# Markov Chain Monte Carlo

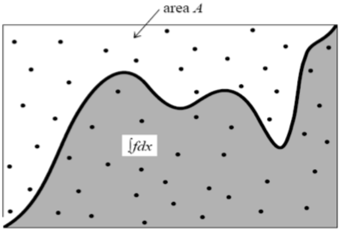
Leonid Zhukov

School of Applied Mathematics and Information Science
**National Research University Higher School of Economics**

05.12.2013

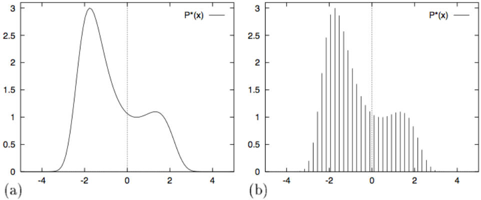
# Monte Carlo integration

- Numerical method to compute high dimensional integrals $\int\int ... \int f(x_1...x_n)dx_1...dx_n$
- Rectangular, Trapzoidal, Simpson's rule..
- Numerical integration using random numbers (samples)
- Compute $I = \int_a^b f(x)dx$



# Monte Carlo integration

- Expectation value $E_p[z] = \int zp(z)dz = \lim_{n\to\infty} \frac{1}{n}\sum_1^n z_i \approx \frac{1}{n}\sum_1^n z_i$
- Integral $I = \int f(x)dx = \int f(x)\cdot 1\ dx = E_1[f(x)] \approx \frac{1}{n}\sum_1^n f(x_i)$
- Split $f(x) = g(x)p(x)$, where $p(x)$ - PDF, then
  $I = \int f(x)dx = \int g(x)p(x)dx = E_{p(x)}[g(x)] \approx \frac{1}{n}\sum_i^n g(x_i)$
- need to be able to generate random samples
  from complex $p(x)$ distributions



# Metropolis-Hastings algorithm

- Metropolis 1953/ Hastings 1970
- Draw samples from any probability distribution $P$
- Idea: construct Markov Chain (transition matrix P) with $\pi(x) = p(x)$ stationary distribution $\pi = \pi P$
- Run the chain, generate random samples $x_0, x_1, x_2 ...$

# Detailed balance

Ergodic Markov Chain
Sufficient condition for $\pi$ to be stationary distribution

- reversibility condition
- $\pi_i P_{ij} = \pi_j P_{ji}$
- $\sum_i \pi_i P_{ij} = \sum_i \pi_j P_{ji} = \pi_j \sum_i P_{ji} = \pi_j$
- $\pi P = \pi$ - stationary distribution

# Continuous time and space Markov chains

DTDS

- $p_j(n)$
- $P_{ij}$
- $\sum_j P_{ij} = 1$
- $p_j(n+1) = \sum_i P_{ij} p_i(n)$
- $\pi_j = \sum_i P_{ij} \pi_i$
- $\pi_i P_{ij} = \pi_j P_{ji}$

CTCS

- $x(t)$
- $Q(x, y)$
- $\int Q(x, y) dy = 1$
- $p_{t+1}(x) = \int Q(x, y) p_t(y) dy$
- $\pi(x) = \int Q(x, y) \pi(y) dy$
- $\pi(x) Q(x, y) = \pi(y) Q(y, x)$

# Transition matrix construction

Need to find $P(x,y)$ such that $\pi(x)P(x,y) = \pi(y)P(y,x)$
Look for $P(x,y) = \alpha(x,y)Q(x,y)$
$Q(x,y)$ - "candidate"desity, $0 \leq \alpha(x,y) \leq 1$

1. if $\pi(x)Q(x,y) = \pi(y)Q(y,x)$, then $\alpha(x,y) = \alpha(y,x) = 1$, $P \equiv Q$
2. if $\pi(x)Q(x,y) > \pi(y)Q(y,x)$, choose $\alpha(y,x) = 1$

   Then $\pi(x)P(x,y) = \pi(x)\alpha(x,y)Q(x,y) =$
   $= \pi(y)\alpha(y,x)\ Q(y,x) = \pi(y)Q(y,x)$

$$\alpha(x,y) = \frac{\pi(y)Q(y,x)}{\pi(x)Q(x,y)}$$

3. if $\pi(x)Q(x,y) < \pi(y)Q(y,x)$, choose $\alpha(x,y) = 1$

$$\alpha(y,x) = \frac{\pi(x)Q(x,y)}{\pi(y)Q(y,x)}$$

# Transition matrix construction

$P(x, y) = \alpha(x, y) Q(x, y)$
$\alpha(x, y)$ - probability of a "forward move"
$\alpha(y, x)$ - probability of "reverse move"

- Metropolis-Hastings:

$$\alpha(x, y) = \min\left[\frac{\pi(y) Q(y, x)}{\pi(x) Q(x, y)}, 1\right]$$

- Metropolis (symmetric $Q(x, y) = Q(y, x)$)

$$\alpha(x, y) = \min\left[\frac{\pi(y)}{\pi(x)}, 1\right]$$

if $\pi(y) < \pi(x)$, move $x \to y$ with probability $\pi(y)/\pi(x)$
if $\pi(y) > \pi(x)$, move $x \to y$ with probability 1

# Metropolis-Hastings algoritm

**Algorithm:** Metropolis-Hastings

**Input**: $P(x)$,$Q(x, y)$,$N$

**Output**: $\{x_0, x_1, ...x_N\}$

initialize $x_0$, $n = 0$

**while** $n < N$ **do**

$x^* \leftarrow Q(x_n, x^*)$

$\alpha(x_n, x^*) = \frac{P(x^*)Q(x^*,x_n)}{P(x_n)Q(x_n,x^*)}$

$u \leftarrow U(0, 1)$

**if** $u \leq \alpha(x^*, x_n)$ **then**

$x_{n+1} = x^*$

**else**

$x_{n+1} = x_n$

$n = n + 1$

return $\{x_0, x_1, ...x_N\}$

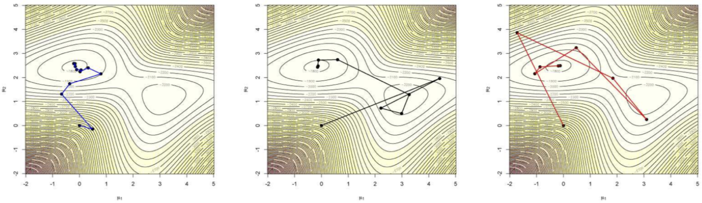
# Random walk

Proposal distribution $Q(x, y)$

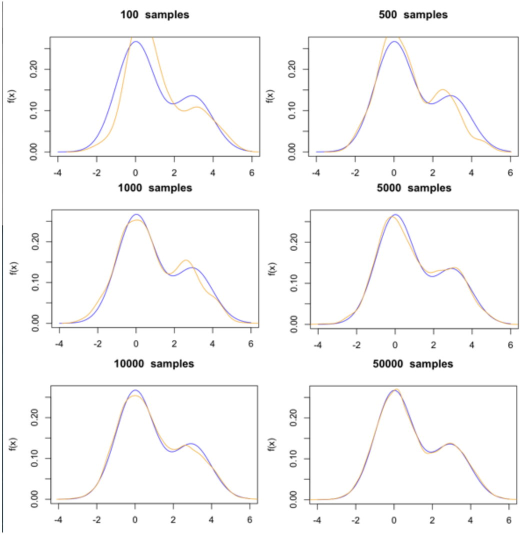
- Metropolis - symmetric: $Q(x, y) = Q(y, x)$, for example Gaussian (normal) distribution centered at $x$:

$$\mathcal{N}(y|x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-x)^2}{2\sigma^2}}$$

- then $\alpha(x, y) = \frac{P(y)}{P(x)}$
- Sample sequence is random walk $y = x + z$, $z \sim \mathcal{N}(0, \sigma)$

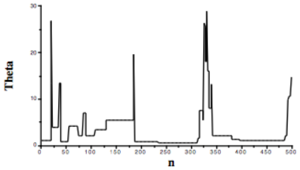


# MCMC convergence

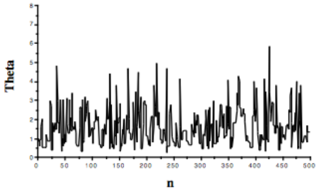


# Sample traces

poor mixing chain




well mixing chain



# Convergence diagnostics

- Autocorrelation sequence $(x_1, ...x_n)$ $k$-th order aurocorrelation

$$\rho_k = \frac{\sum_{t=1}^{n-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{n-k}(x_t - \bar{x})^2}, \quad \bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$$

- Partial $k$-th autocorrelation as a function of lag
- Geweke z-score test
- split sample first 10%, last 50%
- at stationary means are equal, $z$-test. $z_{score} > 2$ still drifting

$$Z_{score} = \frac{\mu_1 - \mu_2}{\sigma}$$

# References


- Monte Carlo sampling methods using Markov chains and their applications. W. K. Hastings, Biometrika, 57,1, 1970, p. 97.
- Understanding the Metropolis-Hastings Algorithm. S. Chib, E. Greenberg, The American Statistician, Vol. 49, No. 4. ,1995, pp. 327-335.
- A guided walk Metropolis algorithm. Paul Gustafson, Statistics and computing (1998) 8, 357-364
- Equation of State Calculations by Fast Computing Machines. Metropolis, N.,Rosenbluth, A. W.. Rosenbluth, M. N., Teller. A. H.. and Teller, E, Journal of Chemichal Physics, 21, 1953, p 1087-1092.