emcee: An Affine-Invariant Sampler

Dustin Lang
Perimeter Institute for Theoretical Physics

Symmetries Graduate School 2023-01-30

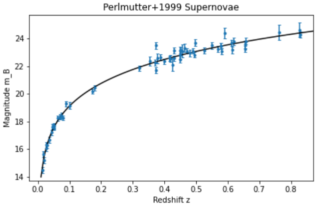
Borrowing heavily from Dan Foreman-Mackey's slides
https://speakerdeck.com/dfm/data-analysis-with-mcmc1
These slides are available at

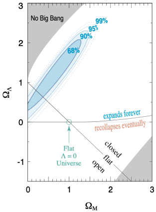
https://github.com/dstndstn/MCMC-talk/emcee-slides

Recap from last week's lecture (1)

- Markov Chain Monte Carlo (MCMC) *draws samples from a probability distribution* when you can *numerically evaluate* the probability function (up to a constant)
- Used extensively in data analysis: *inferring* parameters of models, given observed data
- *Usually* in a Bayesian context; the probability function we run MCMC on is the *posterior* probability:

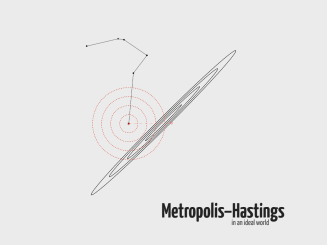
$$\text{posterior}(\text{params}|\text{data}) \propto$$
$$\quad \text{prior}(\text{params}) \times \text{likelihood}(\text{data}|\text{params})$$



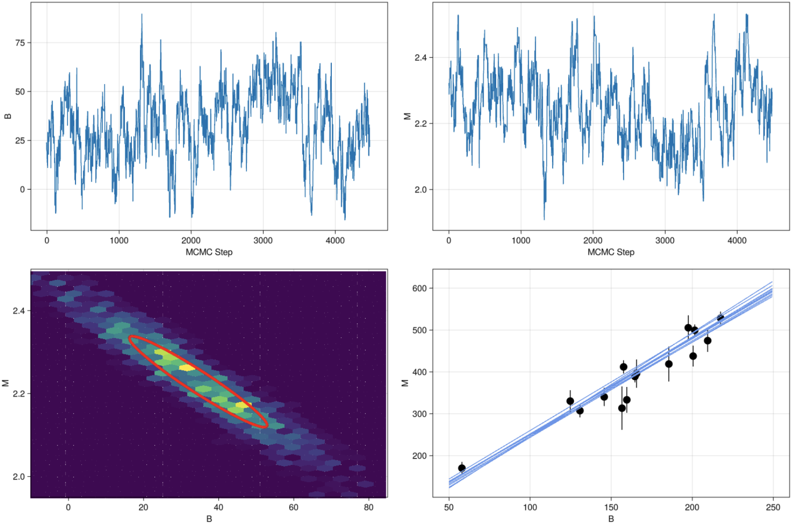


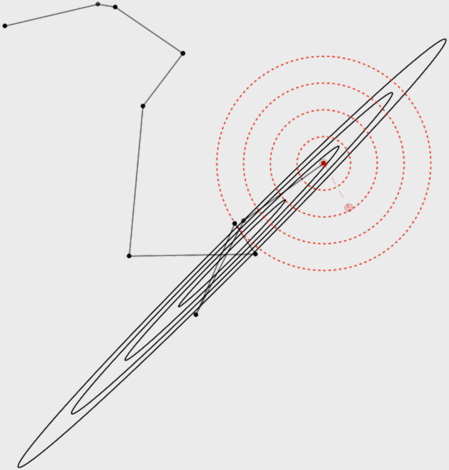
Recap from last week's lecture (2)

- The "classic" Markov Chain Monte Carlo algorithm is *Metropolis–Hastings*, which moves a *walker* or *particle* around the *state space* (*model parameter space*)
- A randomly-drawn *proposed* jump gets *evaluated* (by calling the probability function), and then *accepted*, or not
- A big difficulty is to *customize* the *proposal distribution* to get the algorithm to work efficiently

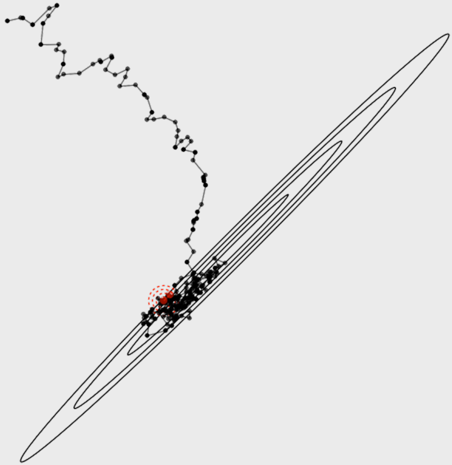


MCMC for model parameter inference





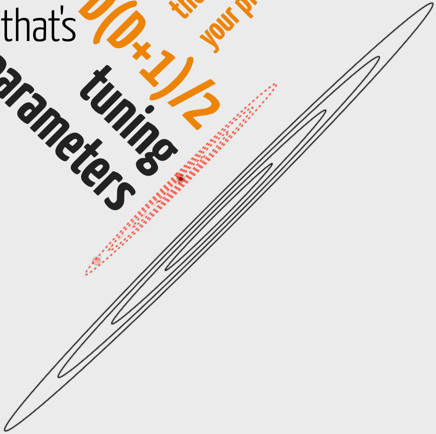
Metropolis–Hastings
in the real world



Metropolis–Hastings
in the real world

that's D(D+1)/2 tuning parameters

the dimension of your problem

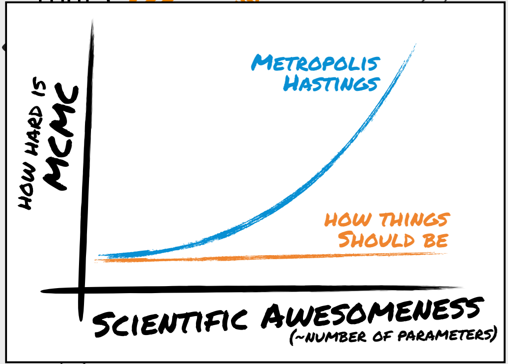


Metropolis–Hastings

in the real world

that's

the dimension of
your problem



Metropolis–Hastings
in the real world

Jonathan Goodman

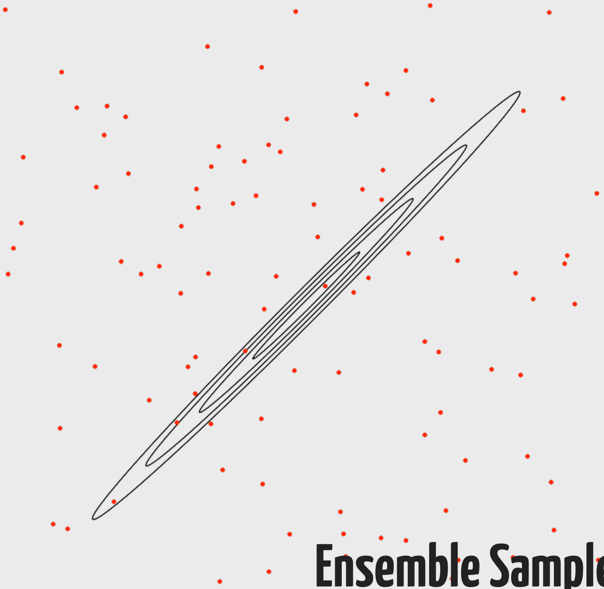
Jonathan Weare

"Ensemble samplers with affine invariance"

(dfm.io/mcmc-gw10)

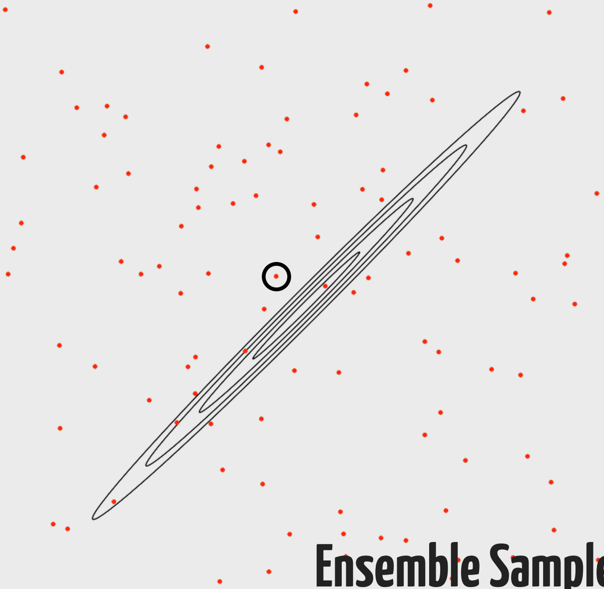
introducing **emcee** the MCMC Hammer

arxiv.org/abs/1202.3665
dan.iel.fm/emcee

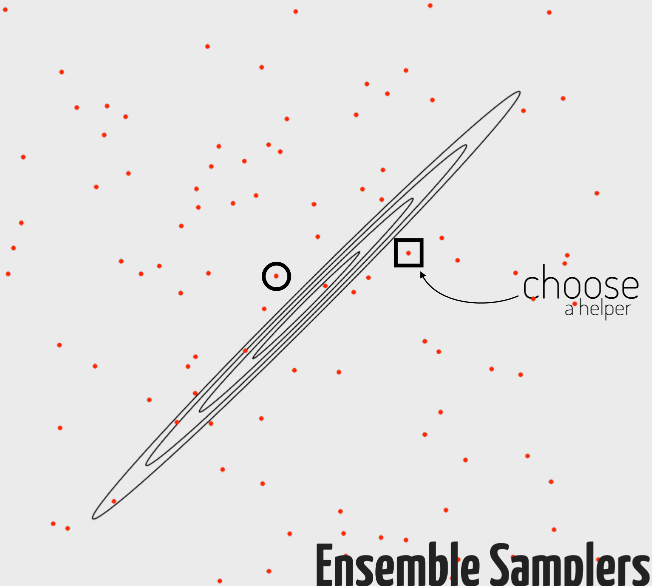


Ensemble Samplers

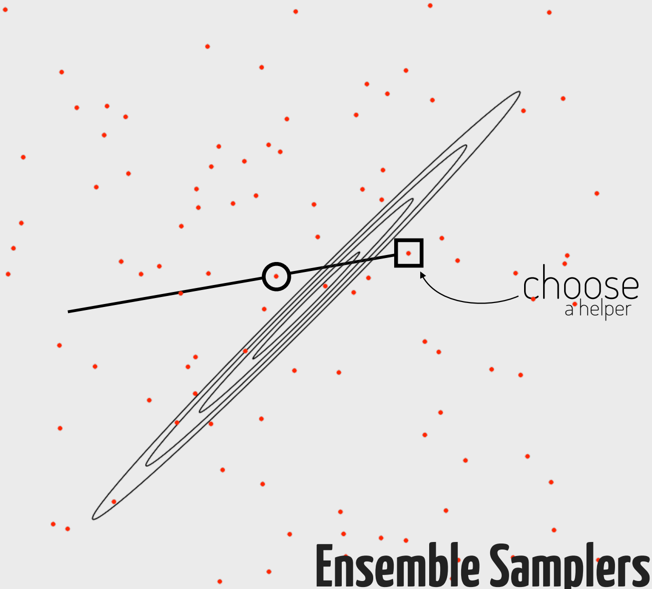
in the **real** world



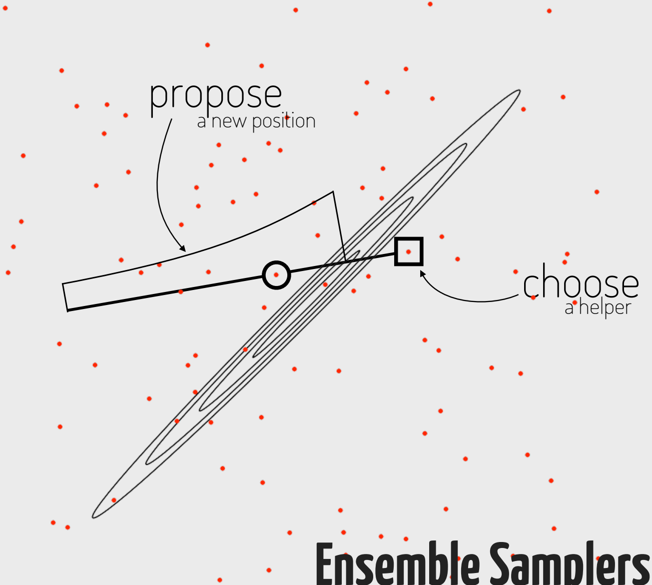
Ensemble Samplers
in the **real** world



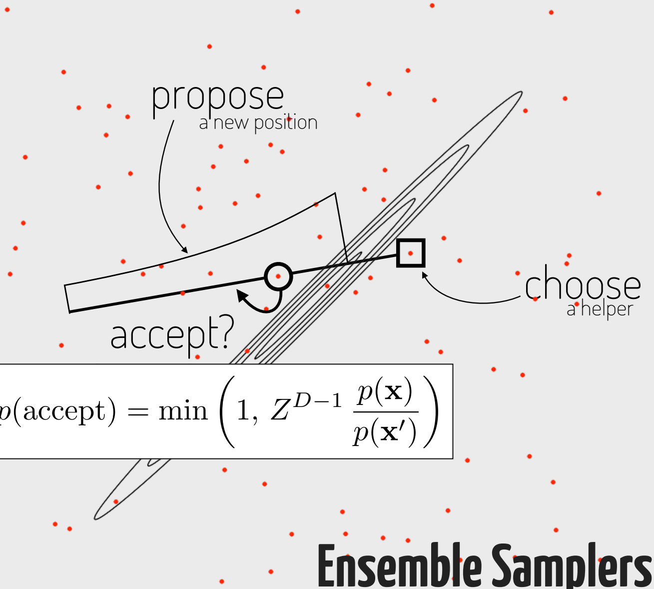
Ensemble Samplers
in the **real** world



Ensemble Samplers
in the **real** world

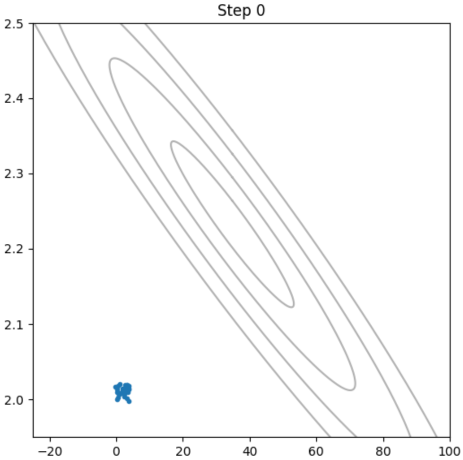


Ensemble Samplers
in the **real** world

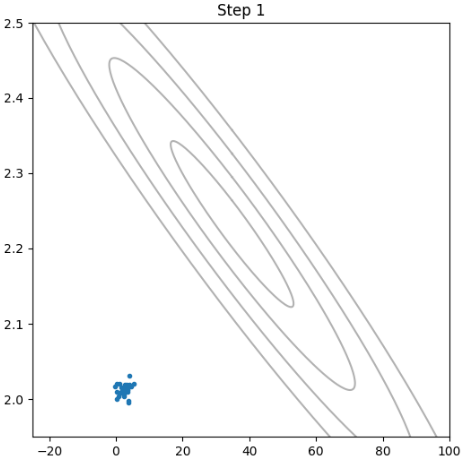


$$p(\text{accept}) = \min\left(1,\ Z^{D-1}\,\frac{p(\mathbf{x})}{p(\mathbf{x}')}\right)$$

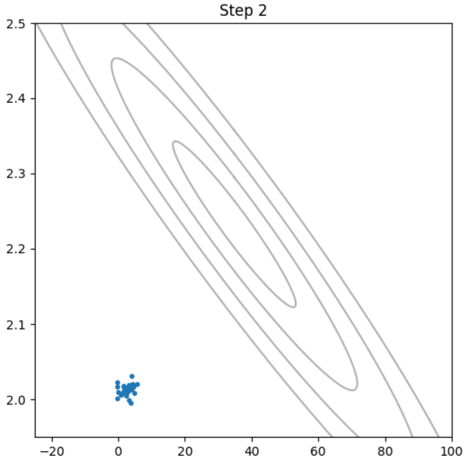
Emcee demo



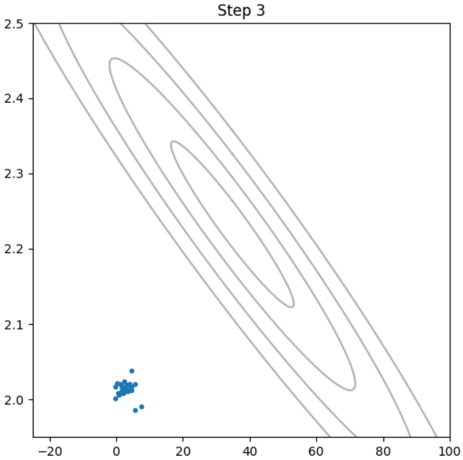
Emcee demo



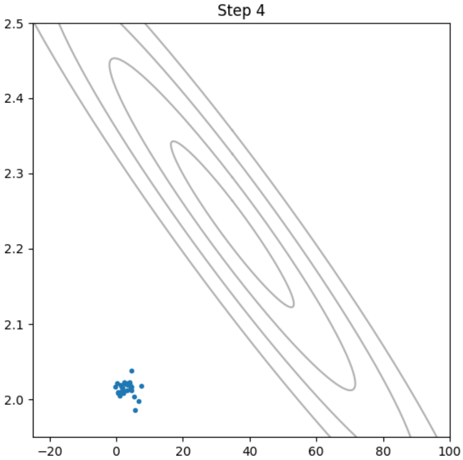
Emcee demo



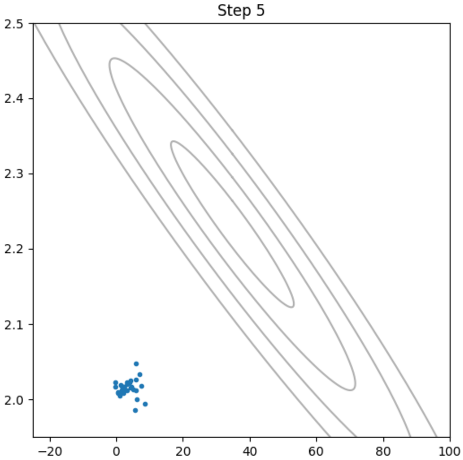
Emcee demo



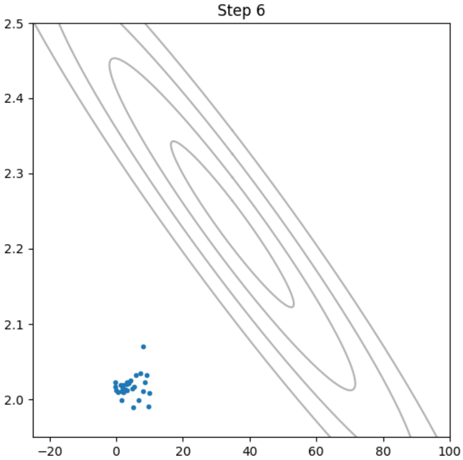
Emcee demo



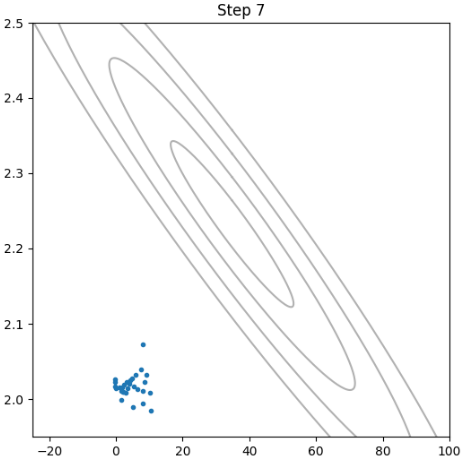
Emcee demo



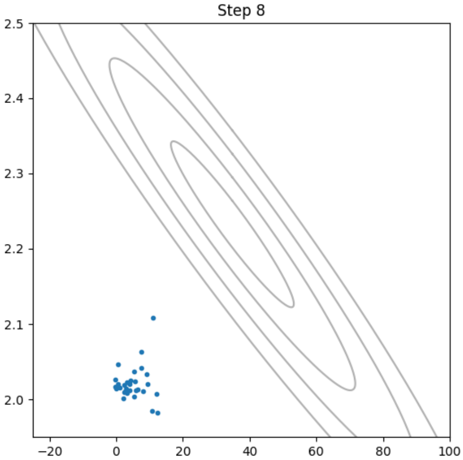
Emcee demo



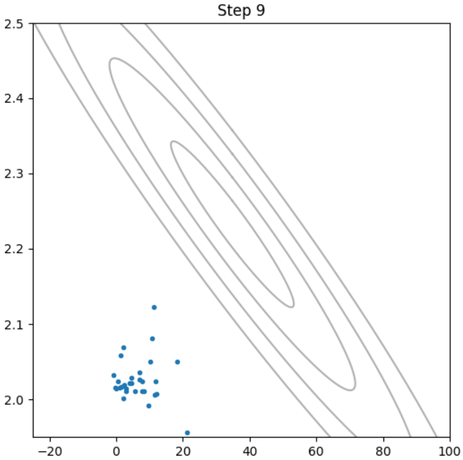
Emcee demo



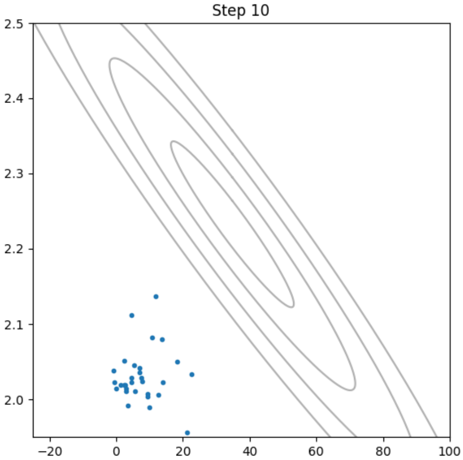
Emcee demo



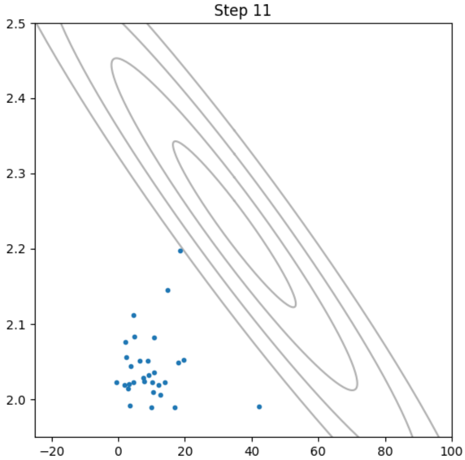
Emcee demo



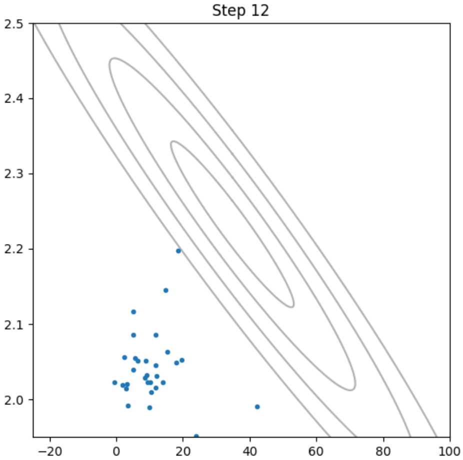
Emcee demo



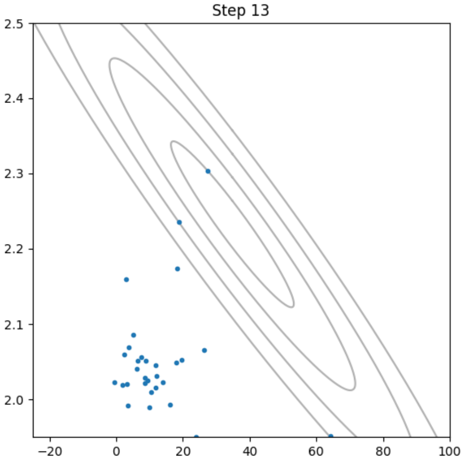
Emcee demo



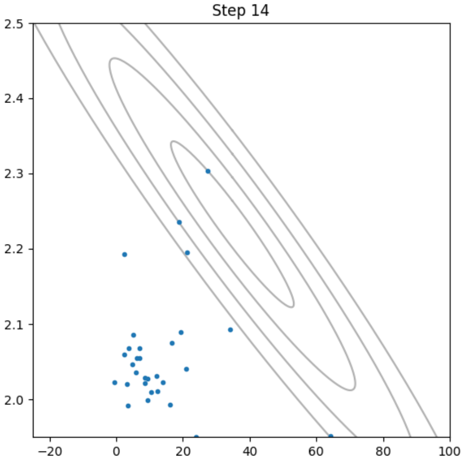
Emcee demo



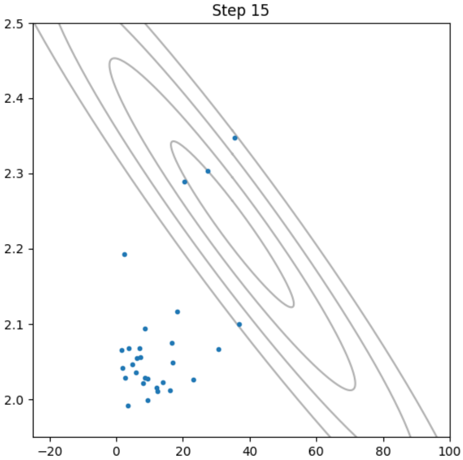
Emcee demo



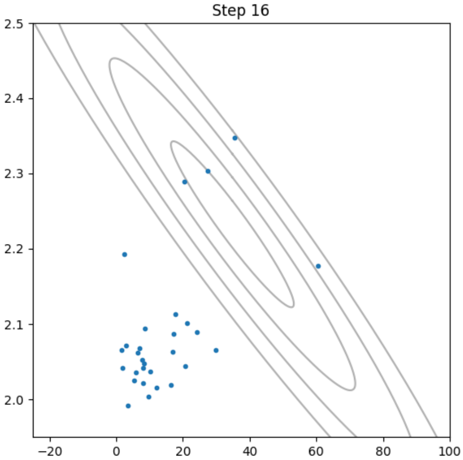
Emcee demo



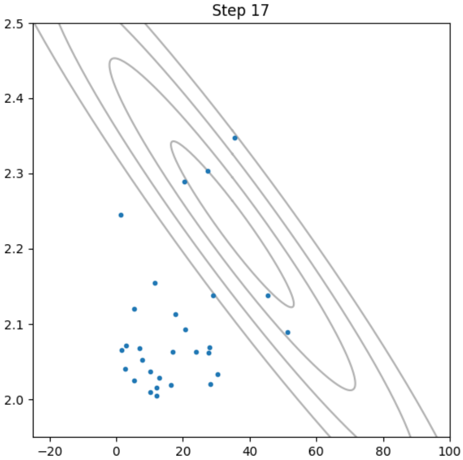
Emcee demo



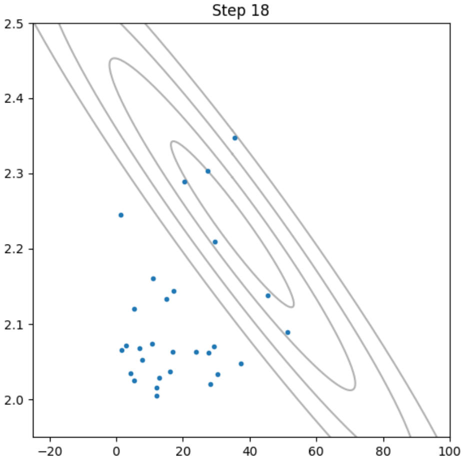
Emcee demo



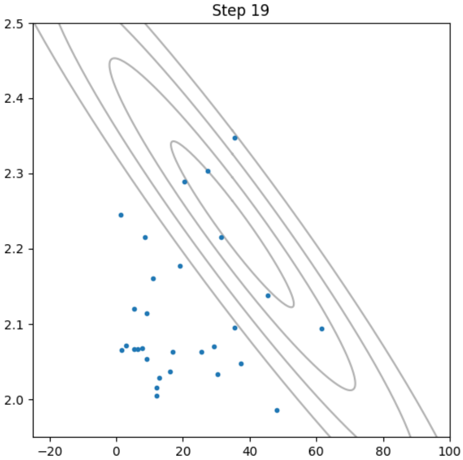
Emcee demo



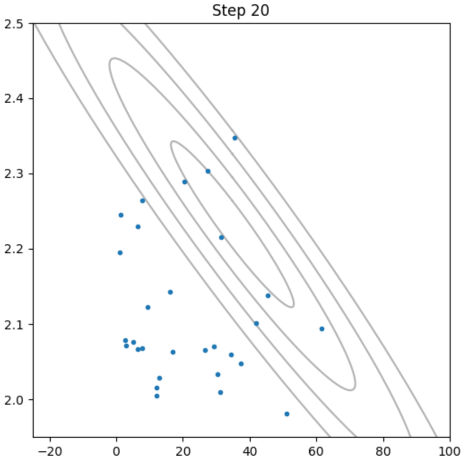
Emcee demo



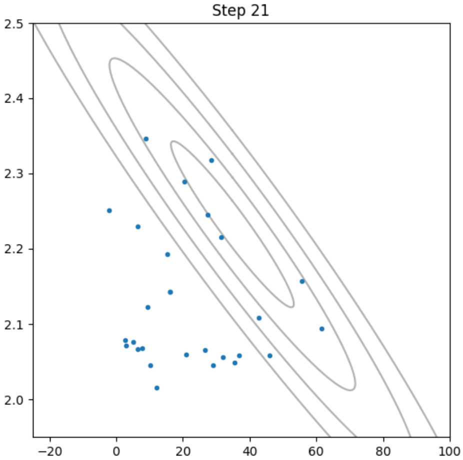
Emcee demo



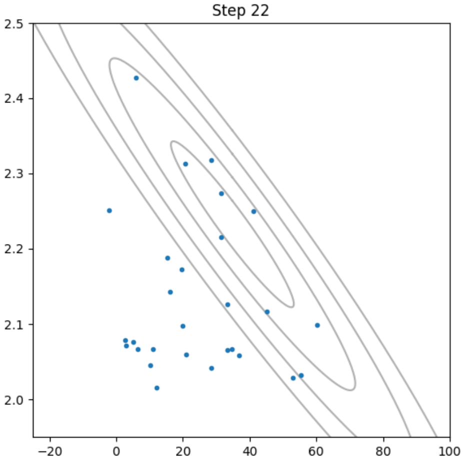
Emcee demo



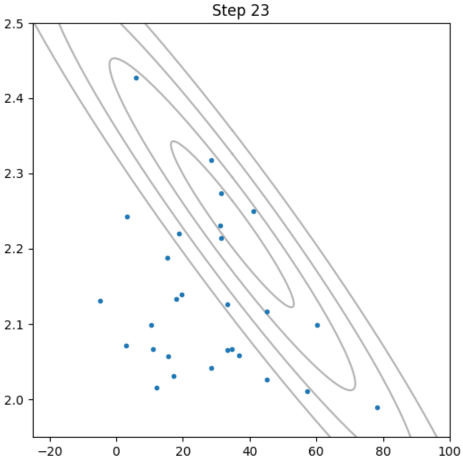
Emcee demo



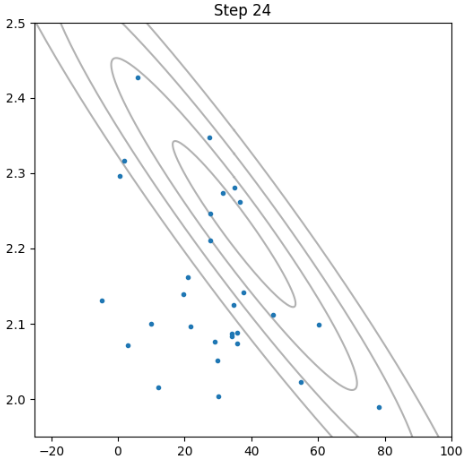
Emcee demo



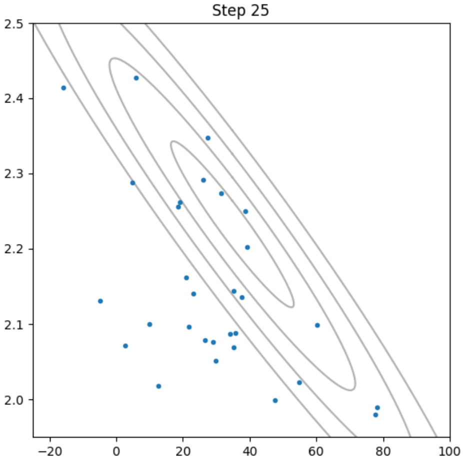
Emcee demo



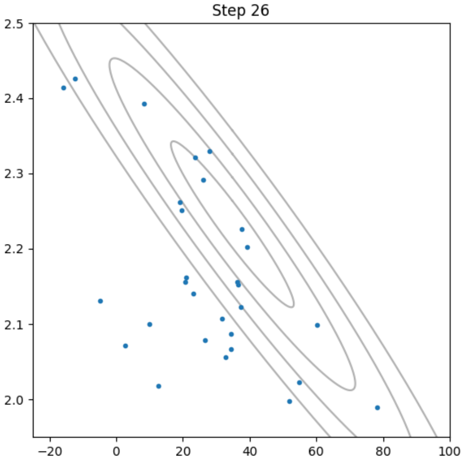
Emcee demo



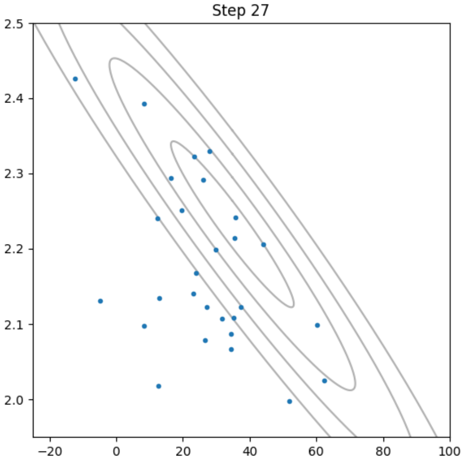
Emcee demo



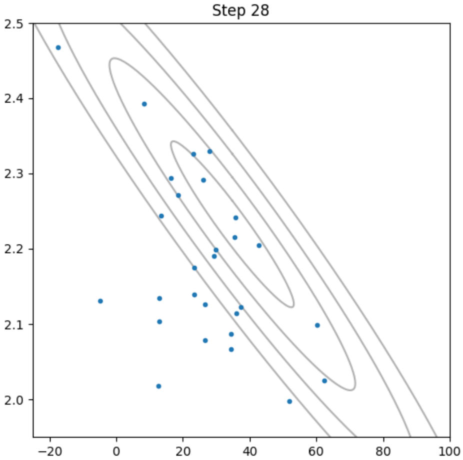
Emcee demo



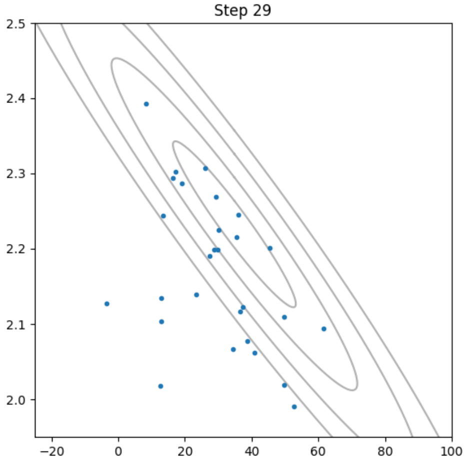
Emcee demo



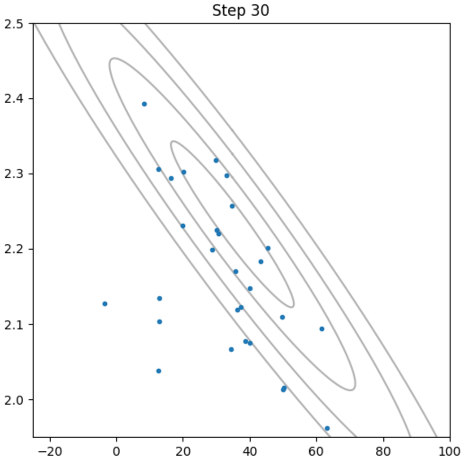
Emcee demo



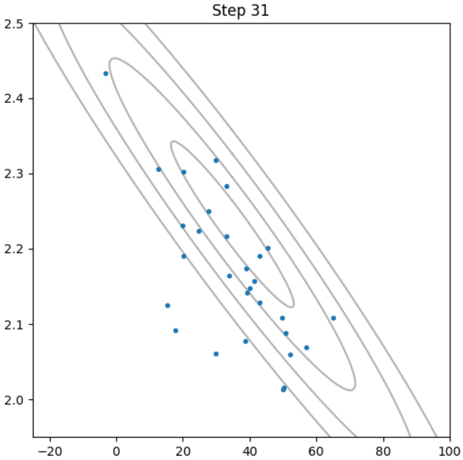
Emcee demo



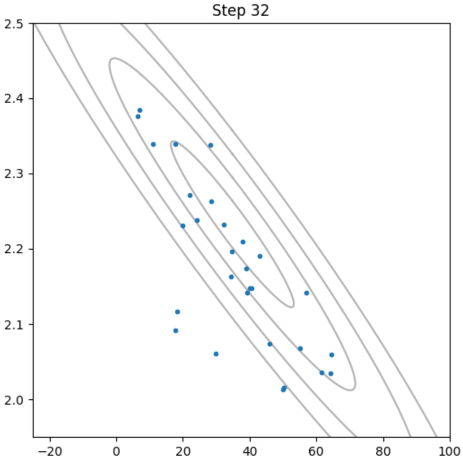
Emcee demo



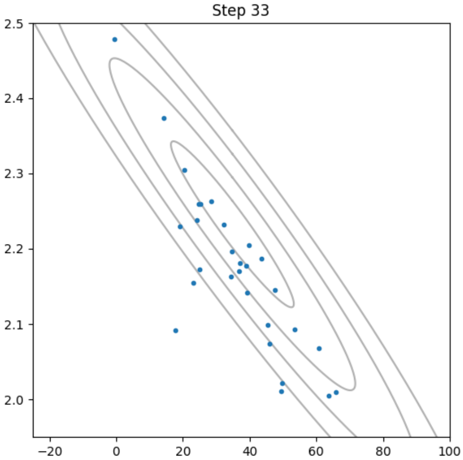
Emcee demo



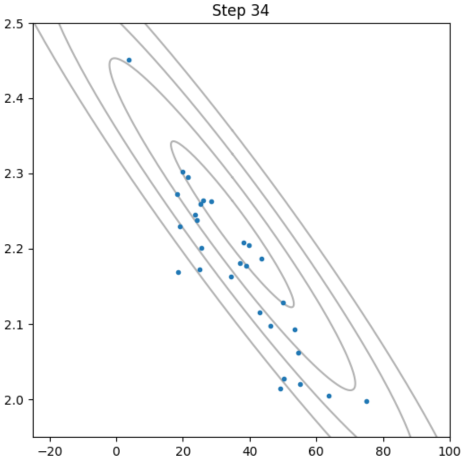
Emcee demo



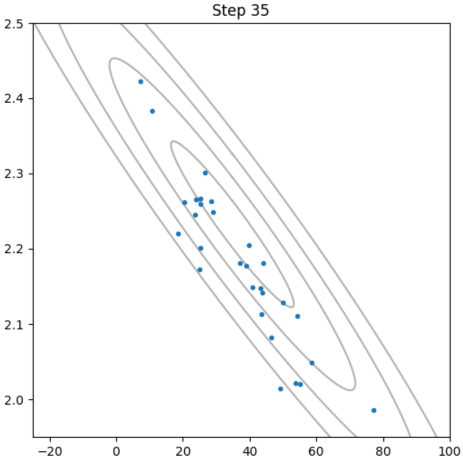
Emcee demo



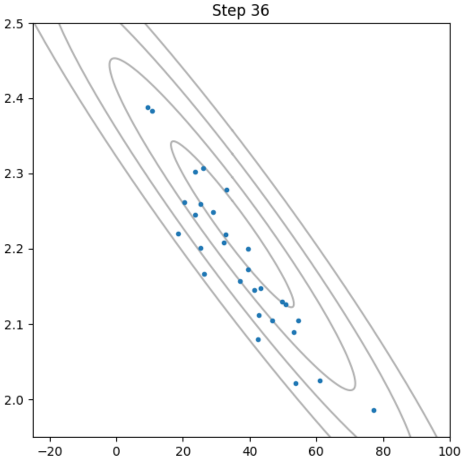
Emcee demo



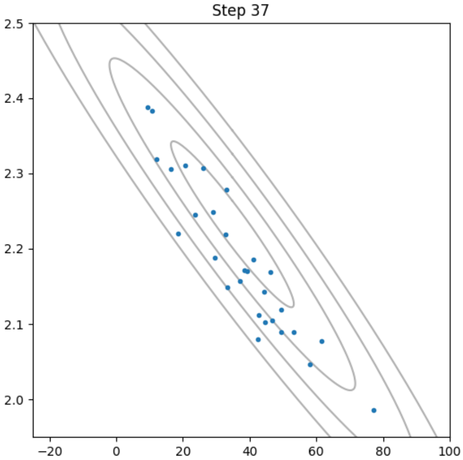
Emcee demo



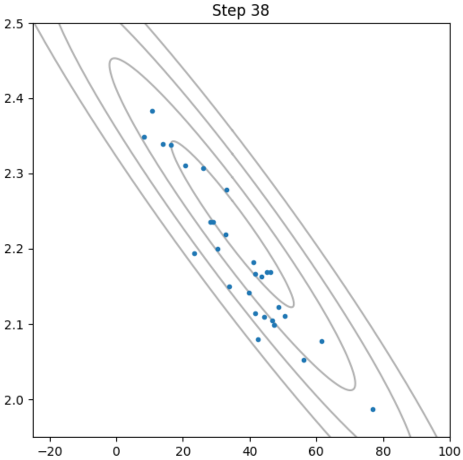
Emcee demo



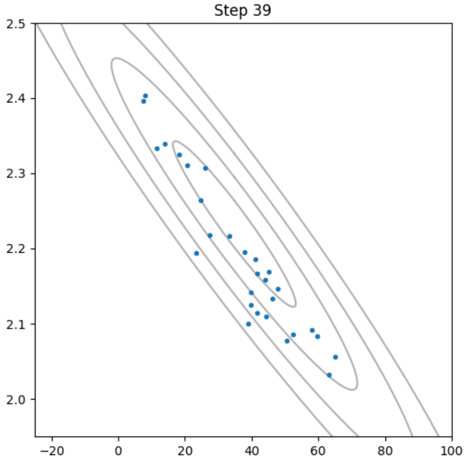
Emcee demo



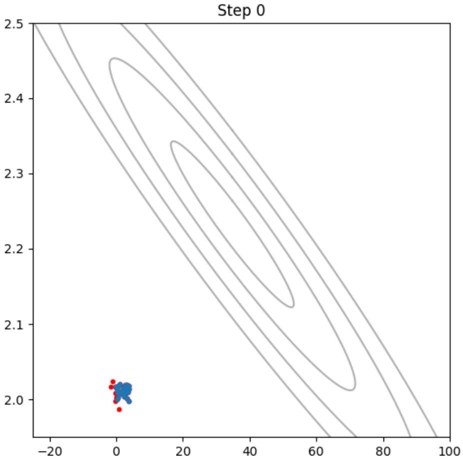
Emcee demo



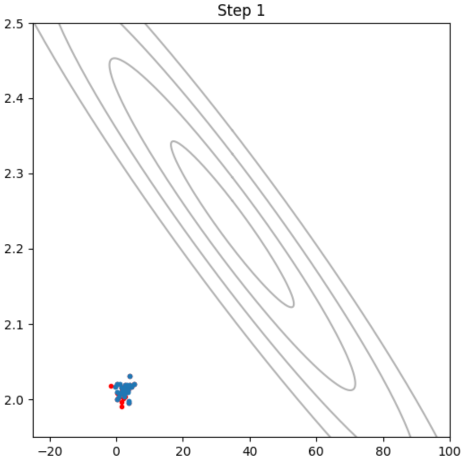
Emcee demo



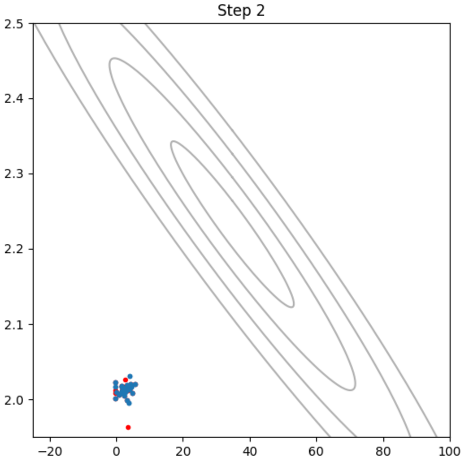
Emcee demo



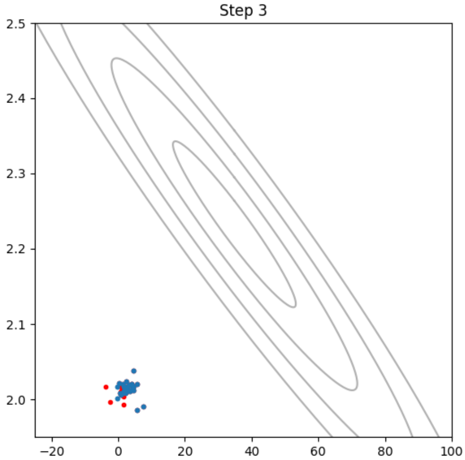
Emcee demo



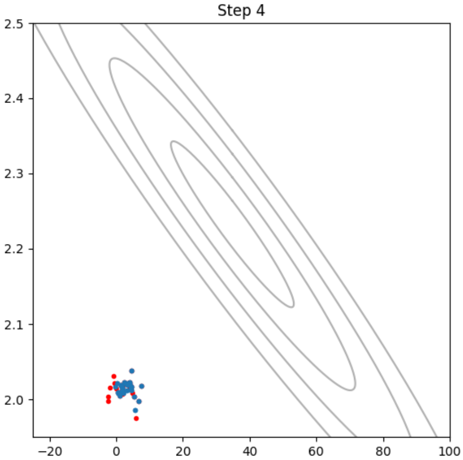
Emcee demo



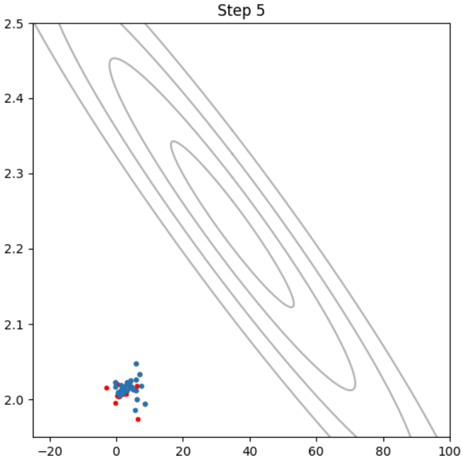
Emcee demo



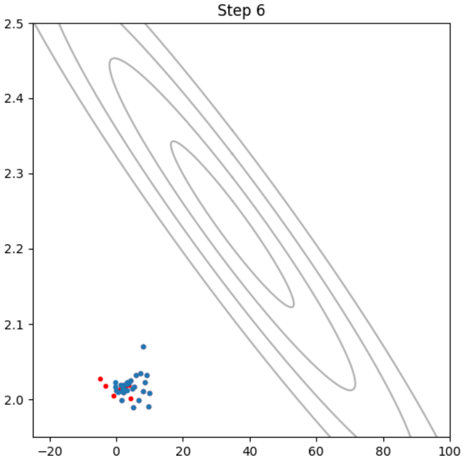
Emcee demo



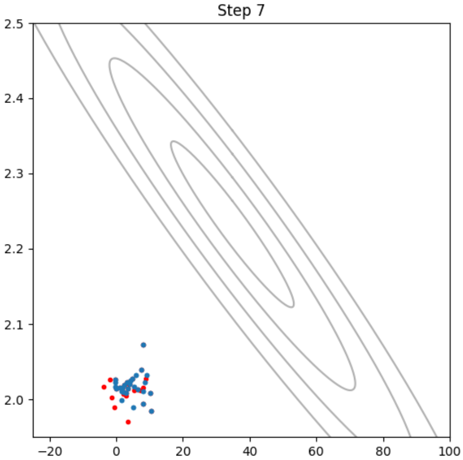
Emcee demo



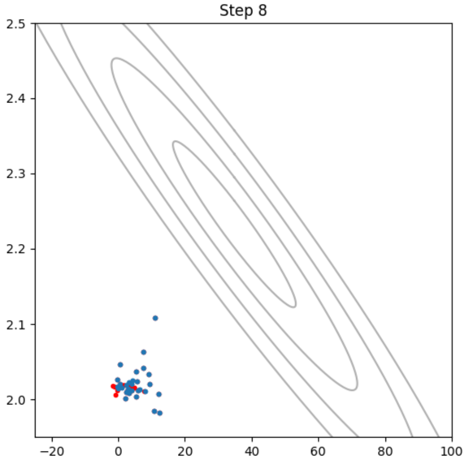
Emcee demo



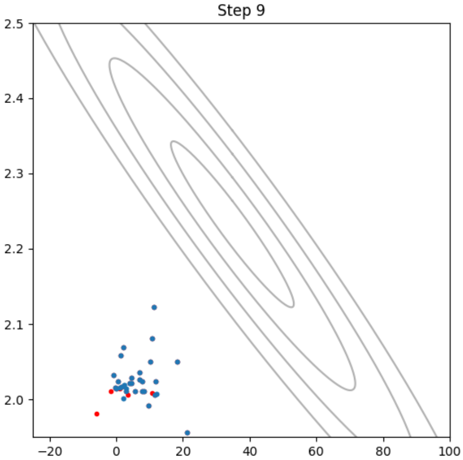
Emcee demo



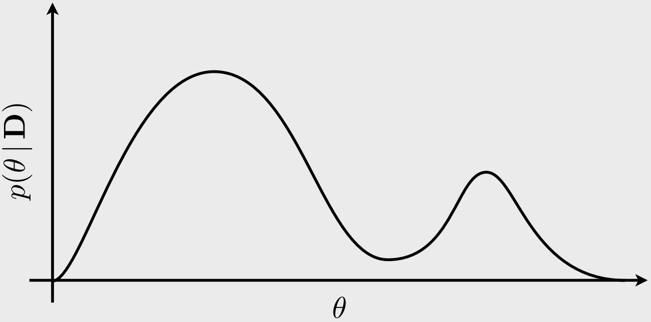
Emcee demo



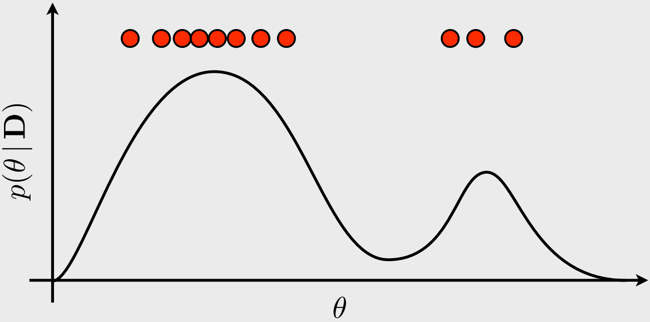
Emcee demo



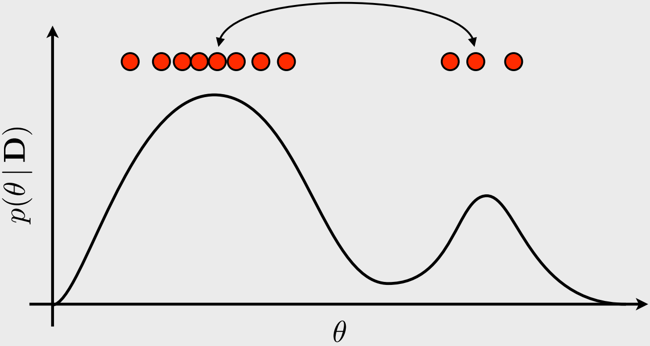
what about multimodal densities?



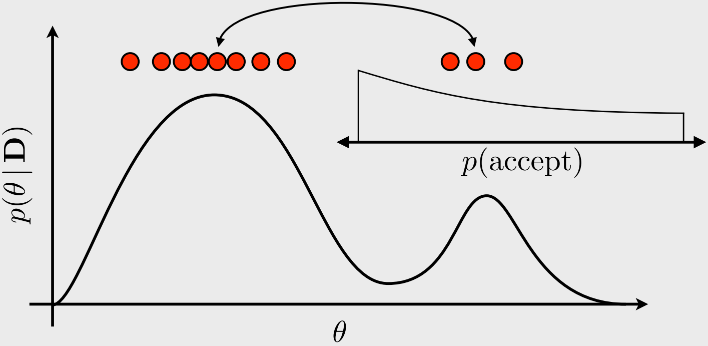
what about multimodal densities?



what about multimodal densities?



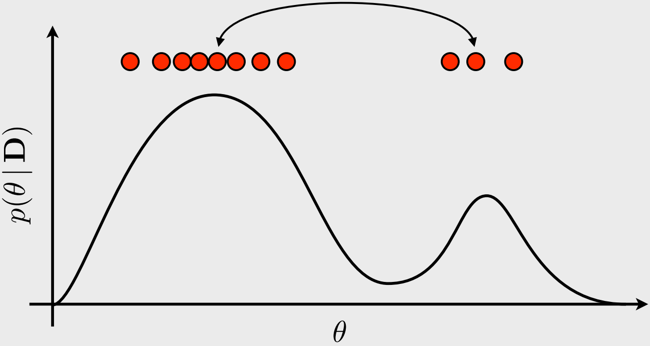
what about multimodal densities?



Differential Evolution move

- emcee allows us to use different *move* types (different *proposal* functions)
- The Differential Evolution (DE) move can improve the sampling for multi-modal distributions
- DE move: randomly select *two* “helpers”
- Propose moving by their vector difference
- (If they are from different modes, this proposes *jumping between modes*)
- Mixing in a fraction of DE moves with the regular “Stretch” move works well!

what about multimodal densities?



Summary

- Traditional Metropolis–Hastings MCMC suffers from a *lack of affine invariance* – requires *tuning parameters* that change for each specific probability function
- *Ensemble samplers* like emcee use the *distribution of the walkers* to achieve *affine invariance*
- $\rightarrow$ much easier to use, and faster sampling
- (Huge side effect: parallelizable!)
- Multi-modal distributions still hard, but *DE Move* can help
- MCMC isn't scary!