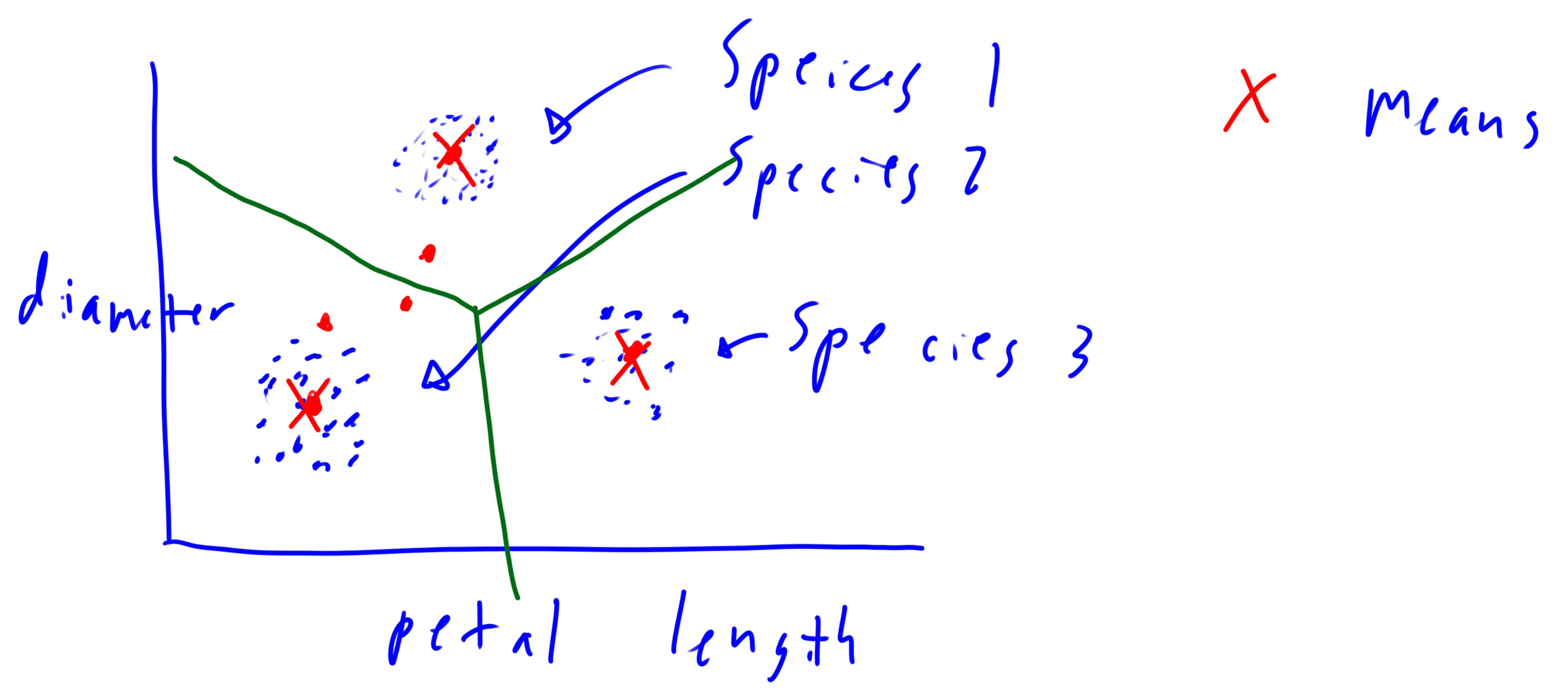


Classification

K means

k clusters

assign $m_i^{(0)}$ randomly $1 \le i \le K$

Assignment

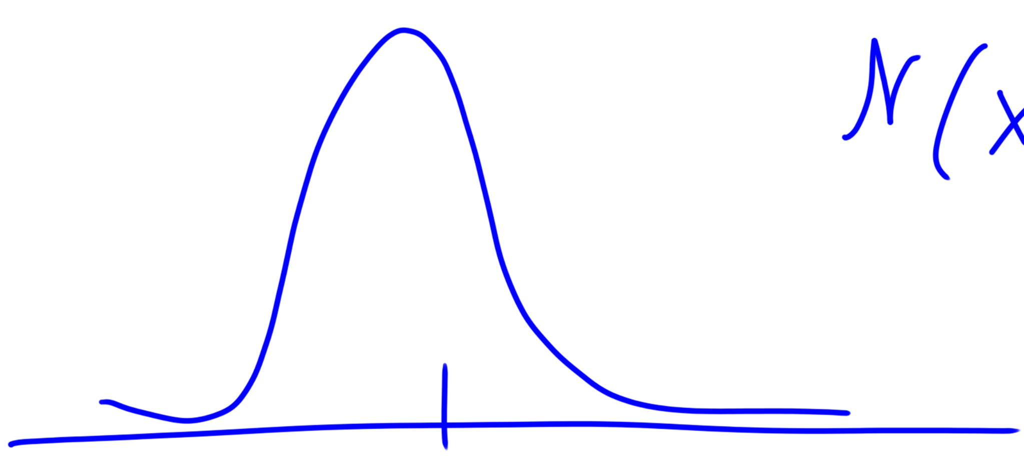
$$S_i^{(t)} = \{ x_p : \| x_p - m_i^{(t)} \|^2 \le \| x_p - m_j^{(t)} \|^2 \ \forall j, \ 1 \le j \le K, \ j \ne i \}$$

Update

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j$$

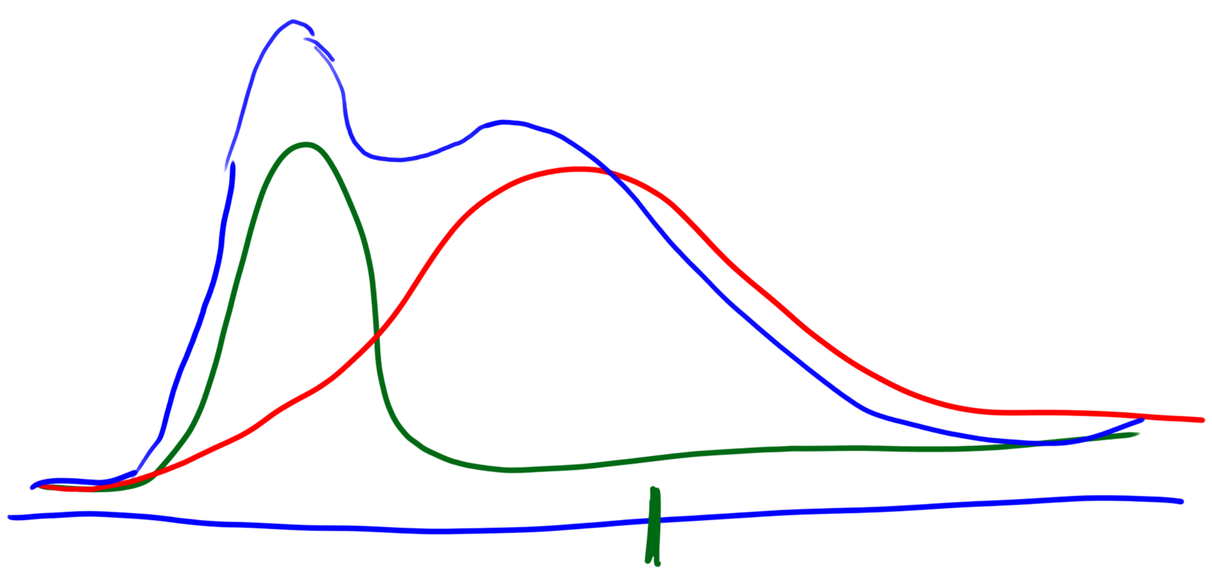
$$= \text{mean}(S_i^{(t)})$$

Gaussian Mixture Model



$N(x \mid \mu, \sigma)$

$$1 = \int_{-\infty}^{\infty} N(x \mid \mu, \sigma)\, dx$$



$$\text{PDF} = \sum_i \alpha_i N(x \mid \mu_i, \sigma_i)$$

$$1 = \int_{-\infty}^{\infty} \sum_i \alpha_i N(x \mid \mu_i, \sigma_i)\, dx$$

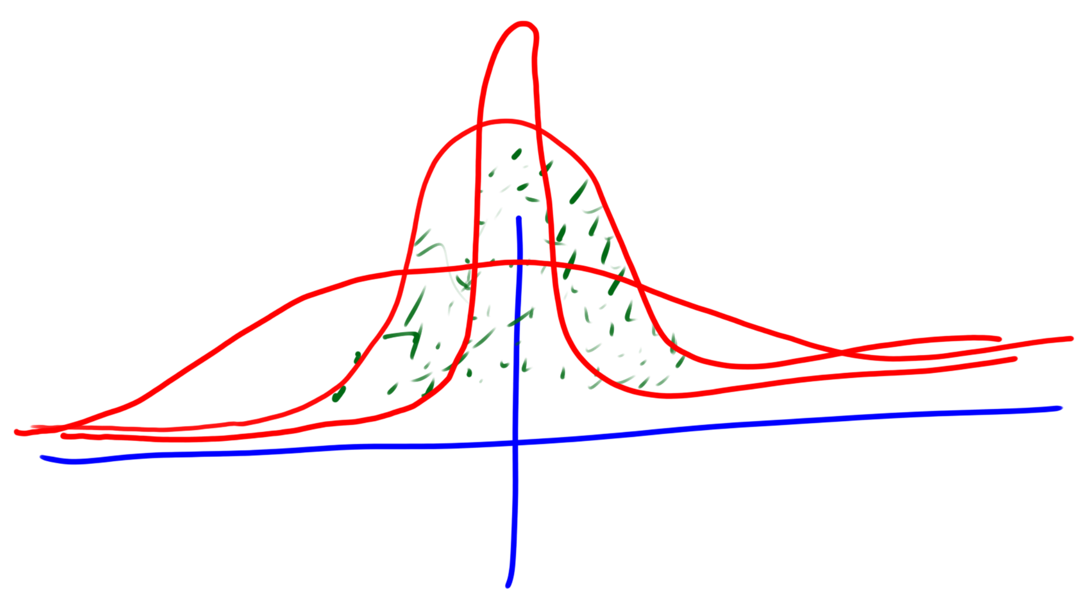
n data points

K clusters

$$P(X \mid \theta) = \sum_{j=1}^{K} \alpha_j N(X \mid \mu_j, \Sigma_j)$$

$$\theta = \{ \alpha_1, \cdots, \alpha_K, \mu_1, \cdots, \mu_K, \Sigma_1, \cdots \Sigma_K \}$$

$\underset{\theta}{\text{maximize}} \ P(X \mid \theta)$



$$P(x_i \mid \theta) = \sum_{j=1}^{K} \alpha_j N(x_i \mid \mu_j, \Sigma_j)$$

$$P(X \mid \theta) = \prod_{i=1}^{n} \sum_{j=1}^{K} \alpha_j N(x_i \mid \mu_j, \Sigma_j)$$

$$\log(P(X \mid \theta)) = \sum_{i=1}^{n} \sum_{j=1}^{K} \alpha_j N(x_i \mid \mu_j, \Sigma_j)$$

Expectation Maximization

iterative

expectation step

Maximization step