Previous

Random Walk for K-SAT: For $k > 2$, we have exponential run time algorithm.

More general version of random walk: Markov chain (Ch 6.2)

There is a sequence of random variables $X_1, X_2, ..., X_n \in S$, $|S| < \infty$, we want the current state only depends on the previous state (like a walk), which bring us the markov property:

$$Pr(X_n|X_{n-1}, X_{n-2}...X_1) = Pr(X_n|X_{n-1})$$

The probability is usually time independent and it is only relevant to the value of states:

$$Pr(X_n = a|X_{n-1} = b) = Pr(X_m = a|X_{m-1} = b) = P_{ba}$$

Once we know the distribution of $X_0$ and the matrix of $P$ (the whole value of $P_{ba}$, for all $a, b$), we will know all distribution of $X_1, X_2, ....$. This give us the way to simulate a random walk problem. We pick a distribution $X_1$, then we can calculate the next distribution of the states $X_2$ with probability matrix $P$.

Makov chain example: Random Walk in a directed graph

Take a directed graph $G_a$, $S$ is the set of vertices of this graph, and we start from vertex $v$, take one step to $v$’s neighbor (must have edge from $v$ to $v'$).

In order to describe this problem, we have $q^t$ means the probability vector of vertices in the graph $G$. $q^t_i$ represents the probability of the position $i$ in $t$ time. For every position $i$ in the time $t+1$, we can use $q^{t+1}_i = \sum_j q^t_j * p_{ji}$, for all $j$ connecting to $i$. In the vector formation, we have $q^{t+1} = q^t * P$. So we have $q^t = q^0 * P^t$.

Long term behavior

What is $q^t$ like as $t$ goes to infinity? We use the eigenvalue to do a standard analysis: $q^0 = \sum_i a_i * v_i$, $v_i$ is the eigenvector of eigenvalue $\lambda_i$. So the $q^t = \sum_i a_i v_i P^t$. Since the $v_i$ is the eigenvector, so every time it multiple by P, it is just multiplied by $\lambda_i$. So $q^t = \sum_i a_i \lambda^t_i v_i$. What is important is we care about the $v$ with the largest eigenvalue $|\lambda_i|$.

Note

There are two properties of $P$:

- for all $i, j$, we have $P_{ij} \geq 0$.
- $\sum_j P_{ij} = 1$, which means $P * 1 = 1$, 1 is the all-one vector. So $P$ has 1 as its eigenvector with value 1.

We now have the right eigenvector of $P$. In a finite dimensional space, $P-I$ is not injective on the right, and it is also injective on the left. So there is a left eigenvector of $P$. Assume we have a $\pi = \pi P$. So we can treat the $\pi$ as a static distribution of $q^t$ and it is a left eigenvector of $P$ with eigenvalue of 1.
So we move one step from $\pi$, we are still in $\pi$.

For any vession $v$, we have:

$$|vP|_1 = \sum_i |\sum_j v_i P_{ij}| \leq \sum_i |v_i| \sum_j P_{ij} = |v|_1$$

This implies for any eigenvection, the eigenvalue $\lambda \leq 1$.

Example:

$g$ is an arbitrary graph. If $g$ is a graph with in-degree($v$) equals to the out-degree($v$), which is $d(v)$ for all $v$, then we have $\pi(v) \propto d(v)$.

Proof:

If $X_0 \sim \pi$. For every vertex, the probability of every out edge is relevant to the out degree of the vertex since the step is uniformly distributed. The probability of using edge from $u$ to $v$ is $P(v = v'|@u)$. According to the distribution, $Pr(@u) = \frac{d(u)}{\|E\}}$ and $Pr(v'|@u) = \frac{1}{d(v)}$ so probability is same for all edges. And the probability ending at $v$ is the $\pi(v) = \sum_{\text{every edge into } v} P(v) = \frac{d(v)}{\|E\}}$, so the $\pi(v) = \frac{d(v)}{\|E\}}$

Different Situations

Now we have the $q$ depends on the largest eigenvalue $\lambda$ and $\lambda \leq 1$. So we make sure that $q^t$ goes to $\pi$ when $t$ goes to infinity. However, this idea does not always go right. We have several situations that this could go wrong.

1. The graph is not connected

   It will not go to $\pi$ if we cannot reach any state from another. The graph is called irreducible if for any two vertices $u,v$, we can reach $v$ from $u$. (This doesn’t mean $P_{uv} > 0$, it only indicates that there is at least one path from $i$ to $j$: for any state $i,j$, we have chain $x_0 = u, x_1, x_2, \ldots x_n, x_n + 1 = v$, so we have path $u \rightarrow x_1 \rightarrow x_2 \ldots \rightarrow x_n \rightarrow v$ and $P_{u \rightarrow x_1 \rightarrow x_2 \ldots \rightarrow x_n \rightarrow v} > 0$.) And if the graph has some disconnected components, each components will have steady distributions when $\pi$ is all positive or all negative.

2. Period Cycle

   In some specified graph, such as a bi-partied graph, the $q^{2^t}$ is on the left side of the graph and $q^{2^{t+1}}$ is on the right side of the graph. So the change will never stop to jump between two sides of the bi-partied graph.

   We define a periodic chain as it is possible that from state $i$ back to state $i$ after $n$ steps, which is written as: $Pr(X_{t+n} = i|X_t = i) > 0$. We define a graph is a aperiodic if the GCD of all periodic chain step is 1:

   For all state $i$ with $Pr(X_{t+n} = i|X_t = i) > 0$, $GCD(\text{all of } i) = 1$

   If a graph has GCD of all periodic chain larger than 1, then this graph is periodic and it may go to infinity rather than $\pi$.

3. Infinite graph

   If the $|S| = \infty$, $\pi$ cannot be normalized. If we use the random walk on a line to go +1 or -1. We still have a probability distribution, but we cannot have a probability distribution that sums to 1.

Markov Chain Convergence

Theorem:

If we have an irreducible, aperiodic markov chain of some finite $S$, there is going to exist a static $\pi$ and $q^t \to \pi$ when $t \to \infty$ for any starting distribution $q^0$.

Proof:
Firstly, based on previous statement, there exists a $\pi$ with eigenvalue 1. The $\pi$ would be all positive or all negative. We want to show that for all starting distributions, they will all converge to this same state $\pi$.

Consider we have two copies of this Markov chain and we have two starting distribution $X_1, X_2, ..., X_N$ and $Y_1, Y_2, ..., Y_N$. And assume $X \sim \pi$, $Y \sim q^t$ when $t \to \infty$.

Consider these are universal constant $N, C > 0$. No matter what $X_i, Y_i$ are, we have $Pr(X_{i+N} = Y_{i+N} \geq c)$. Since $X_i, Y_i$ are finite. Assuming $X_i = a, Y_i = b$, since the graph is irreducible, there are some paths from $a \to b$ and $b \to a$. So there is $n_i$ steps from $a$ to $a$. So the $X$ can step from $a$ to $a$ in $n_1 + n_2 + ... + n_i$ steps and it takes $k$ steps from $b$ to $a$. Meanwhile, this graph is aperiodic, so the GCD of all period chain is 1. So for every $N$ steps there’s at least a probability of $c$ that $X$ and $Y$ coalesce, that $P(X_n = Y_n) = (1 - c)^{n/N}$. We would finally have $X_n = Y_n$ as $n$ goes to infinity ($P(X_n = Y_n) \to 1$).

So there is always some way to put $X_i$ and $Y_i$ into the same state, and since the graph is finite and we can try infinite times, so there must exist $Pr(X_{i+N} = Y_{i+N} \geq c > 0$). So the $X_i, Y_i$ is independent until $X_n = Y_n$ (before N steps). But after N steps, $X$ and $Y$ finally comes together. $X$ and $Y$ will share the same static state. Since $X$ goes to $\pi$ in the last, $Y$ must go to $\pi$ in the last. In this way, we can prove that all starting distribution end in the same static state $\pi$. 

3