Course Overview

→ Bayesian Inference

→ Markov Chain Monte Carlo

- Tempering
- Adaptive MCMC
- HMC
- Unbiased MCMC

Bayesian Statistics

data: x viewed as a realization of X

model: $\{ f_\theta : \theta \in \Theta \}$

Assume: $X \sim f_\theta$

Q: Is there some true fixed parameters θ that generated the model?

Bayesian: View initially $\theta \sim \pi$ then

update $\pi(d\theta | x) \propto f_\theta(x)\pi(d\theta)$

Advantages

- full dist for inference
- incorporated prior info

Disadvantage

- Prior misspecification
- Computationally difficult

MCMC

- best tool to sample from
  $\pi(d\theta | x)$
- Ravi Kannan et al (1991)

showed mcmc much better than any deterministic method.

<u>Rmk</u>: Pseudo-Bayesian Inference

$$\tilde{f}_\theta(x) \Rightarrow \tilde{\pi}(d\theta|x) = \tilde{f}_\theta(x)\,\pi(d\theta)$$

$\tilde{f}_\theta(x)$ doesn't necessarily need to be a density

<u>Ex</u>: $(X_1, \dots, X_n) = X \sim N(0, \theta^{-1})$

$$\pi(\theta|X) \propto \exp\left\{-\frac{1}{2}\operatorname{Tr}(X\theta) - \frac{1}{2}\log\det\theta\right\}\pi(\theta)$$

Approximate likelihood by

$$\prod_{i=1}^{d} f_{\theta_i}(x_i \mid x_{[-i]})$$

$$\tilde{\pi}(\theta \mid x) \propto \prod_{i=1}^{d} f_{\theta_i}(x_i \mid x_{[-i]})\, \pi(\theta)$$

Markov Chain Monte Carlo

Let π be a measure on $\mathcal{X}(\mathbb{R}^p)$

we want to produce samples from π.

MCMC: design MC s.t.

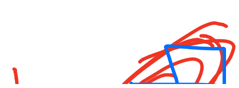
$$\mathbb{P}(X_n \in A) \longrightarrow \pi(A)$$

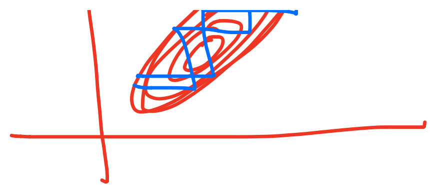
$\forall A \in \mathcal{F}_{\mathcal{X}}$ measurable sets.

Many ways to build these

- Gibbs Sampler (Coordinate desc.)

Could be bad. - high correlation





– Metropolis-Hastings (Gradient descent)

Take $Q(x,y)$ another kernel

Set $\alpha(x,y) = \text{Min}\left\{1, \dfrac{\pi(y)\, Q(y,x)}{\pi(x)\, Q(x,y)}\right\}$

$$P(x,dy) = \alpha(x,y)\, Q(x,y)\, dy + (1-a(x,y))\, \delta_x(dy)$$

$$a(x) = \int_{\mathcal{X}} \alpha(x,y)\, Q(x,y)\, dy$$

Q: How do we choose Q ?