Chapter 15

Random Walks II

By Sariel Har-Peled, January 20, 2012[①]

> “Then you must begin a reading program immediately so that you man understand the crises of our age," Ignatius said solemnly. "Begin with the late Romans, including Boethius, of course. Then you should dip rather extensively into early Medieval. You may skip the Renaissance and the Enlightenment. That is mostly dangerous propaganda. Now, that I think about of it, you had better skip the Romantics and the Victorians, too. For the contemporary period, you should study some selected comic books.”
>
> “You’re fantastic.”
>
> “I recommend Batman especially, for he tends to transcend the abysmal society in which he’s found himself. His morality is rather rigid, also. I rather respect Batman.”
>
> – John Kennedy Toole, A confederacy of Dunces.

15.1 The 2SAT example

Let $G = G(V, E)$ be a undirected connected graph. For $v \in V$, let $\Gamma(v)$ denote the neighbors of v in G. A random walk on G is the following process: Starting from a vertex v_0, we randomly choose one of the neighbors of v_0, and set it to be v_1. We continue in this fashion, such that $v_i \in \Gamma(v_{i-1})$. It would be interesting to investigate the process of the random walk. For example, questions like: (i) how long does it take to arrive from a vertex v to a vertex u in G? and (ii) how long does it take to visit all the vertices in the graph.

15.1.1 Solving 2SAT

Consider a 2SAT formula F with m clauses defined over n variables. Start from an arbitrary assignment to the variables, and consider a non-satisfied clause in F. Randomly pick one of the clause variables, and change its value. Repeat this till you arrive to a satisfying assignment.

Consider the random variable X_i, which is the number of variables assigned the correct value (according to the satisfying assignment) in the current assignment. Clearly, with probability (at least) half $X_i = X_{i-1} + 1$.

Thus, we can think about this algorithm as performing a random walk on the numbers $0, 1, \ldots, n$, where at each step, we go to the right probability at least half. The question is, how long does it take to arrive to n in such a settings.

Theorem 15.1.1 *The expected number of steps to arrive to a satisfying assignment is $O(n^2)$.*

Proof: Consider the random walk on the integer line, starting from zero, where we go to the left with probability $1/2$, and to the right probability $1/2$. Let Y_i be the location of the walk at the i step. Clearly, $\mathbf{E}[Y_i] \geq \mathbf{E}[X_i]$. In fact, by defining the random walk on the integer line more carefully, one can ensure that $Y_i \leq X_i$. Thus, the expected number of steps till Y_i is equal to n is an upper bound on the required quantity.

To this end, observe that the probability that in the ith step we have $Y_i \geq n$ is

$$\sum_{m=n/2}^{i} \frac{1}{2^i} \binom{i}{i/2 + m} > 1/3,$$

for $i > \mu = c'n^2$, where c' is a large enough constant. To see that, observe that if we get $i/2 + m$ times $+1$, and $i - (i/2 + m) = i/2 - m$ times -1, then we have that $Y_i = (i/2 + m) - ((i/2) - m) = 2m \geq n$.

Next, if X_i fails to arrive to n at the first μ steps, we will reset $Y_\mu = X_\mu$ and continue the random walk, repeating this process as many phases as necessary. The probability that the number of phases exceeds i is $\leq (2/3)^i$. As such, the expected number of steps in the walk is at most

$$\sum_i c'n^2 i \left(\frac{2}{3}\right)^i = O(n^2),$$

as claimed. ■

15.2 Markov Chains

Let S denote a state space, which is either finite or countable. A ***Markov chain*** is at one state at any given time. There is a ***transition probability*** P_{ij}, which is the probability to move to the state j, if the Markov chain is currently at state i. As such, $\sum_j \mathsf{P}_{ij} = 1$ and $\forall i, j, 0 \leq \mathsf{P}_{ij} \leq 1$. The matrix $\mathbf{P} = \left\{\mathsf{P}_{ij}\right\}_{ij}$ is the ***transition probabilities matrix***.

The Markov chain start at an initial state X_0, and at each point in time moves according to the transition probabilities. This form a sequence of states $\{X_t\}$. We have a distribution over those sequences. Such a sequence would be referred to as a ***history***.

Similar to Martingales, the behavior of a Markov chain in the future, depends only on its location X_t at time t, and does not depends on the earlier stages that the Markov chain went through. This is the ***memorylessness property*** of the Markov chain, and it follows as P_{ij} is independent of time. Formally, the memorylessness property is

$$\mathbf{Pr}\left[X_{t+1} = j \,\middle|\, X_0 = i_0, X_1 = i_1, \ldots, X_{t-1} = i_{t-1}, X_t = i\right] = \mathbf{Pr}\left[X_{t+1} = j \,\middle|\, X_t = i\right] = \mathsf{P}_{ij}.$$

The initial state of the Markov chain might also be chosen randomly.

For states $i, j \in \mathtt{S}$, the *t-step transition probability* is $\mathtt{P}_{ij}^{(t)} = \mathbf{Pr}\Big[X_t = j \;\Big|\; X_0 = i\Big]$. The probability that we visit j for the first time, starting from i after t steps, is denoted by

$$\mathtt{r}_{ij}^{(t)} = \mathbf{Pr}\Big[X_t = j \text{ and } X_1 \neq j, X_2 \neq j, \ldots, X_{t-1} \neq j \;\Big|\; X_0 = i\Big].$$

Let $\mathtt{f}_{ij} = \sum_{t>0} \mathtt{r}_{ij}^{(t)}$ denote the probability that the Markov chain visits state j, at any point in time, starting from state i. The expected number of steps to arrive to state j starting from i is

$$\mathtt{h}_{ij} = \sum_{t>0} t \cdot \mathtt{r}_{ij}^{(t)}.$$

Of course, if $\mathtt{f}_{ij} < 1$, then there is a positive probability that the Markov chain never arrives to j, and as such $\mathtt{h}_{ij} = \infty$ in this case.

Definition 15.2.1 A state $i \in \mathtt{S}$ for which $\mathtt{f}_{ii} < 1$ (i.e., the chain has positive probability of never visiting i again), is a ***transient*** state. If $\mathtt{f}_{ii} = 1$ then the state is ***persistent***.

If a state is persistent, but $\mathtt{h}_{ii} = \infty$ are called ***null persistent***. If the state i is persistent and $\mathtt{h}_{ii} \neq \infty$ then it is ***non null persistent***.

Example 15.2.2 Consider the state 0 in the random walk on the integers. We already know that in expectation the random walk visits the origin infinite number of times. Let figure out the probability $\mathtt{r}_{00}^{(2n)}$. To this end, consider a walk $X_0, X_1, \ldots, X_{2n}$ that starts at 0 and return to 0 only in the $2n$ step. Let $S_i = X_i - X_{i-1}$, for all i. Clearly, we have $S_i \in -1, +1$ (i.e., move left or move right). Assume the walk starts by $S_1 = +1$ (the case -1 is handled similarly). Clearly, the walk $S_2, \ldots, S_{2n-1}$ must be prefix balanced; that is, the number of 1s is always bigger (or equal) for any prefix of this sequence.

Strings with this property are known as ***Dyck words***, and the number of such words of length $2m$ is the ***Catalan number*** $C_m = \frac{1}{m+1}\binom{2m}{m}$. As such, the probability of the random walk to visit 0 for the first time (starting from 0 after $2n$ steps, is

$$\mathtt{r}_{00}^{(2n)} = 2\frac{1}{n}\binom{2n-2}{n-1}\frac{1}{2^{2n}} = \Theta\left(\frac{1}{n} \cdot \frac{1}{\sqrt{n}}\right) = \Theta\left(\frac{1}{n^{3/2}}\right).$$

(the 2 here is because the other option is that the sequence starts with -1), using the fact that $\binom{2n}{n}\Theta\left(2^{2n}/\sqrt{n}\right)$.

It is not hard to show that $\mathtt{f}_{00} = 1$ (this requires a clever track). On the other hand, we have that

$$\mathtt{h}_{00} = \sum_{t>0} t \cdot \mathtt{r}_{00}^{(t)} \geq \sum_{n=1}^{\infty} 2n\mathtt{r}_{00}^{(2n)} = \sum_{n=1}^{\infty} \Theta\left(1/\sqrt{n}\right) = \infty.$$

Namely, 0 (and in fact all integers) are null persistent.

In finite Markov chains, there are no null persistent states (this requires a proof, which is left as exercise). There is a natural directed graph associated with a Markov chain. The states are the vertices, and the transition probability $\mathtt{P}_{ij}$ is the weight assigned to the edge $(i \to j)$. Note that we include only edges with $\mathtt{P}_{ij} > 0$.

Definition 15.2.3 A ***strong component*** of a directed graph G is a maximal subgraph C of G such that for any pair of vertices i and j in the vertex set of C, there is a directed path from i to j, as well as a directed path from j to i.

Definition 15.2.4 A strong component C is said to be a ***final strong component*** if there is no edge going from a vertex in C to a vertex not in C.

In a finite Markov chain, there is positive probability to arrive from any vertex on C to any other vertex of C in a finite number of steps. If C is a final strong component, then probability is 1, since the Markov chain can never leave C once it enters it. It follows that a state is persistent if and only if it lies in a final strong component.

Definition 15.2.5 A Markov chain is ***irreducible*** when its underlying graph consists of single strong component.

Clearly, if a Markov chain is irreducible, then all states are persistent.

Definition 15.2.6 Let $\mathbf{q}^{(t)} = \left(q_1^{(t)}, q_2^{(t)}, \ldots, q_n^{(t)}\right)$ be the *state probability vector* (also called the distribution of the chain at time t), to be the row vector whose ith component is the probability that the chain is in state i at time t.

The key observation is that

$$\mathbf{q}^{(t)} = \mathbf{q}^{(t-1)}\mathbf{P} = \mathbf{q}^{(0)}\mathbf{P}^t.$$

Namely, a Markov chain is fully defined by $\mathbf{q}^{(0)}$ and $\mathbf{P}$.

Definition 15.2.7 A ***stationary distribution*** for a Markov chain with the transition matrix $\mathbf{P}$ is a probability distribution π such that $\pi = \pi\mathbf{P}$.

In general, stationary distribution does not necessarily exist. We will mostly be interested in Markov chains that have stationary distribution. Intuitively it is clear, that if a stationary distribution exists, then the Markov chain, given enough time, will converge to the stationary distribution.

Definition 15.2.8 The ***periodicity*** of a state i is the maximum integer T for which there exists an initial distribution $\mathbf{q}^{(0)}$ and positive integer a such that, for all t if at time t we have $q_i^{(t)} > 0$ then t belongs to the arithmetic progression $\left\{a + ti \,\middle|\, i \geq 0\right\}$. A state is said to be ***periodic*** if it has periodicity greater than 1, and is ***aperiodic***state!aperiodic otherwise. A Markov chain in which every state is aperiodic is ***aperiodic***.

A neat trick that forces a Markov chain to be aperiodic, is to shrink all the probabilities by a factor of 2, and make every state to have a transition probability to itself equal to 1/2. Clearly, the resulting Markov chain is aperiodic.

Definition 15.2.9 An ***ergodic*** state is aperiodic and (non-null) persistent.

An ***ergodic*** Markov chain is one in which all states are ergodic.

The following theorem is the fundamental fact about Markov chains that we will need. The interested reader, should check the proof in [Nor98].

Theorem 15.2.10 (Fundamental theorem of Markov chains) *Any irreducible, finite, and aperiodic Markov chain has the following properties.*

(i) *All states are ergodic.*

(ii) *There is a unique stationary distribution* π *such that, for* $1 \le i \le n$, $\pi_i > 0$.

(iii) *For* $1 \le i \le n$, $\mathtt{f}_{ii} = 1$ *and* $\mathtt{h}_{ii} = 1/\pi_i$.

(iv) *Let* $N(i,t)$ *be the number of times the Markov chain visits state i in t steps. Then*

$$\lim_{t \to \infty} \frac{N(i,t)}{t} = \pi_i.$$

Namely, independent of the starting distribution, the process converges to the stationary distribution.

Bibliography

[Nor98] J. R. Norris. *Markov Chains*. Statistical and Probabilistic Mathematics. Cambridge Press, 1998.