[-\frac{1}{2}(\vec{x}-\vec{\mu})^T\vec{\Sigma}^{-1}(\vec{x}-\vec{\mu})\right]}$.

9. $p(\vec{x};\vec{\mu},\vec{\theta}^{-1}) = \dfrac{1}{\sqrt{(2\pi)^n|\theta^{-1}|}} e^{\left[-\frac{1}{2}(\vec{x}-\vec{\mu})^T\vec{\theta}(\vec{x}-\vec{\mu})\right]}$

10. $\because p(\vec{x};\vec{\mu},\vec{\Sigma}) = \dfrac{1}{\sqrt{(2\pi)^n|\vec{\Sigma}|}} e^{\left[-\frac{1}{2}(\vec{x}-\vec{\mu})^T|\vec{\Sigma}|^{-1}(\vec{x}-\vec{\mu})\right]}$, $\vec{\mu} = \begin{pmatrix}\mu_X\\ \mu_Y\end{pmatrix}$,

$\vec{\Sigma} = \begin{pmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y\\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix}$, let $\vec{x} = \begin{pmatrix}x\\ y\end{pmatrix}$

$\therefore p(\vec{x};\vec{\mu},\vec{\Sigma}) = \dfrac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}-\frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right.\right.}$

Then marginal probability density function p_Y:

$$p_Y = \int_{-\infty}^{+\infty} p(\vec{x};\vec{\mu},\vec{\Sigma})\,dx$$

because $\dfrac{(x-\mu_X)^2}{\sigma_X^2} - \dfrac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} = \left(\dfrac{x-\mu_X}{\sigma_X} - \dfrac{\rho(y-\mu_Y)}{\sigma_Y}\right)^2 - \dfrac{\rho^2(y-\mu_Y)^2}{\sigma_Y}$

$$\text{Then } p_Y = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} e^{\left(-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X} - \frac{\rho(y-\mu_Y)}{\sigma_Y}\right)^2 + \frac{(1-\rho^2)(y-\mu_Y)^2}{\sigma_Y^2}\right]\right)} dx$$

$$= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X} - \frac{\rho(y-\mu_Y)}{\sigma_Y}\right)^2\right]} dx$$

$$\text{let } t = \frac{1}{\sqrt{1-\rho^2}}\left[\frac{x-\mu_X}{\sigma_X} - \frac{\rho(y-\mu_Y)}{\sigma_Y}\right]$$

$$\text{Then, } p_Y = \frac{1}{2\pi\sigma_Y} e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}} \int_{-\infty}^{+\infty} e^{-\frac{t}{2}} dt$$

$$\because \int_{-\infty}^{+\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}$$

$$\therefore p_Y = \frac{1}{\sqrt{2\pi}\sigma_Y} e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}}$$

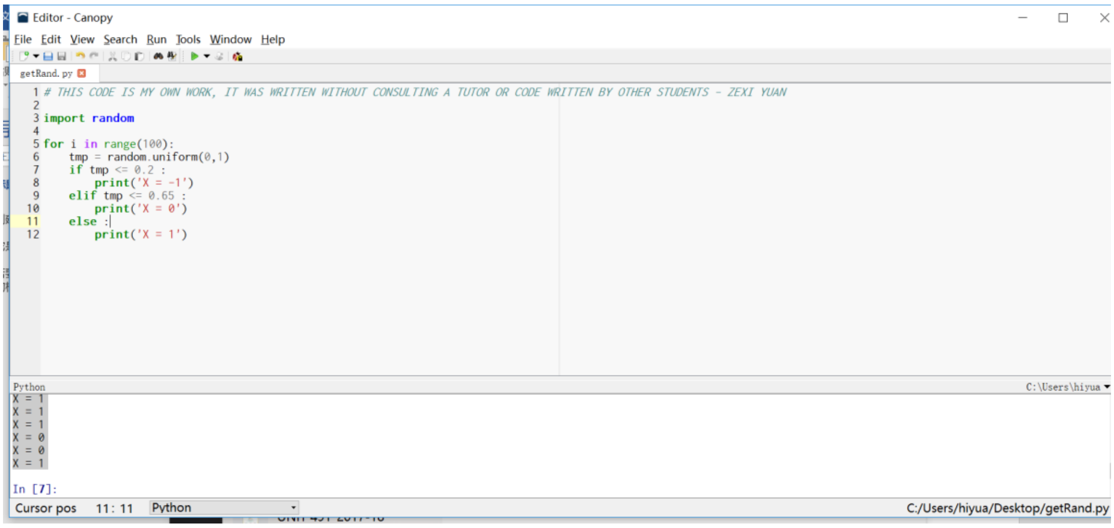
$$\therefore P(Y) = \int_{-\infty}^{y} p_Y \, dy = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\sigma_Y} e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}} dy.$$

ii. $P(X|Y) = \int_{-\infty}^{x} \frac{p(\vec{x};\vec{\mu},\vec{\Sigma})}{p_Y} dx.$

$\because$ we know p_Y and $p(\vec{x};\vec{\mu},\vec{\Sigma})$ from question i.

$$\therefore P(X|Y) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} - \frac{(1-\rho^2)(y-\mu_Y)^2}{\sigma_Y^2}\right]} dx$$

$$= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{x-\mu_X}{\sigma_X} - \frac{\rho(y-\mu_Y)}{\sigma_Y}\right]^2} dx$$



getRand.py

```
# THIS CODE IS MY OWN WORK, IT WAS WRITTEN WITHOUT CONSULTING A TUTOR
OR CODE WRITTEN BY OTHER STUDENTS - ZEXI YUAN

import random

for i in range(100):
    tmp = random.uniform(0,1)
    if tmp <= 0.2 :
        print('X = -1')
    elif tmp <= 0.65 :
        print('X = 0')
    else :
        print('X = 1')
```

Result :

%run "C:/Users/hiyua/Desktop/getRand.py"

X = -1

X = -1

X = 1

X = 0

X = 1

X = 0

X = 1

X = 0

X = 0

X = 0

X = 0

X = 0

X = 1

X = -1
X = 0
X = -1
X = 1
X = 0
X = 1
X = 0
X = -1
X = 0
X = 1
X = -1
X = -1
X = 0
X = -1
X = 0
X = 1
X = 1
X = 0
X = 1
X = 1
X = 0
X = 0
X = -1
X = 0
X = 0
X = -1
X = 0
X = 0
X = 1
X = 0
X = -1
X = 0
X = 0
X = 1
X = 1
X = 0
X = -1
X = 1
X = -1
X = 0
X = 1
X = -1
X = 0
X = 0

X = 1
X = 0
X = 0
X = -1
X = -1
X = 0
X = 1
X = 0
X = 1
X = -1
X = 0
X = 0
X = -1
X = 0
X = 0
X = 1
X = 1
X = 1
X = 0
X = 0
X = 0
X = 0
X = 0
X = 0
X = 1
X = 0
X = 0
X = 0
X = 1
X = 0
X = 1
X = 1
X = -1
X = 0
X = 0
X = 0
X = 1
X = 1
X = 1
X = 1
X = 0
X = 0
X = 1