Wrocław, May 14, 2021

Exercise list no 10

Simulations and algorithmic applications of Markov chains

Ising model. Let $G = (V, K)$. Ising model provides a way of choosing a random element $\mathbf{e}$ from $\{-1,+1\}^V$, which we will call configuration. For given configuration $\mathbf{e}(v) \in \{-1,+1\}$ is the value of the spin in vertex v. The model gives us the following formula

$$P(Y = \mathbf{e}) = \pi_{G,\beta}(\mathbf{e}) \equiv \pi(\mathbf{e}) = \frac{1}{Z_{G,\beta}} exp(-\beta H(\mathbf{e})), \quad \text{where} \quad H(\mathbf{e}) = -\sum_{\tilde{k}=(v_1,v_2) \in K} \mathbf{e}(v_1)\mathbf{e}(v_2).$$

$H(\mathbf{e})$ is called the energy of the configuration, while β is called the inverse of the temperature. We mostly consider $\beta > 0$ and we we call the model ferromagnet then, but we can also take $\beta < 0$ and then the model is called antiferromagnet.

Definition 1 *Let $Z \sim \pi$. A probability measure π on $\mathbb{E} = S^V$ is* ***monotone*** *if for $\mathbf{e} \preceq \mathbf{e}'$, all $v \in V$ such that $\pi(\mathbf{e}_{-v}) > 0$ and $\pi(\mathbf{e}'_{-v}) > 0$ we have*

$$P(Z(v) \le s | Z_{-v} = \mathbf{e}_{-v}) \ge P(Z(v) \le s | Z_{-v} = \mathbf{e}'_{-v}) \text{ for all } s \in S. \tag{1}$$

Definition 2 *Let $Z \sim \pi$. A probability measure π on $\mathbb{E} = S^V$ is* ***anti-monotone*** *if for $\mathbf{e} \preceq \mathbf{e}'$, all $v \in V$ such that $\pi(\mathbf{e}_{-v}) > 0$ and $\pi(\mathbf{e}'_{-v}) > 0$ we have*

$$P(Z(v) \le s | Z_{-v} = \mathbf{e}_{-v}) \le P(Z(v) \le s | Z_{-v} = \mathbf{e}'_{-v}) \text{ for all } s \in S. \tag{2}$$

Exercise 1 Consider Ising model on graph $G = (V, K)$ and pick any vertex v. Let $\mathbf{e} \in \{-1,+1\}^{V\setminus\{v\}}$ be some configuration assignign spins to all vertices except v. We denote $\mathbf{e}^+$ to be the configuration $\mathbf{e}$ extended by assigning positive spin to v and $\mathbf{e}^-$ to be the configuration $\mathbf{e}$ extended by assigning negative spin to v. Show that

$$\frac{\pi(\mathbf{e}^+)}{\pi(\mathbf{e}^-)} = exp(2\beta(k_+(v,\mathbf{e}) - k_-(v,\mathbf{e}))),$$

where $k_+(v,\mathbf{e})$ denotes the number of neighbours of v with positive spin and $k_-(v,\mathbf{e})$ denotes the number of neighbours of v with negative spin.

Exercise 2 Consider Ising model on graph $G = (V, K)$ with the same notations as in Exercise 1. Assume that configuration Y has been chosen according to distribution π. We now 'forget' the spin of one particular vertex v and are interested in its distribution conditioned on all the other spins in configuration Y. Show that

$$\pi(Y(v) = +1 \mid Y(V \setminus \{v\}) = \mathbf{e}) = \frac{exp(2\beta(k_+(v,\mathbf{e}) - k_-(v,\mathbf{e})))}{1 + exp(2\beta(k_+(v,\mathbf{e}) - k_-(v,\mathbf{e})))}.$$

Exercise 3 Consider Ising model on graph $G = (V, K)$ with the same notations as in Exercise 1 and pick any vertex v. We introduce the ordering in the configurations as follows: for two configurations $\mathbf{e}, \mathbf{e}' \in \{-1,+1\}^{V\setminus\{v\}}$ we say that $\mathbf{e} \preceq \mathbf{e}'$ if $\mathbf{e}(w) \le \mathbf{e}'(w)$ for all $w \in V \setminus \{v\}$. Show that if $\mathbf{e} \preceq \mathbf{e}'$, then

$$\pi(Y(v) = +1 \mid Y(V \setminus \{v\}) = \mathbf{e}) \ \le \ \pi(Y(v) = +1 \mid Y(V \setminus \{v\}) = \mathbf{e}').$$

Exercise 4 Prove that if $\beta > 0$, then measure given by Ising model is monotone and that if $\beta < 0$, then measure given by Ising model is anti-monotone.

Exercise 5 Recall the definition of the birth and death process on a state space $E = \{0, 1, \ldots, N\}$ with the transition probability matrix

$$\mathbf{P} = \begin{bmatrix} 1-p_0 & p_0 & 0 & 0 & 0 & 0 & \ldots & 0 \\ q_1 & r_1 & p_1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & q_2 & r_2 & p_2 & 0 & 0 & \ldots & 0 \\ & & \ddots & \ddots & \ddots & & & \\ 0 & 0 & 0 & q_i & r_i & p_i & 0\ldots & 0 \\ & & & & \ddots & \ddots & \ddots & \\ 0 & \ldots & & & & q_{N-1} & r_{N-1} & p_{N-1} \\ 0 & \ldots & & & & 0 & q_0 & 1-q_0 \end{bmatrix},$$

where $p_0 > 0, q_0 > 0$ and for all $i \in \{1, \ldots, N-1\}$: $p_i > 0, q_i > 0, p_i + q_i + r_i = 1$. We say that the Markov chain X with partial ordering $\preceq$ is stochastically monotone if

$$\forall(\mathbf{e}_i \preceq \mathbf{e}_j)\forall(A \in \mathcal{U}) \quad P(\mathbf{e}_i, A) \leq P(\mathbf{e}_j, A),$$

where $A \in \mathcal{U}$ if we have that $(\mathbf{e}_i \preceq \mathbf{e}_j, i \in A) \Rightarrow (j \in A)$ (in other words: if we have two states and smaller one belongs to A than the larger one also has to belong to A). When is the birth and death process defined above stochastically monotone with regards to linear ordering $\leq$?

Exercise 6 Updating function is monotone with regards to partial ordering $\preceq$ if

$$\forall(\mathbf{e}_i \preceq \mathbf{e}_j) \;\; \forall(u \in [0,1]) \qquad \phi(\mathbf{e}_i, u) \preceq \phi(\mathbf{e}_j, u)$$

Markov chain with the transition matrix $\mathbf{P}$ is realizable monotone if there exists monotone updating function for it. What is the relation between stochastic monotonicity and realizable monotonicity? Give conditions under which birth and death process from the previous exercise is monotone realizable and give monotone updating function.

Exercise 7 Recall the construction of the Gibbs algorithm for the knapsack problem for stationary distribution $\pi(x) = \frac{1}{Z} e^{\beta \sum_{i=1}^{d} x_i c_i}$, where Z is normalizing constant, c_i are values of items, x_i are indicators whether the item was taken in the backpack, β is some constant. Let W be the size of the knapsack. Simulate the behaviour of such algorithm for (a) $c_i = w_i = i^2, W = 5000, d = 100$, (b) $c_i = 100^2 - i^2, w_i = i^2, W = 5000, d = 100$. Find possibly best knapsack using this algorithm. Test out how the algorithm behaves for different β.

Exercise 8 Recall the construction of Metropolis algorithm in the Moodle notebook link for the travelling salesman problem. We have that for a permutation of travelling through cities the cost of travel is $f(\sigma) = \sum_{i=1}^{n-1} M(\sigma_i, \sigma_{i+1}) + M(\sigma_n, \sigma_1)$. Suggest an appropriate $\pi(\sigma)$ based on $f(\sigma)$. Build a solution to the problem presented in the notebook for the 13 US cities. Test out how the algorithm behaves for different β.