10.2 Metropolis - Hastings Algorithm

Let E be a finite state space, π a distribution on E with $\pi_x > 0$ f.a $x \in E$.

Choose
- arbitrary irreducible and aperiodic transition matrix $Q \in [0,1]^{|E| \times |E|}$ satisfying $q_{xx'} = 0 \Leftrightarrow q_{x'x} = 0$ f.a $x, x' \in E$ called proposal function.
- $A \in [0,1]^{|E| \times |E|}$ with $a_{xx'} = \dfrac{s_{xx'}}{1 + t_{xx'}}$ with

$$t_{xx'} = \begin{cases} \dfrac{\pi_x q_{xx'}}{\pi_{x'} q_{x'x}} & , \text{ if } q_{xx'} > 0 \\ 0 & , \text{ otherwise} \end{cases} \quad \text{and}$$

S arbitrary $|E| \times |E|$-matrix with $0 < s_{xx'} \le 1 + \min\{t_{xx'}, t_{x'x}\}$ f.a $x, x' \in E$ called acceptance function.

Then define the matrix $\mathbb{P}$ by

$$p_{xx'} = q_{xx'} a_{xx'} \quad \text{for } x \neq x' \quad \text{and} \quad p_{xx} = 1 - \sum_{\substack{y \in E \\ y \neq x}} q_{xy} a_{xy}$$

f.a. $x, x' \in E$.

10.2.1 Remarks

1. Transition prob. are calculated based on proposal and acceptance rates.
2. If $q_{xx'} = 0$, the MC will never transition from x to x'.
3. For the choice $s_{xx'} = 1 + \min\{t_{xx'}, t_{x'x}\}$ and symmetrical Q we obtain $a_{xx'} = \min\left\{1, \dfrac{\pi_{x'}}{\pi_x}\right\}$.

   This is the classical Metropolis algorithm.

10.2.2 Theorem

$\mathbb{P}$ as defined above is a transition matrix, the corresponding MC is ergodic with stationary distribution π.

Implementation : generate realization $x_0, x_1, \dots$ of the MC.

1. choose some $x_0 \in E$
2. for current state x_k draw $y \in E$ acc. to distribution $(q_{x_k 1}, \dots, q_{x_k |E|})$
3. compute acceptance probability $a_{x_k y}$
4. draw $r \in [0,1]$ randomly from uniform distribution
5. set $x_{k+1} = \begin{cases} y & , r \leq a_{x_k y} \quad \text{(accepted)} \\ x_k & , r > a_{x_k y} \quad \text{(rejected)} \end{cases}$
6. return to step 2

→ after a large number of steps return accepted states as samples.

10.2.3 Remark

The speed of convergence depends on the eigenvalues of $\mathbb{P}$.

10.2.4 Ex

Parameter estimation (often used in Bayesian inference methods)

10.3 Observables

10.3.1 Def

Given a MC (X_n) on state space E, we call a function $f: E \to \mathbb{R}$ an observable.

The empirical average is $S_n(f) = \frac{1}{n} \sum_{k=1}^{n} f(x_k)$ for samples $x_1, \dots, x_n \in E$.

10.3.2 Remarks

1. Observables can be seen as measurements modeled by the MC.
2. The empirical average $S_n(f)$ can be used to approx. the expectation value $E_\pi(f) = \sum_{x \in E} f(x) \pi_x$ if samples $x_1, \dots, x_n$ are chosen acc to π.

10.3.3 Ex

For the Hard-Core model (10.1) the function $m: E \to \mathbb{R}$ counting how many vertices of a configuration have value 1 is an observable.

10.3.4 Remark

The Metropolis-Hastings Algo allows us to approx. expectation values for observables even if the distribution π is not explicitly know. It is enough to know the ratio $\frac{\pi_x}{\pi_y}$ (cp 10.2.4).