Section 6 (Stat 171)

- TF: Jiaze Qiu
  - Email: jiazeqiug.harvard.edu
  - Section/OH: Wed 9-12 pm (ET)
- TF: Fernando Vicente
  - Email: fernando_vicente@fas.harvard.edu
  - Section/OH: Mon 7-9 pm and Thu 8-10 pm (ET)

- Acknolegement: This handout is partially based on notes created by Lisa Ruan and Christy Huo.

1 Review

We use $\pi(\dot{})$ to denote the target density.

- Metropolis-Hastings: pick proposal distribution $q(x, \cdot)$, and do the following.
  1. Pick initial state x_0.
  2. From $t = 0$ to $T - 1$:
     - draw $x^* \sim q(x_t, \cdot)$.
     - Let x_{t+1} be x^* with probability $\min\{1, \frac{\pi(x^*)q(x^*,x_t)}{\pi(x_t)q(x_t,x^*)}\}$, and be x_t otherwise.

  Under mild conditions, the empirical distributions of samples $x_{0:T}$ will approach π as $T \to \infty$. *You only need to be able to evaluate the density up to a normalizing constant in order to implement this.*

- Gibbs sampler: for π a distribution $\mathbb{R}^d$, we have d conditional distributions, denoted $\pi(x_j|x_{[-j]})$, i.e., distribution of a single coordinate conditional on all other coordinates. Here $x_{[-j]} = x_1, \ldots, x_{j-1}, x_{j+}, \ldots, x_d)$.
  1. Pick initial state $x^{(0)}$.
  2. Let $x^{(t+1)} = x^{(t)}$ - this is temporary; we will update all the elements of $x^{(t+1)}$
  3. From $t = 0$ to $T - 1$: update $x_1^{(t+1)} \sim \pi(x_1|x_[-1] = x_[-1]^{(t+1)}$

  Under mild conditions, the empirical distribution of samples of $x^{(0:T)}$ will approach π as $T \to \infty$.

- Total Variance (TV) distance: $d_{TV}(\pi, \pi') = \sup_{A \in \mathcal{S}} |\pi(A) - \pi'(A)|$ for π, π' probability distributions on state space $\mathcal{S}$. We let $\nu(n) = \max_{i \in \mathcal{S}} d_{TV}(P^n(i, \cdot), \pi)$ represents the worst case TV-distance to stationary distribution π over the choice of X_o.
- ϵ-Mixing time: $\inf\{n : \nu(n) < \epsilon\}$
- Spectral conditions for convergence of Markov Chain
  - Finite state space: $|\mathcal{S}| < \infty$
  - $\{X_n : n \geq 1\}$ with transition matrix P.
  - P is ergodic.
  - P is reversible with respect to the stationary distribution π.

  Then the chain has geometric convergence to the stationary distribution with rate driven by eigenvalues of the transition matrix P.

2 Problems

1. Total variation distance Prove that if the state space is final, i.e. $|\mathcal{S}| < \infty$, then the total variation distance can be calculated as:

$$d_{TV}(\pi, \pi') = \frac{1}{2} \sum_{i \in \mathcal{S}} |\pi_i - \pi'_i|$$

where $\pi_i = \pi(\{i\})$ and $\pi'_i = \pi'(\{i\})$.

2. Finite Markov Chain: Ergodic and Regular We want to prove that a **finite** Markov Chain is ergodic if and only if its transition matrix $\mathbf{P}$ is regular.
(a) Prove that $\mathbf{P}$ regular implies ergodic.

(b) Prove that if i is an aperiodic state, there exists a $N > 0$ such that $P_{ii}^{n} > 0$ for all $n \geq N$

(c) Prove that an ergodic chain implies a regular transition matrix