Optimization-Based Sampler RandomizeThenOptimize.jl

Zheng Wang

Massachusetts Institute of Technology
Department of Aeronautics and Astronautics

December 15, 2016

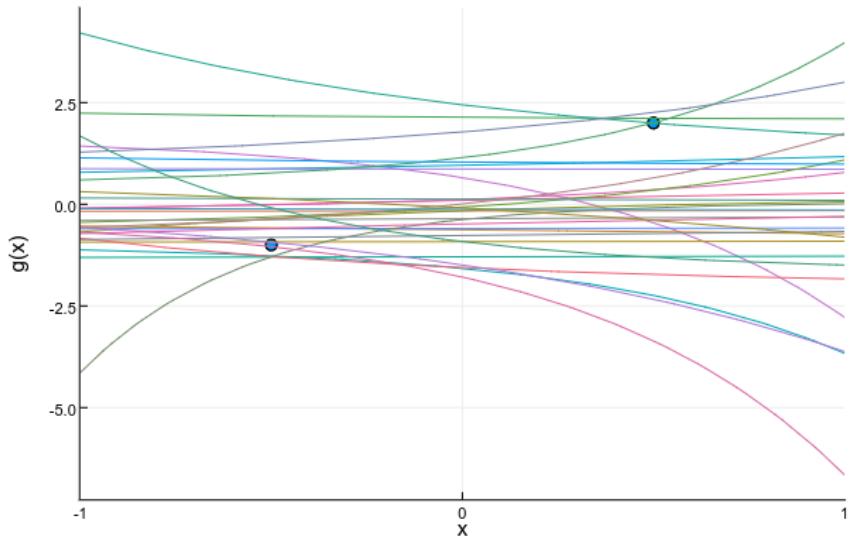
6.338 Final Project

Motivation

Consider the model with parameters θ

$$g(x;\theta) = \theta_1 + \theta_2\, e^{\theta_3 x}$$

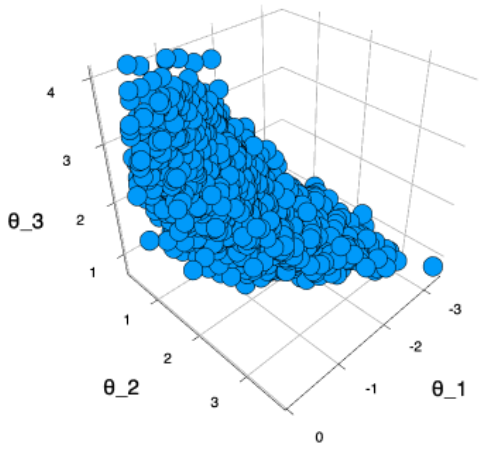
Given noisy measurements y

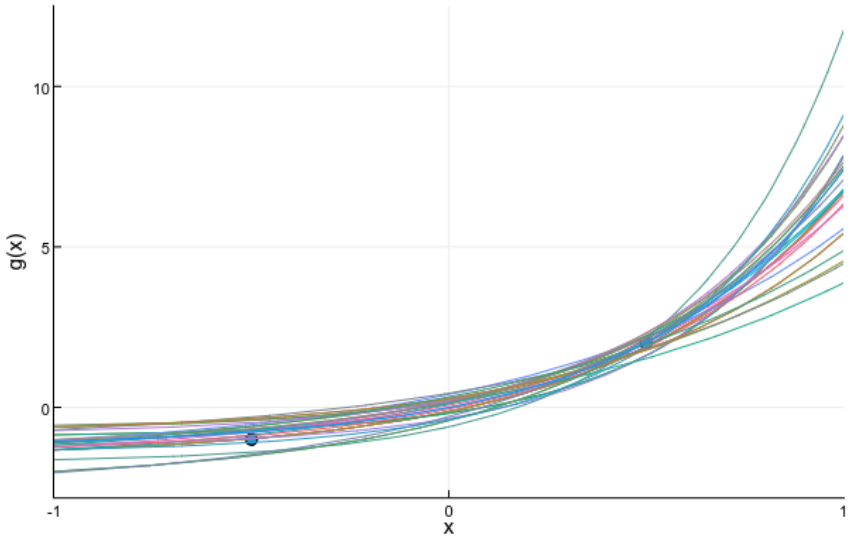


What can we say about θ?

Motivation

Describe our knowledge on θ using a distribution



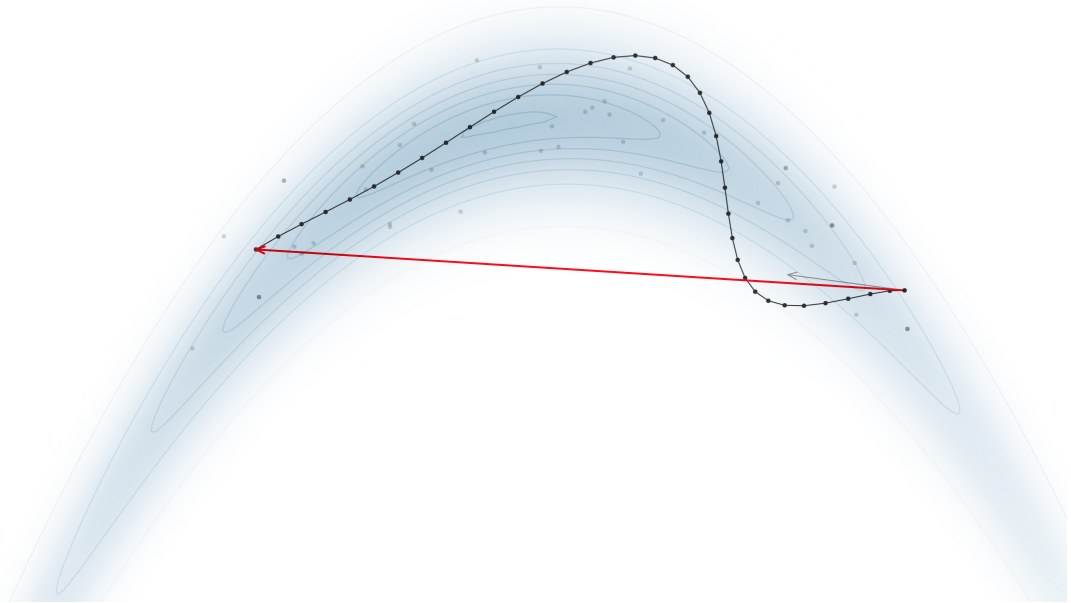


Numerical task: sample from distribution

Sampling algorithms

Markov chain Monte Carlo

- Random-walk Metropolis – "drunken walk"
- Hamiltonian Monte-Carlo (HMC) – "flicking a marble"



▸ MCMC Demo

Randomize-then-Optimize Julia Package

RandomizeThenOptimize.jl

- Implements Randomize-then-Optimize (RTO)
- Uses `NLopt` for optimization
- Available on GitHub

▸ GitHub Link

▸ IJulia Notebook

Experiments

Experiment 1: $\mathbb{R}^3 \to \mathbb{R}^2$

$$f(\theta) = \begin{bmatrix} \theta_1 + \theta_2\, e^{\theta_3 x_1} \\ \theta_1 + \theta_2\, e^{\theta_3 x_2} \end{bmatrix}$$

Experiment 2: $\mathbb{R}^2 \to \mathbb{R}$

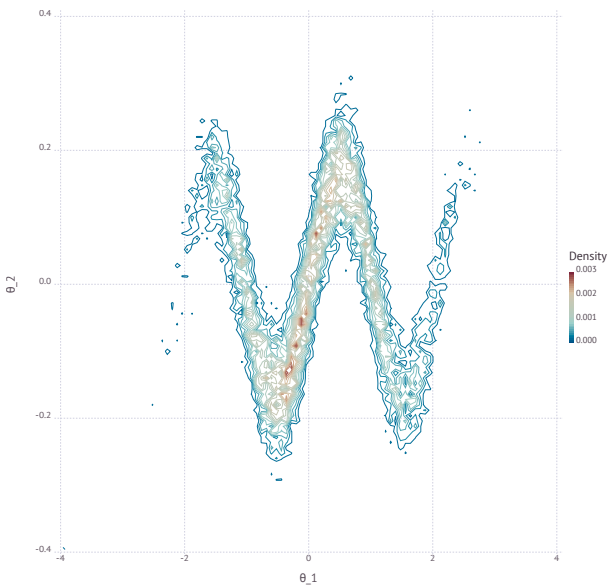
$$f(\theta) = a \sin(b\,\theta_1) - c\,\theta_2$$

Experiment 3: $\mathbb{R}^{30} \to \mathbb{R}^{20}$

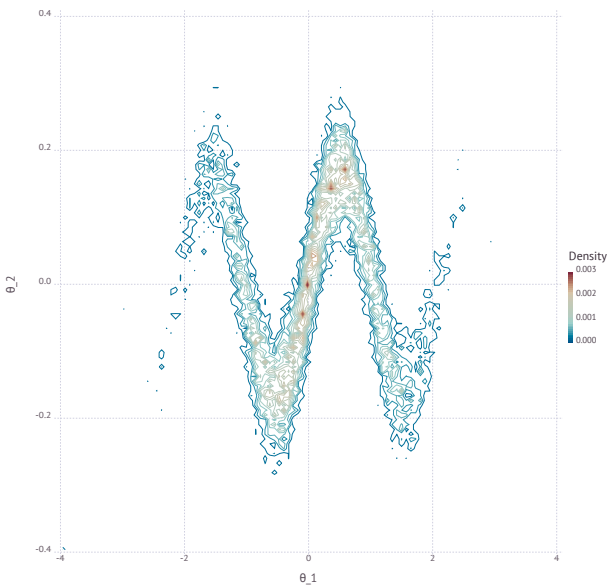
$$f(\theta) = A\,\theta$$

Sinusoid Distributions

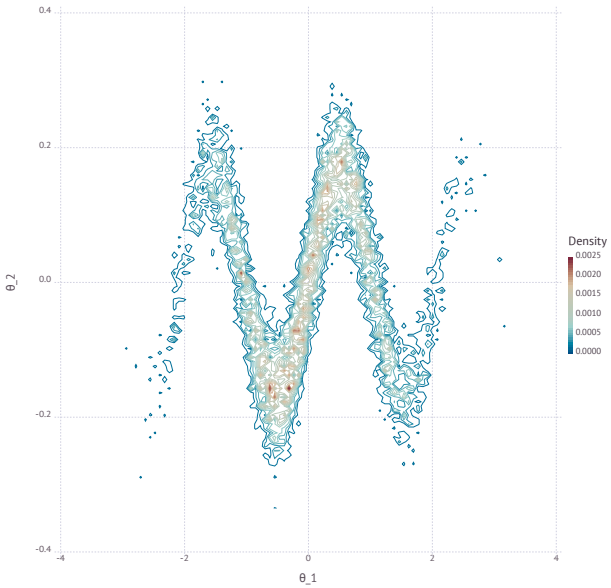
Sanity check: all samplers give the same distribution



MALA



HMC



RTO

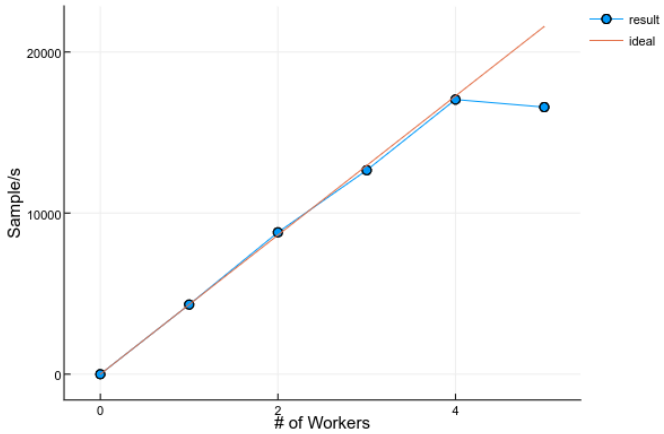
Sampling Efficiency

Effective sample size (ESS) measures quality of an MCMC chain

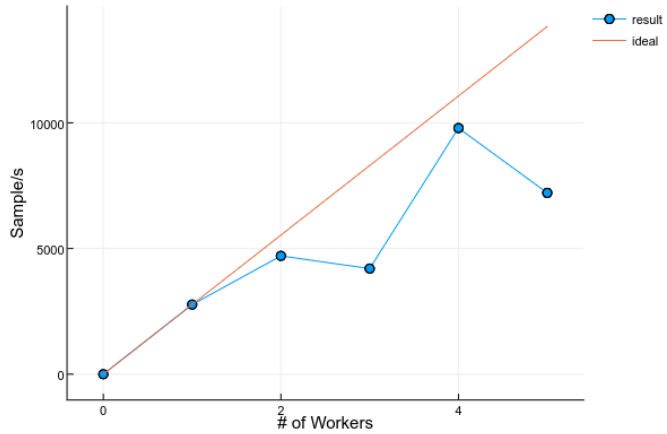
ESS per second

	MALA	HMC	RTO
Experiment 1	900	525	1531
Experiment 2	452	199	1670
Experiment 3	230	212	241

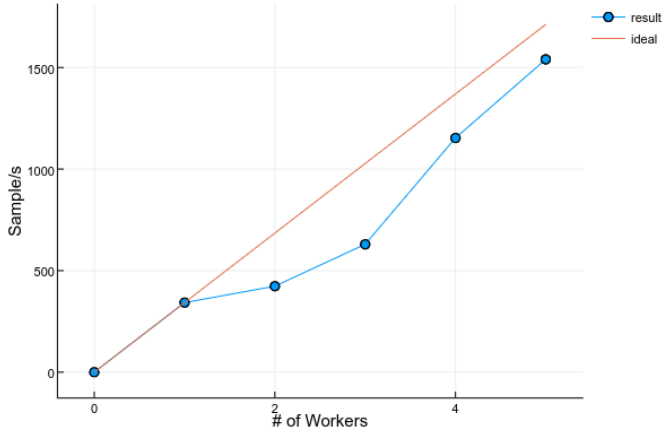
Parallel Scaling



Experiment 1



Experiment 2



Experiment 3