18.445 HOMEWORK 1 SOLUTIONS

Exercise 1.2. A graph G is *connected* when, for two vertices x and y of G, there exists a sequence of vertices $x_0, x_1, \ldots, x_k$ such that $x_0 = x, x_k = y$, and $x_i \sim x_{i+1}$ for $0 \le i \le k-1$. Show that random walk on G is irreducible if and only if G is connected.

Proof. Let P denote the transition matrix of random walk on G. The random walk is irreducible if for any vertices x and y there exists an integer k such that $P^k(x,y) > 0$. Note that $P^k(x,y) > 0$ if and only if there exist vertices $x_0 = x, x_1, \ldots, x_k = y$ such that $\prod_{i=0}^{k-1} P(x_i, x_{i+1}) > 0$, i.e., $x_i \sim x_{i+1}$ for all $0 \le i \le k-1$. Therefore, the random walk is irreducible if and only if G is connected. $\square$

Exercise 1.3. We define a graph to be a *tree* if it is connected but contains no cycles. Prove that the following statements about a graph T with n vertices and m edges are equivalent:

(a) T is a tree.
(b) T is connected and $m = n - 1$.
(c) T has no cycles and $m = n - 1$.

Proof. The equivalence can be easily seen from Euler's formula $m = n + l - 2$ where l denotes the number of faces of the graph, because any two of the following conditions will imply the other:

(1) T is connected $\iff m = n + l - 2$;
(2) T has no cycles $\iff l = 1$;
(3) $m = n - 1$.

Since this simple equivalence is a special case (and sometimes the starting point of the proof) of Euler's formula, it should be proved without the use of the more general theorem. We provide a long yet elementary proof here. All three parts of the following proof are based on a simple operation, namely, removing one edge and one vertex at a time. We assume without loss of generality that G has at least one edge. First we need a claim.

Claim: If each vertex of a graph G has degree at least 2, then G contains a cycle.

Start from any vertex x_0 of G and we can find $x_1 \sim x_0$. Suppose we already find distinct $x_0, \ldots, x_i$ such that $x_0 \sim x_1 \sim \cdots \sim x_i$. Since x_i has degree at least 2, we can find $x_{i+1} \ne x_{i-1}$ such that $x_i \sim x_{i+1}$. If $x_{i+1} = x_j$ for some $j < i - 1$, then we form a cycle. Otherwise we continue the process. The process must end because G is finite, so G contains a cycle.

(a) implies (b): Since T is connected and contains no cycles, the claim implies that there exists a vertex of degree 1 in T. We delete this vertex and the attached edge from T, and the remaining object T' is still a connected graph with no cycles. We continue this process until the remaining graph has only one edge and thus two vertices. Since at each step we delete one edge and one vertex, it follows that $m = n - 1$.

(b) implies (c): If there exists a vertex of degree 1 in T, we delete this vertex and the attached edge from T. Then the remaining object T' is still a connected graph with $m' = n' - 1$ where m' is the number of edges and n' is the number of vertices. We continue this process until the remaining graph has no edges, or every vertex has degree at least 2. The second case cannot happen because otherwise $n' \le m'$ which is a contradiction. In the first case, T cannot contain a cycle, because otherwise when we first delete an edge and one of its vertex in a cycle, the remaining object is no longer a graph.

(c) implies (a): If there exists a vertex of degree 1 in T, we delete this vertex and the attached edge from T. The remaining object T' is still a graph with no cycles and $m' = n' - 1$. Note that if T is not connected, then T' is not connected. We continue this process until the remaining graph has no edges, or every vertex has degree at least 2. The second case contradicts the claim because T has no cycles. In the first case, because the relation $m = n - 1$ is preserved, the remaining graph contains exactly one vertex and is thus connected. We conclude that T is connected. $\square$

Date: February 21, 2015. Prepared by Cheng Mao.

Exercise 1.4. Let T be a tree. A *leaf* is a vertex of degree 1.

(a) Prove that T contains a leaf.
(b) Prove that between any two vertices in T there is a unique simple path.
(c) Prove that T has at least 2 leaves.

Proof. Part (a) is established by the claim in the previous proof.

For (b), since T is connected and has no cycles, for vertices x and y in T, there is a simple path $x \sim x_1 \sim \cdots \sim x_k \sim y$ between them. Suppose there exists another simple path $x \sim y_1 \sim \cdots \sim y_m \sim y$ between them. Then $x \sim x_1 \sim \cdots \sim x_k \sim y \sim y_m \sim \cdots \sim y_1 \sim x$ contains a cycle, which is a contradiction.

For (c), let x_0 be a leaf and $x_1 \sim x_0$. Suppose we already have a simple path $x_0 \sim \cdots \sim x_i$. If x_i is a leaf, then we are done; otherwise, there exists $x_{i+1} \neq x_{i-1}$ such that $x_i \sim x_{i+1}$. Since T has no cycles, $x_{i+1} \notin \{x_0, \ldots, x_i\}$. The process must end because T is finite, so we will eventually find another leaf x_i. □

Exercise 1.11. Here we outline another proof, more analytic, of the existence of stationary distributions. Let P be the transition matrix of a Markov chain on a finite state space Ω. For an arbitrary initial distribution μ on Ω and $n > 0$, define the distribution ν_n by

$$\nu_n = \frac{1}{n}(\mu + \mu P + \cdots + \mu P^{n-1}).$$

(a) Show that for any $x \in \Omega$ and $n > 0$,

$$|\nu_n P(x) - \nu_n(x)| \leq \frac{2}{n}.$$

(b) Show that there exists a subsequence $(\nu_{n_k})_{k\geq 0}$ such that $\lim_{k\to\infty} \nu_{n_k}(x)$ exists for every $x \in \Omega$.
(c) For $x \in \Omega$, define $\nu(x) = \lim_{k\to\infty} \nu_{n_k}(x)$. Show that ν is a stationary distribution for P.

Proof. *(a).* We have

$$\begin{aligned}|\nu_n P(x) - \nu_n(x)| &\leq \frac{1}{n}|\mu P(x) + \mu P^2(x) + \cdots + \mu P^n(x) - \mu(x) - \mu P(x) - \cdots - \mu P^{n-1}(x)| \\ &= \frac{1}{n}|\mu P^n(x) - \mu(x)| \leq \frac{2}{n}.\end{aligned}$$

(b). Since $\nu_n \in [0,1]^{|\Omega|}$ which is compact, there exists a subsequence $(\nu_{n_k})_{k\geq 0}$ which converges at every $x \in \Omega$.

(c). Since the set of probability distribution $\{(a_1, \ldots, a_{|\Omega|}) \in \mathbb{R}^{|\Omega|} : \sum_{i=1}^{|\Omega|} a_i = 1, a_i \geq 0\}$ is closed, the limit ν is a probability distribution. Moreover, Part (a) tells us that for every $x \in \Omega$,

$$|\nu P(x) - \nu(x)| = \lim_{k\to\infty} |\nu_{n_k} P(x) - \nu_{n_k}(x)| \leq \lim_{k\to\infty} \frac{2}{n_k} = 0.$$

Therefore, ν is stationary. □

Exercise 2.10. (Reflection Principle). Let (S_n) be the sample random walk on $\mathbb{Z}$. Show that

$$\mathbb{P}(\max_{1\leq j\leq n} |S_j| \geq c) \leq 2\mathbb{P}(|S_n| \geq c).$$

Proof. If $|S_j| = c$ for some $j \leq n$, then by symmetry $S_n \geq c$ with probability at least $1/2$. Therefore,

$$\frac{1}{2}\mathbb{P}(\max_{1\leq j\leq n} |S_j| \geq c) \leq \mathbb{P}(|S_n| \geq c),$$

so the conclusion follows. □