GMM: EM Algorithm

1. Choose **k** random points to be cluster centers (or estimate using k-means...)
2. For each data point, calculate the **probability** of belonging to each cluster
3. Using these probability weights, recalculate the **means + variances** (and weights)
4. Repeat 2 and 3 until **distributions converge**

The M-step in EM Algorithm

Via MLE we get the following estimates:

$$\mu_k = \frac{1}{N_k} \sum_{n=1}^{N} r_{nk} x_n \qquad \Sigma_k = \frac{1}{N_k} \sum_{n=1}^{N} r_{nk} (x_n - \mu_k)(x_n - \mu_k)^T \qquad w_k = \sum_{n=1}^{N} r_{nk}$$

the higher the responsibility of a data point for a cluster is, the more influence it has on what the mean and variance is

Note: $N_k = \sum_{n=1}^{N} r_{nk}$ is now based on soft assignments now

if data points are unlikely to belong to cluster k, the N_k small,
if data points are likely to belong to cluster k, then N_k large