Markov Chain Monte Carlo II

Leonid E. Zhukov

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

Lecture plan

- MCMC Convergence
- Simulated Annealing
- Gibbs sampler
Convergence diagnostics

- Autocorrelation sequence $(x_1, \ldots, x_n)$ k-th order autocorrelation
  \[
  \rho_k = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{n-k} (x_t - \bar{x})^2}, \quad \bar{x} = \frac{1}{n} \sum_{t=1}^{n} x_t
  \]

- 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. $z_{\text{score}} > 2$ still drifting
  \[
  Z_{\text{score}} = \frac{\mu_1 - \mu_2}{\sigma}
  \]
Simulated annealing

- "Hill climbing" optimization algorithm
- accepts probability of downhill move
- probability decreases with time (process)
- Metropolis sampling

\[ \alpha(x_n, x^*) = \min \left[ 1, \left( \frac{p(x^*)}{p(x_n)} \right)^{1/T(t)} \right] \]

- "Cooling" schedule

\[ T(t) = T_0 \left( \frac{T_f}{T_0} \right)^{t/n} \]

- \( T_f \) - final "temperature"
The Gibbs sampler

- joint multivariate density \( p(x) = p(x_1, \ldots, x_n) \)
- compute samples \( X_1, \ldots, X_k \)
- compute marginal \( p(x_1) = \int \ldots \int p(x_1 \ldots x_n) dx_1 \ldots dx_n \)
- can compute univariate \( p(x_1 | x_2 \ldots x_n) \)
- generate sequence of univariate conditionals
Bivariate case

- **distribution**
  \[ p(x, y) \]

- **marginal distribution**
  \[ p(x) = \int p(x, y) dy \]
  \[ p(y) = \int p(x, y) dx \]

- **conditional probability**
  \[ p(x|y) = \frac{p(x, y)}{p(y)} \]
  \[ p(y|x) = \frac{p(x, y)}{p(x)} \]

- **marginal from conditional distribution**
  \[ p(x) = \int p(x|y)p(y)dy = E_{p(y)}[p(x|y)] \]
  \[ p(y) = \int p(y|x)p(x)dx = E_{p(x)}[p(y|x)] \]
Gibbs sampler: bivariate case

- Given: $p(x|y), p(y|x)$
- Choose $y_0$, $t = 0$
- do "sampler scan"
  - $x_t \sim p(x|y = y_t)$
  - $y_{t+1} \sim p(y|x = x_t)$
- repeat $k$-times, Gibbs sequence $(x_0, y_0), (x_1, y_1)\ldots (x_k, y_k)$
Gibbs sampler

- \( p(x_1, \ldots x_n) \)

- single iteration:
  
  \[
  x_{1}^{t+1} = p(x_1 | x_2^t \ldots x_n^t) \\
  x_{2}^{t+1} = p(x_2 | x_1^{t+1}, x_3^t \ldots x_n^t) \\
  x_{j}^{t+1} = p(x_j | x_1^{t+1}, x_2^{t+1}, \ldots x_{j-1}^t, x_{j+1}^t, \ldots x_n^t) 
  \]

- samples

\[ X^1 \ldots X^k, \ X = (x_1, \ldots x_n) \]
MCMC connection

- Metropolis-Hastings

\[ \alpha(x, x^*) = \frac{p(x^*) Q(x^*, x)}{p(x) Q(x, x^*)} = \frac{p(x^*) Q(x|x^*)}{p(x) Q(x^*|x)} \]

- Use conditional \( p(x|x^*) \) as a candidate density \( Q(x|x^*) \)

\[ (x^t, y^t) \rightarrow (x^{t+1}, y^t) \]

\[ \alpha = \frac{p(x^{t+1}, y^t) p(x^t|y^t)}{p(x^t, y^t) p(x^{t+1}|y^t)} = \frac{p(x^{t+1}, y^t)}{p(x^t, y^t)} \frac{p(x^t, y^t)}{p(y^t)} \frac{p(y^t)}{p(x^{t+1}, y^t)} = 1 \]

- MCMC algorithm with acceptance probability 1