Sampling Based Inference

Il-Chul Moon
Dept. of Industrial and Systems Engineering
KAIST
icmoon@kaist.ac.kr

Weekly Objectives

- Learn basic sampling methods
  - Understand the concept of Markov chain Monte Carlo
  - Able to apply MCMC to the parameter inference of Bayesian networks
  - Know the mechanism of rejection sampling
  - Know the mechanism of importance sampling
- Learn sampling based inference
  - Understand the concept of Metropolis-Hastings Algorithm
  - Know the mechanism of Gibbs sampling
- Know a case study of sampling based inference
  - Understand the latent Dirichlet allocation model
  - Know the collapsed Gibbs sampling
  - Know how to derive Gibbs sampling formula for LDA

$q(z^t|z^*)P(z^*) < q(z^*|z^t)P(z^t)$ → Movement from z^t to z^* is too often

Metropolis-Hastings Algorithm

We want the stationary distribution, $\pi(z)$, of our MCMC sampling to be P(z)

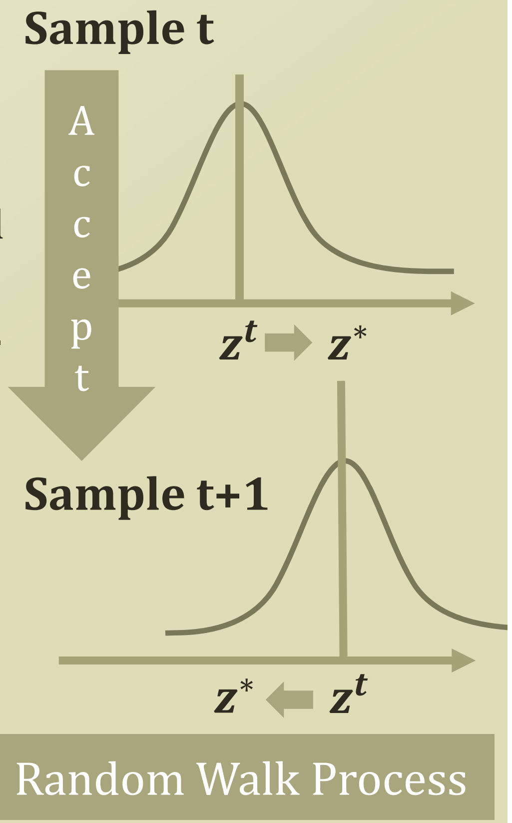
Reversible Markov chain
$\pi_i T_{i,j} = \pi_j T_{j,i}$

q is not well-designed to be the reversible MC, so we adjust by ***r***

- General algorithm of MCMC
  - Current value: z^t
  - Propose a candidate $z^* \sim q(z^*|z^t)$ where q_t is a proposal distribution
    - Same as importance and rejection samplings, yet the difference is the Markov property idea in the proposal distribution
  - With an acceptance probability, α
    - Accept → $z^{t+1} = z^*$
    - Reject → $z^{t+1} = z^t$
- Metropolis-Hastings algorithm
  - Given the general algorithm of MCMC
  - Consider a ratio, $r(z^*|z^t) = \frac{q(z^t|z^*)P(z^*)}{q(z^*|z^t)P(z^t)}$, we want this to be 1
    - $q(z^t|z^*)P(z^*)r_{z^* \to z^t} = q(z^*|z^t)P(z^t)r_{z^t \to z^*}$
    - $r(z^*|z^t) < 1 \to q(z^t|z^*)P(z^*) < q(z^*|z^t)P(z^t)$
      - Increase $r_{z^* \to z^t}$=1, degrease $r_{z^t \to z^*} = r(z^*|z^t)$
    - $r(z^*|z^t) > 1 \to q(z^t|z^*)P(z^*) > q(z^*|z^t)P(z^t)$
      - Decrease $r_{z^* \to z^t}$=$r(z^t|z^*)$, increase $r_{z^t \to z^*} = 1$
  - Acceptance probability $\alpha(z^*|z^t) = \min\{1, r(z^*|z^t)\}$

Random Walk M-H Algorithm

- $T_{t,*}^{MH} = q(z^*|z^t)\alpha(z^*|z^t)$
  - Transition probability to satisfy the balance equation with $\mathrm{P}(z)$ as the stationary distribution
  - $\alpha(z^*|z^t) = \min\{1, \frac{q(z^t|z^*)P(z^*)}{q(z^*|z^t)P(z^t)}\}$
  - Here, we already assumed, so far, the easy calculation of $P(z)$
  - What we miss is the definition of $q(z^*|z^t)$, but this is a proposal that any probability distribution can be
    - Surely, there are better and worse proposal probability distributions.
    - Choosing $q(z^*|z^t)$ determines the type of M-H algorithm
- Random walk M-H algorithm
  - $T_{t,*}^{MH} = q(z^*|z^t)\alpha(z^*|z^t)$
  - $z^* \sim N(z^t, \sigma^2)$
  - $q(z^*|z^t) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{(z^*-z^t)^2}{2\sigma^2})$



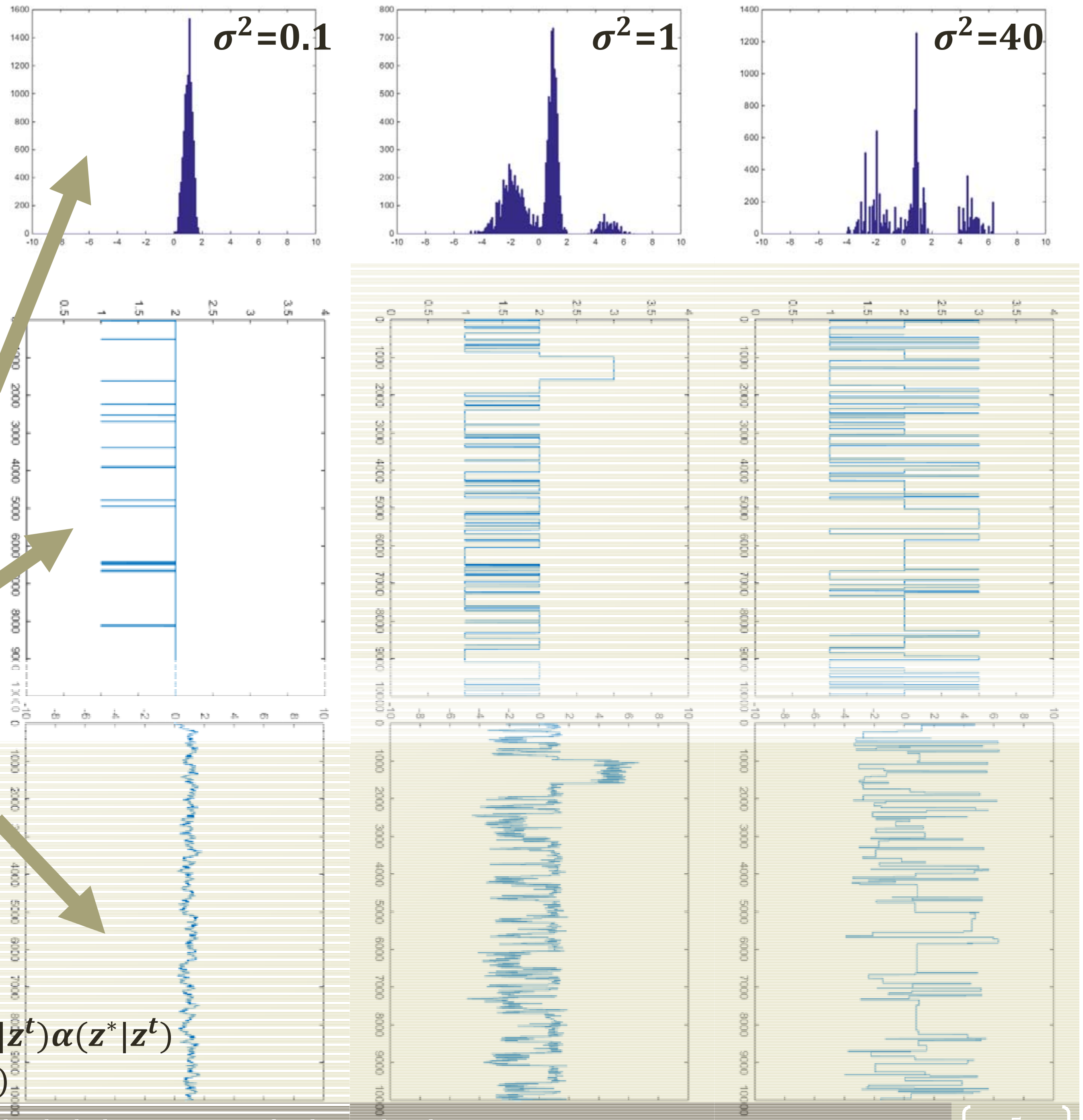
Result of Random Walk M-H

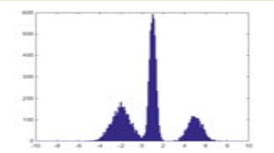
Sampling Result of Random Walk M-H

Overall Sampling

Latent Mode Selection Sampling

Observed Variable Sampling





Target Mixture Distribution $T^{MH}_{t,*} = q(z^*|z^t)\alpha(z^*|z^t)$

$z^* \sim N(z^t, \sigma^2)$