

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?

$$\begin{aligned}
 P(\text{Apple}) &= P(\text{Apple} | \text{Red}) * P(\text{Red}) + P(\text{Apple} | \text{Blue}) * P(\text{Blue}) + P(\text{Apple} | \text{Green}) * P(\text{Green}) \\
 &= 0.3 * 0.2 + 0.5 * 0.2 + 0.3 * 0.6 \\
 &= 0.06 + 0.1 + 0.18 \\
 &= 0.34
 \end{aligned}$$

b) If select orange, what is the probability that it came from the green box?

$$\begin{aligned}
 P(\text{Green} | \text{Orange}) &= P(\text{Green Orange}) / P(\text{Orange}) \\
 &= 0.3 * 0.6 / (0.4 * 0.2 + 0.5 * 0.2 + 0.3 * 0.6) \\
 &= 0.18 / 0.36 \\
 &= 0.5
 \end{aligned}$$

Q2.

Question 2.

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

We know probability density: $f(x|\mu, \sigma^2) = \frac{1}{\sqrt{\sigma^2} \sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$

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

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

So,

$$\frac{\partial \text{Log } L(\mu, \sigma^2)}{\partial \mu} = \frac{-\sum_{i=1}^n (x_i - \mu)(-1)}{\sigma^2} = \frac{\sum_{i=1}^n (x_i - \mu)}{\sigma^2} \stackrel{\text{let the function equals 0}}{=} 0$$

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

$$\frac{\partial \text{Log } L(\mu, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^4} \stackrel{\text{let the function equals 0}}{=} 0$$

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

Q3

Question 3.

\therefore data points are drawn independently from distribution

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

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

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 $\frac{1}{2} \sum_{n=1}^N (y_n - \theta^T x_n)^2$

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

Q4

Question 4

The Bayes's theorem: $p(A|B) = \frac{p(A) \times p(B|A)}{p(B)}$
 \downarrow prior prob \rightarrow posterior prob
 \downarrow prior prob

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

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

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

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

$$= \left(\frac{\alpha}{2\pi}\right)^{M+1/2} \cdot \left(\frac{1}{2\pi}\right)^{M+2/2} \exp\left(-\frac{\alpha}{2} w^T w\right) \cdot \frac{1}{2\pi} \exp\left(-\frac{\alpha}{2} w^T w\right)$$

$$= \left(\frac{\alpha}{2\pi}\right)^{M+1/2} \cdot \left(\frac{1}{2\pi}\right)^{M+2/2} \exp\left(-\frac{\alpha}{2} w^T w\right) \cdot \frac{1}{2\pi} \exp\left(-\frac{\alpha}{2} w^T w\right)$$

use log:

$$= \log\left(\frac{\alpha}{2\pi}\right)^{M+1/2} + \log\left(\frac{1}{2\pi}\right)^{M+2/2} - \frac{\alpha}{2} w^T w - \frac{1}{2\pi} \sum_{n=1}^N (x_n^T w - y_n)^2 + \frac{\alpha}{2} w^T w$$

$$x = [x_1 \dots x_n]^T \Rightarrow$$

$$y = [y_1 \dots y_n]^T$$

$$w$$

So, we can see that $\frac{1}{2\pi} \sum_{n=1}^N (x_n^T w - y_n)^2 + \frac{\alpha}{2} w^T w$

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

-5 You should include your data splitting schemes, your code related choices and comments in your report