S&DS 351: Stochastic Processes - Homework 5

Bryan SebaRaj

Professor Ilias Zadik

March 3, 2025

1. (20 points) Let $G = (V, E)$ be a connected simple graph. Let $d(i)$ denote the degree of vertex i, which varies for different vertices. Let π be the uniform distribution on the vertex set V. Let the base chain be the random walk on G. Apply the Metropolis method to modify the chain so that the stationary distribution is the uniform distribution π. Find the resulting transition matrix.

2. (Metropolis for optimization) Consider the knapsack problem: Given m items with weights $w_1, \ldots, w_m$ and values $v_1, \ldots, v_m$, and a total weight budget W, the goal is to find the subset of items with maximal value subject to a weight constraint. This can be formulated as a constrained optimization problem:

$$\max \sum_{i=1}^{m} x_i v_i$$

$$\text{s.t.} \quad \sum_{i=1}^{m} x_i w_i \leq W$$

$$x_i \in \{0, 1\}.$$

Here the maximization is over the decision variable $x = (x_1, \ldots, x_m) \in \{0, 1\}^m$, where x_i indicates the ith item is included or not. This is a hard problem to solve fast.

(a) (5 points) Consider the following Markov chain. Starting from the initial state $(0, 0, \ldots, 0)$ (an empty knapsack), if the current state is $x = (x_1, \ldots, x_m)$, in the next step update it as follows: Choose an item J uniformly at random and replace x_J by $1 - x_J$. If this satisfies the constraint, update x accordingly; otherwise, do not update x. Identify the state space of this Markov chain and its transition rule.

(b) (5 points) Show that the stationary distribution of this chain is the uniform distribution over the feasible set

$$C = \{(x_1, \ldots, x_m) : \sum_{i=1}^{m} x_i w_i \leq W, x_i \in \{0, 1\}\}.$$

(c) (5 points) Recall the goal is to maximize the value of the selected items. Fix some parameter $\beta > 0$. Define a distribution π over the feasible set C such that

$$\pi(x) \propto \exp(\beta f(x)), \quad x \in C$$

where $f(x) = \sum_{i=1}^{m} x_i v_i$ is the objective function. If we choose a large β, π is close to the uniform distribution over the maximizers. Use the chain in part (a) as the base chain and apply Metropolis method to produce a modified chain with stationary distribution π. Find the transition rule.

Chang Problems

1.26. (15 points) Let π_0 and ρ_0 be probability mass functions on $\mathcal{S}$, and define $\pi_1 = \pi_0 P$ and $\rho_1 = \rho_0 P$, where P is a probability transition matrix. Show that $|\pi_1 - \rho_1| \leq |\pi_0 - \rho_0|$. That is, in terms of total variation distance, π_1 and ρ_1 are closer to each other than π_0 and ρ_0 were.

2.1. (5 points) For a branching process $\{G_t\}$ with $G_0 = 1$, define the probability generating function of G_t to be ψ_t, that is,

$$\psi_t(z) = \mathbb{E}[z^{G_t}] = \sum_{k=0}^{\infty} z^k P(G_t = k).$$

With ψ defined as in (2.1), show that $\psi_1(z) = \psi(z)$, $\psi_2(z) = \psi(\psi(z))$, $\psi_3(z) = \psi(\psi(\psi(z)))$, and so on.

2.3. (5 points) Consider a branching process with offspring distribution Poisson(2), that is, Poisson with mean 2. Calculate the extinction probability p to four decimal places.

2.7. Consider an irreducible, time-reversible Markov chain $\{X_t\}$ with $X_t \sim \pi$, where the distribution π is stationary. Let A be a subset of the state space. Let $0 < \alpha < 1$, and define on the same state space a Markov chain $\{Y_t\}$ having probability transition matrix Q satisfying, for $i \neq j$,

$$Q(i,j) = \begin{cases} \alpha P(i,j) & \text{if } i \in A \text{ and } j \notin A, \\ P(i,j) & \text{otherwise.} \end{cases}$$

Define the diagonal elements $Q(i,i)$ so that the rows of Q sum to 1.

(a) (8 points) What is the stationary distribution of $\{Y_t\}$, in terms of π and α?

(b) (2 points) Show that the chain $\{Y_t\}$ is also time-reversible.

(c) (5 points) Show by example that the simple relationship of part (1) need not hold if we drop the assumption that X is reversible.

2.12. [Metropolis-Hastings method] For simplicity, let us assume that π is positive, so that we won't have to worry about dividing by 0. Choose any probability transition matrix $Q = (Q(i,j))$ [again, suppose it is positive], and define $P(i,j)$ for $i \neq j$ by

$$P(i,j) = Q(i,j) \min\left(1, \frac{\pi(j)Q(j,i)}{\pi(i)Q(i,j)}\right),$$

and of course define $P(i,i) = 1 - \sum_{j \neq i} P(i,j)$.

(a) (5 points) Show that the probability transition matrix P has stationary distribution π.

(b) (5 points) Show how the Metropolis method we have discussed is a special case of this Metropolis-Hastings method.

3.9. (15 points) Derive the recursion,

$$\beta_{t-1}(x_{t-1}) = \sum_{x_t} A(x_{t-1}, x_t) B(x_t, y_t) \beta_t(x_t).$$

for the "backward" probabilities. Show that it is appropriate to start the calculations by setting

$$\beta_n(x_n) = 1 \quad \text{for all } x_n \in \mathcal{X}.$$