Stat 202C Homework #1 (10 points)

Due: May 2, Wednesday 11:55pm on CCLE.

Problem 1 (Markov chain basics). Consider three transition matrices for the 5-family metaphor.

$$K_0 = \begin{pmatrix} 0.2, & 0.8, & 0.0, & 0.0, & 0.0 \\ 0.3 & 0.0, & 0.4, & 0.0, & 0.3 \\ 0.1, & 0.4, & 0.0, & 0.5, & 0.0 \\ 0.0, & 0.0, & 0.4, & 0.2, & 0.4 \\ 0.0, & 0.0, & 0.3, & 0.6, & 0.1 \end{pmatrix}$$

$$K_1 = \begin{pmatrix} 0.3, & 0.1, & 0.3, & 0.0, & 0.3 \\ 0.2 & 0.4, & 0.4, & 0.0, & 0.0 \\ 0.0, & 0.6, & 0.3, & 0.1, & 0.0 \\ 0.0, & 0.0, & 0.0, & 0.5, & 0.5 \\ 0.0, & 0.0, & 0.0, & 0.4, & 0.6 \end{pmatrix}$$

$$K_2 = \begin{pmatrix} 0.0, & 0.0, & 0.0, & 0.3, & 0.7 \\ 0.0 & 0.0, & 0.0, & 0.5, & 0.5 \\ 0.0, & 0.0, & 0.0, & 0.8, & 0.2 \\ 0.0, & 0.4, & 0.6, & 0.0, & 0.0 \\ 0.5, & 0.0, & 0.5, & 0.0, & 0.0 \end{pmatrix}$$

1. Are K_0, K_1 and K_2 irreducible and aperiodic?
2. Printout the 5 eigen-values and 5 left eigen-vectors of the 3 matrices.
3. How many and what are the invariant probabilities for each matrix?
4. Printout the matrix K_0^{200} and see whether it become the dream matrix we discussed in class (i.e. each row is π).

Problem 2 (Convergence analysis). For Markov chain K_0 in question 1, suppose we start with an initial probabilities $\nu = (0, 0, 1, 0, 0)$ i.e. we know for sure that the initial state is at $x_0 = 3$. So, at step n, the Markov chain state follows a distribution $\mu_n = \nu \cdot K_0^n$. We compute the distance between μ_n and π by TV-norm,

$$d_{\mathrm{TV}}(n) = ||\pi - \mu_n||_{\mathrm{TV}} = \frac{1}{2}\sum_{i=1}^{5} |\pi(i) - \mu_n(i)|;$$

or KL-divergence,

$$d_{\mathrm{KL}}(n) = \sum_{i=1}^{5} \pi(i) \log \frac{\pi(i)}{\mu_n(i)}.$$

1. Plot $d_{\mathrm{TV}}(n)$ and $d_{\mathrm{KL}}(n)$ for the first 200 steps.

2. Calculate the contraction coefficient for K_0. Note that contraction coefficient is the maximum TV-norm between any two rows in the transition kernel,

$$C(K_0) = \max_{x,y} ||K_0(x,\cdot) - K_0(y,\cdot)||_{\mathrm{TV}}.$$

One can prove that for any two initial probability ν_1, ν_2

$$||\nu_1 \cdot K_0 - \nu_2 \cdot K_0||_{\mathrm{TV}} \le C(K_0)||\nu_1 - \nu_2||_{\mathrm{TV}}$$

As $||\nu_1 - \nu_2||_{\mathrm{TV}} \le 1$, if $C(K) < 1$ then the convergence rate could be upper bounded by

$$A(n) = ||\nu_1 \cdot K_0^n - \nu_2 \cdot K_0^n||_{\mathrm{TV}} \le C^n(K_0)||\nu_1 - \nu_2||_{\mathrm{TV}} \le C^n(K_0), \quad \forall \nu_1, \nu_2.$$

Plot the bound $C^n(K_0)$ over $n = 1, ..., 100$.

3. There is another bound – the Diaconis-Hanlon bound below,

$$B(n) = ||\pi - \nu K_0^n||_{\mathrm{TV}} \le \sqrt{\frac{1-\pi(x_0)}{4\pi(x_0)}} \lambda_{\mathrm{slem}}^n$$

where $x_0 = 1$ is the initial state and $\pi(x_0)$ is a target probability at $x_0 = 1$, and λ_{slem} is the second largest eigen-value modulus. Plot the real convergence rate $d_{\mathrm{TV}}(n)$ in comparison with $A(n)$ and $B(n)$.

Problem 3. In a finite state space Ω, suppose at step t a Markov chain MC has state X following a probability ν. By applying the Markov kernel P once, its state in $t+1$ is Y which follows a probability $\mu = \nu \cdot P$. We know that P observes the detailed balance equations with invariant probability π, i.e.

$$\pi(x)P(x,y) = \pi(y)P(y,x), \quad \forall x, y \in \Omega.$$

$P(x,y)$ is the transition (conditional) probability from state x to y. Suppose we denote by $Q(y,x)$ the reverse transition probability of P at time t,

$$Q(y,x) = \frac{P(x,y)\nu(x)}{\mu(y)}.$$

Show that the Kullback-Leibler divergence decreases monotonically,

$$KL(\pi||\nu) - KL(\pi||\mu) = E[KL(P(y,x)||Q(y,x))] \geq 0.$$

[KL-divergence is $KL(p_1||p_2) = \sum_x p_1(x) \log \frac{p_1(x)}{p_2(x)}$.]

(Hint: Compute the joint probability of two consecutive states (x, y) according to two chains: the stationary chain $\pi(x, y)$ and the current chain. Then compute the KL-divergence between the two. Then you can derive the result.)

Problem 4. A Markov chain returning time $\tau_{\text{ret}}(i)$ is the minimum steps that a Markov chain returns to state i after leaving i. Suppose we consider a random walk in the countable set of non-negative numbers $\Omega = \{0, 1, 2, ..., \}$. At a step, the Markov chain state $x_t = n$, it has probability α to go up (i.e. $x_{t+1} = n+1$) and probability $1-\alpha$ to return to $x_{t+1} = 0$. Calculate the probability for returning to state 0 in finite step

$$Prob(\tau_{\text{ret}}(0) < \infty) = \sum_{\tau(0)=1}^{\infty} Prob(\tau(0)).$$

Calculate the expected return time

$$E[\tau_{\text{ret}}(0)].$$