Exercise 1  PageRank and Mixing Time

The total variation distance between two probability measures $\mu$ and $\nu$ on a set $\Omega$ (which we assume to be finite) is defined by

$$\|\mu - \nu\|_{TV} = \max_{A \subseteq \Omega} |\mu(A) - \nu(A)|.$$  

1. Show that

$$\frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)| = \sum_{\mu(x) \geq \nu(x)} (\mu(x) - \nu(x)).$$

2. Show that for $B = \{x, \mu(x) \geq \nu(x)\}$ and all $A \subset \Omega$, we have:

$$\mu(A) - \nu(A) \leq \mu(B) - \nu(B).$$

3. Deduce that

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$  

A coupling of $\mu$ and $\nu$ is a pair $(X, Y)$ of random variables defined on $\Omega$ such that the marginal distribution of $X$ is $\mu$, i.e., $P(X = x) = \mu(x)$, and the marginal distribution of $Y$ is $\nu$, i.e., $P(Y = y) = \nu(y)$.

4. Show that

$$\|\mu - \nu\|_{TV} \leq \inf \{P(X \neq Y) : (X, Y) \text{ is a coupling of } \mu \text{ and } \nu\}.$$  

5. Show that

$$\sum_{x \in \Omega} \mu(x) \wedge \nu(x) = 1 - \|\mu - \nu\|_{TV} \in [0, 1].$$

6. Find a coupling $(X, Y)$ such that $\|\mu - \nu\|_{TV} = P(X \neq Y)$. 

A Markov chain with state space $\Omega$ and transition matrix $P$ is a sequence of random variables $X_0, X_1, X_2, \ldots$ such that

$$P(X_{t+1} = y | X_t = x, X_{t-1} = x_{t-1}, \ldots, X_0 = x_0) = P(X_{t+1} = y | X_t = x) = P(x, y).$$

A chain is called irreducible if for all $x, y \in \Omega$ there exists a $t$ such that $P^t(x, y) > 0$. Let $T(x) = \{t \geq 1, P^t(x, x) > 0\}$. The period of $x$ is the greatest common divisor of $T(x)$. A chain is called aperiodic if all its states have period 1.
7. Show that if a chain is irreducible, then all its states have the same period.

8. Show that a chain with matrix $P$ is irreducible and aperiodic if and only if there is a positive integer $t$ such that $P^t$ is strictly positive.

9. Show that, in this case, there exists a unique stationary distribution $\pi$ on $\Omega$ satisfying $\pi = \pi P$.

We write $P^t(x, \cdot)$ for the distribution of the state $X_t$ at time $t$ with deterministic initial condition $X_0 = x \in \Omega$ and $\mu P^t$ for the case that the distribution of the initial state $X_0$ is $\mu$. We define

$$d(t) = \max_{x \in \Omega} \|P^t(x, \cdot) - \pi\| \text{TV},$$

$$\bar{d}(t) = \max_{x,y \in \Omega} \|P^t(x, \cdot) - P^t(y, \cdot)\| \text{TV}.$$  

10. Show that

$$d(t) \leq \bar{d}(t) \leq 2d(t).$$

11. Show that

$$d(t) = \sup_{\mu} \|\mu P^t - \pi\| \text{TV},$$

$$\bar{d}(t) = \sup_{\mu, \nu} \|\mu P^t - \nu P^t\| \text{TV},$$

where $\mu$ and $\nu$ are distributions on $\Omega$.

12. Show that

$$\|\mu P - \nu\| \text{TV} \leq \mu - \nu \| \text{TV}.$$  

Deduce that $d(t)$ and $\bar{d}(t)$ are nonincreasing in $t$.

13. For a coupling $(X_s, Y_s)$ of $P^s(x, \cdot)$ and $P^s(y, \cdot)$, show that

$$\|P^{t+s}(x, \cdot) - P^{t+s}(y, \cdot)\| \text{TV} = \frac{1}{2} \sum_w \mathbb{E} \left[ P^t(X_s, w) - P^t(Y_s, w) \right] .$$

14. Deduce that $\bar{d}$ sub-multiplicative: $\bar{d}(t+s) \leq \bar{d}(t) \bar{d}(s)$.

The mixing time is defined by

$$t_{\text{mix}}(\epsilon) = \min\{t, d(t) \leq \epsilon\} \quad \text{and} \quad t_{\text{mix}} = t_{\text{mix}}(1/4).$$

15. Show that $d(k t_{\text{mix}}(\epsilon)) \leq 2^{-k}$ and $t_{\text{mix}}(\epsilon) \leq \lfloor \log_2 \epsilon^{-1} \rfloor t_{\text{mix}}$.

We define a coupling of Markov chains with transition matrix $P$ as the random process $(X_t, Y_t)_{t=0}^\infty$ having the property that $(X_t)$ and $(Y_t)$ are Markov chains with transition matrix $P$. (The two chains can have different initial states.) Every coupling can be modified in such a way that if $X_s = Y_s$, then $X_t = Y_t$ for all $t \geq s$. We then define the coupling time by

$$\tau_{\text{couple}} = \min\{t, X_t = Y_t\}.$$  

16. As an example, consider the random walk on $\{0, 1, \ldots, n\}$ that increments and decrements with probability $1/2$ and stays on its current position with probability $1/2$ on the edges $0$ and $n$. By constructing a coupling, show that $P^t(x, n) \leq P^t(y, n)$ if $x \leq y$. 

2
17. Show that
\[ \| P^t(x, \cdot) - P^t(y, \cdot) \|_{TV} \leq \mathbb{P}_{x, y}(\tau_{\text{couple}} > t). \]

18. Consider the Markov chain describing the PageRank algorithm (which chooses a random node with probability $1 - \alpha$ in every step). Show that its mixing time satisfies $t_{mix} \leq \lceil \alpha^{-1} \rceil$. What can we deduce on the convergence speed of PageRank?