What makes a good sampler?

A BOB Challenge!

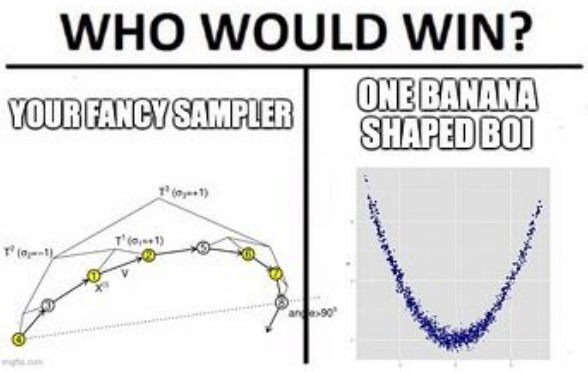
Ingredients / Teams / Directories to work in

1. The problem set (benchmarks)?

2. How do we compare methods? (asymptotically exact vs biased, parallel computation). How can we visualise/compare the efficiency?

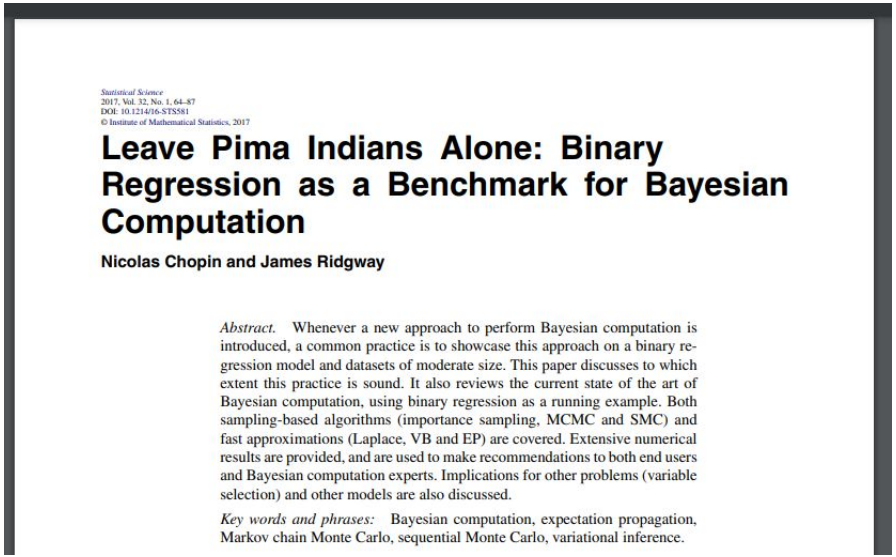
Benchmarks

mschauer



stan-dev/
posteriordb-r

2 Contributors · 1 Issue · 7 Stars · 2 Forks



Statistical Science
2017, Vol. 32, No. 1, 64–87
DOI: 10.1214/16-STS581
© Institute of Mathematical Statistics, 2017

Leave Pima Indians Alone: Binary Regression as a Benchmark for Bayesian Computation

Nicolas Chopin and James Ridgway

Abstract. Whenever a new approach to perform Bayesian computation is introduced, a common practice is to showcase this approach on a binary regression model and datasets of moderate size. This paper discusses to which extent this practice is sound. It also reviews the current state of the art of Bayesian computation, using binary regression as a running example. Both sampling-based algorithms (importance sampling, MCMC and SMC) and fast approximations (Laplace, VB and EP) are covered. Extensive numerical results are provided, and are used to make recommendations to both end users and Bayesian computation experts. Implications for other problems (variable selection) and other models are also discussed.

Key words and phrases: Bayesian computation, expectation propagation, Markov chain Monte Carlo, sequential Monte Carlo, variational inference.

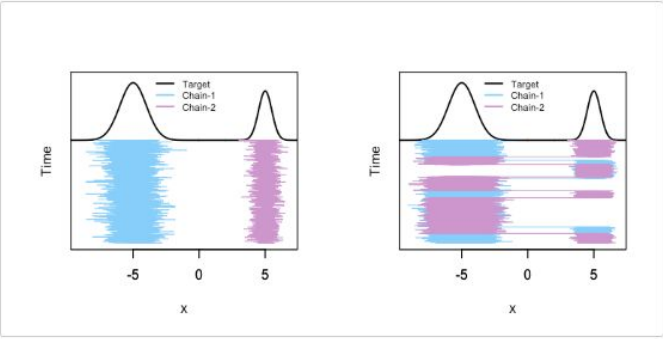
Metrics (comparing samplers)

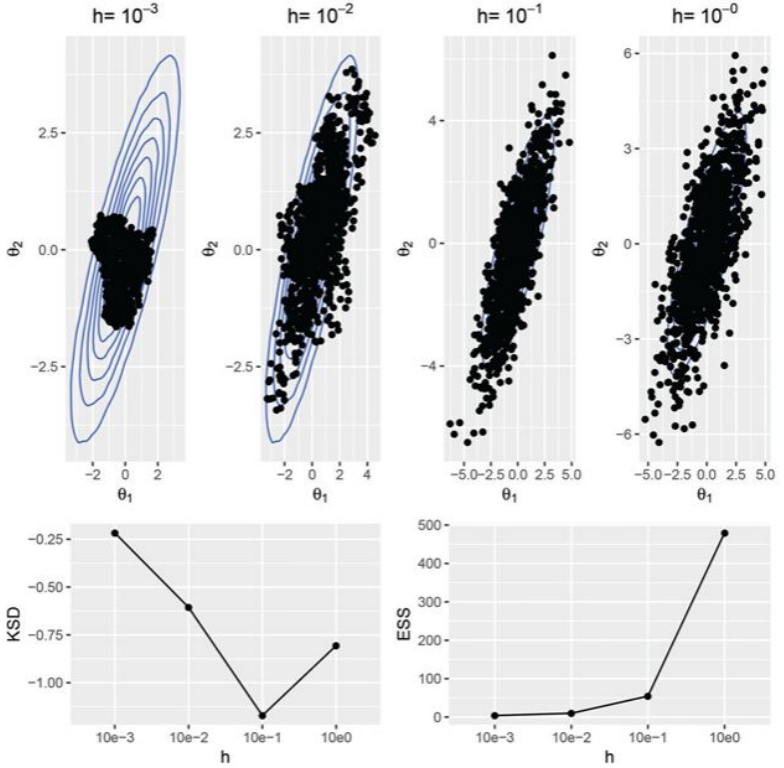
Effective Sample Size

The effective sample size of N samples generated by a process with autocorrelations ρ_t is defined by

$$N_{\text{eff}} = \frac{N}{\sum_{t=-\infty}^{\infty} \rho_t} = \frac{N}{1 + 2\sum_{t=1}^{\infty} \rho_t}.$$

multichainACF (Agarwal, Vats)





Kernel Stein Discrepancy (Gorham, Mackey)

Plots / Animations

1. plot(comp_cost vs metric)
2. Scaling dimension
3. Visualising behaviour (trajectories)