

Dirichlet Process Mixtures

Variational Inference

Jessy Lin, Uma Roy, Ajinkya Ghorpade

6.882 Spring 2018

Table of contents

1. Demo
2. Empirical Comparison

Demo

Empirical Comparison

Gibbs Sampler Convergence Diagnostics

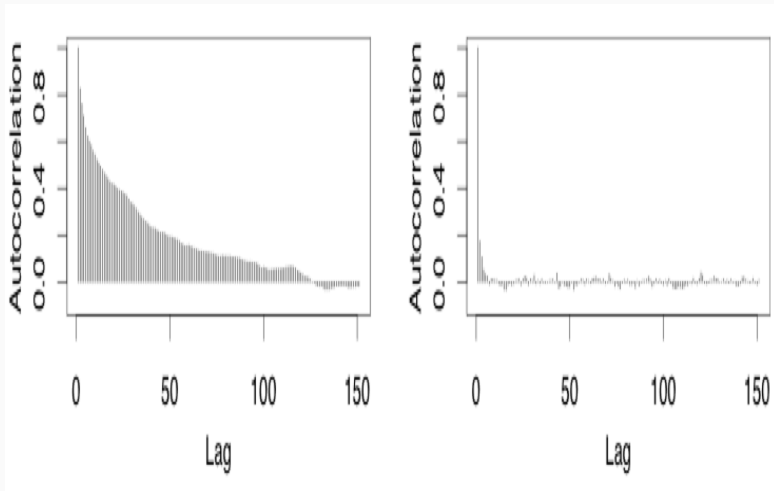


Figure 1: Autocorrelation plots on the size of the largest component for the truncated DP Gibbs sampler (left) and collapsed Gibbs sampler (right) in an example dataset of 50-dimensional Gaussian data.

Variational Inference Convergence

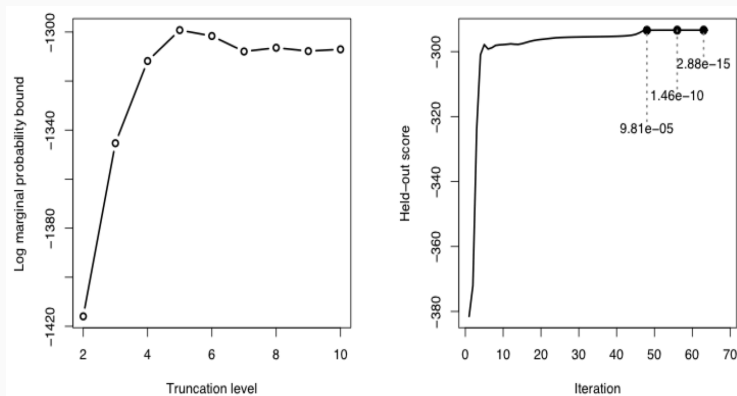


Figure 2: The optimal bound on the log probability as a function of the truncation level (left). There are five clusters in the simulated 20-dimensional DP mixture of Gaussian data set which was used. Held-out probability as a function of iteration of variational inference for the same simulated data set (right). The relative change in the log probability bound of the observations is labeled at selected iterations.

Convergence Time Comparison

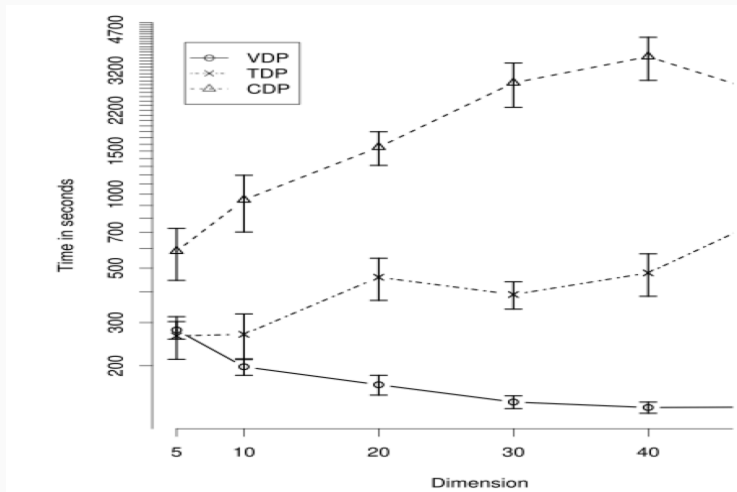


Figure 3: Mean convergence time and standard error across ten datasets per dimension for variational inference (VDP), TDP Gibbs sampling, and the collapsed Gibbs sampler (CDP).

Accuracy Comparison

Dim	Mean held out log probability (Std err)		
	Variational	Collapsed Gibbs	Truncated Gibbs
5	-147.96 (4.12)	-148.08 (3.93)	-147.93 (3.88)
10	-266.59 (7.69)	-266.29 (7.64)	-265.89 (7.66)
20	-494.12 (7.31)	-492.32 (7.54)	-491.96 (7.59)
30	-721.55 (8.18)	-720.05 (7.92)	-720.02 (7.96)
40	-943.39 (10.65)	-941.04 (10.15)	-940.71 (10.23)
50	-1151.01 (15.23)	-1148.51 (14.78)	-1147.48 (14.55)

Table 1: Average held-out log probability for the predictive distributions given by variational inference, TDP Gibbs sampling, and the collapsed Gibbs sampler.

Image Analysis

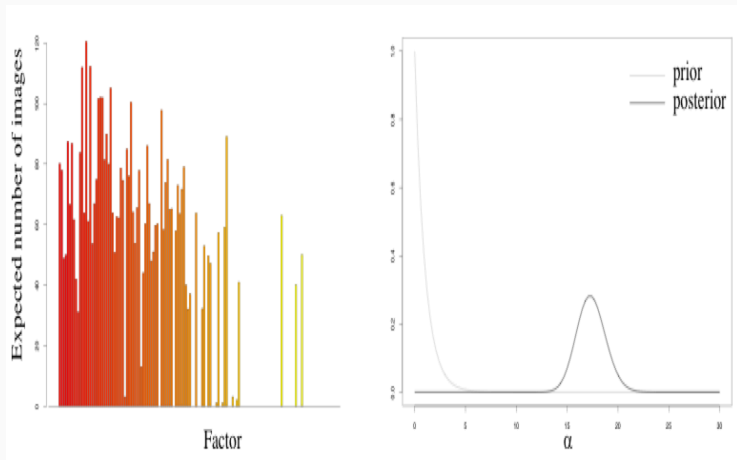


Figure 4: The expected number of images allocated to each component in the variational posterior (left). The posterior uses 79 components to describe the data. The prior for the scaling parameter α and the approximate posterior given by its variational distribution (right)

Image Analysis

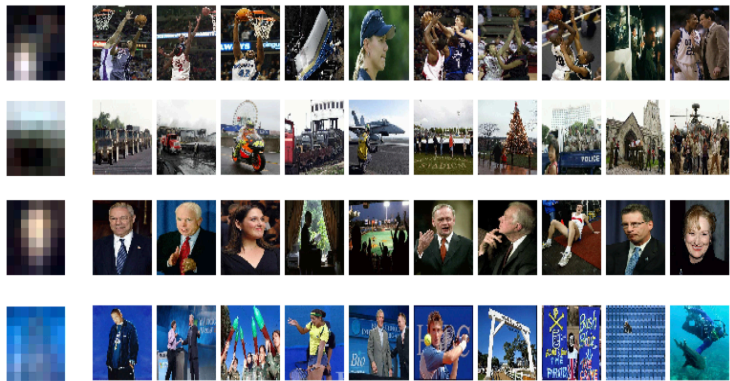


Figure 5: Four sample clusters from a DP mixture analysis of 5000 images from the Associated Press. The left-most column is the posterior mean of each cluster followed by the top ten images associated with it. These clusters capture patterns in the data, such as basketball shots, outdoor scenes on gray days, faces, and pictures with blue backgrounds.