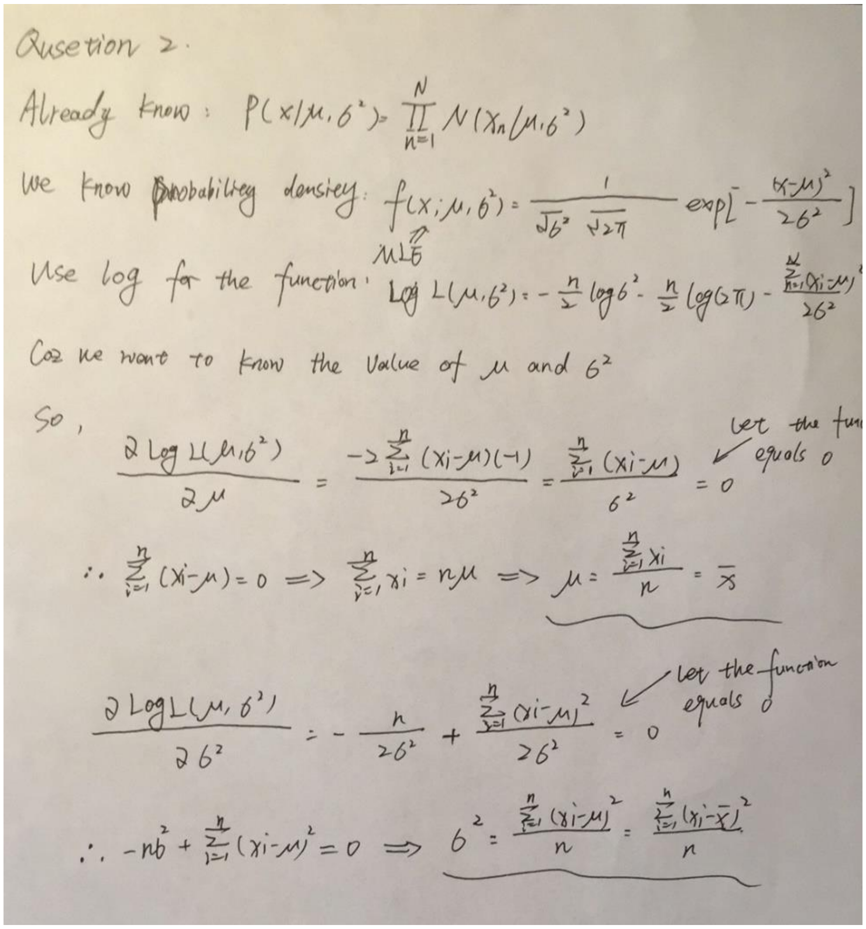
Q1 As we know, the information is in the table below:

	Red Box	Blue Box	Green Box
Apple	3	1	3
Oranges	4	0	3
Limes	3	1	4

a) What is the probability of selecting an apple?
P(Apple) = P(Apple|Red)*P(Red) + P(Apple|Blue)*P(Blue) + P(Apple|Green)*P(Green)
= 0.3 * 0.2 + 0.5 * 0.2 + 0.3 * 0.6
= 0.06 + 0.1 + 0.18
= 0.34

b) If select orange, what is the probability that it came from the green box?
P(Green|Orange) = P(Green Orange)/P(Orange)
= 0.3 x 0.6 / (0.4 * 0.2 + 0.5 * 0.2 + 0.3 * 0.6)
= 0.18 / 0.36
= 0.5

Q2.



Qusetion 2.

Already know: $P(x|\mu,\sigma^2) = \prod_{n=1}^{N} N(x_n|\mu,\sigma^2)$

We know probabilitey densiey: $f(x;\mu,\sigma^2) = \frac{1}{\sqrt{\sigma^2}\sqrt{2\pi}} \exp[-\frac{(x-\mu)^2}{2\sigma^2}]$

MLE

Use log for the function: $\log L(\mu,\sigma^2) = -\frac{n}{2}\log\sigma^2 - \frac{n}{2}(\log 2\pi) - \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{2\sigma^2}$

Coz we want to know the value of μ and σ^2

So,

$$\frac{\partial \log L(\mu,\sigma^2)}{\partial \mu} = \frac{-2\sum_{i=1}^{n}(x_i-\mu)(-1)}{2\sigma^2} = \frac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma^2} = 0$$

Let the fun
equals 0

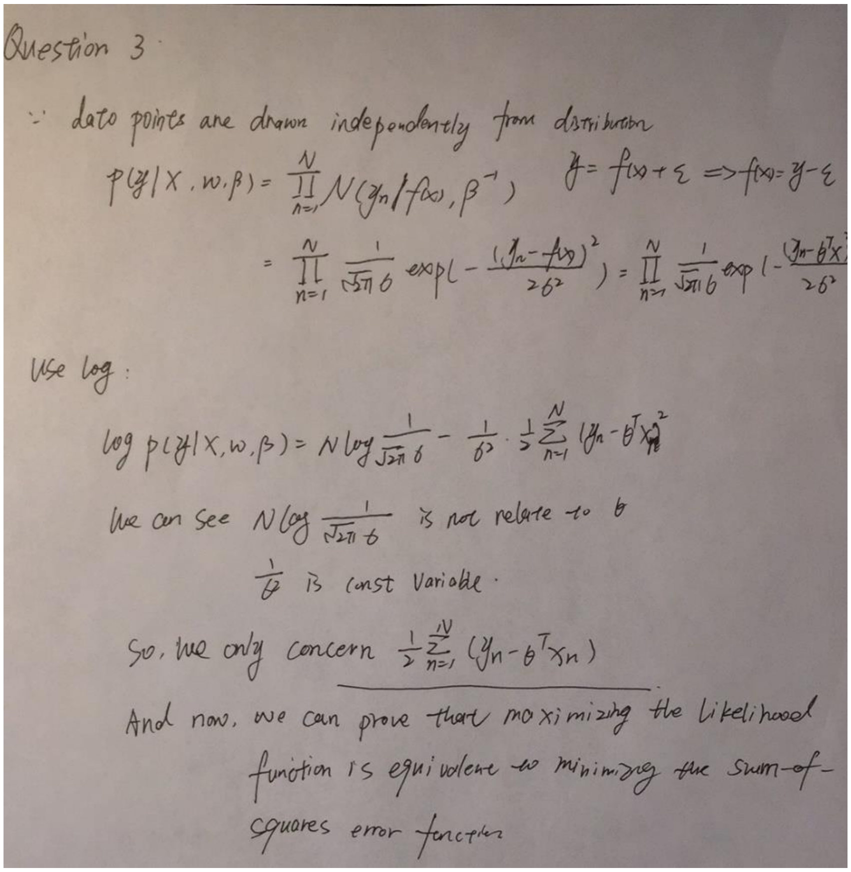
$$\therefore \sum_{i=1}^{n}(x_i-\mu) = 0 \Rightarrow \sum_{i=1}^{n} x_i = n\mu \Rightarrow \underline{\mu = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}}$$

$$\frac{\partial \log L(\mu,\sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2} = 0$$

Let the function
equals 0

$$\therefore -n\sigma^2 + \sum_{i=1}^{n}(x_i-\mu)^2 = 0 \Rightarrow \underline{\sigma^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}}$$

Q3



Question 3 :

$\because$ data points are drawn independently from distribution

$$p(y|X,w,\beta) = \prod_{n=1}^{N} \mathcal{N}(y_n | f(x), \beta^{-1}) \qquad y = f(x) + \varepsilon \Rightarrow f(x) = y - \varepsilon$$

$$= \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(y_n - f(x))^2}{2\sigma^2}) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(y_n - \theta^T x}{2\sigma^2}$$

use log :

$$\log p(y|X,w,\beta) = N \log \frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{\sigma^2} \cdot \frac{1}{2} \sum_{n=1}^{N} (y_n - \theta^T x_n)^2$$

We can see $N \log \frac{1}{\sqrt{2\pi}\sigma}$ is not relate to θ

$\frac{1}{\sigma^2}$ is const variable.

So, we only concern $\underline{\frac{1}{2} \sum_{n=1}^{N} (y_n - \theta^T x_n)}$.

And now, we can prove that maximizing the likelihood function is equivalent to minimizing the sum-of-squares error function

Q4

Question 4

prior prob

The Bayes's theorem: $p(A|B) = \frac{p(A) \times p(B|A)}{p(B)}$ → postirior prob

→ prior prob

We Note that: posterior distribution of w: $p(w|x, y, \alpha, \beta)$

prior distribution of w: $p(w|\alpha) = (\frac{\alpha}{2\pi})^{(M+1)/2} \exp(-\frac{\alpha}{2} w^T w)$

We can get that: $p(w|x, y, \alpha, \beta) = \frac{p(y|x, w, \beta) \times p(w|\alpha)}{p(x, y, \alpha, \beta)}$

$\propto p(y|x, w, \beta) \times p(w|\alpha)$

~~$= (\frac{\alpha}{\sigma})^{\frac{M+1}{2}} \cdot (\frac{1}{2\pi})^{\frac{M+2}{2}} \exp(-\frac{f(x_n, w)}{2\sigma^2} +$~~

$= (\frac{\alpha}{\sigma})^{\frac{M+1}{2}} \cdot (\frac{1}{2\pi})^{\frac{M+2}{2}} \exp(\frac{(f(x, \vec{w}) - y)^2}{2\sigma^2} + \frac{\alpha}{2} w^T w)$

use log: $= \underbrace{\log (\frac{\alpha}{\sigma})^{\frac{M+1}{2}} (\frac{1}{2\pi})^{\frac{M+2}{2}}}_{\text{const varaible}} \cdot \frac{1}{2\sigma^2} \sum_{n=1}^{N} (f(x, \vec{w}) - y_n)^2 + \frac{\alpha}{2} w^T w$

$x = (x_1 \cdots x_n)^T$

$y = (y_1 \cdots y_n)^T$

$\vec{w}$

$\Rightarrow$

So, we can see that $\underbrace{\frac{1}{2\sigma^2} \sum_{n=1}^{N} (f(x, \vec{w} - y_n)^2 + \frac{\alpha}{2} w^T w}$

we can prove that MAP is equivalent to minimizing the regularized sum-of-squares error function