Monte Carlo vs Variational Inference

Monte Carlo

$$\mathbb{E}_{p(x)}\, f(x) \approx \frac{1}{M}\sum_{s=1}^{M} f(x_s)$$

$$x_s \sim p(x)$$

Unbiased estimate (larger M => better accuracy)

$$\mathbb{E}_{p(x)} \frac{1}{M}\sum_{s=1}^{M} f(x_s) = \mathbb{E}_{p(x)}\, f(x)$$

Monte Carlo vs Variational Inference

Monte Carlo

$$\mathbb{E}_{p(x)} f(x) \approx \frac{1}{M} \sum_{s=1}^{M} f(x_s)$$

$$x_s \sim p(x)$$

Unbiased estimate (larger M => better accuracy)

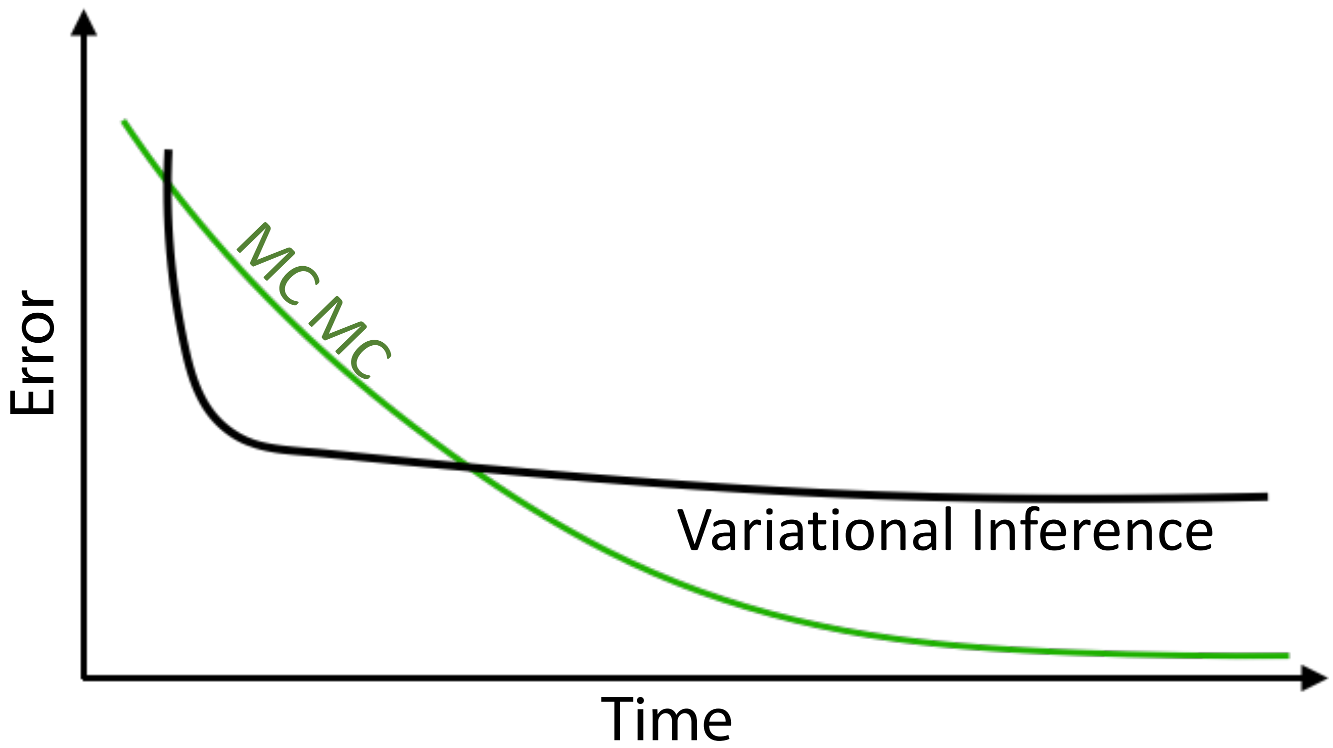
Variational Inference (week 3)

$$p(x) \approx q(x)$$

$$\mathbb{E}_{p(x)} f(x) \approx \mathbb{E}_{q(x)} f(x)$$

Monte Carlo vs Variational Inference

Schematic illustration



NOTE: turns out variational inference is the better choice if we have limited time and resources.
However, with more abundance of this, MCMC fares better as we reach the limit of the actual expected value(s) from the monte carlo method.

Methods

Best

- Full inference $p(T, \theta | X)$
  - Mean field $q(T)q(\theta) \approx p(T, \theta | X)$
  - MCMC $T_s, \Theta_s \sim p(T, \Theta \mid X)$
- EM algorithm $q(T),\ \theta = \theta_{\mathrm{MP}}$
  - Variational EM $q_1(T_1) \ldots q_d(T_d),\ \theta = \theta_{\mathrm{MP}}$
  - MCMC EM $T_s \sim p(T \mid \Theta, X), \Theta = \Theta_{\mathrm{MP}}$

Worst

Summary of Markov Chain Monte Carlo

Pros

- Easy to implement
- Easy to parallelize
- Unbiased estimates (wait more => more accuracy)

Cons

- Usually slower than alternatives