Stats 202C: Homework #1

Professor S.C. Zhu
Assignment: 1-4

Eric Chuu
UID: 604406828

April 28, 2017

Problem 1

(1) The period of a state i is given by

$$d_i = \gcd\{n \geq 1 : K_{ii}^n > 0\} \tag{1}$$

We also know that a state i is aperiodic if $d_i = 1$.

For K_0, we draw the transition graph diagram and see that all states can be reached in finite time with nonzero probability, so it is irreducible. Since the Markov chain is irreducible, then all states have the same period. Thus, it suffices to find the period of any state. If we begin in state 1, we can return to state 1 in two steps by going to state 2 and then back to state 1. We can also return to state 1 in three steps by going to state 2 then state 3, then returning back to state 1 directly. Then, $d_1 = 1$, and state 1 is aperiodic, and we conclude that K_0 is aperiodic.

For K_1, we can see from the transition graph diagram, $d_i = 1$ for $i = 1, \ldots, 5$, so K_1 is aperiodic. We can see that K_1 is reducible by considering states 5 and 4. If we are in states 4 or 5, the probability of going into states 1, 2, or 3 is 0.

For K_2, we draw the transition graph diagram and see that all states can be reached in finite time with nonzero probability, so it is irreducible. For each state, we can see that $d_i = 2$ for $i = 1, \ldots, 5$, so K_2 is periodic.

(2) The eigenvalues and left eigen-vectors for each matrix are given below.

(3) For each matrix, there are (insert number here) invariant probabilities.

(4) The eigen-values for K_0, K_0^{50}, K_0^{200} are plotted below.

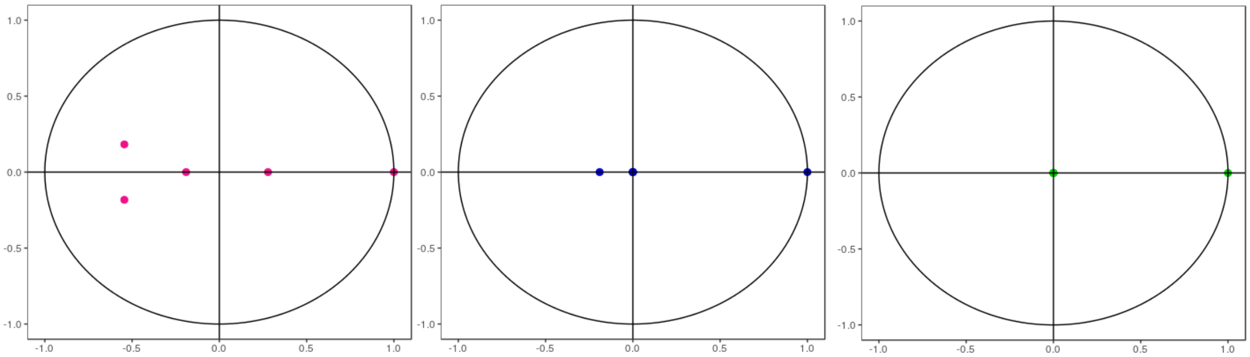

Figure 1: Left: Eigenvalues for K_0. Center: Eigenvalues for K_0^{50}. Right: Eigenvalues for K_0^{200}

(5) We compute the matrix K_0^{200}, and we see that it becomes the "ideal" matrix, where each row is π.

$$K_0^{200} = \begin{bmatrix} 0.809 & 0.2512 & 0.3071 & 0.2389 & 0.1219 \\ 0.809 & 0.2512 & 0.3071 & 0.2389 & 0.1219 \\ 0.809 & 0.2512 & 0.3071 & 0.2389 & 0.1219 \\ 0.809 & 0.2512 & 0.3071 & 0.2389 & 0.1219 \\ 0.809 & 0.2512 & 0.3071 & 0.2389 & 0.1219 \end{bmatrix}$$

$\square$

Problem 2

(1) The plot of $d_{\text{TV}}(n), d_{\text{KL}}(n)$ for the first 200 steps is shown below.
(2) We calculate the contraction coefficient K_0 using

$$C(K_0) = \max_{x,y} \|K_0(x,\cdot) - K_0(y,\cdot)\|_{\text{TV}} \tag{2}$$

For two initial probbailities ν_1, ν_2,

$$\|\nu_1 \cdot K_0 - \nu_2 \cdot K_0\|_{\text{TV}} \leq C(K_0)\|\nu_1 - \nu_2\|_{\text{TV}} \tag{3}$$

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

$$A(n) = \|\nu_1 \cdot K_0^n - \nu_2 \cdot K_0^n\|_{\text{TV}} \leq C^n(K_0)\|\nu_1 - \nu_2\|_{\text{TV}} \leq C^n(K_0), \quad \forall \nu_1, \nu_2 \tag{4}$$

The plot for the bound $C^n(K_0)$ over $n = 1, \dots, 100$ is shown below.
(3) The Diaconis-Hanlon bound is given by

$$B(n) = \|\pi - \nu K_0^n\|_{\text{TV}} \leq \sqrt{\frac{1 - \pi(x_0)}{4\pi(x_0)}}\lambda_{\text{slem}}^n \tag{5}$$

where $x_0 = 1$ is the initial state and $\pi(x_0)$ is the target probability at $x_0 = 1$ and λ_{slem}^n is the second largest eigen-value modulus. The plot of the real convergence d_{TV} in comparison with $A(n), B(n)$ is given below.

We see that both converge to 0 very quickly, with KL-Divergence converging to 0 slightly faster. Bound $A(n)$, calculated using the contraction coefficient is shown in the turquoise curve below, while Bound $B(n)$, the Diaconis-Hanlon bound, is shown in the purple curve below. We can see that $B(n)$ provides a tighter bound on the convergence than does $A(n)$.

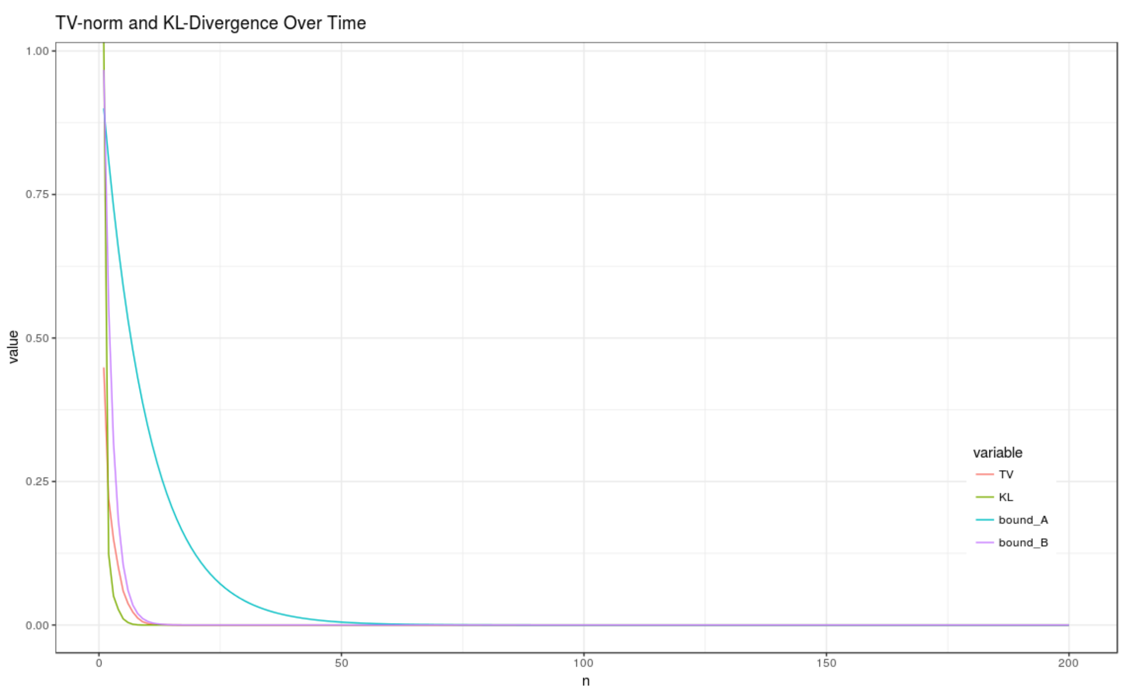

Figure 2: TV-norm and KL-Divergence for $n = 200$ steps. Bound A and Bound B refer to $A(n), B(n)$, respectively, both as calculated above.

Problem 3

Show that the Kullback-Leibler divergence decreases monotonically,

$$KL\left(\pi||\nu\right) - KL\left(\pi||\mu\right) = \mathrm{E}\left[KL\left(P\left(y,x\right)||Q\left(y,x\right)\right] \geq 0 \tag{6}$$

Solution