Probability on Graphs

Shend Thaqi

May 7, 2023

Contents

1 Basics of graph theory

Definition 1 (Graph).
A graph $\mathcal{G} = (V, E)$ is a consists of

1. $V \neq \emptyset$ and
2. $E \subseteq \{e \subseteq V \mid |e| = 2\}$

We call elements of V vertices or nodes and elements of E edges or bonds. We write $v \sim u$ if $(u, v) \in E$. We further define

1. degree of a vertex is the number of edges incident to the vertex.
2. a locally finite graph is a graph where all degrees are finite.
3. a regular graph is a graph for which all degrees are equal.

Remark (Further definitions of graphs). 1. A graph $\mathcal{G} = (V, E)$ where E consists of ordered pairs is called ordered directed graph.

2. There are definitions of graphs that allow multiple edges $\{a, b\}$ and self-loops.

Definition 2 (Product of graphs).
Let $\mathcal{G}_1, \mathcal{G}_2$ be graphs. We define

$$\mathcal{G}_1 \boxtimes \mathcal{G}_2 := (V_1 \times V_2, \{((x_1, y_1), (x_2, y_2)) \mid \text{either } (x_1 = x_2,\ (y_1, y_2) \in E_2) \text{ or } (y_1 = y_2,\ (x_1, x_2) \in E_1)\})$$

as the product of $\mathcal{G}_1$ and $\mathcal{G}_2$.

Definition 3 (Path).
Let $\mathcal{G}$ be a graph. A path is a (possibly infinite) sequence of vertices $v_1, v_2, \ldots$ such that $(v_n, v_{n+1}) \in E$.

1. If a path does not visit any vertex more thatn once, it is called a vertex simple path.
2. If a path does not visit any edge more than once, it is called an edge simple path.

Definition 4 (Network).
A network is a graph G combined with a funciton $c : E \to [0, \infty)$. It will be convenient to extend c to be a function from the set of all pairs in V such that $c(v, u) \equiv 0$ if $(v, u) \notin E$. We call a network connected if $(V, \{e \mid c(e) > 0\})$ is connected. A network is locally finite if for any $v \in V$, we have $\sum_{w \neq v} c(v, w) < \infty$.

Lemma 1 (Example).
$V = \mathbb{Z}, E = \{(v, u) \mid u \neq v \in V\}$ with $c(v, u) = \frac{1}{|u-v|^2}$.

Definition 5 (Cycle).
A cycle is a path starting and ending in the same point. A cycle is simple if the only vertex appearing more than once is the first (and last) vertex.

Definition 6 (Induced graphs and networks).
Let $G = (V, E)$ be a graph and let $K \subseteq V$. The induced graph is the graph $G_K = (K, \{e \in E \mid e \subseteq K\})$. The induced network is defined similarly.

Definition 7 (Connected component).
Let G be a graph a graph or a network. A connected component is a set K of vertices such that G_K is connected, but for every $K \subsetneq H$, G_H is not connected.

Definition 8 (Contraction of graphs).
Let $G = (V, E, c)$ be a graph (or network) and let $K \subseteq G$. The connected graph (or network) G/K is the graph $(V \setminus \cup \{z\}, \{e \in E \mid e \subseteq V \setminus K\} \cup \{(v, z) \mid v \notin K, \exists u \in K : (v, u) \in E\})$. For all $u, v \notin K$, take conductance $c'(v, u) = c(v, u)$ for $v \notin K$ take $c(v, u) = \sum_{u \in K} c(v, u)$.

Definition 9 (Forest and tree).
A forest is a graph with no (simple) cycles. A tree is a connected forest.

Remark.
A contraction of an edge lets it "disappear" in a sense that the vertices are merged.

Definition 10 (Homomorphisms)**.**
Let G_1, G_2 be graphs. A function $\phi : V_1 \to V_2$ is called a homomorphism if

1. for all $(u, v) \in E_1$, we have that $(\phi(u), \phi(v)) \in E_2$

ϕ is called isomorphism if ϕ is bijective, and both ϕ, ϕ^{-1} are homomorphisms. G_1, G_2 are called isomorphic if such an isomorphism exists. An isomorphism ϕ is called an automorphism if $G_1 = G_2$.

Definition 11 (Transitivity)**.**
A graph $G = (V, E)$ is called (vertex) transitive if for all $v, u \in V$ there is an automorhpism ϕ such that $\phi(v) = u$. G is called edge transitive if for every $e_1, e_2 \in E$ there is an automorphism $\phi(e_1) = e_2$.

Remark.
There are graphs that are vertex transitive but not edge transitive.

Remark.
Good source of example for transitive graphs: Let G be a graph, and let $Y \subseteq G \setminus \{\text{unit}\}$. Take the graph $V = G$ and $(u, v) \in E$ if there is $y \in Y$ such that $u = yv$ or $v = yu$.

1.1 Random walks on networks

Definition 12 (Simple rnadom walk on networks)**.**
Let $G = (V, E, c)$ be a network and assume that G is locally finite. The simple random walk on G is the (discrete time) Markov chain with transition matrix

$$\Pi(x, y) = \frac{c(x, y)}{\sum_z c(x, z)} \cdot \mathbb{1}\left[\sum_z c(x, z) > 0\right] + 1 \cdot \mathbb{1}\left[\sum_z c(x, z) = 0\right]$$

A random walk on a locally finite graph $G = (V, E)$ is a random walk on the network $(V, E, 1)$.

Remark.
Question: Does $\mathbb{Z}^d$ have a transient subgraph?
Answer: No, every subgraph of a recurrent graph is recurrent.
The main tool to provide the proof for the answer is based on the connection between the random walk and electrical networks.

Lemma 2.
Let $G = (V, E, c)$ be a finite connected network. For every $x \in V$, let p^x be the distribution of the random walk on G started on x. Let $a, z \in V$. Let τ_a (τ_z) be the first time hitting a (z), and let $\tau_{a,z}$ be first time hitting either of them. Then, $\tau_a, \tau_z, \tau_{a,z} < \infty$.

Fact.
Let $G = (V, E, c)$ be a finite connected network. We consider the simple random walk defined as above. Then, every state x is recurrent. We have $\mathbb{P}_x(\tau_a < \infty) = \mathbb{P}_x(\tau_z < \infty) = 1$.

Proof. For every x, let γ_x be a path from x to a. Let h_x be the probability that the random walk takes the path γ_x immediately starting in x. Let $h = \min_{x \in V} h_x > 0$. Let $\mathcal{F}_n) = \sigma(X_1, \ldots, X_N)$, $\mathcal{F} = \sigma(X_1, X_2, \ldots)$. Then, $\{\tau_a < \infty\} \in \mathcal{F}$. By Kolmogorov's 0-1-law, $\mathbb{P}_x(\tau_a < \infty \,|\, \mathcal{F}) \in \{0, 1\}$. We have $\mathbb{P}_x(\tau_a < \infty \,|\, \mathcal{F}) = \lim_{n \to \infty} \mathbb{P}_x(\tau_a < \infty \,|\, \mathcal{F}_n)$. We have $\mathbb{P}_x(\tau_a < \infty \,|\, \mathcal{F}_n) \geq h$. Thus, $\mathbb{P}_x(\tau_a < \infty \,|\, \mathcal{F}) \geq h$. It follows from the 0-1-law, $\mathbb{P}_x(\tau_a < \infty \,|\, \mathcal{F}) = 1$ and thus $\mathbb{P}_x(\tau_a < \infty) = \mathbb{E}\left[\mathbb{P}_x(\tau_a < \infty \,|\, \mathcal{F})\right]$. □

Definition 13 (Voltage function)**.**
The voltage function is defined as $v : V \to \mathbb{R}, x \mapsto v(x) := p^x(\tau_{a,z} = \tau_a)$.

Lemma 3 (Properties of the voltage function)**.**
Let $v : V \to \mathbb{R}, x \mapsto p^x(\tau_{a,z} = \tau_a)$ be the voltage function. Then, the following statements hold:

- $v(a) = 1$
- $v(z) = 0$

- For $x \neq a, z$, we have $v(x) = \sum_{y \in V} \mathbb{P}_x(X_1 = y) v(y) = \sum_{y \in V} \frac{c(x,y)}{\sum_{z \in V} c(x,z)} v(y) = \sum_{y \in V} c(x,y) \frac{v(y)}{c(x)}$ (harmonicity)

Proof. To show the harmonicity, we use the Markov property. Since $x \neq a, z$, $p^x(\tau_a \geq 1) = 1$. The Markov property implies

$$\begin{aligned}
p^x(\tau_a \leq \tau_z) &= \mathbb{E}_x\left[p^x(\tau_a < \tau_z)\right] \\
&= \sum_y p^x(X_1 = y) \cdot p^y(\tau_a < \tau_z) \\
&= \sum_y p^x(X_1 = y) \cdot v(y)
\end{aligned}$$

□

Lemma 4 (Uniqueness of voltage function)**.**
The voltage function v is the only function satisfying the properties above.

Proof. Assume v_1, v_2 satisfy the properties of the voltage function. We show $h \equiv v_1 - v_2 = 0$.

1. By properties 1 and 2, $h(a) = h(z) = 0$.
2. For $x \neq a, z$, we have $h(x) = sum_y \frac{c(x,y)}{c(x)} h(y)$ ($h(x)$ is a weighted average over its neighbors). Assume $h(x) \neq 0$. We assume Œ$h(x) > 0$. Let $M = \max\{h(x) \,|\, x \in V\} > 0$. Let $H = \{x \,|\, h(x) = M\}$. Note that $a, z \notin H$. Let x be an element of H with a neighbor w outside of H. Then, the following are true
   (a) $x \in H$, thus $h(x) = M$.
   (b)

$$\begin{aligned}
h(x) &= \sum_y \frac{c(x,y)}{c(x)} h(y) \\
&= \frac{c(x,w)}{c(x)} h(w) + \sum_{y \neq w} \frac{c(x,y)}{c(x)} h(y) \\
&\leq h(w) \frac{c(x,w)}{c(x)} + M \sum_{y \neq w} \frac{c(x,y)}{c(x)} \\
& h(w) \frac{c(x,w)}{c(x)} + M\left(1 - \frac{c(x,w)}{c(x)}\right) \\
&= M + \underbrace{(h(w) - M)}_{<0} \cdot \underbrace{\frac{c(x,w)}{c(x)}}_{>0} < M
\end{aligned}$$

   This is a contradiction to $h \neq 0$.

□

Remark.
The proof is similar to the proof of uniqueness of the solution to the Dirichlet problem.

Remark.
In the theory of electrical networks, the (physical) voltage function satisfies the same three properties (up to normalization of the boundary conditions). Therefore, everything that is known about physical electrical networks has the potential of being useful in the understanding of random walks on networks.

2 Chapter two

3 Chapter three

4 Chapter four

5 Chapter five

6 Chapter six

A Measure theory

Theorem 1 (Layer-cake representation).
Let $f : \Omega \to \mathbb{R}$ such that $f \geq 0, f' \geq 0$. Then

$$\mathbb{E}[f(X)] = \int_0^\infty \mathbb{P}(f(X) > t)\,\mathrm{d}t = \int_{f^{-1}(0)}^{f^{-1}(\infty)} \mathbb{P}(X > t) f'(t)\,\mathrm{d}t$$

A.1 Convergence

Theorem 2 (Fatou's lemma).
Let $(X_n)_n$ consist of non-negative random variables. Then,

$$\mathbb{E}[\varliminf_{n\to\infty} X_n] \leq \varliminf_{n\to\infty} \mathbb{E}[X_n]$$

Theorem 3 (Monotone convergence theorem).
Let $(X_n)_n$ consist of non-negative increasing random variables. Then,

$$\mathbb{E}[\lim_{n\to\infty} X_n] = \lim_{n\to\infty} \mathbb{E}[X_n]$$

Theorem 4 (Dominated convergence theorem).
Let $(X_n)_n$ consist of converging random variables that are absolutely dominated by $Y \in L^1(\Omega, \mathbb{P})$. Then,

$$\mathbb{E}[\lim_{n\to\infty} X_n] = \lim_{n\to\infty} \mathbb{E}[X_n]$$

B Probability theory

B.1 Calculus of probability

Theorem 5 (Chebychev's inequality).
Let $\phi : \mathbb{R} \to [0, \infty)$ be increasing. Then

$$\mathbb{P}(X \geq \lambda) \leq \frac{\mathbb{E}[\phi(X)]}{\phi(\lambda)}$$

Remark.
Chebychev's inequality can be used to find sharp estimations by considering parametrized functions such as $\phi(\xi) = \exp(t\xi))$ or $\phi(\xi) = (\xi + t)^2$ and optimizing t.

Definition 14 (Limit of events).
We define

$$\varlimsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_n$$
$$\varliminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_n$$

Theorem 6 (Calculus of probability).

Theorem 7 (Borel-Cantelli).
Let $A_n \Omega$. Then

1. If $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, then $\mathbb{P}(\varlimsup_{n\to\infty} A_n) = 0$ (A_n happens finitely).
2. If $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$ and $(A_n)_n$ are independent, then $\mathbb{P}(\varlimsup_{n\to\infty} \mathbb{P}(A_n) = 1)$

B.2 Conditional expectation

Definition 15 (Conditional expectation)**.**
Y is called $\mathcal{G}$-conditional expectation of X if

1. Y is $\mathcal{G}$-measurable
2. for each $A \in \mathcal{G}$, $\mathbb{E}[Y \mathbb{1}_A] = \mathbb{E}[X \mathbb{1}_A]$

We write $Y = \mathbb{E}[X|\mathcal{G}]$.

B.3 Radon-Nikodym theorem

Theorem 8 (Radon-Nikodym)**.**

B.4 Stochastic processes

Definition 16 (Martingale)**.**

Theorem 9 (Doob's maximal inequality)**.**

B.4.1 Stopping times

Theorem 10 (Optional stopping)**.**

C Markov chains

Definition 17 (Discrete-time Markov chain)**.**
A discrete-time Markov chain is a sequence $X = (X_n)_n$ of random variables with

1. $\mathbb{P}(X_{n+1} = y \mid X_1 = x_1, \ldots, X_n = x_n) = \mathbb{P}(X_{n+1} = y \mid X_n = x_n)$

One might also use the following characterization which extends more easily to $X = (X_t)_{t \geq 0}$:

1. $\mathbb{E}_x[g \circ \theta_s \mid \mathcal{F}_s] = \mathbb{E}_{X_s}[g]$ for all measurable and bounded g

where $\mathbb{P}_x(X_0 = x) = 1$.

Definition 18 (Time-homogeneity)**.**
A Markov chain X is called time-homogeneous if

$$\mathbb{P}(X_{n+1} = y \mid X_n = x) = \mathbb{P}(X_1 = y \mid X_0 = x)$$

Definition 19 (Stationarity)**.**
A Markov chain X is called stationary if

$$\mathbb{P}(X_k = x_0, \ldots, X_{n+k} = x_n) = \mathbb{P}(X_0 = x_0, \ldots, X_n = x_n)$$

Definition 20 (Standard stopping times)**.**
For a (right-continuous) family of random variables $X = (X_t)_{t \in \mathcal{T}}$, where we allow $\mathcal{T} = \mathbb{R}, \mathbb{N}$, we write

1. $\tau_x = \inf\{t > 0 \mid X_t = x\}$
2. $\sigma_x = \inf\{t > 0 \mid X_t \neq x\}$
3. $R_x^n = 0$, $R_x^{n+1} = (\tau_x \circ \theta_{\sigma_x} + \sigma_x) \circ \theta_{R_x^n} + R_x^n$

Notice that a process X is a Markov chain if $(R_x^{n+1} - R_x^n) =_d R_x^1$ and $(R_x^{n+1} - R_x^n)_n$ forms independent and identically distributed random variables. (So called recurrence times)

Remark.
The standard approach to prove statements about probabilities of Markov chains is as follows: ....

Definition 21 (Invariant measure)**.**
A measure π is called invariant for the Markov chain X if $\pi\Pi = \pi$

1. $\pi(y) = \sum_x \pi(x)p(x, y)$

Definition 22 (Reversible measure)**.**
A measure π is called reversible for the Markov chain X if

1. $\pi(x)p(x, y) = \pi(y)p(y, x)$

Definition 23 (Irreducible Markov chain)**.**
A Markov chain X is called irreducible if for all x, y there is $n \geq 1$ such that $\mathbb{P}_x(X_n = y) > 0$.

Remark.
Irreducibility means there is the probability that a chain starting in x might reach y. It does not have to occur, though.

Definition 24 (Recurrence and transience)**.**
x is called recurrent if $\mathbb{P}_x(R_x^1 < \infty) = 1$, otherwise transient.

Theorem 11 (Invariant distribution implies recurrence)**.**
An irreducible Markov chain with an invariant distribution is recurrent.

Theorem 12 (Existence of an invariant measure)**.**
An irreducible and recurrent Markov chain has an invariant distribution which is unique and stricly positive.

Theorem 13 (Characterization of invariant distribution)**.**
Let X be a Markov chain with an invariant distribution π. Then

$$\pi(y) = \lim_{n\to\infty} \mathbb{P}_x(X_n = y)$$

regardless of the choice for x.

Theorem 14 (Cycle characterization of reversibility)**.**
A discrete-time Markov chain X is reversible if and only if for all states $x_1, \dots, x_N$ with $x_1 = x_N$, we have

$$\prod_{n=1}^{N} p(x_n, x_{n+1}) = \prod_{n=1}^{N} p(x_{N-n+1}, p_{N-n})$$

Proof. □