ENM 360: Introduction to Data-driven Modeling

Lecture #25: Variational inference

Paris Perdikaris
November 24, 2020

Penn
UNIVERSITY *of* PENNSYLVANIA

Bayesian inference

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta|)p(\theta)}{p(\mathcal{D})} = \frac{p(\mathcal{D}|\theta|)p(\theta)}{\int p(\mathcal{D}|\theta|)p(\theta)d\theta}$$

<u>MCMC:</u> cleverly construct a Markov Chain whose stationary distribution approximates $p(\theta|\mathcal{D})$

- It's asymptotically exact, meaning that if we run the chain long enough, we should get samples from the posterior distribution

- Tuning MCMC can be tricky

- In practice requires thousands of samples, and each draw involves multiple (likely tens, perhaps hundreds) of evaluations of the un-normalised log posterior.

- This is fine for many models, but sometimes it just takes too long.

- It is hard to scale to large data-sets and models with many parameters and singular posteriors (e.g. deep learning).

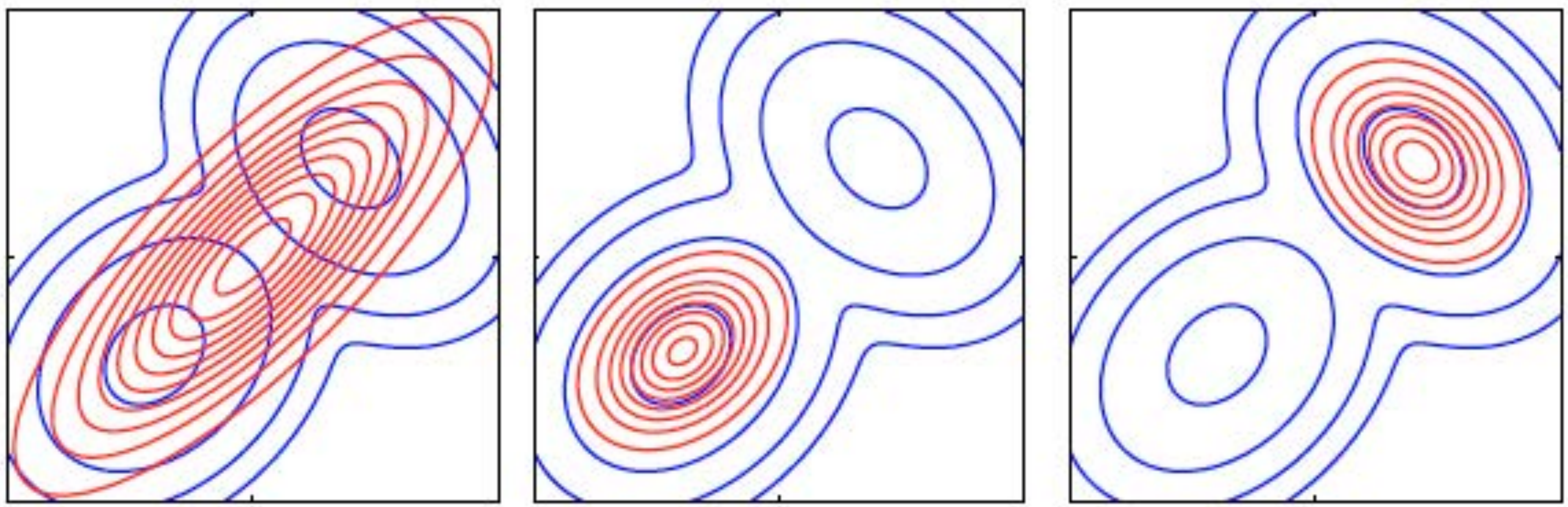
Variational inference

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta|)p(\theta)}{p(\mathcal{D})} = \frac{p(\mathcal{D}|\theta|)p(\theta)}{\int p(\mathcal{D}|\theta|)p(\theta)d\theta}$$

- Variational inference is another way to do Bayesian inference.
- The idea is that we'll approximate the posterior distribution with a family of distributions that is easy to work with.
- It will provide us with a set of tools for transforming the sampling problem (integration) to an optimization problem, that can be scaled to large models (i.e. with many parameters) and large data-sets.
- It also tends to favor approximations that underestimate the variance, and it usually will result in approximate distributions that get the means right but underestimate the variance.

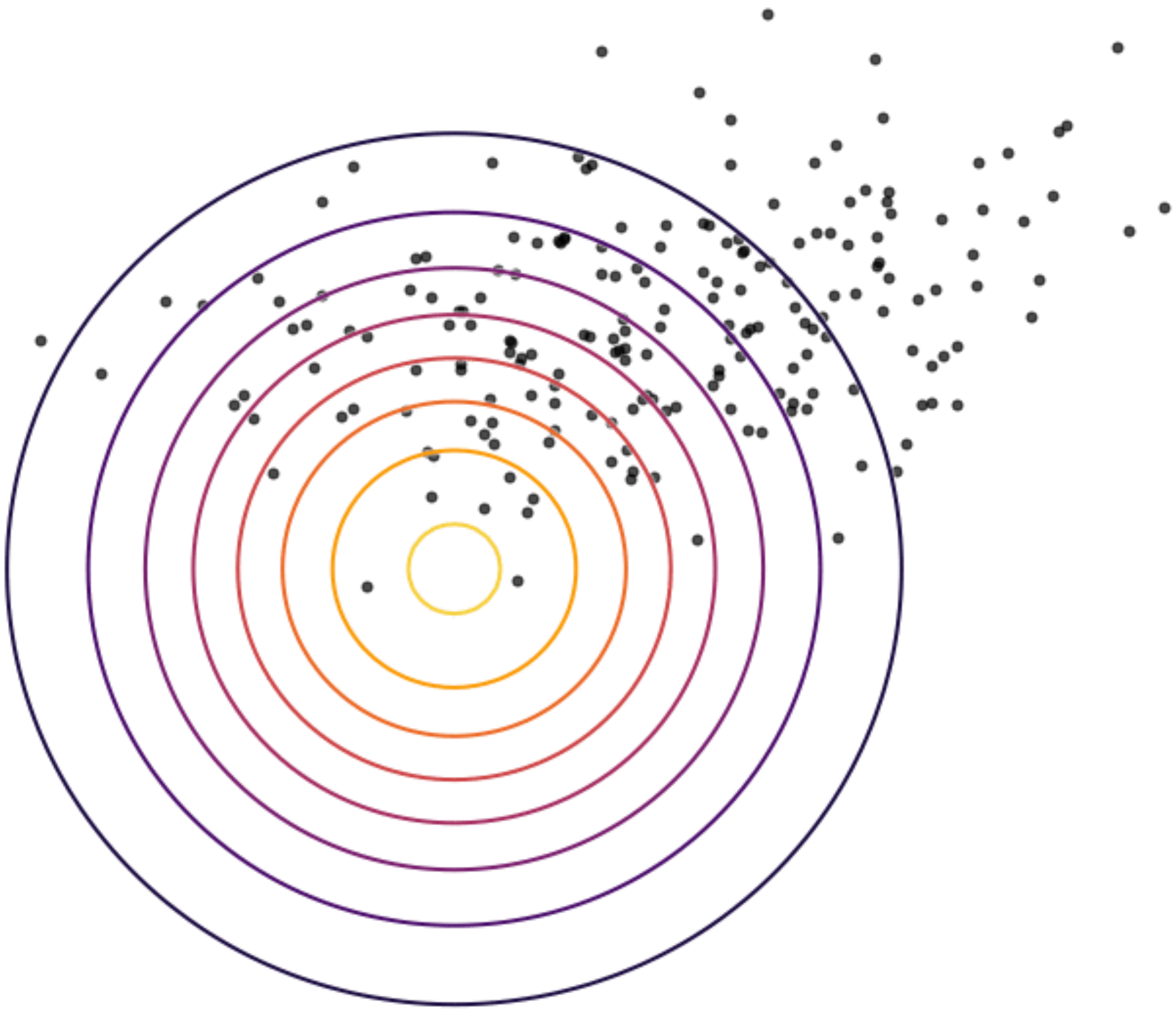
Variational inference

$\mathbb{KL}[q_\phi(x)||p(x)]$ $\mathbb{KL}[p(x)||q_\phi(x)]$ $\mathbb{KL}[p(x)||q_\phi(x)]$



KL Reverse KL Reverse KL

Variational inference



Probabilistic programming

Stan

http://mc-stan.org/

PyMC3

https://github.com/pymc-devs/pymc3

Edward

http://edwardlib.org/

https://github.com/uber/pyro