Problem Ex.1

$$\begin{aligned}\hat{y}(x) &= \underset{k}{argmax}\, P(Y=k|x) \\ &= \underset{k}{argmax}\, f_k(x)\pi_k \\ &= \underset{k}{argmax} \log f_k(x) + \log \pi_k \\ \log f_k(x) &= -\frac{1}{2}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k) - \frac{p}{2}\log 2\pi - \frac{1}{2}\log|\Sigma_k| \\ &= x^T\Sigma_k^{-1}\mu_k - \frac{1}{2}\mu_k^T\Sigma_k^{-1}\mu_k - \frac{1}{2}(\log|\Sigma_k| + p\log 2\pi + x^T\Sigma_k^{-1}x)\end{aligned}$$

Noticed that Σ_k is constant for each k in LDA, so

$$\hat{y}(x) = \underset{k}{argmax}(\log \pi_k + x^T\Sigma_k^{-1}\mu_k - \frac{1}{2}\mu_k^T\Sigma_k^{-1}\mu_k)$$

Problem Ex.2

(1).LDA classifies to class 2 if and only if $\log\frac{f_2(x)\pi_2}{f_1(x)\pi_1} > 0$.

$$\begin{aligned}\log\frac{f_2(x)\pi_2}{f_1(x)\pi_1} &= \log\frac{N_2}{N_1} + x^T\Sigma^{-1}(\hat{\mu_2}-\hat{\mu_1}) - \frac{1}{2}(\hat{\mu_2}^T\Sigma^{-1}\hat{\mu_2} - \hat{\mu_1}^T\Sigma^{-1}\hat{\mu_1}) \\ &= \log\frac{N_2}{N_1} + x^T\Sigma^{-1}(\hat{\mu_2}-\hat{\mu_1}) - \frac{1}{2}(\hat{\mu_2}+\hat{\mu_1})^T\Sigma^{-1}(\hat{\mu_2}-\hat{\mu_1})\end{aligned}$$

So the LDA rules classifies x to class 2 when

$$x^T\Sigma^{-1}(\hat{\mu_2}-\hat{\mu_1}) > \frac{1}{2}(\hat{\mu_2}+\hat{\mu_1})^T\Sigma^{-1}(\hat{\mu_2}-\hat{\mu_1}) - \log N_2/N_1$$

(2).In this particular case,

$$\nabla_{\beta_0} = -2\sum_{i=1}^{N}(y_i - \beta_0 - x_i^T\beta) = 0$$

$$\beta_0 = \frac{1}{N}\sum_{i=1}^{N}(y_i - x_i^T\beta)$$

$$\nabla_{\beta} = -2\sum_{i=1}^{N}(y_i - \beta_0 - x_i^T\beta)x_i = 0$$

$$\beta_0\sum_{i=1}^{N}x_i + \beta^T\sum_{i=1}^{N}x_ix_i^T = \sum_{i=1}^{N}y_ix_i$$

Noticed that,

$$\sum_{i=1}^{N} y_i = -N_1\frac{N}{N_1} + N_2\frac{N}{N_2} = 0$$

$$(\sum_{i=1}^{N} x_i^T)\beta = \beta^T(N_1\hat{\mu_1} + N_2\hat{\mu_2})$$

So $\beta_0 = -\frac{1}{N}\beta^T(N_1\hat{\mu_1} + N_2\hat{\mu_2})$. At one hand,

$$\sum_{i=1}^{N} y_i x_i = -\frac{N}{N_1}N_1\hat{\mu_1} + \frac{N}{N_2}N_2\hat{\mu_2} = N(\hat{\mu_2} - \hat{\mu_1})$$

$$\begin{aligned}\sum_{i=1}^{N} \beta_0 x_i &= -\frac{1}{N}\beta^T(N_1\hat{\mu_1} + N_2\hat{\mu_2})(N_1\hat{\mu_1} + N_2\hat{\mu_2})^T \\ &= -\frac{1}{N}\beta^T(N_1^2\hat{\mu_1}\hat{\mu_1}^T + N_2^2\hat{\mu_2}\hat{\mu_2}^T + N_1N_2\hat{\mu_1}\hat{\mu_2}^T + N_1N_2\hat{\mu_2}\hat{\mu_1}^T)\end{aligned}$$

So we got that:

$$\beta_0\sum_{i=1}^{N} x_i + \beta^T\sum_{i=1}^{N} x_i x_i^T = (\sum_{i=1}^{N} x_i x_i^T - \frac{N_1^2}{N}\hat{\mu_1}\hat{\mu_1}^T - \frac{N_2^2}{N}\hat{\mu_2}\hat{\mu_2}^T - \frac{N_1N_2}{N}\hat{\mu_1}\hat{\mu_2}^T + \frac{N_1N_2}{N}\hat{\mu_2}\hat{\mu_1}^T)\beta$$

At the other hand,

$$\begin{aligned}(N-2)\Sigma &= \sum_{x\in C_1}(x-\hat{\mu_1})(x-\hat{\mu_1})^T + \sum_{x\in C_2}(x-\hat{\mu_2})(x-\hat{\mu_2})^T \\ &= \sum_{i=1}^{N} x_i x_i^T - N_1\hat{\mu_1}\hat{\mu_1}^T - N_2\hat{\mu_2}\hat{\mu_2}^T \\ N\Sigma_B &= \frac{N_1N_2}{N}(\hat{\mu_1}-\hat{\mu_2})(\hat{\mu_1}-\hat{\mu_2})^T \\ &= \frac{N_1N_2}{N}(\hat{\mu_1}\hat{\mu_1}^T + \hat{\mu_2}\hat{\mu_2}^T - \hat{\mu_1}\hat{\mu_2}^T - \hat{\mu_2}\hat{\mu_1}^T)\end{aligned}$$

Noticed that $N = N_1 + N_2$, now we got:

$$(N-2)\Sigma + N\Sigma_B = \sum_{i=1}^{N} x_i x_i^T - \frac{N_1^2}{N}\hat{\mu_1}\hat{\mu_1}^T - \frac{N_2^2}{N}\hat{\mu_2}\hat{\mu_2}^T - \frac{N_1N_2}{N}\hat{\mu_1}\hat{\mu_2}^T + \frac{N_1N_2}{N}\hat{\mu_2}\hat{\mu_1}^T$$

$$(N-2)\Sigma + N\Sigma_B = \beta_0\sum_{i=1}^{N} x_i + \beta^T\sum_{i=1}^{N} x_i x_i^T$$

So,

$$((N-2)\Sigma + N\Sigma_B)\beta = N(\hat{\mu_2} - \hat{\mu_1}) \tag{1}$$

(3).Noticed that $(\hat{\mu_1} - \hat{\mu_2})^T\beta \in \mathbb{R}$.

$$\Sigma_B\beta = \frac{N_1N_2}{N^2}(\hat{\mu_1} - \hat{\mu_2})[(\hat{\mu_1} - \hat{\mu_2})^T\beta]$$

So $\Sigma_B\beta$ is in the direction of $\hat{\mu_1} - \hat{\mu_2}$. From (1), we conclude that $\Sigma\beta \propto (\hat{\mu_2} - \hat{\mu_1})$.

Problem Ex.3

Let A denotes that the patient has flu, B denotes he/she has a fever and C denotes cough. Considering the circumstance, there should be $B \subset A$ and $C \subset A$.

$$
\begin{aligned}
P(B|A) &= \frac{P(B)}{P(A)} = 0.5 \\
P(BC|A) &= \frac{P(BC)}{P(A)} = 0.25 \\
P(B) &= P(C) = 0.25, P(BC) = 0.125 \\
Corr(B, C) &= \frac{P(BC) - P(B)P(C)}{\sqrt{P(B)P(\bar{B})P(C)P(\bar{C})}} \\
&= \frac{1}{3}
\end{aligned}
$$