Intro to MCMC

- What is Bayes Stat?

Statistics about uncertainty of the parameters given data.

y : the observed data.

θ : the parameters (the values we are interested in)

$L(y|\theta)$: the likelihood function

(the density function of y given θ)

$\pi(\theta)$: the prior function of θ.

$\boxed{\text{Posterior distribution of } \theta}$

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(y|\theta)\pi(\theta)}{p(y)} \propto p(y|\theta)\pi(\theta)$$

- The uncertainty of θ is measured by the posterior distribution of θ.

If we know the posterior distribution of θ, we fully specify the uncertainty of θ.

1. When we know the partial density function of the distribution.

ex)

$$p(x) \propto \exp\left\{-\frac{x^2}{2}\right\} \Rightarrow N(0, 1)$$

$$p(x) \propto \exp\left\{-\frac{(x-2)^2}{4}\right\} \Rightarrow N(2, 2)$$

$$p(x) \propto \exp\{-x\} \Rightarrow \exp(1)$$

2. What if the density function is complex ?

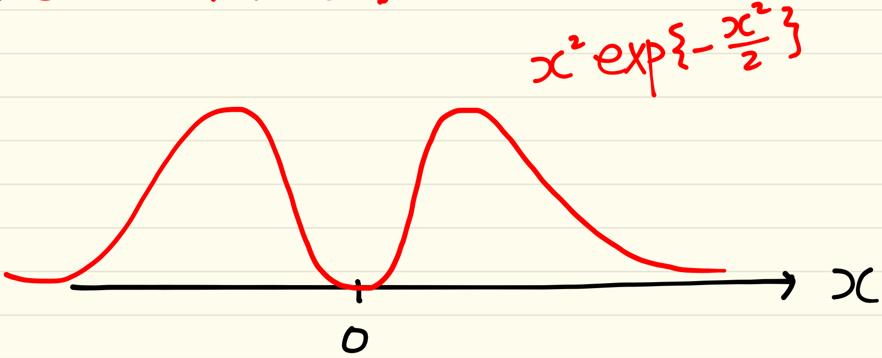
ex)

$$p(x) \propto x^2 \exp\left\{-\frac{x^2}{2}\right\}$$

2. What if the density function is complex ?

ex) $p(x) \propto x^2 \exp\{-\frac{x^2}{2}\}$?

We can draw it !



3. What if the dimension of θ is greater than 2 ?

It's impossible to draw the density function.

4. Alternatively, What if we can generate a billion samples from the distribution ?

Problems Solved !

- Why do we need MCMC?

To generate random samples from the complex posterior distribution of θ

- What is MCMC?

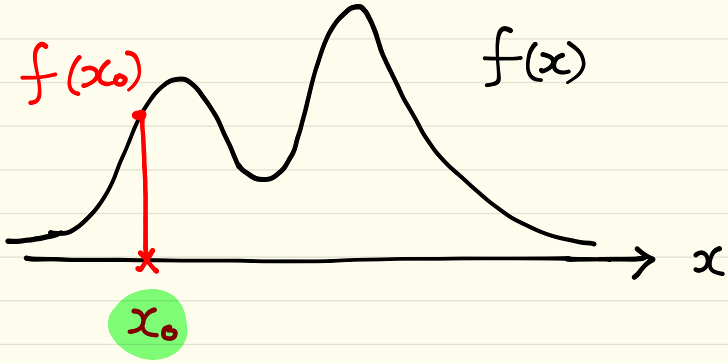
Markov Chain Monte Carlo.

- What is Markov Chain Monte Carlo?

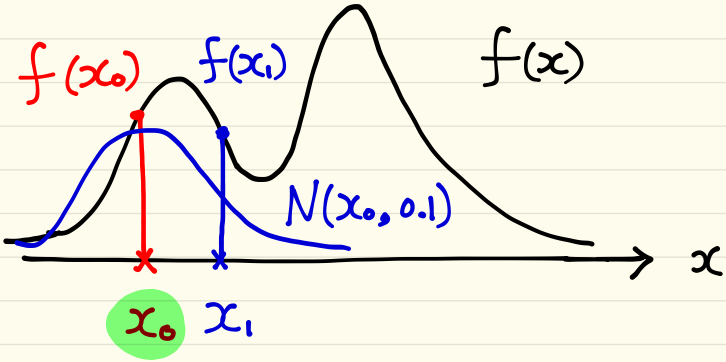
↳ random number generation

MCMC : random number generation using Markov Chain.

- Intuition of MCMC (Metropolis-Hastings Algorithm)

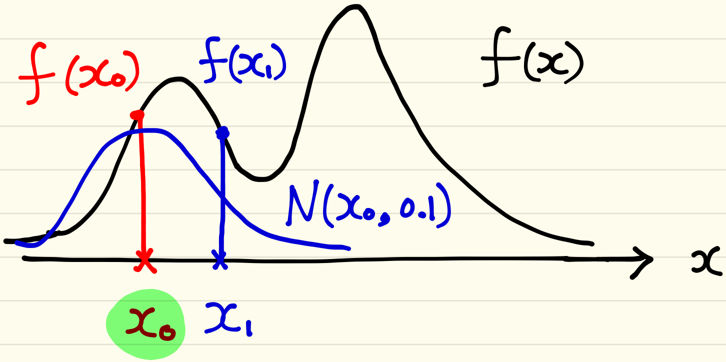


- Intuition of MCMC



1. Generate x_1 based on x_0 ; ex) $x_1 | x_0 \sim N(x_0, 0.1)$

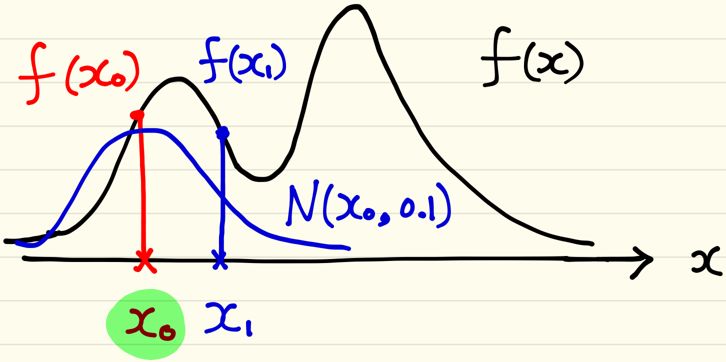
- Intuition of MCMC



1. Generate x_1 based on x_0 ; ex) $x_1 | x_0 \sim N(x_0, 0.1)$
2. Can we say x_1 is from $f(x)$?

   Accept x_1 or Reject x_1?

• Intuition of MCMC

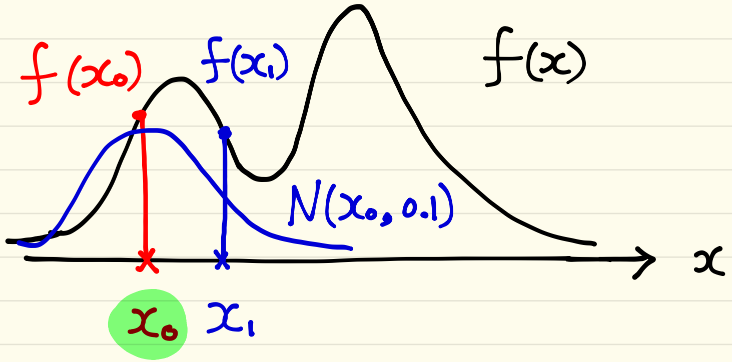


1. Generate x_1 based on x_0 ; ex) $x_1 | x_0 \sim N(x_0, 0.1)$
2. Can we say x_1 is from $f(x)$?

   Accept x_1 or Reject x_1 ?

3. If $f(x_1)$ is large enough compared to $f(x_0)$, Accept x_1 !

- Intuition of MCMC

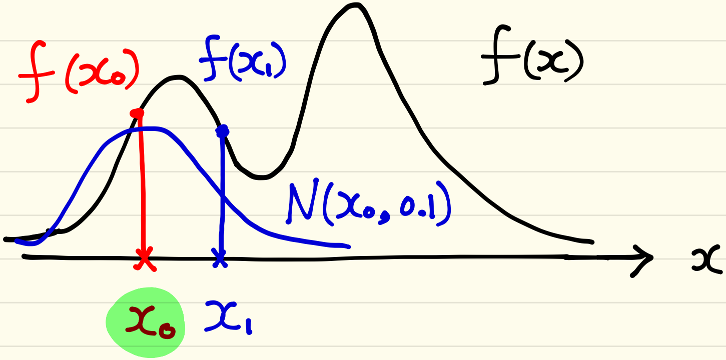


1. Generate x_1 based on x_0 ; ex) $x_1 | x_0 \sim N(x_0, 0.1)$
2. Can we say x_1 is from $f(x)$?

   Accept x_1 or Reject x_1 ?

3. If $f(x_1)$ is large enough compared to $f(x_0)$, Accept x_1 !
4. Go to step 1. by replacing x_0 with x_1, if accepted.

- Formal description of Metropolis-Hastings Alg



proposal dist.
$q(x_1|x_0)$

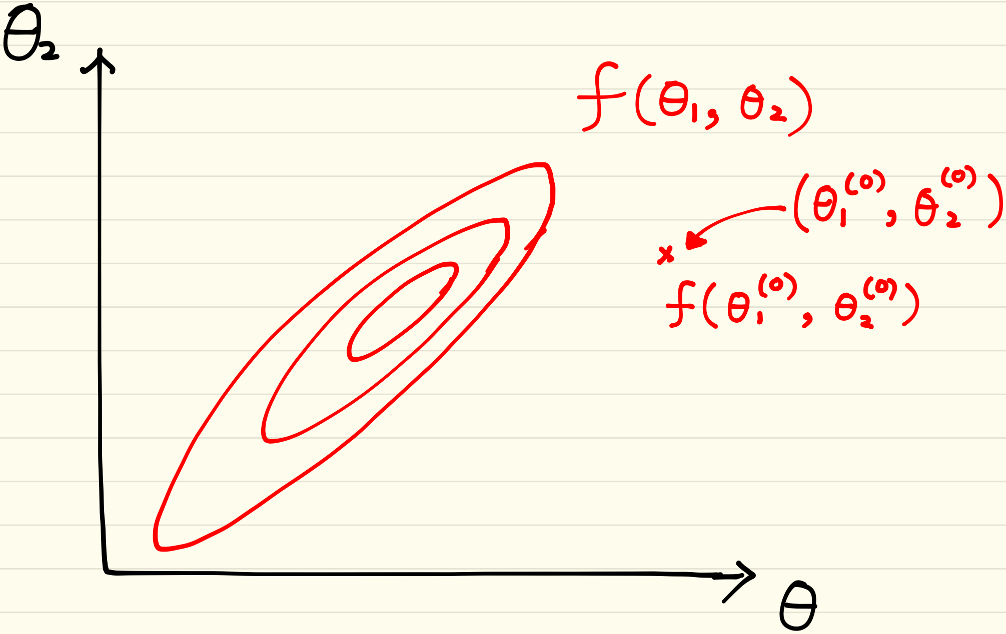
1. Generate x_1 based on x_0; ex) $x_1|x_0 \sim \boxed{N(x_0, 0.1)}$
2. Can we say x_1 is from $f(x)$?

   Accept x_1 or Reject x_1?

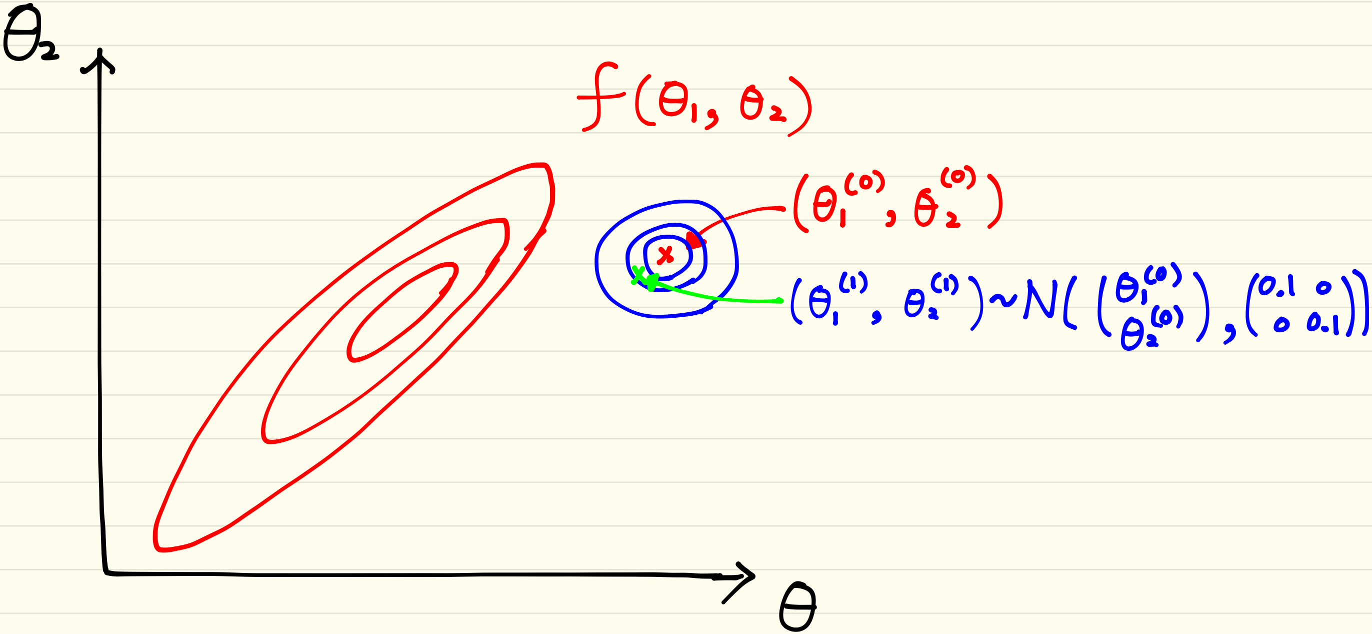
   If $\dfrac{f(x_1)\,q(x_0|x_1)}{f(x_0)\,q(x_1|x_0)} > u$, $u \sim \text{Unif}(0,1)$

3. If $f(x_1)$ is large enough compared to $f(x_0)$, Accept x_1!
4. Go to step 1. by replacing x_0 with x_1, if accepted.

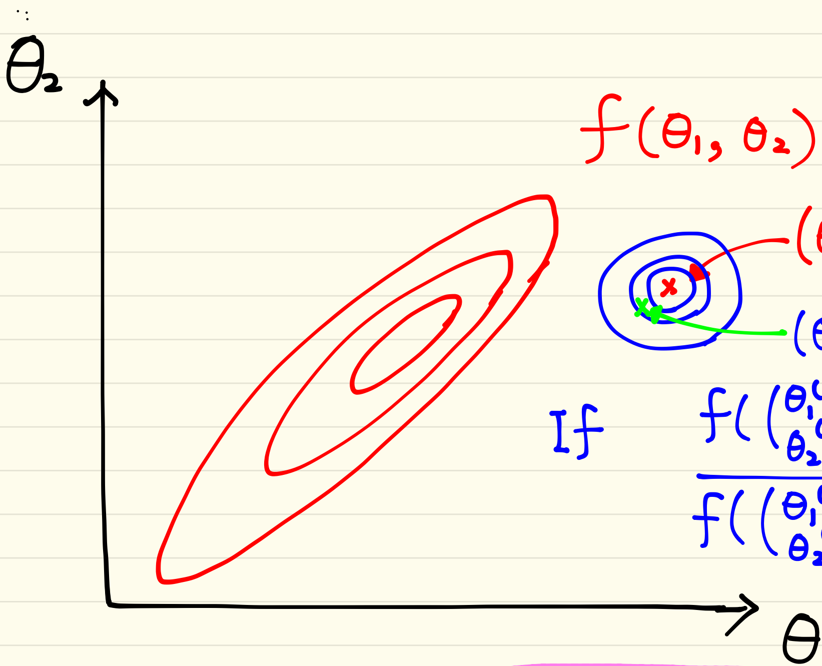
Examples in 2 dim.



Examples in 2 dim.



Examples in 2 dim.



$(\theta_1^{(0)}, \theta_2^{(0)})$

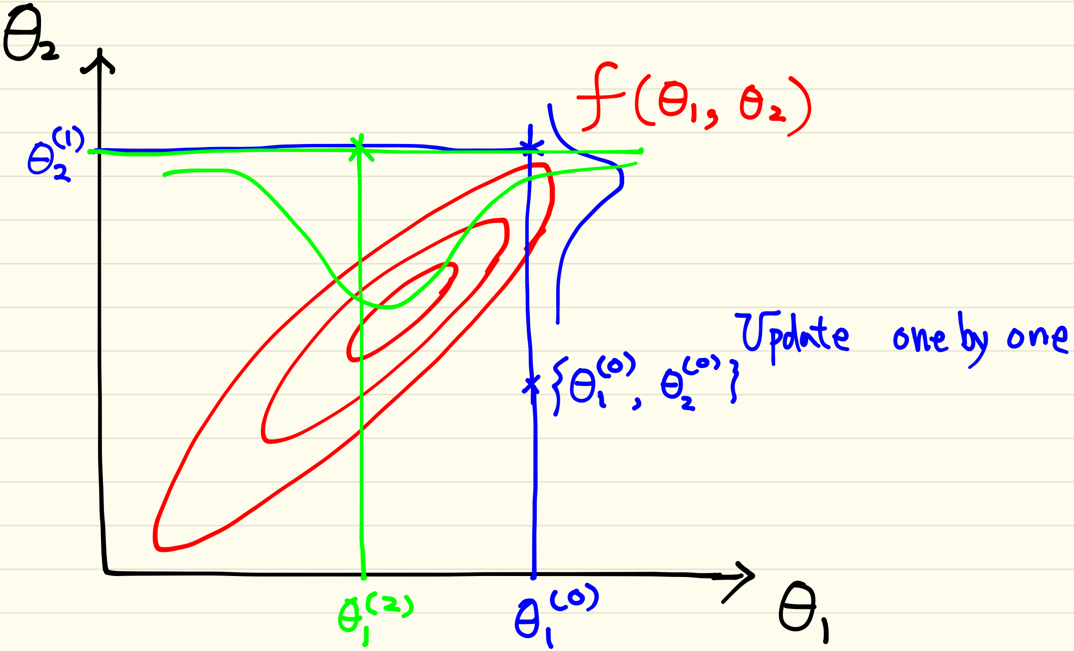
$(\theta_1^{(1)}, \theta_2^{(1)}) \sim N\left(\begin{pmatrix} \theta_1^{(0)} \\ \theta_2^{(0)} \end{pmatrix}, \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix} \right)$

If $\dfrac{f\left(\begin{pmatrix} \theta_1^{(1)} \\ \theta_2^{(1)} \end{pmatrix}\right)}{f\left(\begin{pmatrix} \theta_1^{(0)} \\ \theta_2^{(0)} \end{pmatrix}\right)} > u$, where

$u \sim Unif(0,1)$

Jump to $(\theta_1^{(1)}, \theta_2^{(1)})$.

Examples in 2 dim. (Gibbs Sampler)



Intrinsic Scatter Model.

$$y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \ , \quad \text{where} \quad \varepsilon_i \sim N(0, \sigma_i^2 + \sigma^2)$$

for $i = 1, \cdots, n.$

parameter $\theta = \{\beta_0, \beta_1, \sigma^2\}.$

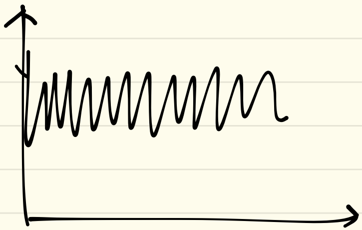
$$L(y_i \mid \theta) \sim N(\beta_0 + \beta_1 X_i, \ \sigma_i^2 + \sigma^2)$$

$$\beta_i \sim \text{Cauchy dist} ; \ \frac{1}{\pi} \frac{1}{\beta^2 + 1}$$

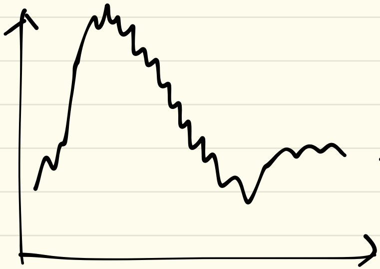
$$\sigma^2 \sim \text{Inv-Gamma}(a_0, b_0) ; \ (\sigma^2)^{-(a_0 - 1)} \exp\left\{-\frac{b_0}{\sigma^2}\right\}$$

Diagonostic of MCMC Chains.

① Trace Plot



→ O.K



→ No good

Diagonostic of MCMC Chains.

② Acceptance Rate.

Higher Acc rate → good