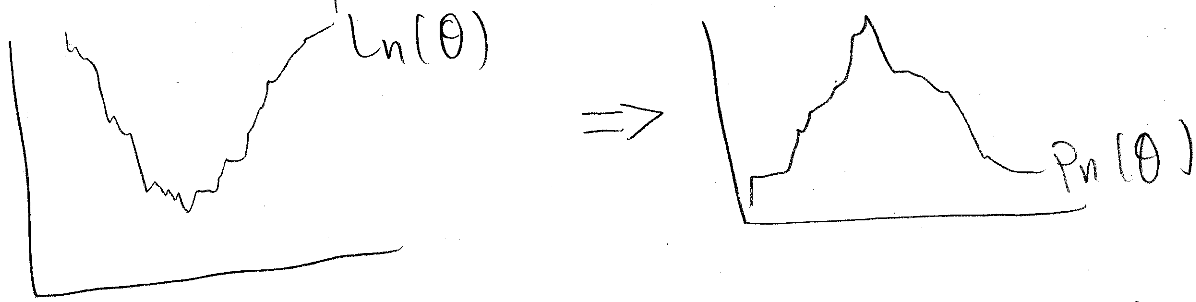
Recall:

$L_n(\theta) = -\frac{n}{2} g_n(\theta)' W(\theta) g_n(\theta)$ - criterion function

$p_n(\theta) = C \exp\{L_n(\theta)\}$ - pseudo posterior



Can motivate $p_n(\theta)$ as out come of Bayesian learning.

(1) $\hat{\theta} = \int \theta p_n(\theta) d\theta$ • Bayesian estimator

$\sqrt{n}(\hat{\theta} - \theta^*) \xrightarrow{P} 0 \Rightarrow \hat{\theta}$ and θ^* are asymptotically equivalent

Can compute integral (1) by MCMC approach. Draw $S = (\theta^{(1)}, \ldots, \theta^{(k)})$

Then $\frac{1}{n} \sum_{i=1}^{n} \theta^{(i)} \to \hat{\theta}$.

• can use Bayesian posterior for inference if information equality holds.

Defn Metropolis (random walk): Given quasi-posterior density $p_n(\theta)$, known up to a constant, we can generate $(\theta^{(0)}, \dots, \theta^{(T)})$

Charles Geyer was a good program for this

1. Choose some starting value $\theta^{(0)}$
2. Generate $\xi^{(j)} = \theta^{(j)} + \eta^{(j)}$, $\eta^{(j)} \sim N(0, \sigma^2 I)$
3. Update $\theta^{(j+1)}$ from $\theta^{(j)}$, $j = 1, \dots, T-1$, using

$$\theta^{(j+1)} = \begin{cases} \xi^{(j)} & \text{with probability } \rho(\theta^{(j)}, \xi^{(j)}) \\ \theta^{(j)} & \text{with probability } 1 - \rho(\theta^{(j)}, \xi^{(j)}) \end{cases}$$

$$\text{where } \rho(x,y) = \min\left\{ \frac{p_n(y)}{\underbrace{p_n(x)}_{\text{constant of integration cancels out here.}}}, 1 \right\}$$

Implication:

$$\frac{1}{B} \sum_{i=1}^{B} f(\theta^{(i)}) \xrightarrow{P} \int_\Theta f(\theta) p_n(\theta) d\theta$$

Simulated annealing

$$\lim_{\lambda \to \infty} \frac{\int_\Theta \theta e^{\lambda L_n(\theta)} \pi(\theta) d\theta}{\int_\Theta e^{\lambda L_n(\theta)} \pi(\theta) d\theta} = \underset{\theta \in \Theta}{\operatorname{argmax}} L_n(\theta)$$

In large samples, setting $\lambda = 1$ is "good enough."