k-means clustering with spesific covariance matrix for each k.

Prob density for $\vec{x}$

$$P(\vec{x}) = \sum_{k=1}^{k} \underbrace{\pi_k(\vec{x})}_{\text{Priori prob of } \vec{x} \text{ belonging to group } k} \underbrace{N(\vec{x}, \vec{\mu}_k, \Sigma_k)}_{\text{Multi dim gaussian for each } k}$$

Assuming every $\vec{x}$ belongs to a class π_k's sum to 1

$\sum_{k=1}^{k} \pi_k(\vec{x}) = 1 \rightarrow k-1$ parameters

For each of the k groups we need

$\vec{\mu}_k$: mean of group k, one number for each dim

$\hookrightarrow$ p parameters

The covariance matrix diagonal is σ_i^2 $i = 1, \dots, p$

determined by $\vec{\mu}_k$, no constraints

Also $\Sigma_k^T = \Sigma_k$ so only lower/upper diagonal needs

estimation, given $1 + 2 + 3 + \dots + p-1$ parameters

$$\begin{pmatrix} \sigma_1^2 & a_{11} & a_{21} & \cdots & a_{p-1,1} \\ a_{11} & \sigma_2^2 & a_{22} & \ddots & \vdots \\ a_{21} & a_{22} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a_{p-1,p-1} \\ a_{p-1,1} & \cdots & a_{p-1,p-1} & & \sigma_p^2 \end{pmatrix}$$

$$\sum_{i=1}^{p-1} i = \frac{p(p-1)}{2}$$

And since μ_k, Σ_k are class spesific multiply by k so

$$\underbrace{k-1}_{\pi_k} + \underbrace{k\left(p + \frac{p(p-1)}{2}\right)}_{\vec{\mu}_k, \Sigma_k}$$