Stochastic Gradient Descent Hamiltonian Monte Carlo Applied to Bayesian Logistic Regression

Sta663 Final Project

Gilad Amitai and Beau Coker

Abstract

Hamiltonian Monte Carlo (HMC) is a Markov chain Monte Carlo algorithm for drawing samples from a probability distribution where proposed values are computed using Hamiltonian dynamics to find values of high acceptance probabilities. They allow us to explore sample states more efficiently than random walk proposals, but are limited by the expensive computation of the gradient of the potential energy function. Chen, Fox, and Guestrin propose the method Stochastic Gradient Hamiltonian Monte Carlo (SGHMC), a HMC algorithm that uses a subset of the data to compute the gradient. The authors find that the stochastic gradient is noisy and correct this with a friction term.

In this project, we adapt the SGHMC to be used for Bayesian Logistic regression, implement this method in Python, optimize the code for computational efficiency, validate our approach using simulated data, and apply the algorithm to real world classification problems.

1 Background

Hamiltonian Monte Carlo

2 Description of Algorithm

Stochastic Gradient Hamiltonian Monte Carlo proposes using a subset $\tilde{\mathcal{D}}$ of the entire dataset $\mathcal{D}$ to compute the stochastic gradient

$$\Delta\tilde{U}(\theta) = -\frac{|\mathcal{D}|}{|\tilde{\mathcal{D}}|}\sum_{x\in\tilde{\mathcal{D}}} \Delta \log p(x|\theta) - \Delta \log p(\theta)$$

which can then be used in the Hamiltonian Monte Carlo equations in the stead of the gradient $\Delta U(\theta)$.

$$\Pr\left(y_i = 1|\boldsymbol{x}_i, \boldsymbol{\beta}\right) = \frac{\exp\left\{\boldsymbol{x}_i^T\boldsymbol{\beta}\right\}}{1+\exp\left\{\boldsymbol{x}_i^T\boldsymbol{\beta}\right\}}$$

In the Bayesian framework, with prior $P(\boldsymbol{\beta})$, the posterior distribution will be proportional to $P(\boldsymbol{\beta})\prod_{i=1}^{n}\Pr\left(y_i|\boldsymbol{x}_i\right.$

3 Optimization

blah blah blah

4 Application to Simulated Data

blah blah blah

5 Application to Real Data

blah blah blah

6 Comparative Analysis

blah blah blah

7 Discussion and Conclusion

blah blah blah