Machine Learning in Complex Domains: Assignment 3

Ryan Cotterell, Daniel Deutsch, Dan Crankshaw

4 Blocked Gibbs Sampler

4.1 Analysis Questions

We have the following equation

$$p(z_{d,i} = k, x_{d,i} = c | \mathbf{z} - z_{d,i}, \mathbf{x} - x_{d,i}, \mathbf{c}, \mathbf{w}; \alpha, \beta, \lambda) = \frac{p(\mathbf{z}, \mathbf{x}, \mathbf{c}, \mathbf{w} | \alpha, \beta, \lambda)}{p(\mathbf{z} - z_{d,i}, \mathbf{x} - x_{d,i}, \mathbf{c}, \mathbf{w}; \alpha, \beta, \lambda)}$$

The full join (numerator) is given to us in the assignment. We can factor this according to the joint and every term that does not involve $z_{d,i}$ or $x_{d,i}$ will cancel. This is basically a union of (15) and (18). This yields the following:

$$p(z_{d,i} = k, x_{d,i} = 0 | \mathbf{z} - z_{d,i}, \mathbf{x} - x_{d,i}, \mathbf{c}, \mathbf{w}; \alpha, \beta, \lambda) = \frac{p(x = 0|\lambda) \frac{\Gamma(n^k_{w_{d,i}} + 1 + \beta)}{\Gamma(1 + \sum_w n^k_w + \beta)} \frac{\Gamma(n^d_k + 1 + \alpha)}{\Gamma(1 + \sum_{k'} n^d_{k'} + \alpha)}}{\frac{\Gamma(n^k_{w_{d,i}} + \beta)}{\Gamma(\sum_w n^k_w + \beta)} \frac{\Gamma(n^d_k + \alpha)}{\Gamma(\sum_{k'} n^d_{k'} + \alpha)}}$$

$$p(z_{d,i} = k, x_{d,i} = 0 | \mathbf{z} - z_{d,i}, \mathbf{x} - x_{d,i}, \mathbf{c}, \mathbf{w}; \alpha, \beta, \lambda) \propto \frac{(1 - \lambda)(n^k_{w_{d,i}} + \beta)(n^d_k + \alpha)}{(n^k_* + K\beta)(n^d_* + V\alpha)}$$

By analogy, we get the followoing for the case that $p(x = 1)$.

$$p(z_{d,i} = k, x_{d,i} = 1 | \mathbf{z} - z_{d,i}, \mathbf{x} - x_{d,i}, \mathbf{c}, \mathbf{w}; \alpha, \beta, \lambda) \propto \frac{\lambda(n^k_{w_{d,i}} + \beta)(n^d_k + \alpha)}{(n^k_* + K\beta)(n^d_* + V\alpha)}$$

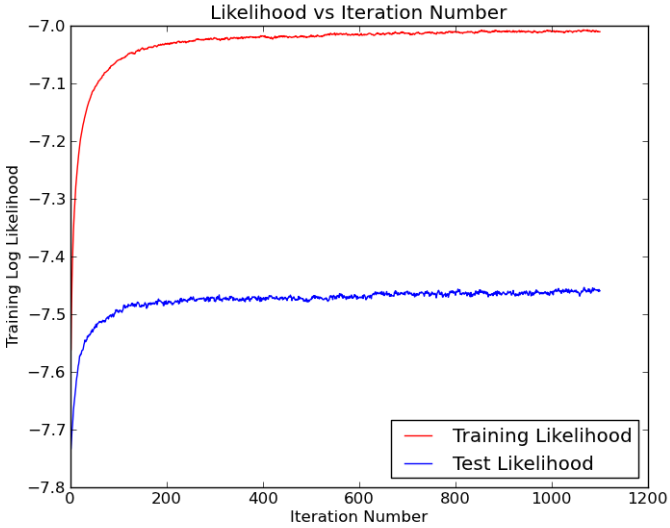
Most of the heavy lifting for this question was done in the appendix so there isn't really any more we can write.

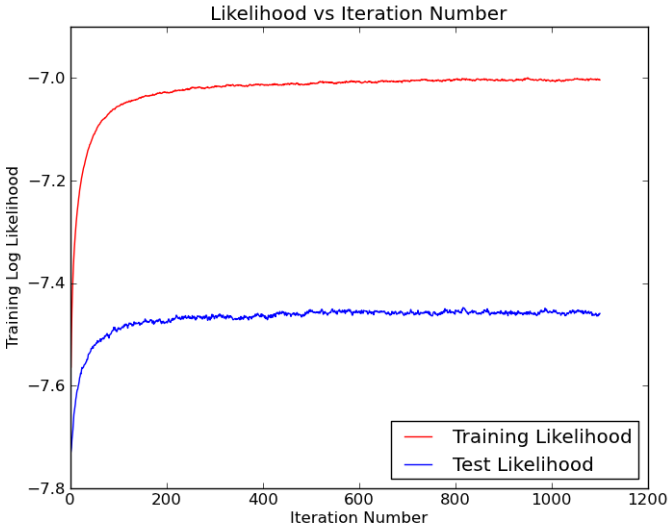
5 Text Analysis with MCLDA

5.1 Empirical Questions

6. (a)
   (b)
   (c)

1.



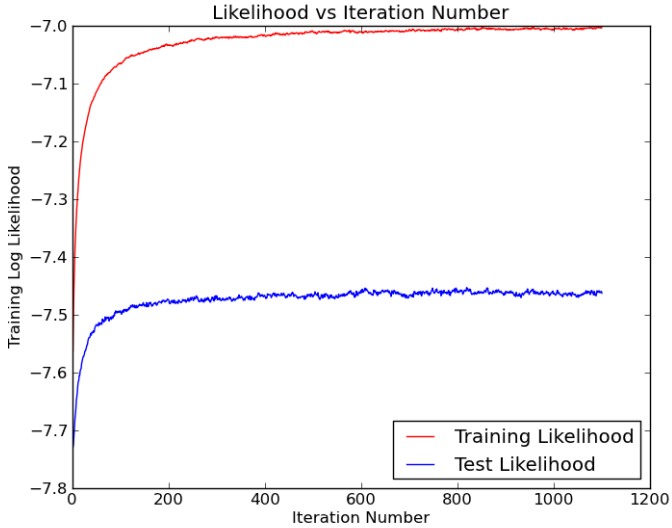


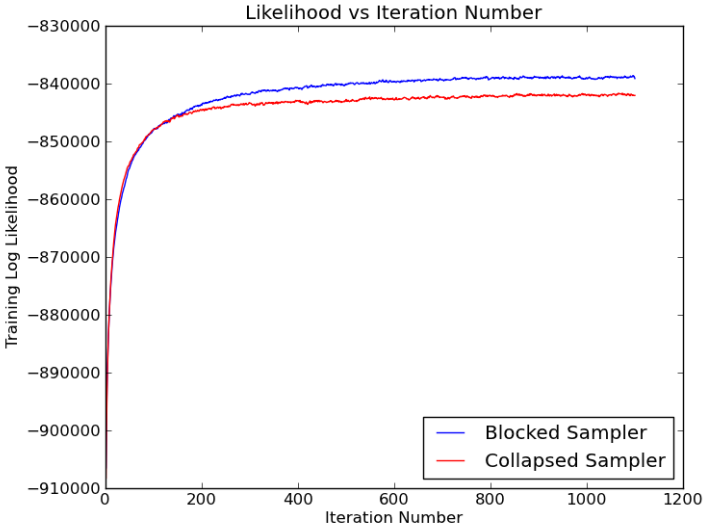
6 Variational Inference

6.1 Analysis Questions

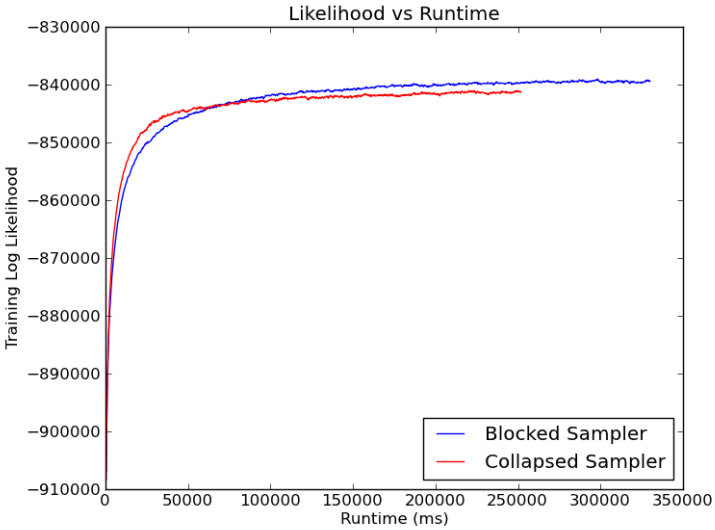
6.3 Empirical Questions

1.
2.
3.
4.


Likelihood vs Iteration Number
Training Log Likelihood
Iteration Number
Training Likelihood
Test Likelihood

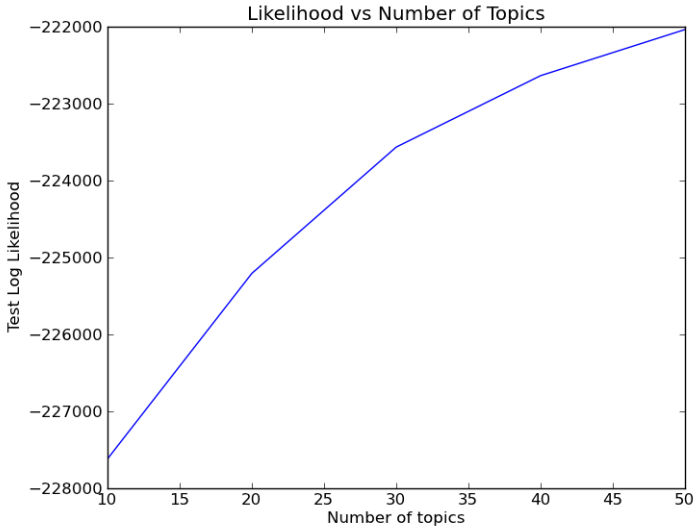

Likelihood vs Iteration Number
Training Log Likelihood
Iteration Number
Blocked Sampler
Collapsed Sampler

2.

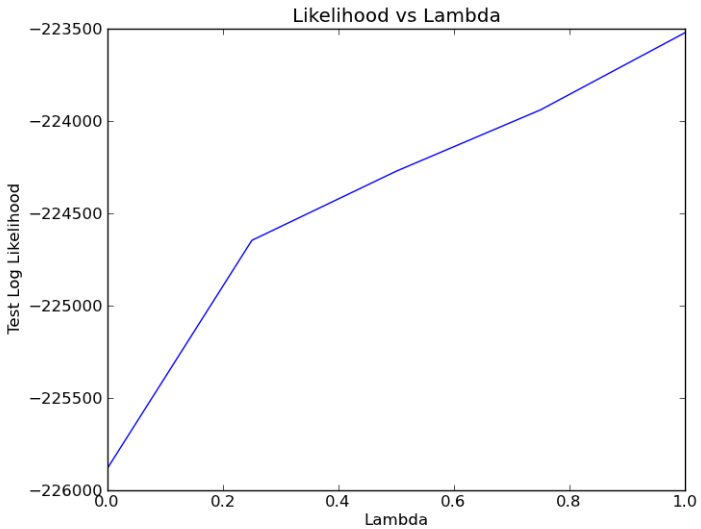

Likelihood vs Runtime
Training Log Likelihood
Runtime (ms)
Blocked Sampler
Collapsed Sampler

3.

4.


Likelihood vs Number of Topics
Test Log Likelihood
−222000
−223000
−224000
−225000
−226000
−227000
−228000
10
15
20
25
30
35
40
45
50
Number of topics

5.


Likelihood vs Lambda
Test Log Likelihood
−223500
−224000
−224500
−225000
−225500
−226000
0.0
0.2
0.4
0.6
0.8
1.0
Lambda