NYU Center for Data Science: DS-GA 1003

Machine Learning and Computational Statistics (Spring 2018)

Brett Bernstein

May 12, 2018

Instructions: Following most lab and lecture sections, we will be providing concept checks for review. Each concept check will:

- List the lab/lecture learning objectives. You will be responsible for mastering these objectives, and demonstrating mastery through homework assignments, exams (midterm and final), and on the final course project.
- Include concept check questions. These questions are intended to reinforce the lab/lectures, and help you master the learning objectives.

You are strongly encourage to complete all concept check questions, and to discuss these (and related) problems on Piazza and at office hours. However, problems marked with a ($\star$) are considered optional.

EM: Concept Check

EM/Mixture Model Objectives

- Be able to set up the GMM problem (multinomial distribution on z, Gaussian $x|z$).
- Be able to give and factorize the joint density $p(x, z)$ for this problem.
- Be able to give pseudocode for the EM Algorithm for GMM (as in slide 29).
- Be able to state the relationship of EM for GMM to k-means.
- Be able to set up problem (in terms of observed Xs, unobserved Zs, and parameter theta).
- Be able to state Jensen's inequality, and define KL divergence (prereqs for ELBO).
- Be able to give EM algorithm pseudocode (as in slide 27).

- Explain how we can compute the two arg maxes needed in this algorithm (i.e. using bound on KL divergence).
- Be able to show that EM gives monotonically increasing likelihood.
- Be able to summarize variations on EM, including generalied EM (address optimization in "M" step), and restriction to a set Q of distributions (addressing optimization in "E" step).

EM Question

Poisson Mixture Model Setup: Consider the poisson mixture model, where each data instance is generated by

1. Drawing an (unobserved cluster) z from a multinomial distribution $(\pi_1, \cdots, \pi_k)$ on k clusters.
2. Drawing a count from a Poisson distribution with PMF:

$$p(x; \lambda_k) = \frac{\lambda_k^x e^{-\lambda_k}}{x!}$$

Problems:

1. Let x, z be the count and cluster assignment for a single instance. Factorize $p(x, z)$.

   Solution.
$$p(x, z) = p(z)p(x|z) = \pi_z \frac{\lambda_z^x e^{-\lambda_z}}{x!}$$

2. For a single data instance, we observe x, and want to know its cluster assignment z. Basic probability review: give an expression for the conditional probability $p(z|x)$ for a single instance (x, z) (just in terms of probability expressions $p(\cdot)$.

   Solution.
$$p(z|x) = \frac{p(x, z)}{p(x)}$$

3. Give an expression for the marginal distribution for a single observed x, $p(x)$ (marginalizing out z), in terms of probability expressions π_k and $p(x; \lambda_k)$.

   Solution.
$$p(x) = \sum_{z=1}^{k} p(x, z) = \sum_{z=1}^{k} \pi_z p(x; \lambda_z)$$

4. Now recall the EM algorithm. In the "E step", we evaluate the responsibilities $\gamma_i^j = p(z = j|x_i)$ for each $j \in \{1, \cdots, k\}$. Give an expression for this responsibility for cluster j and instance i.

Solution.

$$\gamma_i^j = p(z=j|x_i) = \frac{p(x_i;\lambda_j)}{\sum_{z=1}^k \pi_z p(x_i;\lambda_z)} = \frac{\pi_z \frac{\lambda_z^{x_i} e^{-\lambda_z}}{x_i!}}{\sum_{z=1}^k \pi_z \frac{\lambda_z^{x_i} e^{-\lambda_z}}{x_i!}}$$

5. In the "M step", we will update our MLE estimates for π_z and λ_z. Give an expression for π_z^{new}

   Solution.

   $$\pi_z^{new} = \frac{n_z}{n} = \frac{\sum_{i=1}^n \gamma_i^z}{n}$$

   where z_i is the hard cluster assignment.

6. Give an expression for λ_z^{new}. Recall the MLE for a Poisson $\hat{\lambda}_{MLE} = \bar{x}$.

   Solution.

   $$\lambda_z^{new} = \frac{1}{n_z} \sum_{i=1}^n \gamma_i^z x_i$$

7. Let's apply the distributions we just described for the "E step" of a toy problem. Imagine $k=3$, and we have $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 3$. Find $p(z=2|x=1)$ in terms of π_i for i in $\{1,2,3\}$. Hint: Note $p(x)$ is constant for all k, so its straightforward to give proportional expressions for each of $p(z=k|x=1)$ then normalize.

   Solution.

   $$\begin{aligned} p(z=1|x=1) &\propto p(x=1|z=1)p(z=1) = \pi_1 e^{-1} \\ P(z=2|x=1) &\propto p(x=1|z=2)p(z=2) = \pi_2 2e^{-2} \\ P(z=3|x=1) &\propto p(x=1|z=3)p(z=3) = \pi_3 3e^{-3} \end{aligned}$$

   $$P(z=2|X=1) = \frac{\pi_2 2e^{-2}}{\pi_1 e^{-1} + \pi_2 2e^{-2} + \pi_3 3e^{-3}}$$