Homework 4. Due March 6

1. **(10 pts)** Consider the following discrete-time Markov chain modeling virus propagation through a computer network. This model is a discrete-time analog of the model presented in [1]. Suppose there are N computers in the network. Each computer directly is connected to every other computer, i.e., this network is a complete graph. At each discrete moment of time $n = 0, 1, 2, \ldots$ every healthy computer becomes infected with virus with probability proportional to the number of its infected neighbors. At the same time, every infected computer recovers with some fixed probability.

   We are interested in the probability distribution for the number of infected computers i. Therefore, our Markov chain will have $N + 1$ nodes each of which corresponds to $i = 0, 1, \ldots, N$. Let at each moment of time the probability for a healthy computer with i infected neighbors to get infected be ai/N, while the probability for an infected computer to recover be b. Then the stochastic matrix P can be calculated as follows. If at time n there are i infected computers and $N - i$ healthy computers, the probability that k computers will become infected at time $n + 1$ is

   $$p(i,k) := \binom{N-i}{k}\left(\frac{ai}{N}\right)^k\left(1-\frac{ai}{N}\right)^{N-i-k},$$

   and the probability that m infected computers will recover is

   $$q(i,m) := \binom{i}{m}(b)^m(1-b)^{i-m}.$$

   Therefore, if at time n, there are i infected computers, then the probability that at time $n+1$ there will be i' infected computers can be calculated as follows:

   $$\text{Let } \; d := i' - i.$$
   $$P(i,i') = \sum_{k-m=d,\ k\in\{0,1,\ldots,N-i\},\ m\in\{0,1,\ldots,i\}} p(i,k)q(i,m).$$

   (a) Is this Markov chain irreducible?

   (b) Set $N = 40$, $a = 1$ and $b = 0.1$. Calculate P using matlab (the binomial coefficients can be found using the command `nchoosek`).

   (c) Using the command `[V E] = eig(P')` find the equilibrium distribution. It should be $\pi = [1, 0, \ldots, 0]$, i.e., the virus disappears. What is the second largest eigenvalue?

(d) Take the initial distribution $\lambda = \frac{1}{N}[0, 1, \ldots, 1]$. Demonstrate that λP^n evolves toward a bell-shaped distribution ϕ that remains stable for long time. We will call this distribution metastable. Plot this distribution. Note that this distribution is not equilibrium. Compare this distribution with the left eigenvector ψ of P (the right eigenvector of P^T) corresponding to the second largest eigenvalue of P (normalize this eigenvector so that the sum of its positive entries is 1 and superimpose the plots).

(e) Experiment with other values of $a \in (0, 1)$ and $b \in (0, 1)$. Report your observations regarding the existence and stability of the metastable distribution.

2. **(10 pts)**

(a) Read Chorin & Hald [2] about the Ising model and about the Markov chain Monte Carlo (the Metropolis algorithm): 2nd edition, pages 119 – 123, or 3rd edition, pages 150 – 152 and 157 – 161.

(b) Compute the mean magnetization m in the Ising model in 2 dimensions by the Metropolis algorithm (the Markov chain Monte Carlo algorithm presented in [2]), on a 30×30 lattice as a function of $\beta = 1$ and $\beta = 0.2$ starting with all spins up. Make the boundary conditions periodic, i.e., the nearest neighbors of site (i, j) where $0 \le i, j \le 29$ are $(i \pm 1 \mod 30, j \pm 1 \mod 30)$ (see the matlab help for the command `mod`). Compare your results with the exact answer

$$m(\beta) = \begin{cases} \left(1 - [\sinh(2\beta)]^{-4}\right)^{1/8}, & \beta > 1/T_c = 0.4408, \\ 0, & \beta < 1/T_c = 0.4408. \end{cases} \tag{1}$$

(c) Repeat the calculation of m for a collection of values of β in the interval $(0.2, 2)$. The values of β should be denser near the critical value $\beta = 0.4408$. Plot the graph of the computed mean magnetization as a function of β and superimpose it with the graph of $m(\beta)$ given by Eq. (1). Plot the graph of the CPU time or the number of MC steps that your algorithm does to satisfy your stopping criterion versus β. You should observe the blow up of the CPU time near $\beta_c := 1/T_c = 0.4408$.

References

[1] Jeffrey Kephart, Steve White, Directed-Graph Epidemiological Models of Computer Viruses, Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy, pp. 343-359 1991

[2] A. Chorin and O. Hald, Stochastic Tools in Mathematics and Science, 2nd edition, Springer 2009