

Affine-Invariant Samplers

Dustin Lang
Perimeter Institute for Theoretical Physics

PSI Numerical Methods, 2023-01-26

Borrowing heavily from Dan Foreman-Mackey's slides

<https://speakerdeck.com/dfm/data-analysis-with-mcmc1>

And Jonathan Pritchard's slides

<https://www.imperial.ac.uk/media/imperial-college/research-centres/>

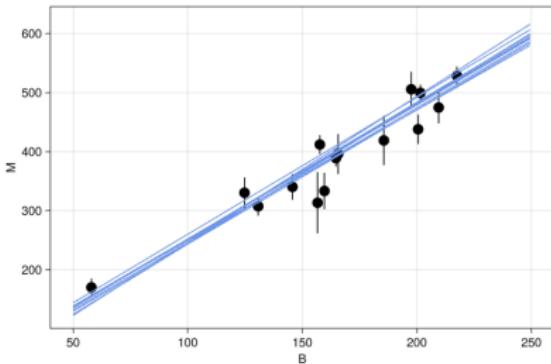
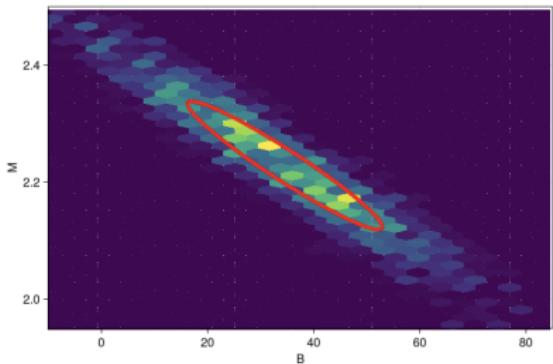
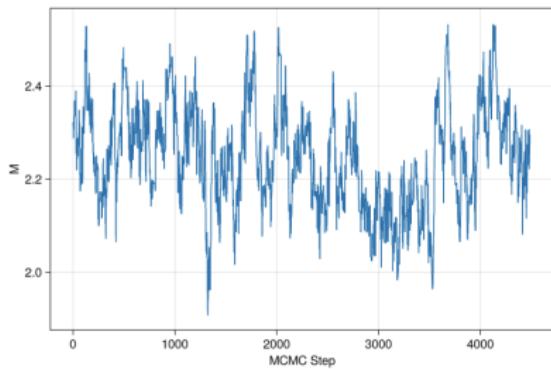
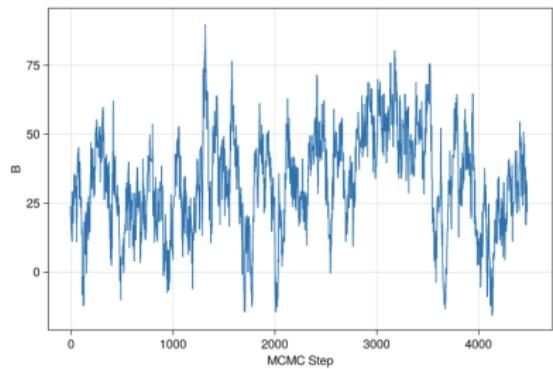
These slides are available at

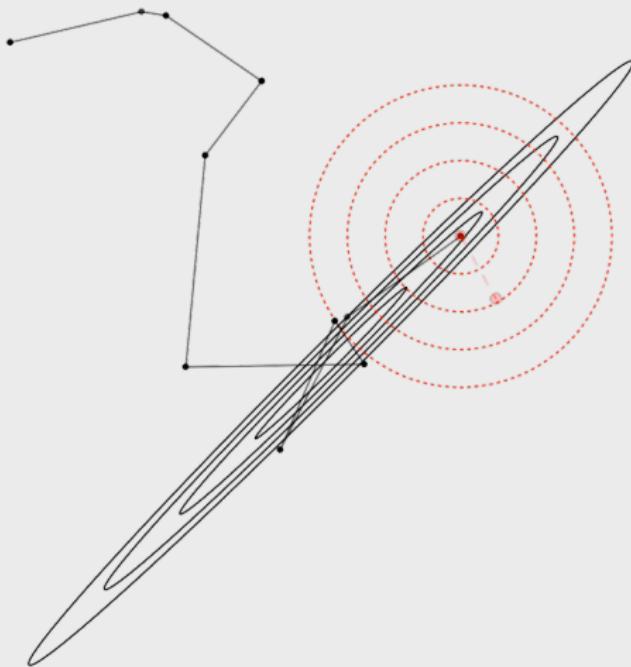
<https://github.com/dstndstn/MCMC-talk>

The MCMC Algorithm

```
function mcmc(logprob_func, logprob_args,
              propose_func, propose_args,
              initial_pos, nsteps)
    p = initial_pos
    logprob = logprob_func(p, logprob_args)
    chain = zeros(Float64, (nsteps, length(p)))
    naccept = 0
    for i in 1:nsteps
        # propose a new position in parameter space
        p_new = propose_func(p, propose_args)
        # compute probability at new position
        logprob_new = logprob_func(p_new, logprob_args)
        # decide whether to jump to the new position
        if exp(logprob_new - logprob) > rand()
            p = p_new
            logprob = logprob_new
            naccept += 1
        end
        # save the position
        chain[i,:] = p
    end
    return chain, naccept/nsteps
end
```

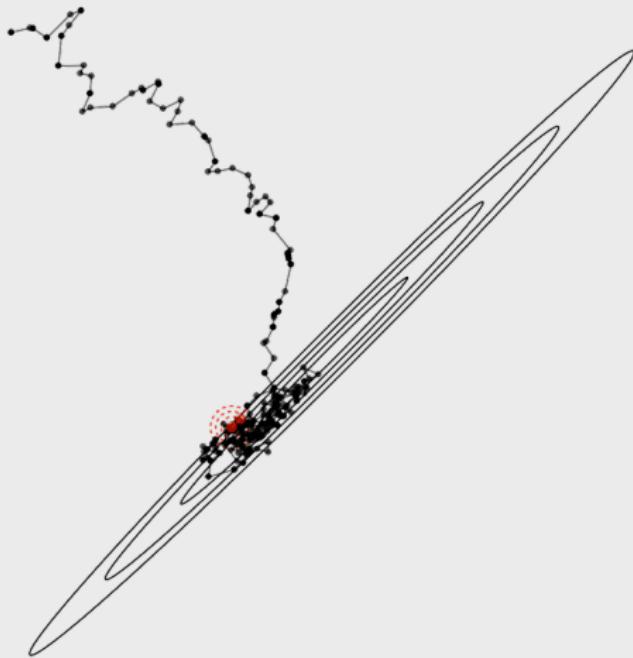
MCMC for model parameter inference





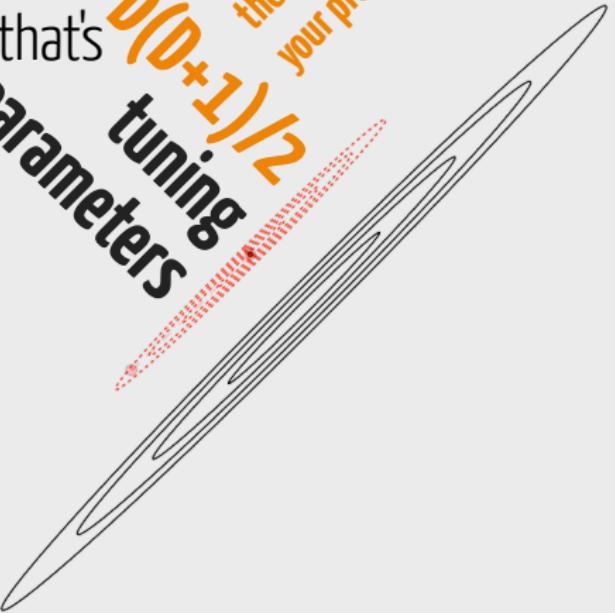
Metropolis–Hastings

in the real world



Metropolis–Hastings

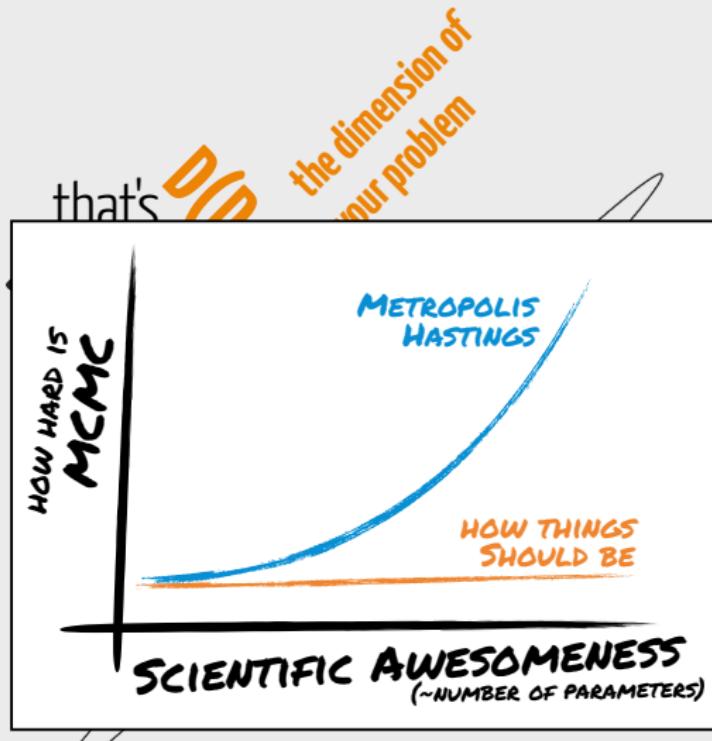
in the real world



that's $D(D+1)/2$ tuning parameters
the dimension of your problem

Metropolis–Hastings

in the real world



Metropolis–Hastings

in the real world



Jonathan Goodman



Jonathan Weare

"Ensemble samplers with affine invariance"

(dfm.io/mcmc-gw10)



introducing **emcee** the MCMC Hammer

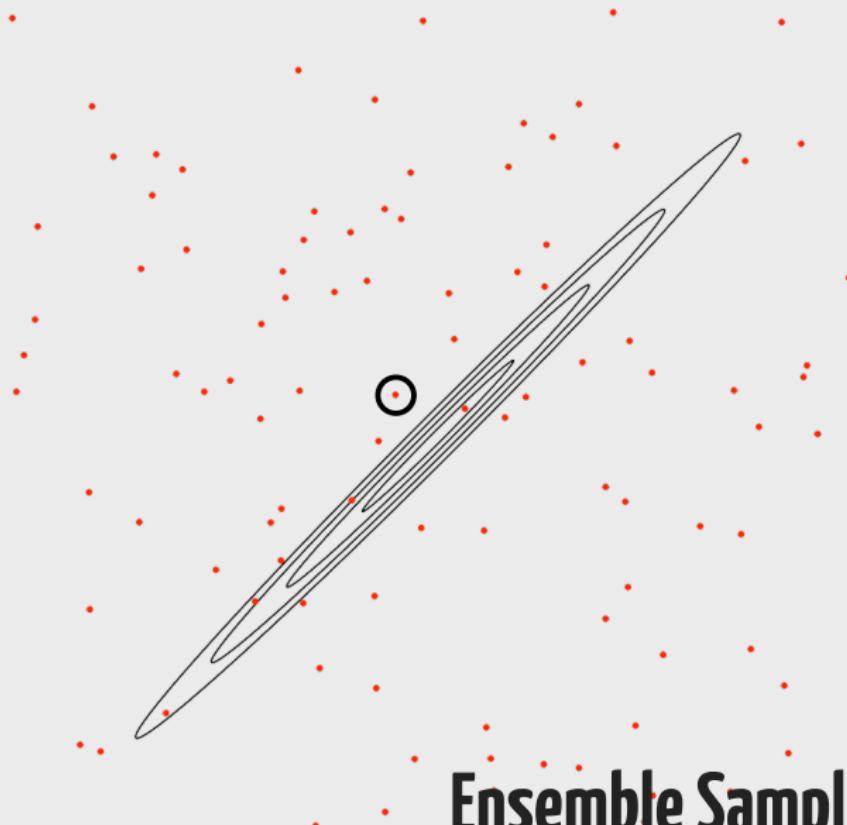
arxiv.org/abs/1202.3665

dan.iel.fm/emcee



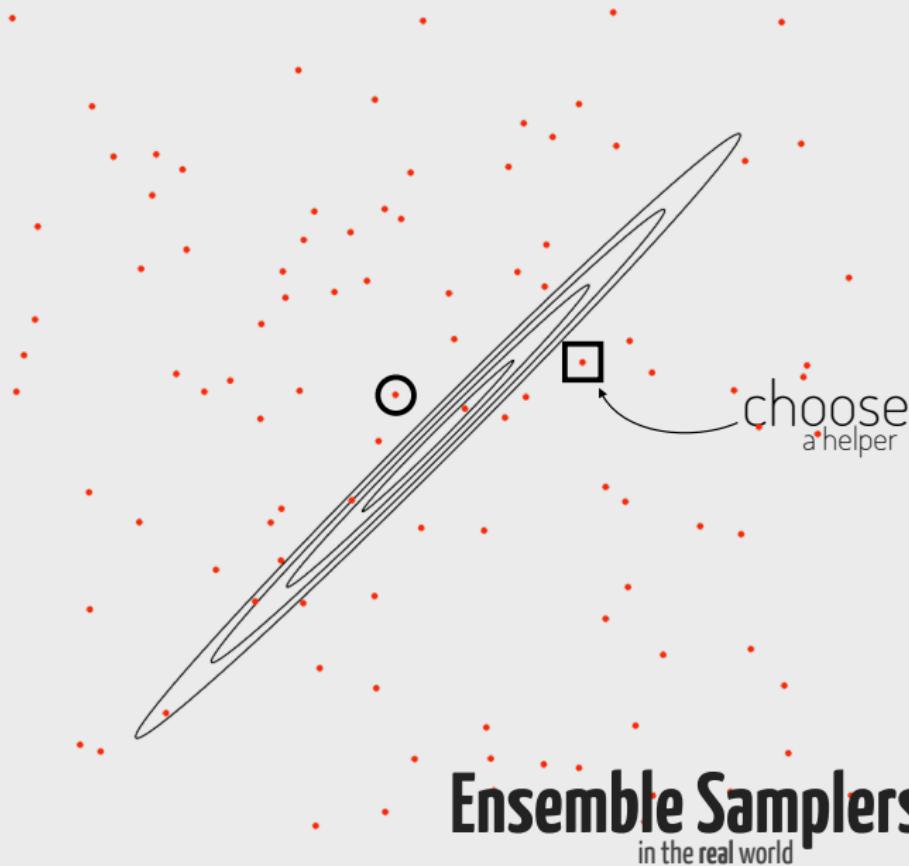
Ensemble Samplers

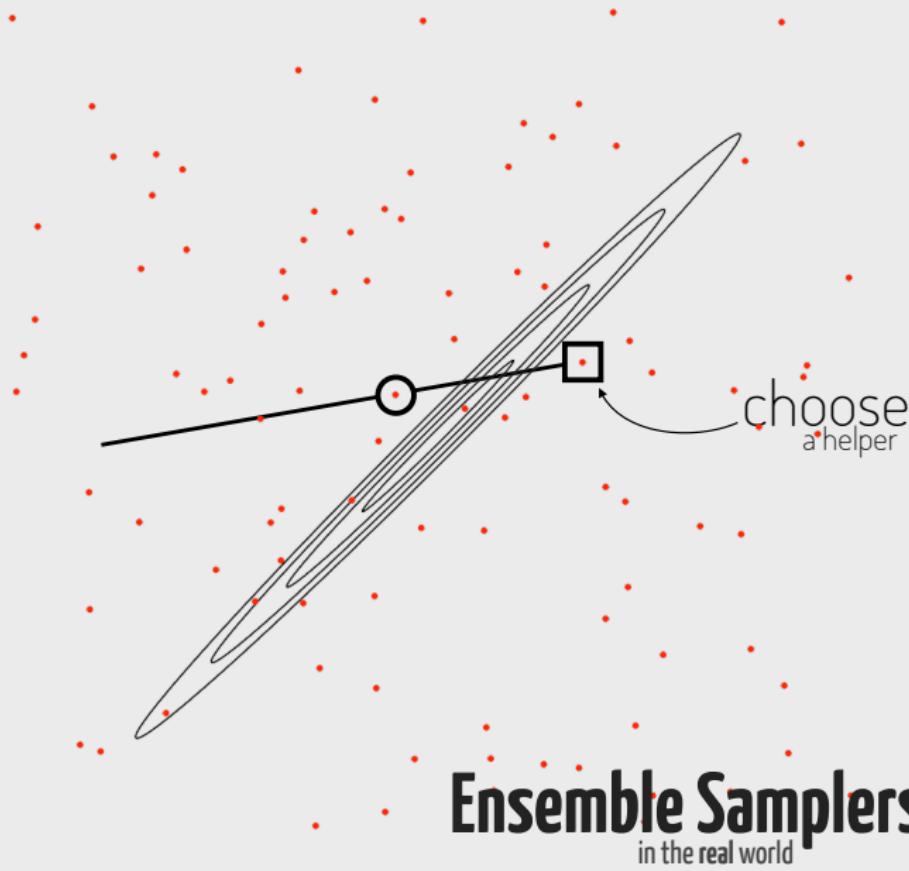
in the real world

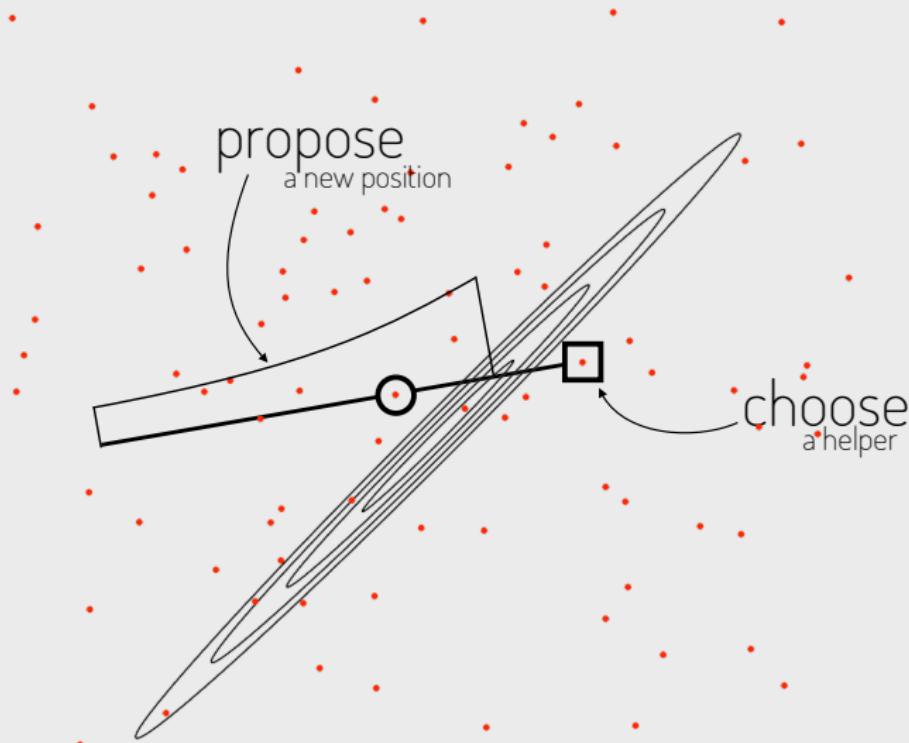


Ensemble Samplers

in the real world







Ensemble Samplers

in the real world

propose
a new position

accept?

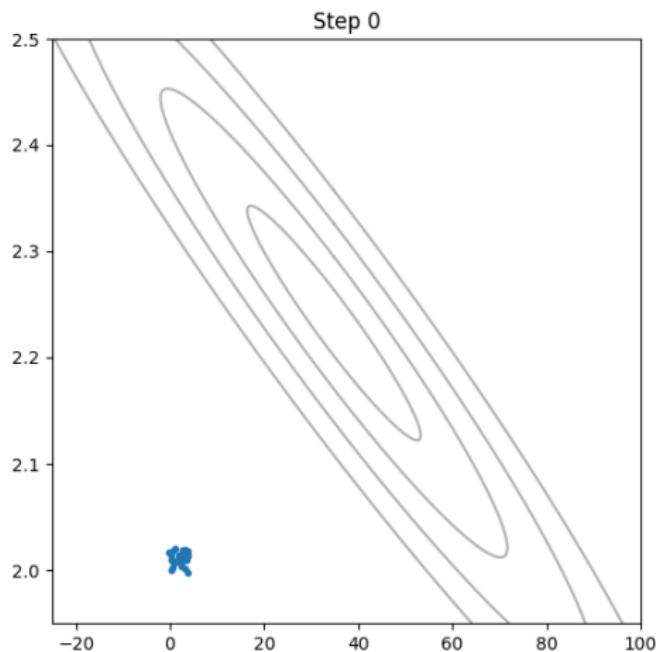
choose
a helper

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

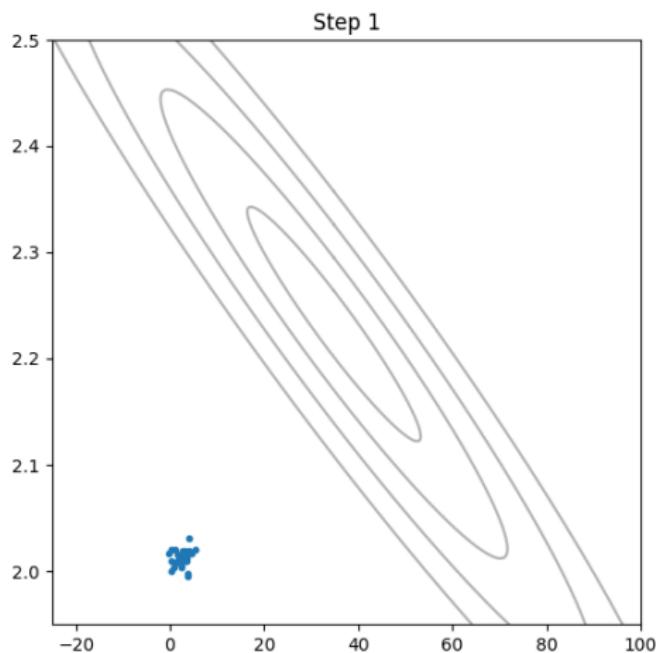
Ensemble Samplers

in the real world

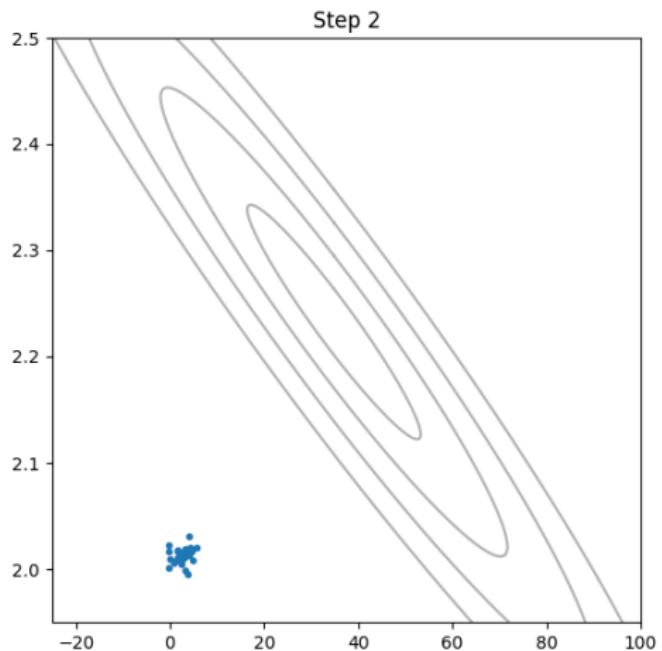
Emcee demo



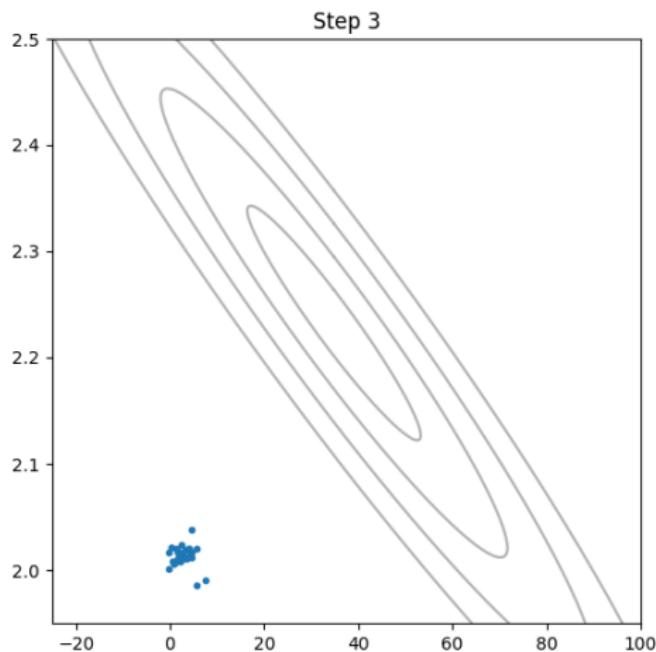
Emcee demo



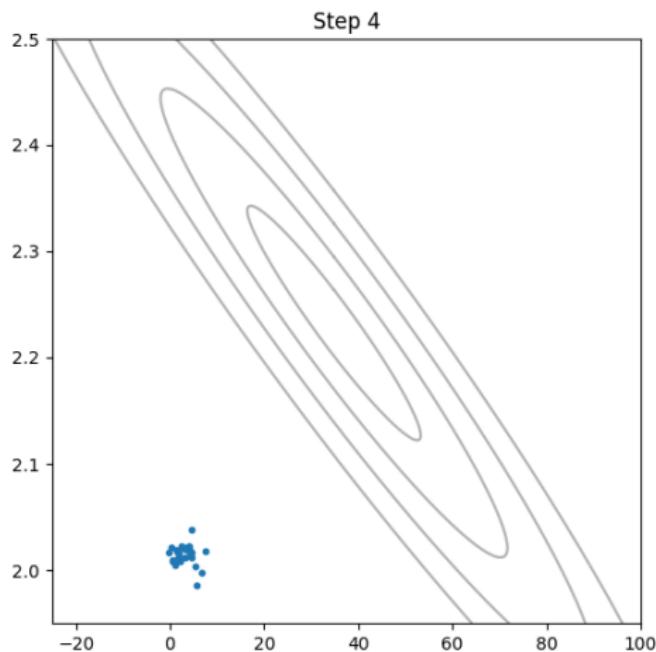
Emcee demo



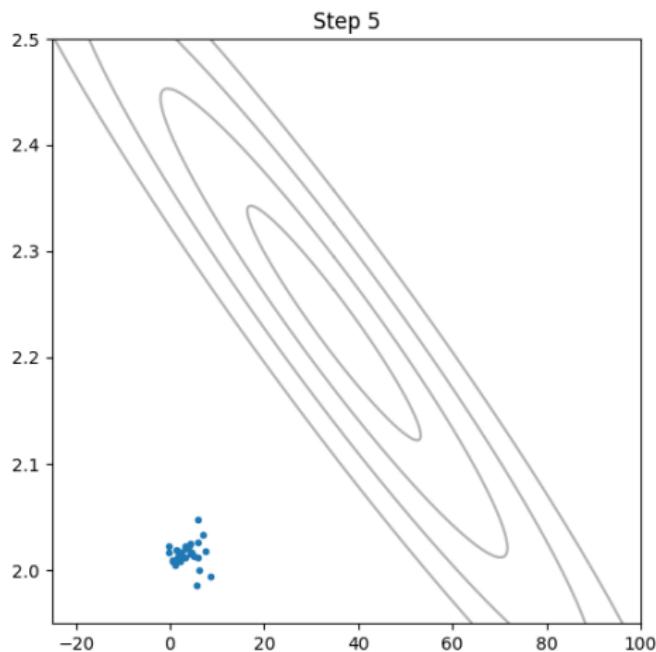
Emcee demo



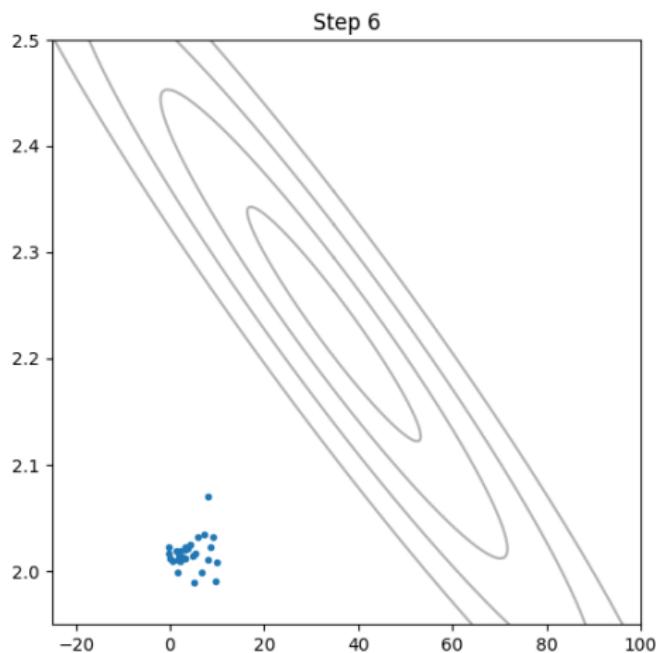
Emcee demo



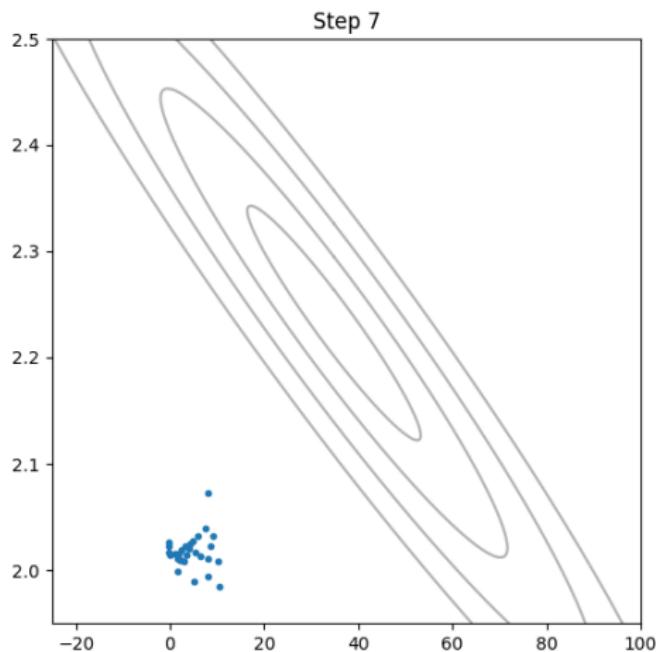
Emcee demo



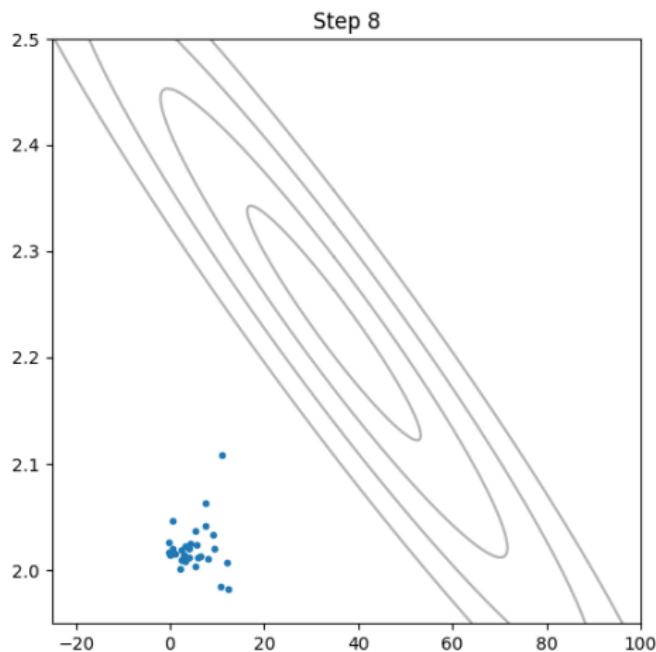
Emcee demo



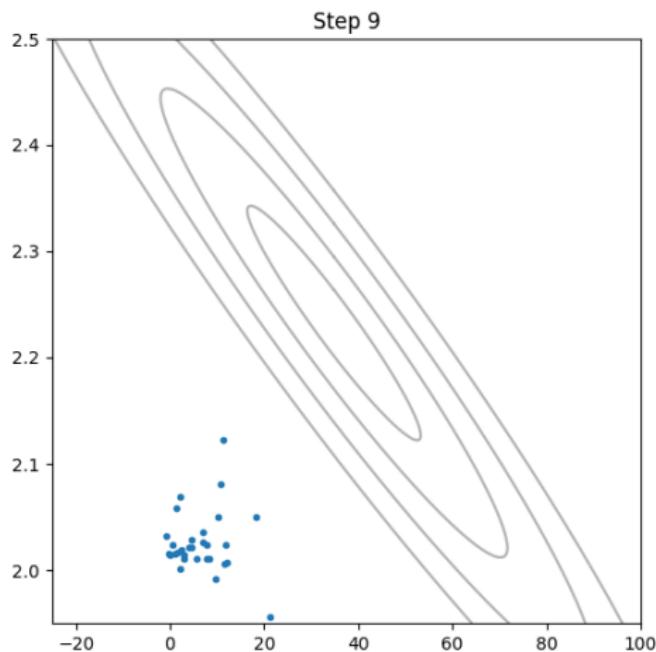
Emcee demo



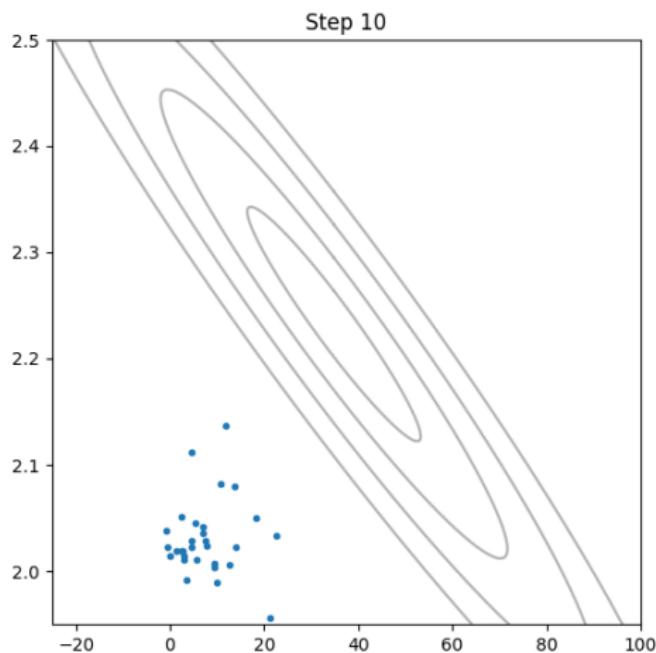
Emcee demo



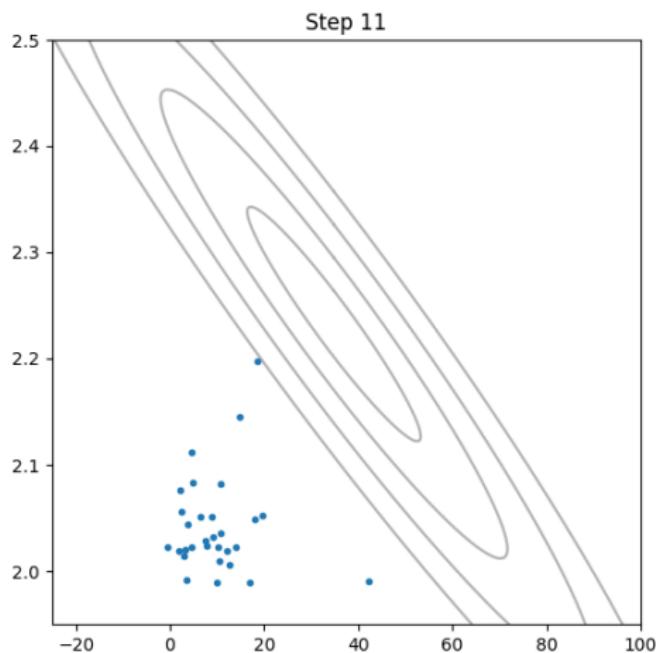
Emcee demo



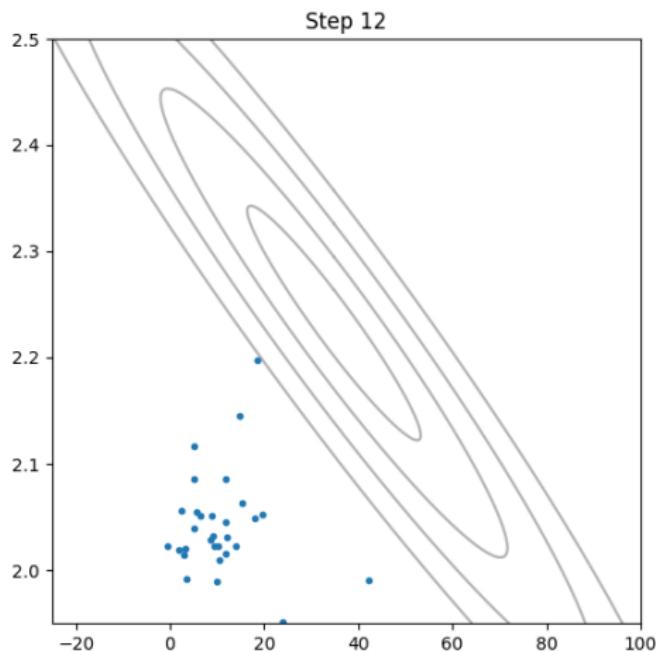
Emcee demo



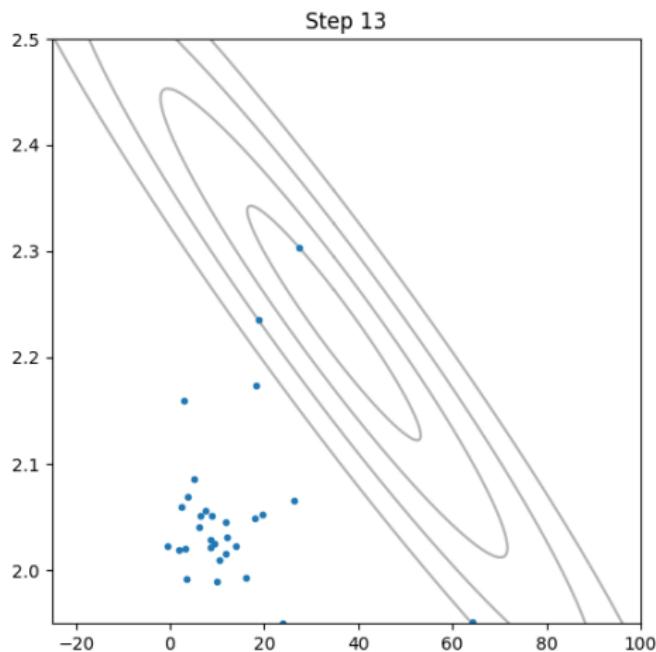
Emcee demo



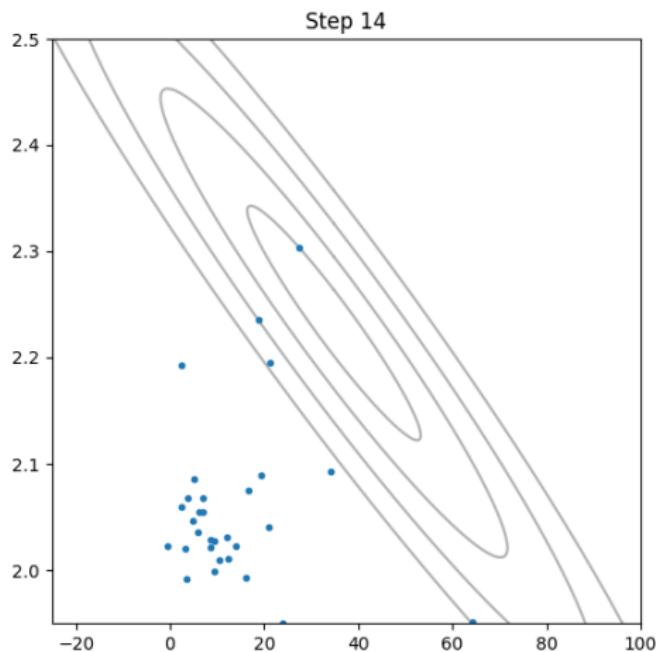
Emcee demo



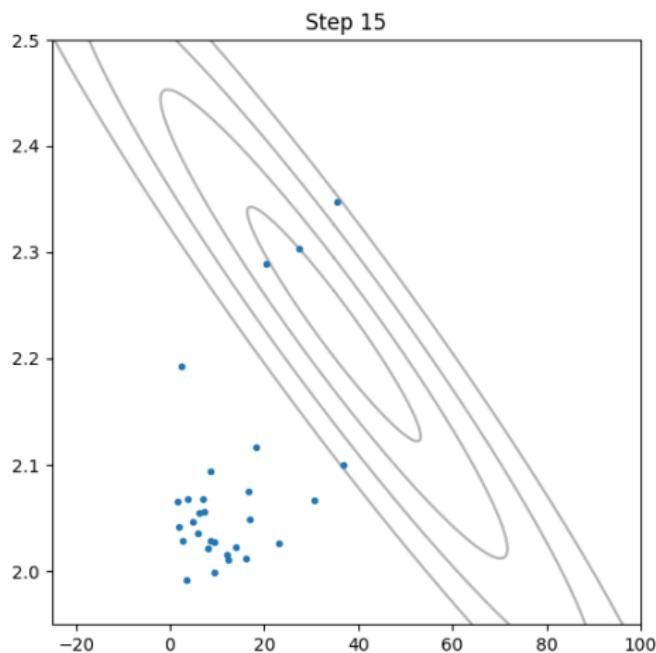
Emcee demo



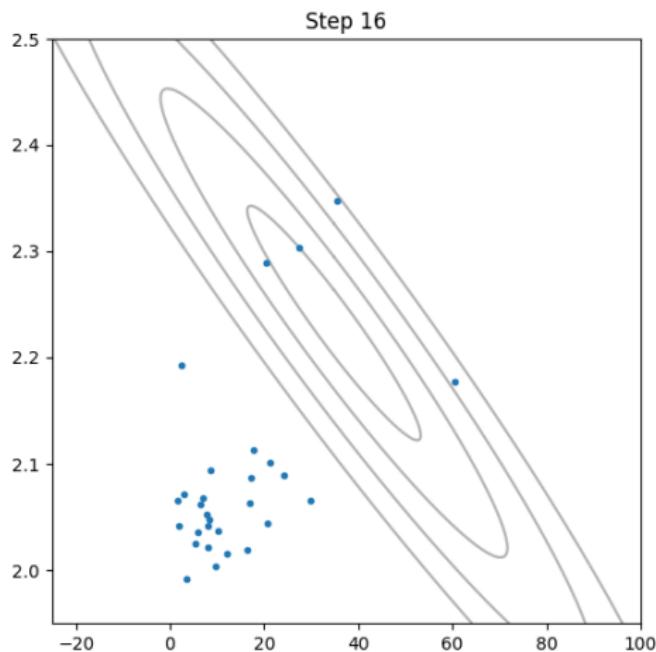
Emcee demo



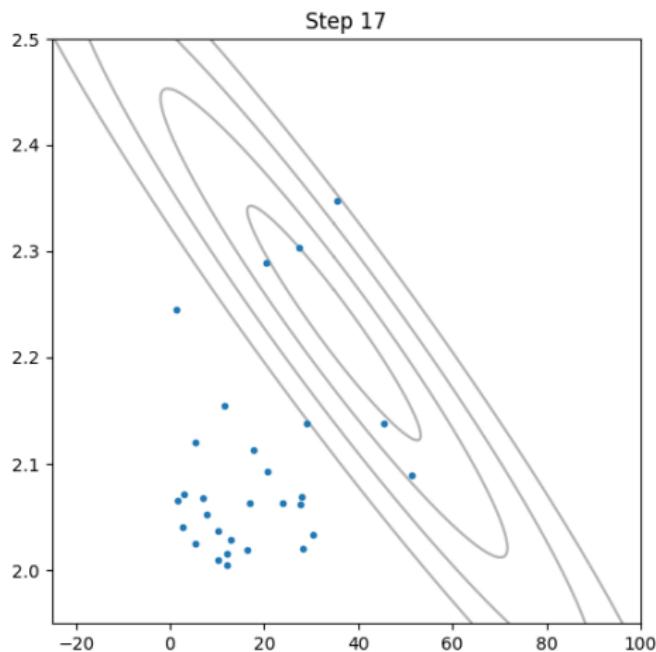
Emcee demo



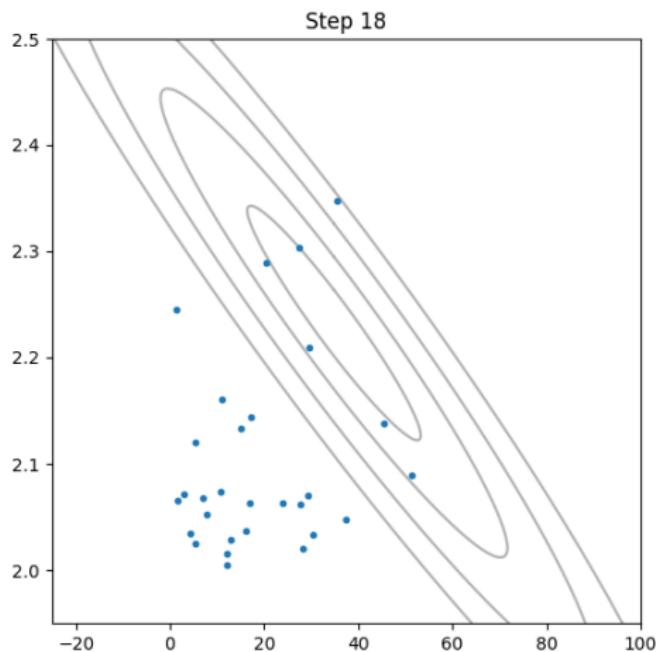
Emcee demo



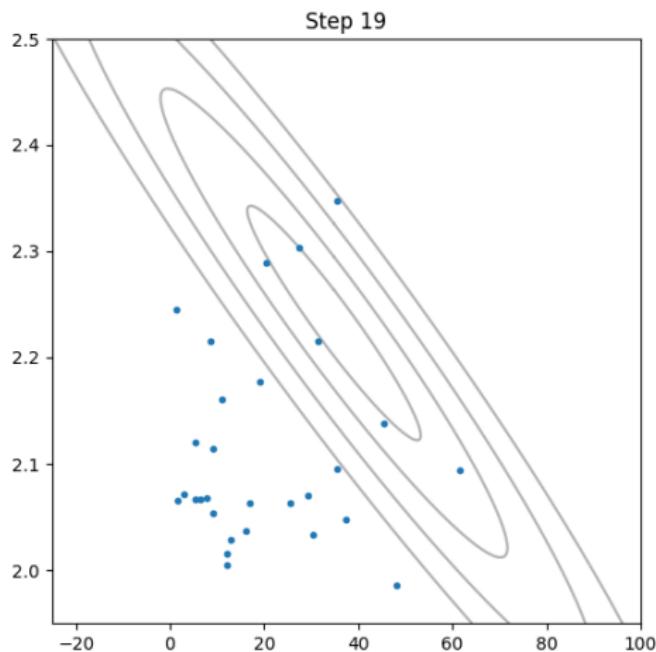
Emcee demo



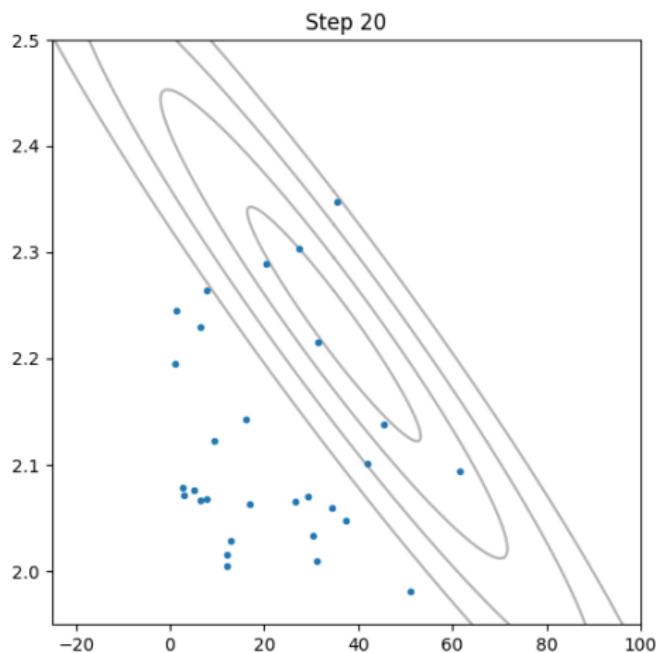
Emcee demo



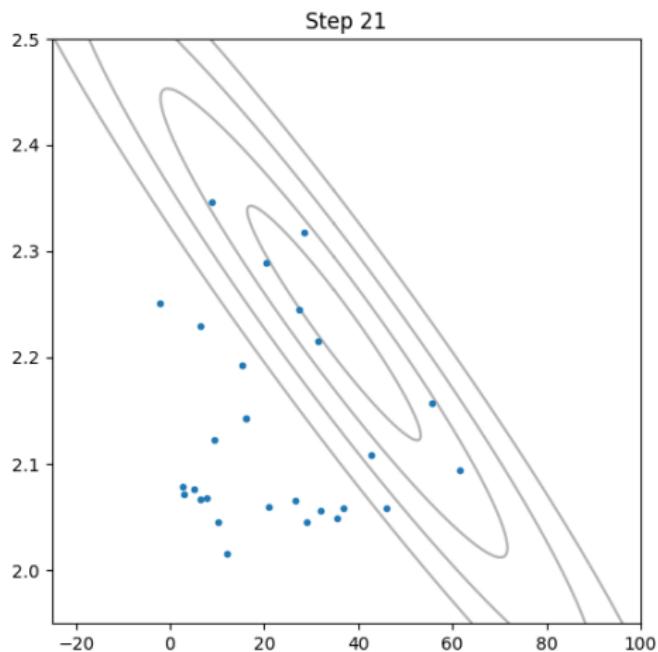
Emcee demo



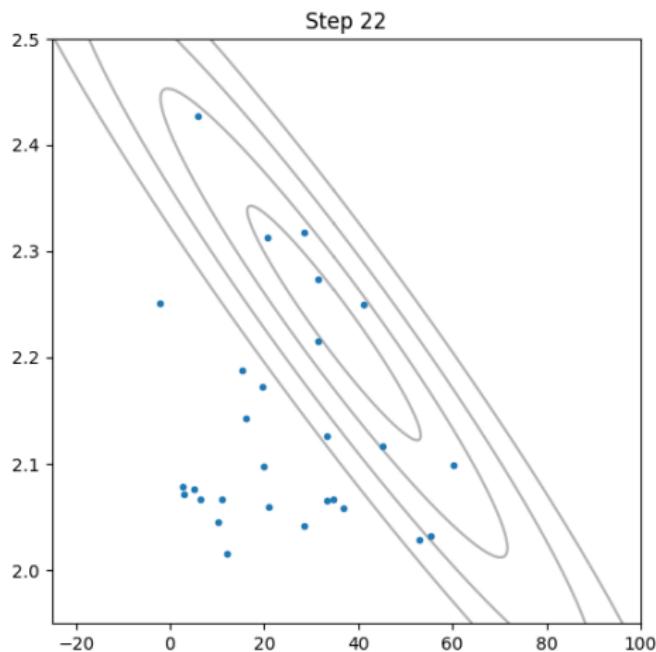
Emcee demo



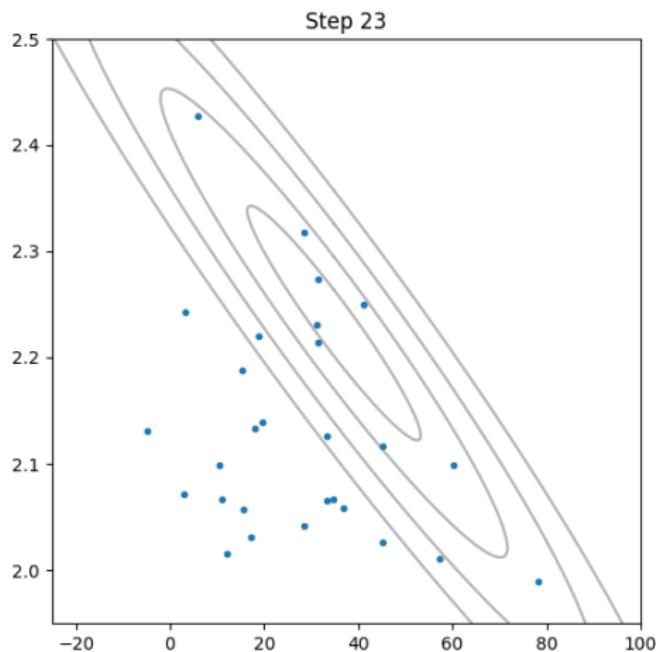
Emcee demo



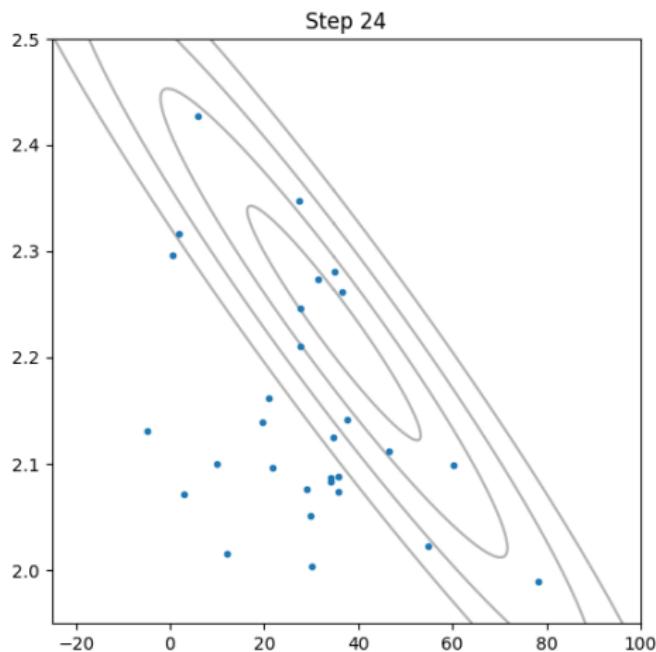
Emcee demo



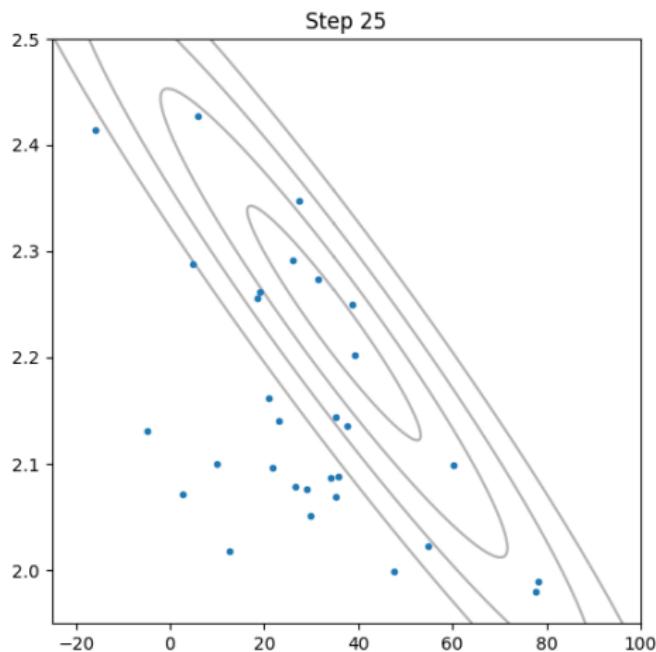
Emcee demo



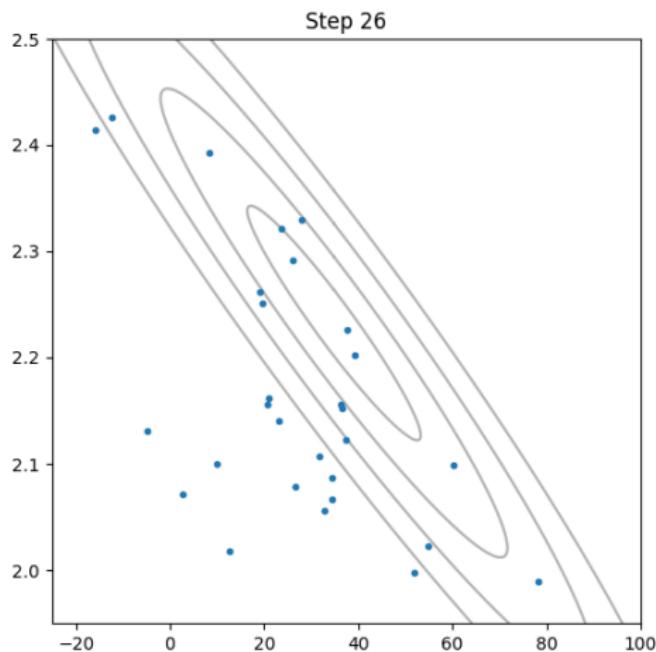
Emcee demo



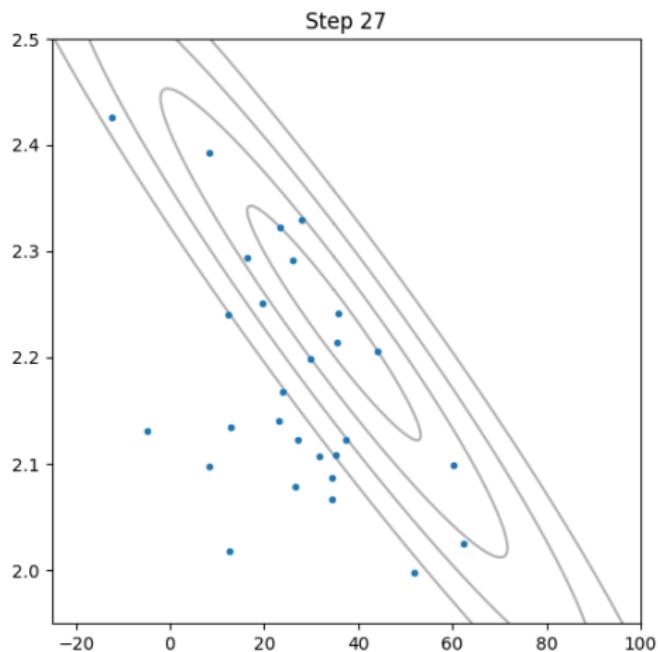
Emcee demo



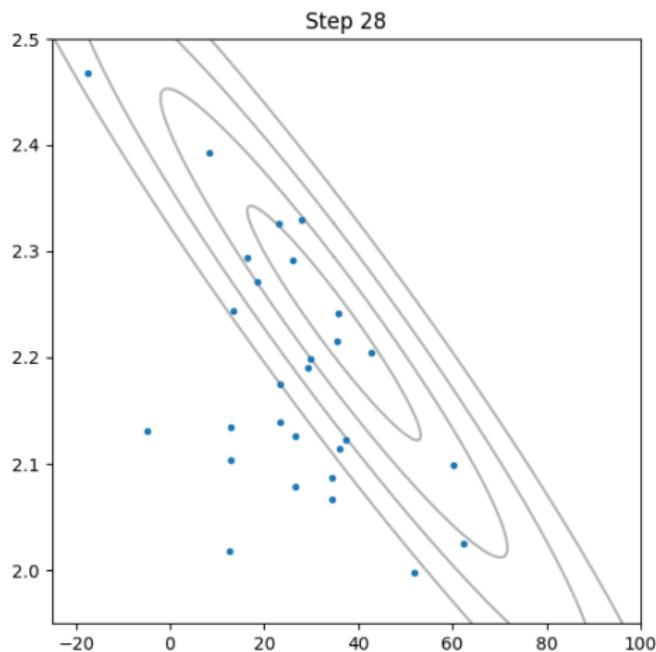
Emcee demo



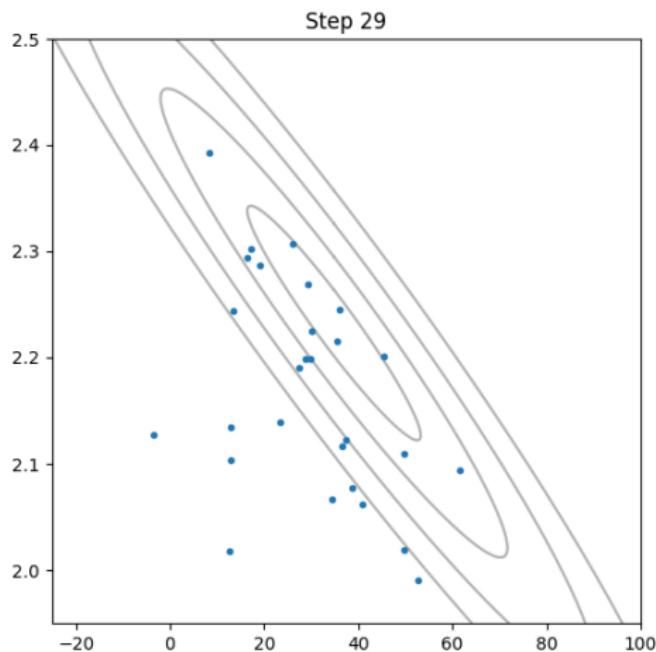
Emcee demo



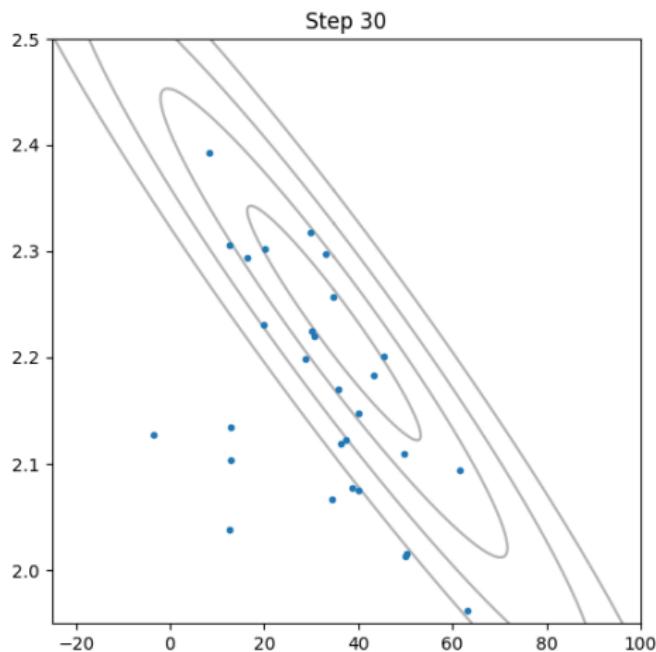
Emcee demo



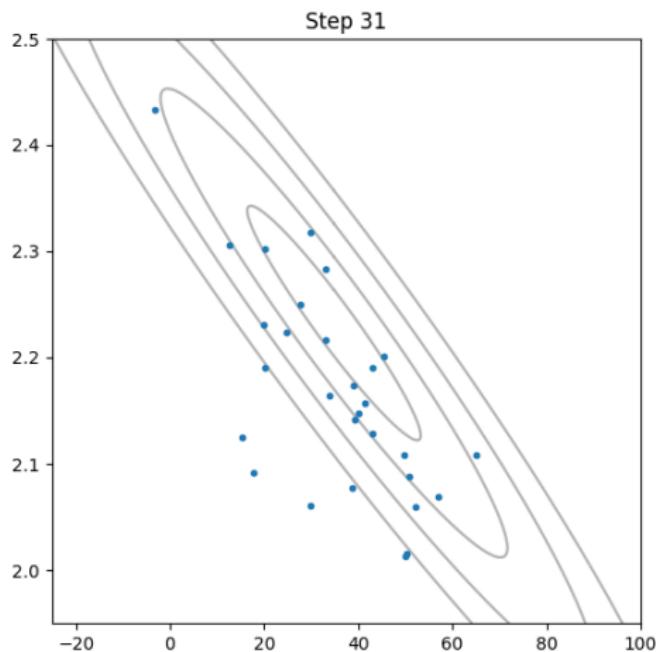
Emcee demo



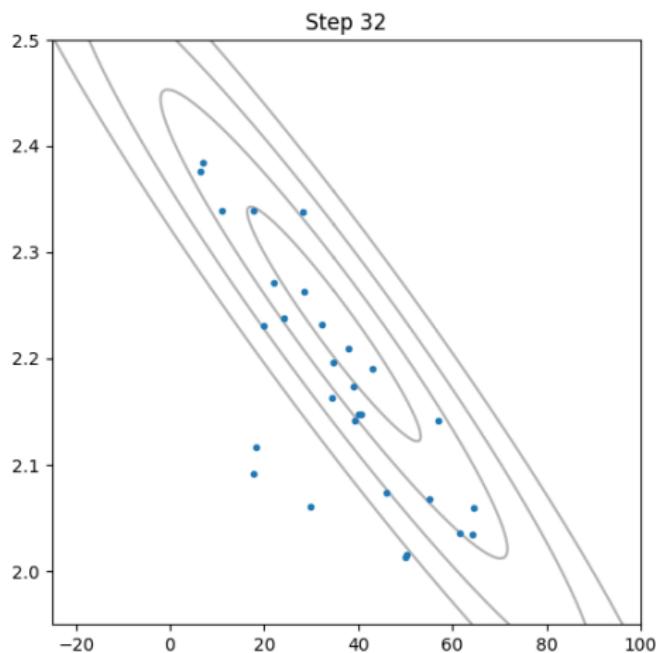
Emcee demo



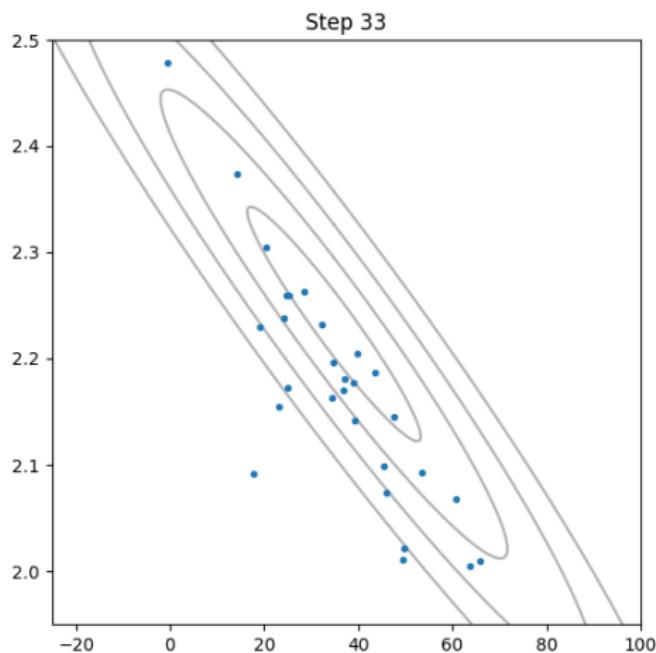
Emcee demo



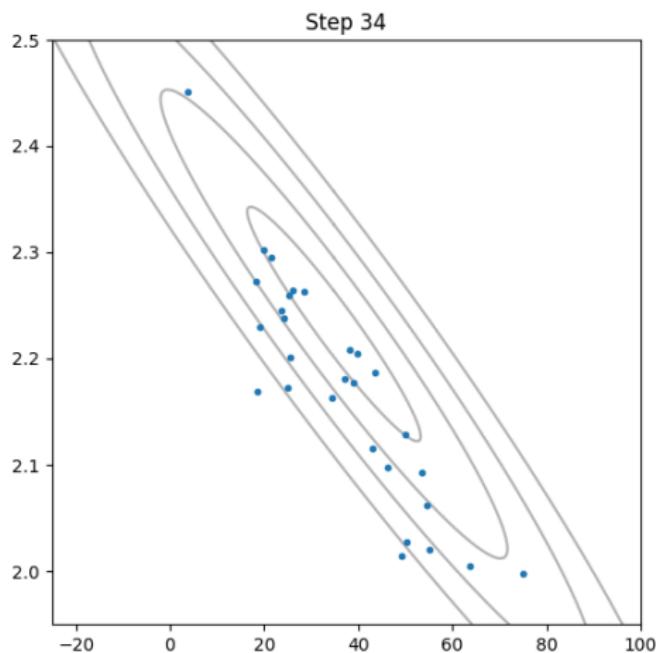
Emcee demo



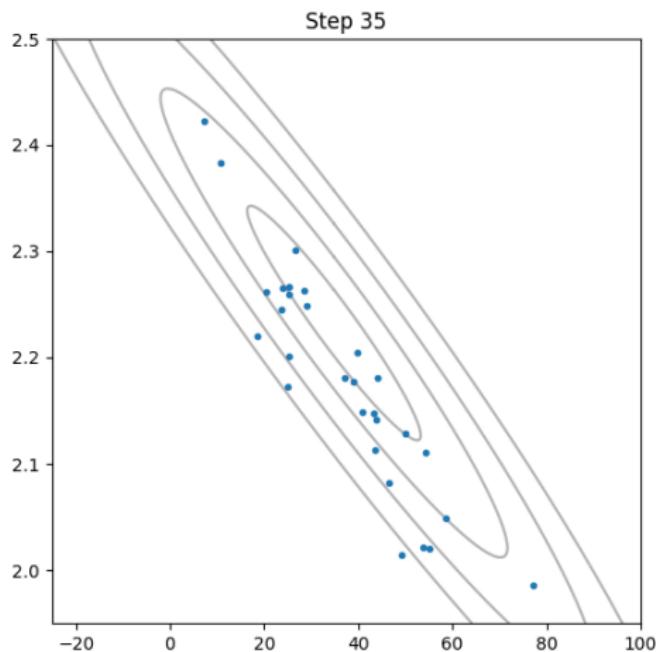
Emcee demo



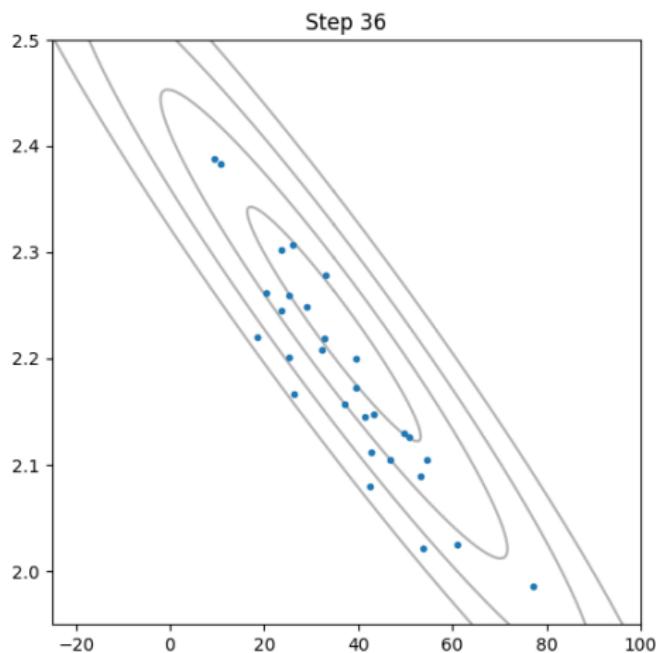
Emcee demo



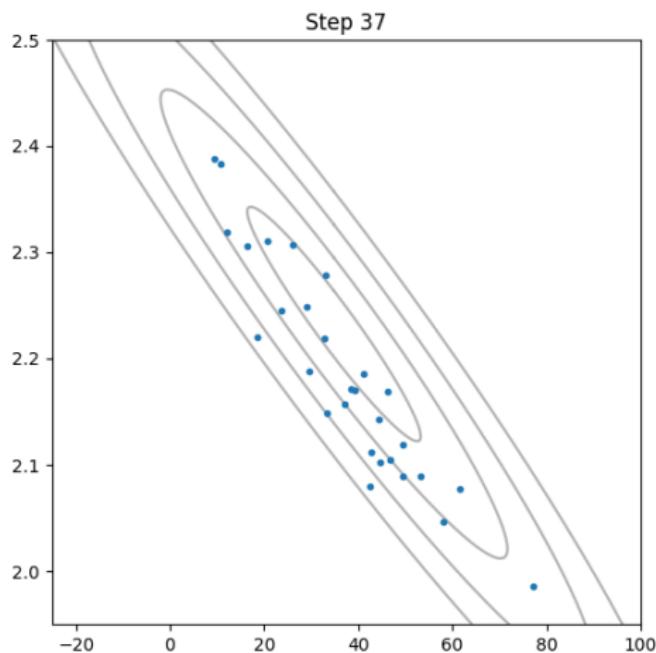
Emcee demo



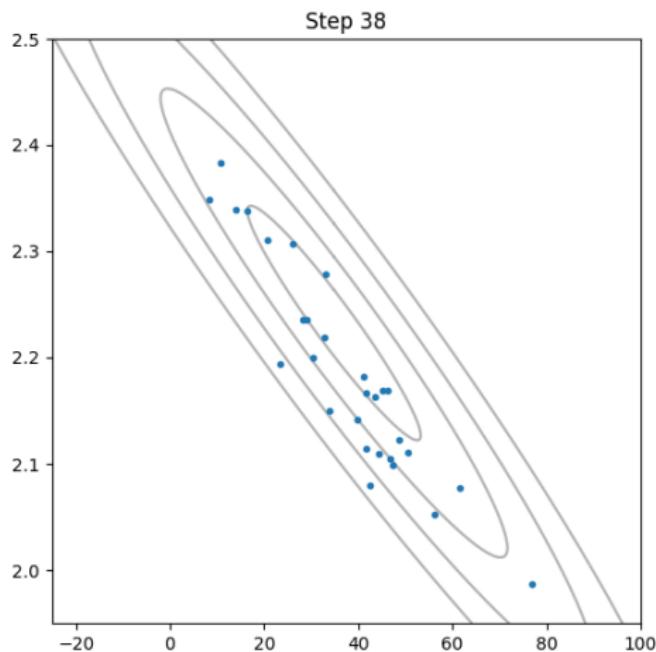
Emcee demo



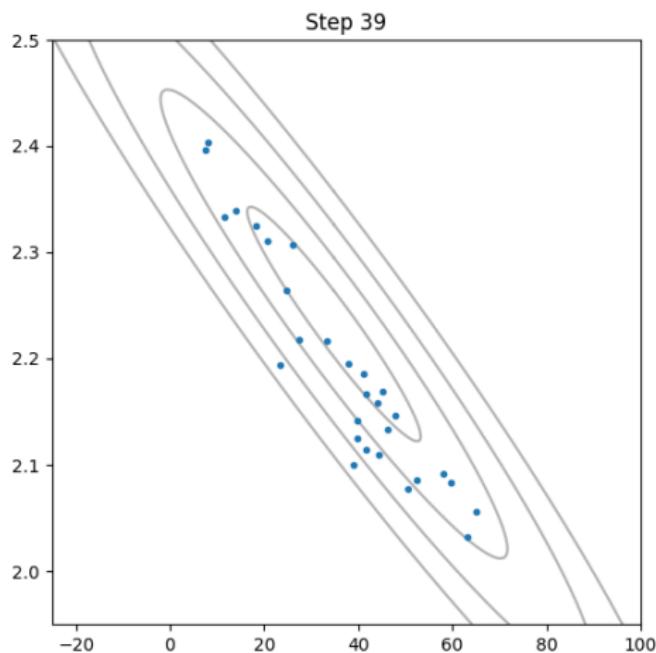
Emcee demo



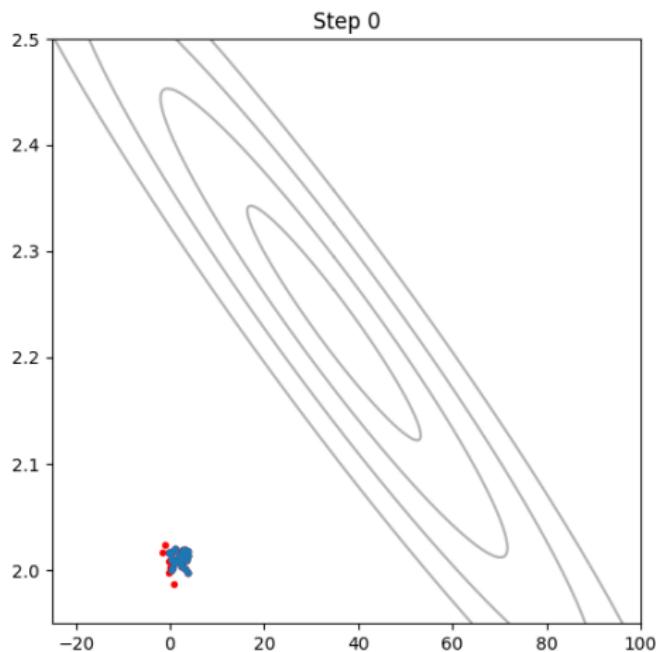
Emcee demo



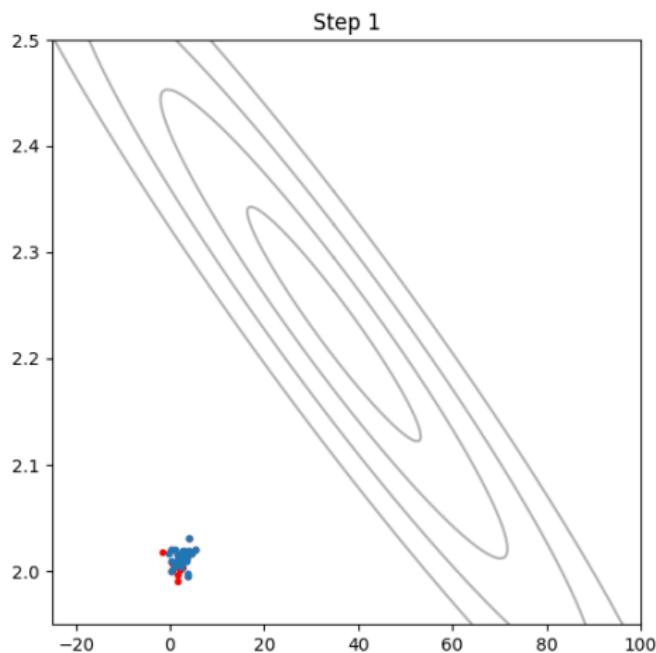
Emcee demo



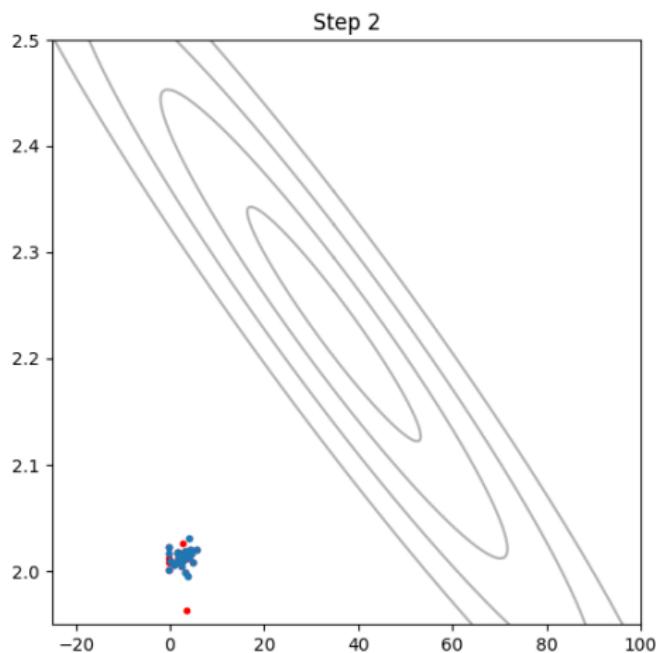
Emcee demo



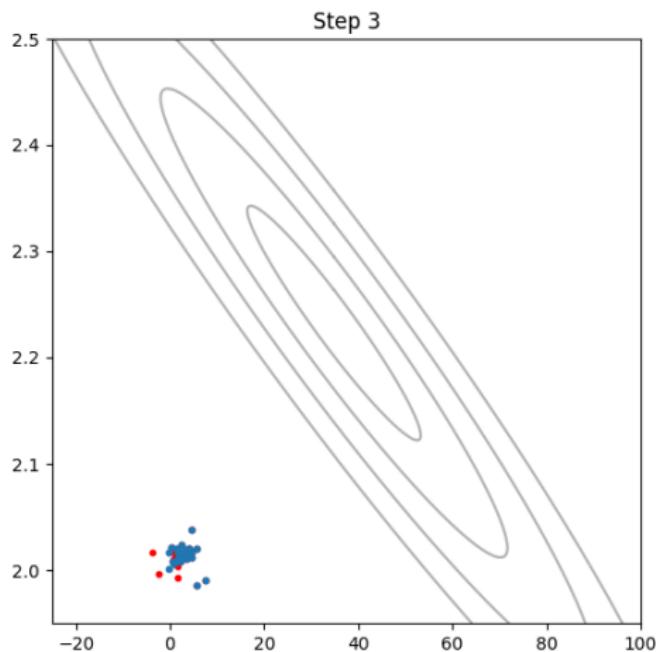
Emcee demo



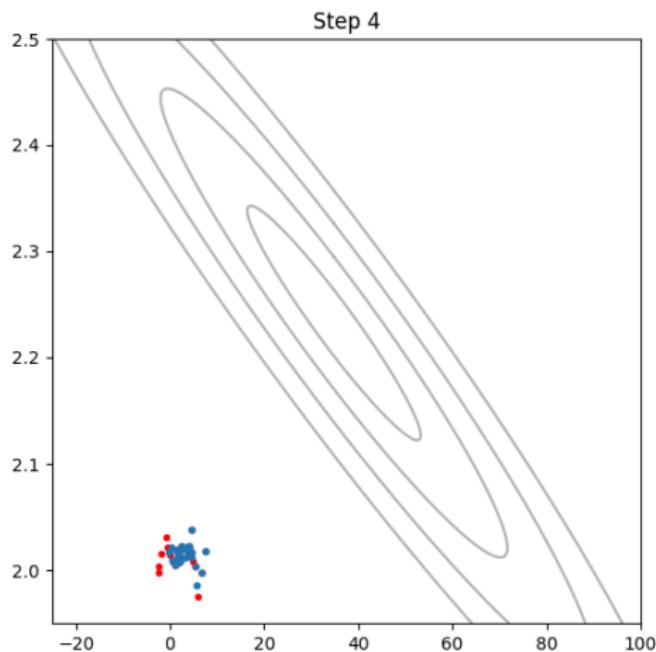
Emcee demo



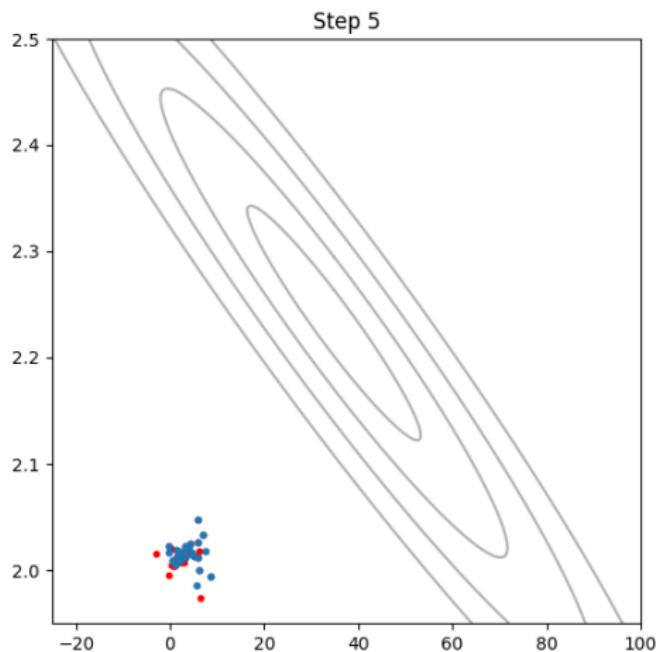
Emcee demo



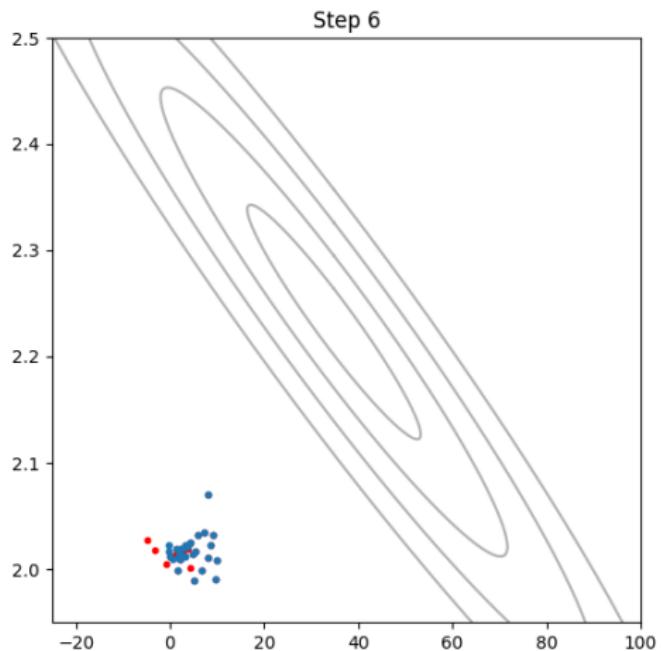
Emcee demo



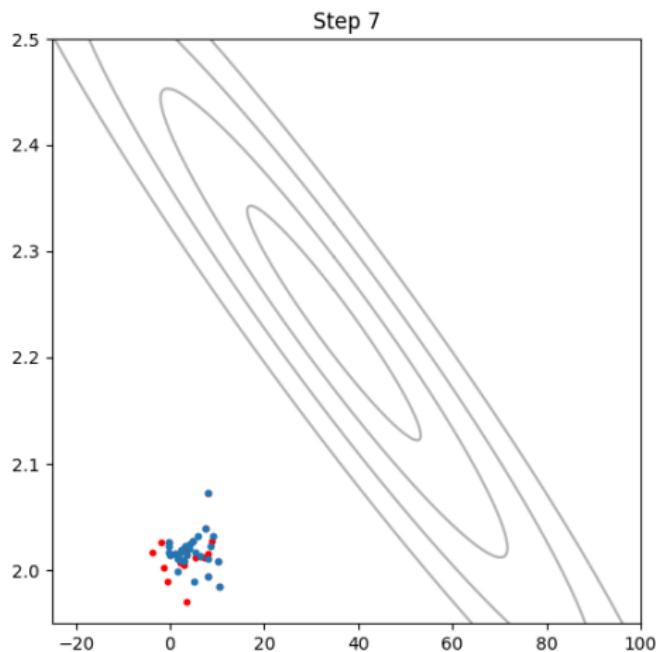
Emcee demo



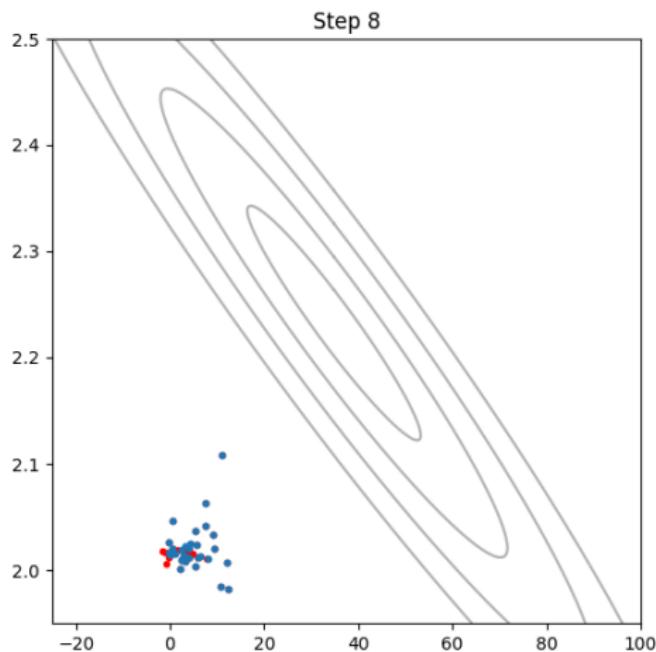
Emcee demo



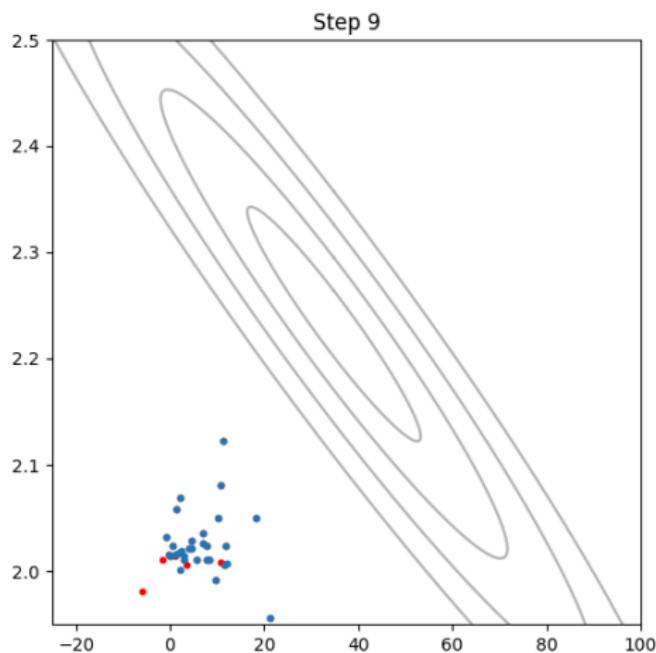
Emcee demo



Emcee demo



Emcee demo



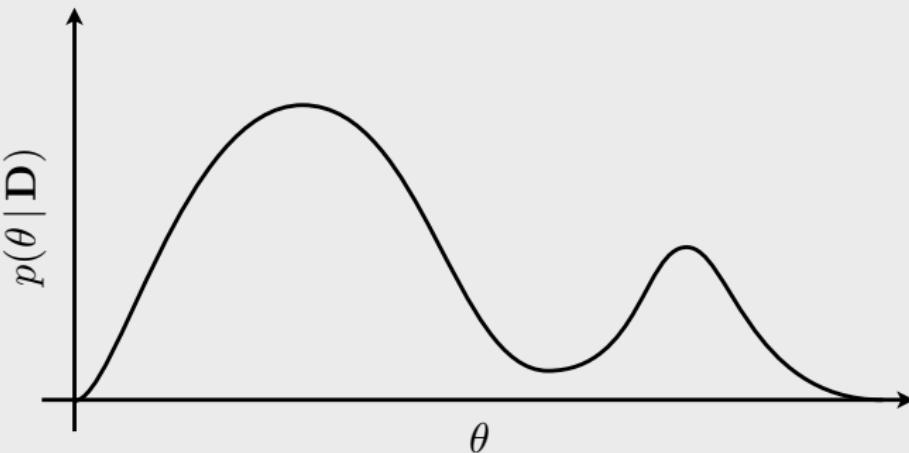
Remember:

**emcee isn't
always
The Right Choice™**

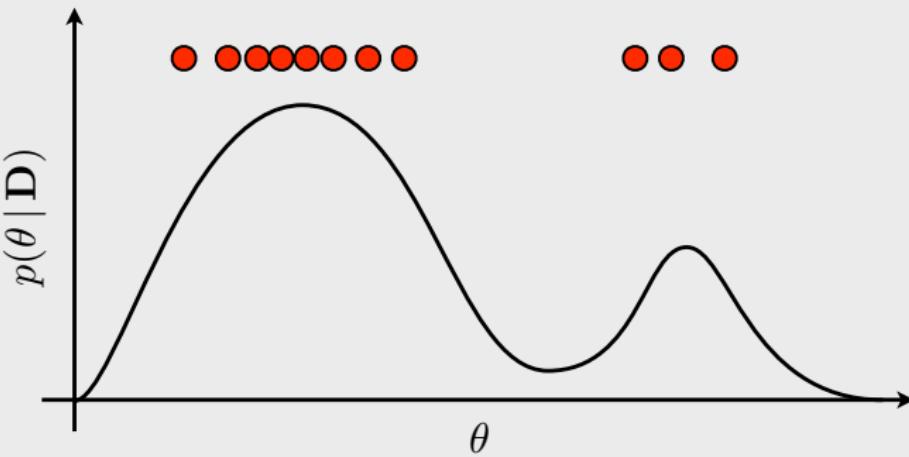


Brendon Brewer's words of wisdom...

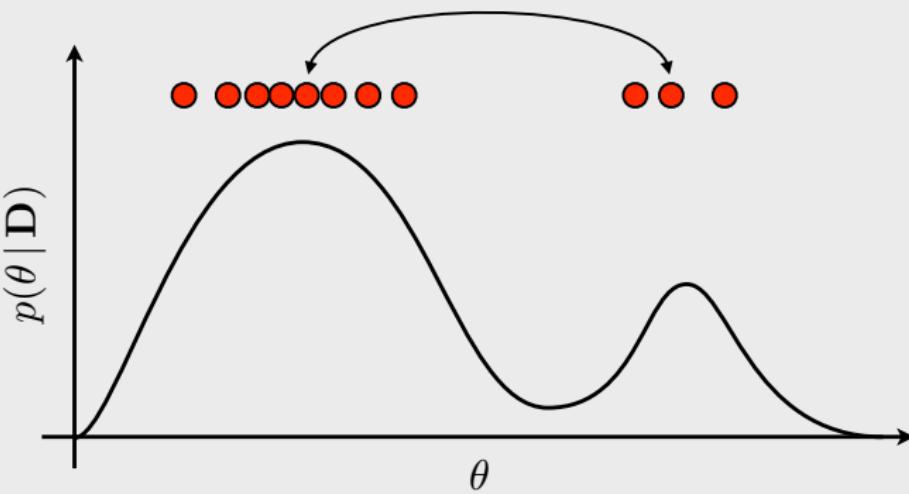
what about multimodal densities?



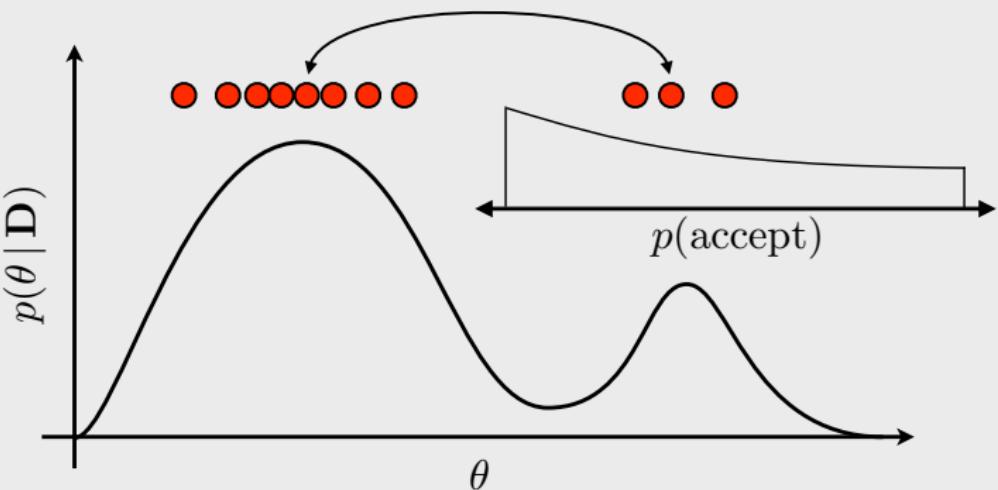
what about multimodal densities?



what about multimodal densities?



what about multimodal densities?



What does it mean?

$$p(\theta | \mathbf{D})$$



B O M B C A T

θ



Nested Sampling

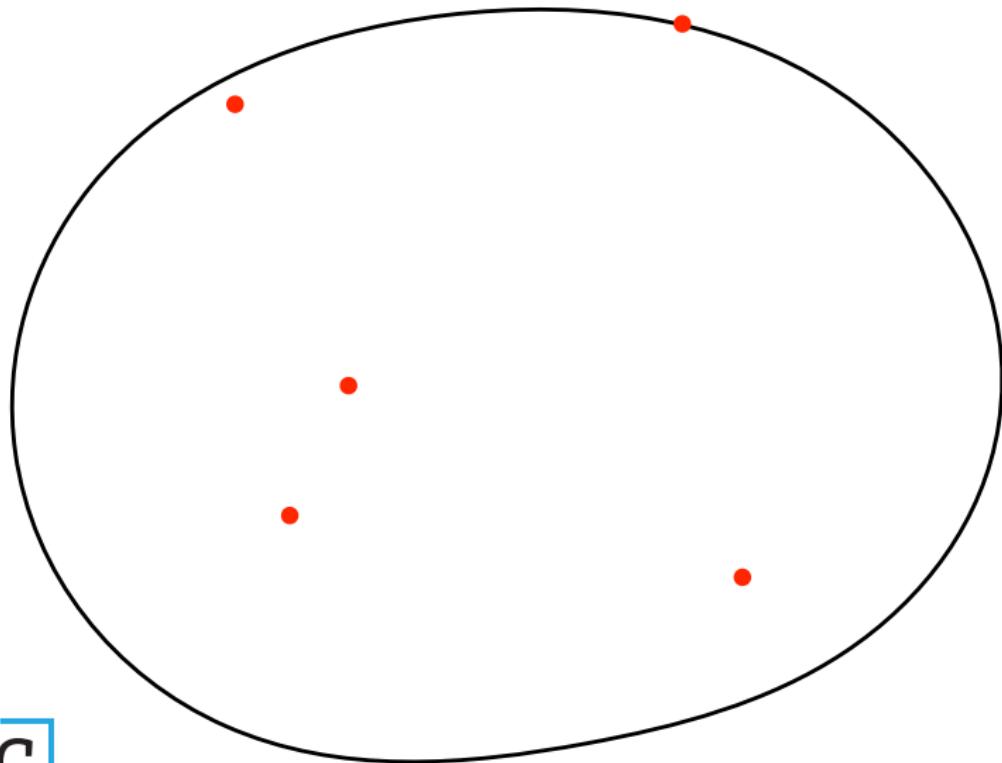
- ▶ An alternative to MCMC (and emcee)
- ▶ (Originally built for integrating the *evidence*)
- ▶ Idea: successively resample points, demanding they be above a minimum likelihood level
- ▶ Think of slowly raising the water level to find terrain contours

five live points chosen uniformly from prior



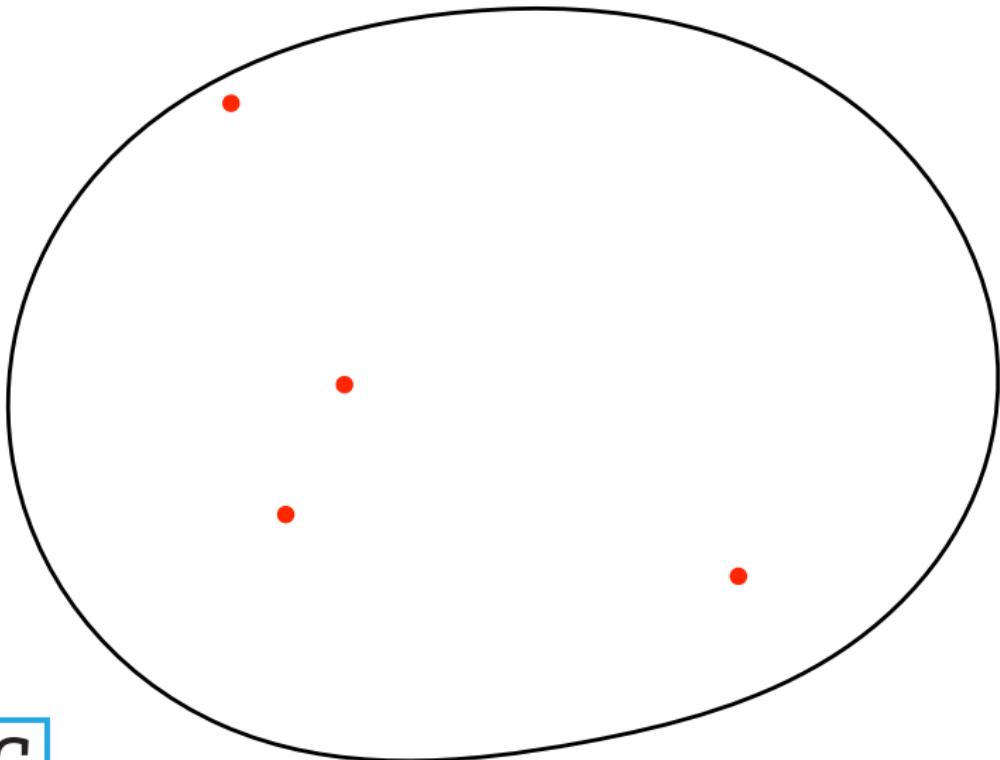
ICIC

Define likelihood contour from lowest point



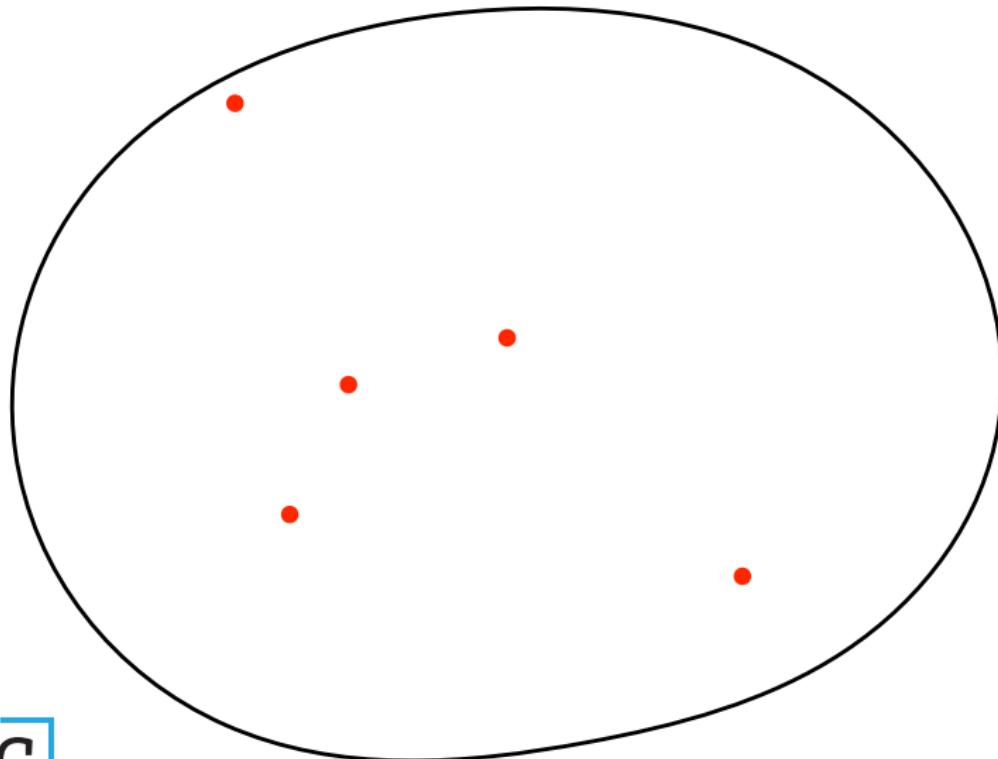
ICIC

Delete that point (storing its value)



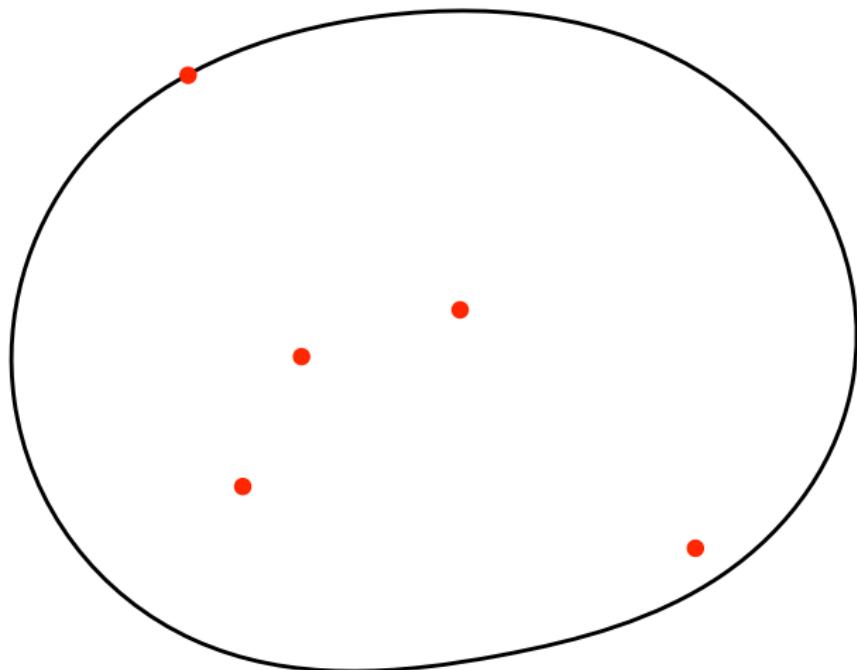
ICIC

Select a new point uniformly sampled subject to requirement $L(X_{\text{new}}) > L(X_{\text{old}})$

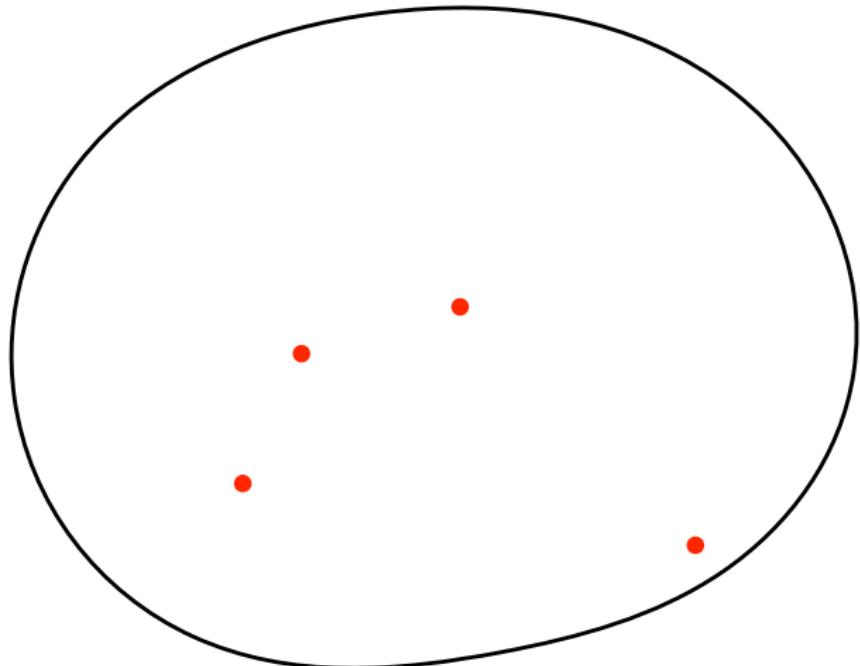


ICIC

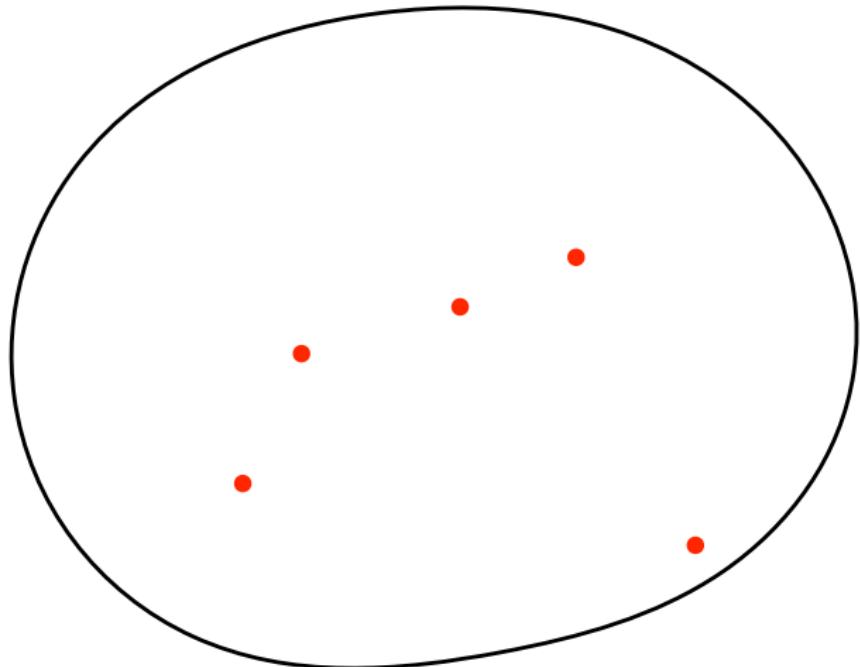
Iterate - contour shrinks by $X \sim \exp(1/n)$



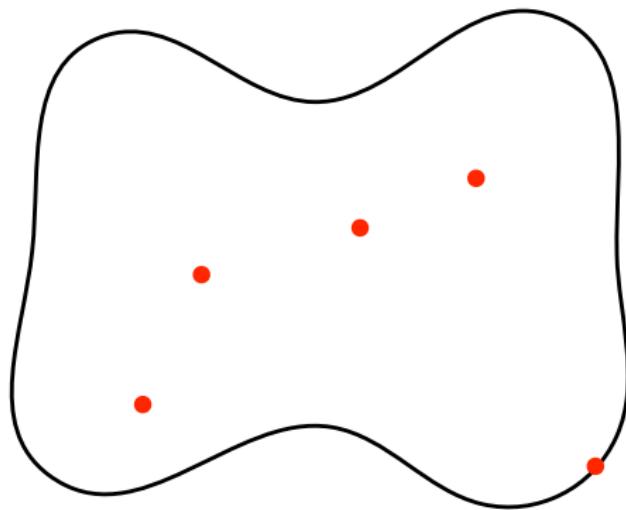
ICIC



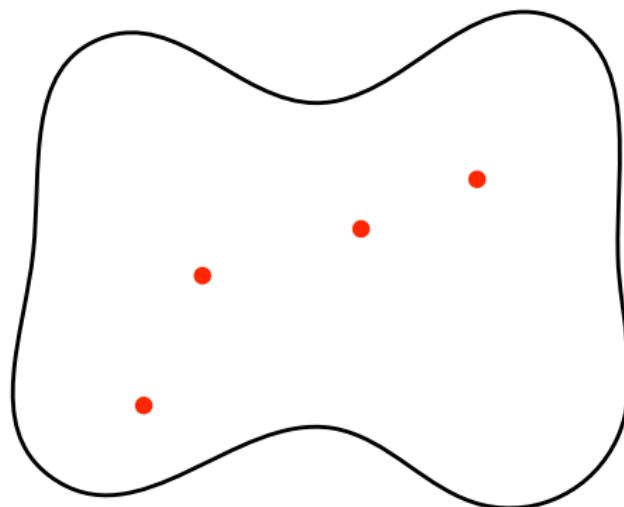
ICIC



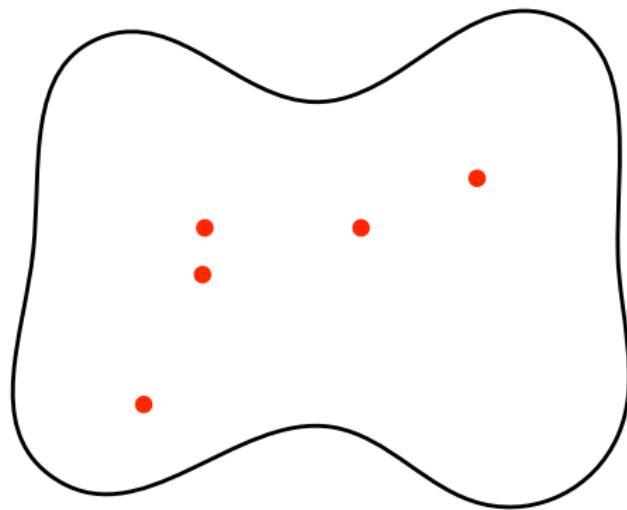
ICIC



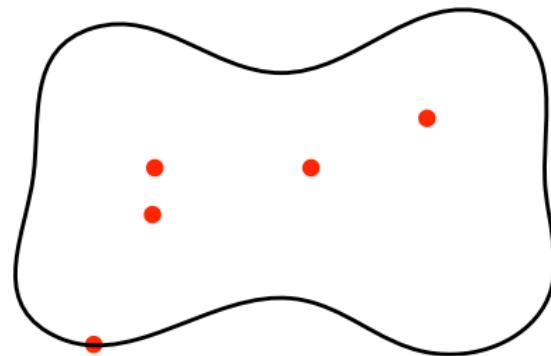
ICIC



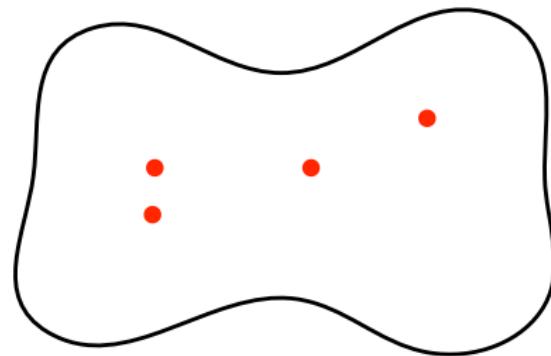
ICIC



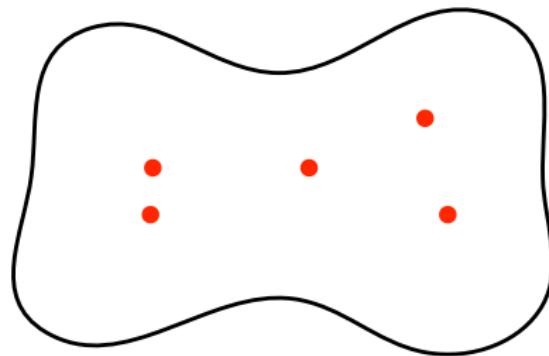
ICIC



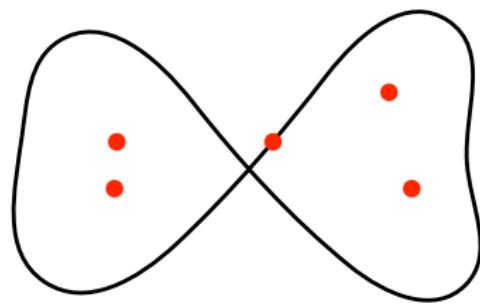
ICIC



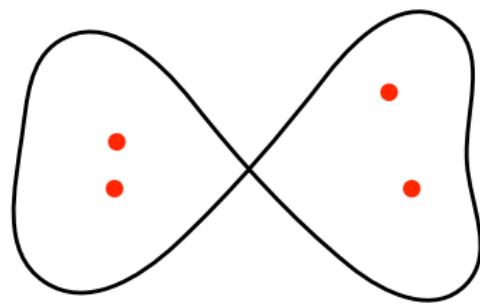
ICIC



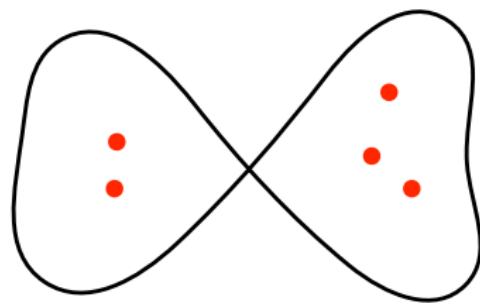
ICIC



ICIC



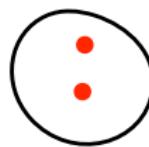
ICIC



ICIC



ICIC



ICIC

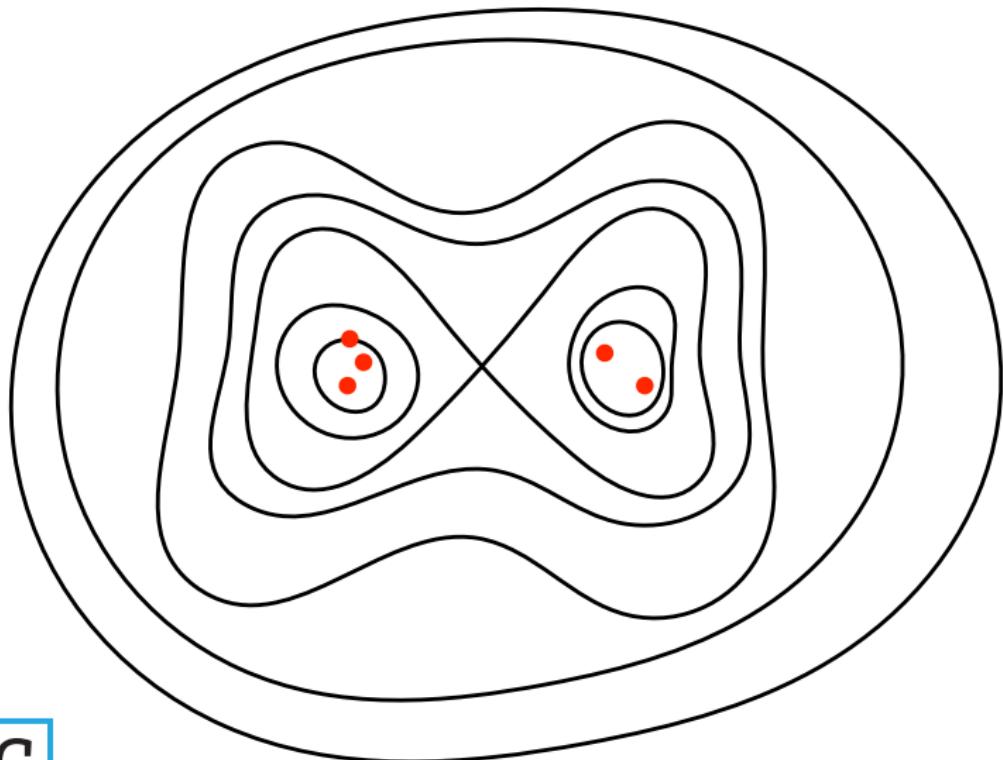


ICIC

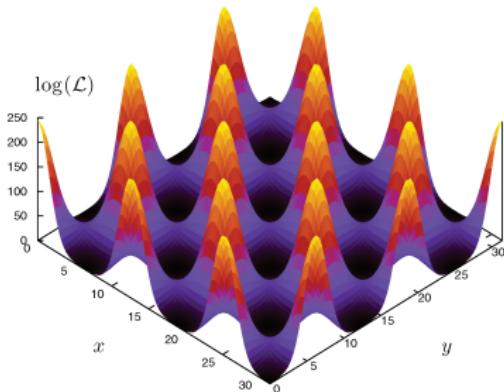


ICIC

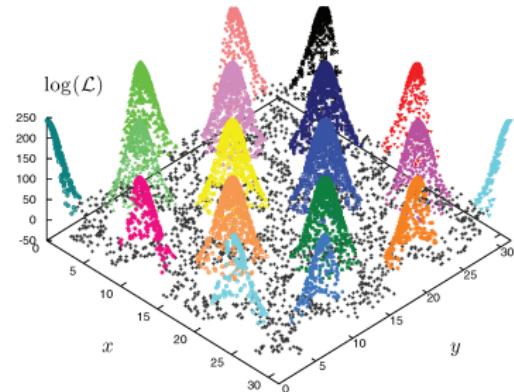
Successive points map likelihood in ordered way



ICIC

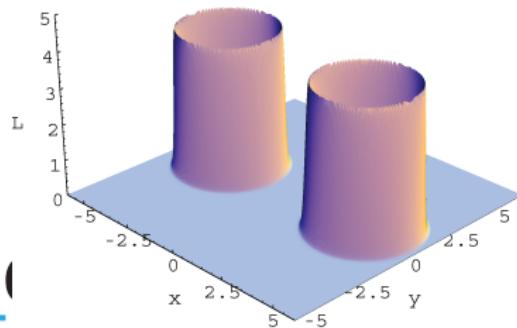


(a)

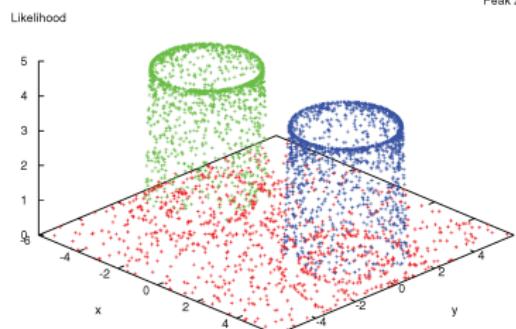


(b)

MultiNest



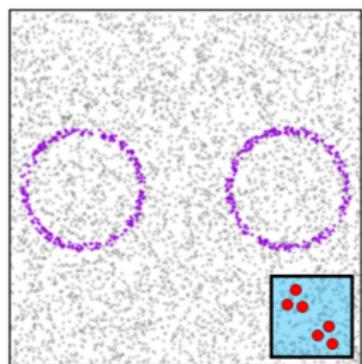
(a)



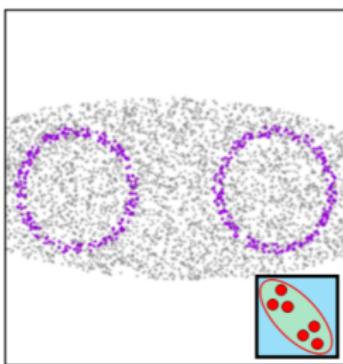
(b)

I

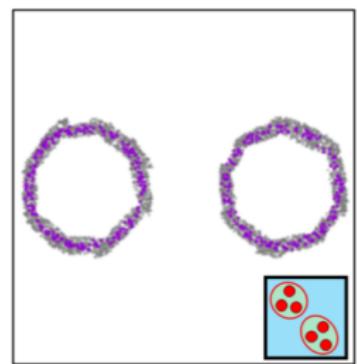
Dynamic Nested Sampling



Unit Cube
(no bound)



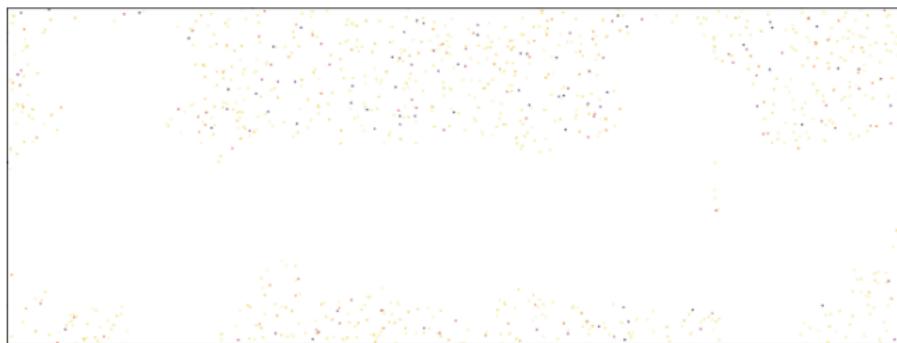
Single
Ellipsoid



Multiple
Ellipsoids

Credit: Joshua Speagle, *dynesty* code

Dynamic Nested Sampling



<https://dynesty.readthedocs.io/en/stable/>
<https://chi-feng.github.io/mcmc-demo/app.html>