Deep Probabilistic Programming - Week 2

Ahmad Salim Al-Sibahi

August 16, 2020

1 Kidney Cancer Model

Bayesian Data Analysis [1] presents in Section 2.7, an example Gamma-Poission model for modelling kidney cancer rate in the U.S. around the 1980's. The data consists of observed death count y_j and population size n_j for each U.S. county j. These data points are presumed to be generated according to the following model:

$$\begin{aligned} \theta_j &\sim \text{Gamma}(\alpha = 20., \beta = 430,000.) \\ y_j &\sim \text{Poisson}(\lambda = 10 n_j \theta_j) \end{aligned}$$

The goal of this exercise is to infer a posterior of the shape:

$$\theta_j | y_j \sim \text{Gamma}(\alpha = \alpha_j, \beta = \beta_j)$$

This should be done using Stochastic Variational Inference in Pyro. There are two ways one could approach this problem, either:

1. Directly, where one declares two parameters for each county α_j, and β_j and optimize them,
2. or, by amortization where one uses observed data to compute the local parameters from global parameters. For this model, we could rewrite the local parameters $\alpha_j = w y_j + k$ and $\beta_j = v n_j + c$ as deterministic calculations from global parameters w, k, v and c. Now, we only need to optimize four parameters instead of $2j$, making inference much more scalable.

Implement both approaches and compare run-time performance and accuracy. Are there any other advantage to amortization than just run-time speedup?

References

[1] Gelman, A., et al. Bayesian Data Analysis, 3 ed. Chapman and Hall/CRC, 2013.