Exercise Sheet 4

19. November 2018

1 Discrete EM: Coin Tosses from Multiple Distributions

The goal is to maximize $Q(\theta, \theta^{old}) = \sum_z P(Z|\chi, \theta^{old}) \cdot logP(\chi, Z|\theta)$
We can define the following terms to simplify the task
Following Bayes' Rule $P(Z|\chi, \theta^{old}) = \frac{P(Z|\theta^{old}) \cdot P(\chi|Z, \theta^{old})}{P(\chi, \theta^{old})}$
$P(Z|\theta^{old}) \cdot P(\chi|Z, \theta^{old}) = \prod_i^N \prod_j^M (\lambda(p_1^{x_j}(1-p_1)^{1-x_j})^{z_i} \cdot ((1-\lambda)(p_2^{x_j}(1-p_2)^{1-x_j})^{1-z_i}$
$P(\chi, \theta^{old}) = \sum_k^N P(Z_k|\theta^{old}) \cdot P(\chi|Z_k, \theta^{old}) = 1$
$P(\chi, Z|\theta) = \prod_i^N \prod_j^M \hat{\lambda}(\hat{p_1}^{x_j}(1-\hat{p_1})^{1-x_j})^{z_i} \cdot ((1-\hat{\lambda})(\hat{p_2}^{x_j}(1-\hat{p_2})^{1-x_j})^{z_i}$
such that
$logP(\chi, Z|\theta) = \sum_i^N \sum_j^M z_i(log\hat{\lambda} + log\hat{p_1}^{x_j} + log(1-\hat{p_1})^{1-x_j}) + (1-z_i)(log(1-\hat{\lambda}) + log\hat{p_2}^{x_j} + log(1-\hat{p_2})^{1-x_j})$
and because
$P(Z|\chi, \theta^{old}) = const.$
the task becomes to find partial extrema of $logP(\chi, Z|\theta) := F$
$\frac{\partial F}{\partial \hat{\lambda}} = \sum_i^N \frac{z_i}{\hat{\lambda}} - \frac{1-z_i}{1-\hat{\lambda}} = 0 \rightarrow \hat{\lambda} = \overline{z}$
$\frac{\partial F}{\partial \hat{p_1}} = \sum_i^N \sum_j^M \frac{z_i \cdot x_j}{\hat{p_1}} - \frac{z_i(1-x_j)}{1-\hat{p_1}} = 0 \rightarrow \hat{p_1} = \overline{x}$ for $z_i = 1$
$\frac{\partial F}{\partial \hat{p_2}} = \sum_i^N \sum_j^M \frac{(1-z_i) \cdot x_j}{\hat{p_2}} - \frac{(1-z_i)(1-x_j)}{1-\hat{p_2}} = 0 \rightarrow \hat{p_2} = \overline{x}$ for $z_i = 0$
concluding in the new parameter vector
$\theta^{new} = (\overline{z}, \overline{x}$ for $z_i = 1, \overline{x}$ for $z_i = 0)^T$