$p(x|c) \sim N(\mu_c, \Sigma_c)$

$$LL(\theta_c) = \log P(D_c|\theta_c) = \sum_{x \in D_c} \log P(x|\theta_c)$$

$$\hat{\theta}_c = \underset{\theta_c}{\arg\max}\, LL(\theta_c) = \underset{\theta_c}{\arg\min} - LL(\theta_c)$$

$$= \underset{\theta_c}{\arg\min} - \sum_{x \in D_c} \log P(x|\theta_c)$$

$$P(x|\theta_c) = P(x|\mu_c, \sigma_c^2) = \frac{1}{\sqrt{(2\pi)^d |\Sigma_c|}} e^{\left(-\frac{1}{2}(x-\mu_c)^T \Sigma_c^{-1} (x-\mu_c)\right)}$$

$$\hat{\mu_c}, \hat{\Sigma_c} = \underset{\mu_c, \Sigma_c}{\arg\min} - \sum_{x \in D_c} \log \left[\frac{1}{\sqrt{(2\pi)^d |\Sigma_c|}} e^{-\frac{1}{2}(x-\mu_c)^T \Sigma_c^{-1} (x-\mu_c)} \right]$$

$$= \underset{\mu_c, \Sigma_c}{\arg\min} - \sum_{x \in D_c} \left[-\frac{d}{2} \log(2\pi) - \frac{1}{2} \log|\Sigma_c| - \frac{1}{2}(x-\mu_c)^T \Sigma_c^{-1} (x-\mu_c) \right]$$

$$= \underset{\mu_c, \Sigma_c}{\arg\min} \sum_{x \in D_c} \left[\frac{1}{2} \log|\Sigma_c| + \frac{1}{2}(x-\mu_c)^T \Sigma_c^{-1} (x-\mu_c) \right]$$

假设 $|D_c| = n$

$$\hat{\mu_c}, \hat{\Sigma_c} = \underset{\mu_c, \Sigma_c}{\arg\min} \sum_{i=1}^{n} \left[\frac{1}{2} \log|\Sigma_c| + \frac{1}{2}(x_i-\mu_c)^T \Sigma_c^{-1} (x_i-\mu_c) \right]$$

$$= \underset{\mu_c, \Sigma_c}{\arg\min} \left[\frac{n}{2} \log|\Sigma_c| + \sum_{i=1}^{n} \frac{1}{2}(x_i-\mu_c)^T \Sigma_c^{-1} (x_i-\mu_c) \right]$$

$$= \underset{\mu_c, \Sigma_c}{\arg\min} \frac{n}{2} \log|\Sigma_c| + \frac{1}{2} tr\left[\Sigma_c^{-1} \sum_{i=1}^{n} (x_i-\bar{x})(x_i-\bar{x})^T \right] + \frac{n}{2}(\mu_c - \bar{x})^T \Sigma_c^{-1} (\mu_c - \bar{x})$$

当 $\mu_c - \bar{x} \neq 0$ 不成立. 故 $\mu_c = \bar{x}$ 时.

$$\hat{\mu_c} = \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{1}{|D_c|} \sum_{x \in D_c} x_i$$

$$\hat{\Sigma_c} = \underset{\Sigma_c}{\arg\min} \frac{n}{2} \log|\Sigma_c| + \frac{1}{2} tr\left[\Sigma_c^{-1} \sum_{i=1}^{n} (x_i-\bar{x})(x_i-\bar{x})^T \right]$$

[illegible] $\frac{n}{2} \log|\Sigma| + \frac{1}{2} tr|\Sigma^{-1} B| \geq \frac{n}{2} \log|B| + \frac{pn}{2}(1 - \log n) \quad \Sigma = \frac{1}{n} B$

$$\Sigma_c = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})^T$$

故 $\Sigma_c = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})^T$

$$= \frac{1}{|D_c|} \sum_{x \in D_c} (x_i - \hat{\mu_c})(x_i - \hat{\mu_c})^T.$$