# $k$-means vs. GMM

Comparison of $k$-means and GMM iterations for the clustering of a set of data points, $\{\mathbf{x_i}\}_{\mathbf{i=1}}^{\mathbf{N}}$.

| | **$k$-means** | **GMMS** |
|---|---|---|
| Init: | Select $K$ cluster centers $(\mathbf{m_1^{(1)}}, \ldots, \mathbf{m_K^{(1)}})$ | $K$ components with means, $\mu_k$ and covariance $\Sigma_k$ and mixing coefficients, $P_k$ |
| E-step: | $S_i^{(t)} = \{\mathbf{x_p} : \|\|\mathbf{x_p} - \mathbf{m_i^{(t)}}\|\|^{\mathbf{2}} \leq \|\|\mathbf{x_p} - \mathbf{m_j^{(t)}}\|\|^{\mathbf{2}} \forall \mathbf{j}\}$ | $\gamma_{nk} = \frac{P_k \mathcal{N}(\mathbf{x_n} \mid \mu_\mathbf{k}, \mathbf{\Sigma_k})}{\sum_{j=1}^{K} P_j \mathcal{N}(\mathbf{x_n} \mid \mu_\mathbf{j}, \mathbf{\Sigma_j})}$ |
| M-step: | $\mathbf{m_i^{(t+1)}} = \frac{\mathbf{1}}{\mathbf{\|S_i^{(t)}\|}} \sum_{\mathbf{x_j} \in \mathbf{S_i^{(t)}}} \mathbf{x_j}$ | $N_k = \sum_{n=1}^{N} \gamma_{nk}$,<br>$\mu_k^{new} = \frac{1}{N_k} \sum_{n=1}^{N} \gamma_{nk} \mathbf{x_n}$,<br>$\Sigma_k^{new} = \frac{1}{N_k} \sum_{n=1}^{N} \gamma_{nk} (\mathbf{x_n} - \mu_\mathbf{k}^{\mathbf{new}})(\mathbf{x_n} - \mu_\mathbf{k}^{\mathbf{new}})^{\mathbf{T}}$,<br>$P_k^{new} = \frac{N_k}{N}$, |
| Stop: | Based on no differences in set allocation. | When the likelihood not increasing fast enough:<br>$\ln \Pr(\mathbf{x} \mid \mu, \mathbf{\Sigma}, \mathbf{P}) = \sum_{\mathbf{n=1}}^{\mathbf{N}} \ln\{\sum_{\mathbf{k=1}}^{\mathbf{K}} \mathbf{P_k} \mathcal{N}(\mathbf{x_n} \mid \mu_\mathbf{k}, \mathbf{\Sigma_k})\}$ |