26/9/19

EE5801: Representation Learning

- Review
- EM for GMM
- NN based approaches
  - unsupervised
  - supervised

- Recall that we want to find the optimal parameters of a GMM.

$$\boxed{\mu_k = \frac{1}{N_k} \sum_{n=1}^{N} \gamma(z_{nk}) \cdot x_n}, \text{ where} \quad -(1)$$

$$\gamma(z_{nk}) = \frac{\pi_k \cdot \mathcal{N}(x_n; \mu_k, \Sigma_k)}{\sum_{j=1}^{K} \pi_j \cdot \mathcal{N}(x_n; \mu_j, \Sigma_j)}, \quad -(2)$$

$$N_k = \sum_{n=1}^{N} \gamma(z_{nk}) \quad -(3)$$

- $$\Sigma_k = \frac{1}{N_k} \sum_{n=1}^{N} \gamma(z_{nk}) \cdot (x_n - \mu_k)(x_n - \mu_k)^T \quad -(4)$$

- $$\pi_k = \frac{N_k}{N} \quad -(5)$$

- E-step is to find $\gamma(z_{nk})$ using (2)
- M-step is to find the optimal parameters - (1), (4), (5)

The EM algorithm:

1. Initialize the parameters of the model $\pi_k, \underline{\mu}_k, \Sigma_k$ for all k.
2. E-step: Compute the posterior probabilities $\gamma(z_{nk})$ using current values of the parameters
3. M-step: Update the parameters using $\gamma(z_{nk})$

$$\underline{\mu}_k^{new} = \frac{1}{N_k} \sum_{n=1}^{N} \gamma(z_{nk}) . \underline{x}_n$$

$$\Sigma_k^{new} = \frac{1}{N_k} \sum_{n=1}^{N} \gamma(z_{nk}) . (\underline{x}_n - \underline{\mu}_k^{new}) . (\underline{x}_n - \underline{\mu}_k^{new})^T$$

$$\pi_k = \frac{N_k}{N}$$

4. Compute log likelihood and check for convergence.
   - If not converged, go to step 2.
   - If converged, output parameters