E.M. for General GMMs

Iterate: On the t'th iteration let our estimates be

$$\lambda_t = \{ \mu_1^{(t)}, \mu_2^{(t)} \ldots \mu_K^{(t)}, \Sigma_1^{(t)}, \Sigma_2^{(t)} \ldots \Sigma_K^{(t)}, p_1^{(t)}, p_2^{(t)} \ldots p_K^{(t)} \}$$

$p_k^{(t)}$ is shorthand for estimate of $P(y=k)$ on t'th iteration

E-step

Compute "expected" classes of all datapoints for each class

$$\mathrm{P}\left(Y_j = k \middle| x_j, \lambda_t\right) \propto p_k^{(t)} \mathrm{p}\left(x_j \middle| \mu_k^{(t)}, \Sigma_k^{(t)}\right)$$

Just evaluate a Gaussian at x_j

M-step

Compute weighted MLE for **μ** given expected classes above

$$\mu_k^{(t+1)} = \frac{\sum_j \mathrm{P}\left(Y_j = k \middle| x_j, \lambda_t\right) x_j}{\sum_j \mathrm{P}\left(Y_j = k \middle| x_j, \lambda_t\right)} \qquad \Sigma_k^{(t+1)} = \frac{\sum_j \mathrm{P}\left(Y_j = k \middle| x_j, \lambda_t\right) \left[x_j - \mu_k^{(t+1)}\right]\left[x_j - \mu_k^{(t+1)}\right]^T}{\sum_j \mathrm{P}\left(Y_j = k \middle| x_j, \lambda_t\right)}$$

$$p_k^{(t+1)} = \frac{\sum_j \mathrm{P}\left(Y_j = k \middle| x_j, \lambda_t\right)}{m}$$

m = #training examples