Discrete Time, Finite, Markov Chain

- A *stochastic process* $\mathbf{X} = \{X(t) : t \in T\}$ is a collection of random variables.
- $X(t) =$ the *state* of the process at time $t = X_t$.
- $\mathbf{X}$ is a *discrete (finite) space* process if for all t, X_t assumes values from a countably infinite (finite) set.
- If T is a countably infinite set we say that $\mathbf{X}$ is a *discrete time* process.

Definition

A discrete time stochastic process $X_0, X_1, X_2, \ldots$ is a *Markov chain* if

$$\begin{aligned}\Pr(X_t = a_t | X_{t-1} = a_{t-1}, X_{t-2} = a_{t-2}, \ldots, X_0 = a_0) \\ = \Pr(X_t = a_t | X_{t-1} = a_{t-1}) = P_{a_{t-1}, a_t}.\end{aligned}$$

Transition probability: $P_{i,j} = \Pr(X_t = j \mid X_{t-1} = i)$
Transition matrix:

$$\mathbf{P} = \begin{pmatrix} P_{0,0} & P_{0,1} & \cdots & P_{0,j} & \cdots \\ P_{1,0} & P_{1,1} & \cdots & P_{1,j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots \\ P_{i,0} & P_{i,1} & \cdots & P_{i,j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots \end{pmatrix}.$$

Probability distribution for a given time t:

$$\bar{p}(t) = (p_0(t), p_1(t), p_2(t), \ldots)$$

$$p_i(t) = \sum_{j \geq 0} p_j(t-1) P_{j,i}$$

$$\bar{p}(t) = \bar{p}(t-1)\mathbf{P}.$$

For any $n \geq 0$ we define the n-step transition probability

$$P^n_{i,j} = \Pr(X_{t+n} = j \mid X_t = i)$$

Conditioning on the first transition from i we have

$$P_{i,j}^n = \sum_{k \geq 0} P_{i,k} P_{k,j}^{n-1}. \tag{1}$$

Let $\mathbf{P^{(n)}}$ be the matrix whose entries are the n-step transition probabilities, so that the entry in the ith row and jth column is $P_{i,j}^n$. Then we have

$$\mathbf{P^{(n)}} = \mathbf{P} \cdot \mathbf{P^{(n-1)}},$$

and by induction on n

$$\mathbf{P^{(n)}} = \mathbf{P^n}.$$

Thus, for any $t \geq 0$ and $n \geq 1$,

$$\bar{p}(t+n) = \bar{p}(t)\mathbf{P^n}.$$

Example

Consider a system with a total of m balls in two containers.
We start with all balls in the first container.
At each step we choose a ball uniformly at random from all the balls and with probability $1/2$ move it to the other container.
Let X_i denote the number of balls in the first container at time i.
$X_0, X_1, X_2, ...$ defines a Markov chain with the following transition matrix:

$$p_{i,j} = \begin{cases} \frac{m-i}{2m} & j = i+1 \\ \frac{i}{2m} & j = i-1 \\ \frac{1}{2} & j = i \\ 0 & |i-j| > 1 \end{cases}$$

Randomized 2-SAT Algorithm

Given an instance of 2-SAT, that is a formula in conjunctive normalform with exactly two variables per clause, find a Boolean assignment that satisfies all clauses.

Algorithm:

1. Start with an arbitrary truth assignment to the variables.
2. **Repeat** till all clauses are satisfied:
    1. Pick an unsatisfied clause C.
    2. Choose one of its literals uniformly at random and change its value.

What the is the expected run-time of this algorithm?
Assume that the formula has a satisfying assignment. Pick one such assignment S.
Let X_i be the number of variables with the correct assignment according to the assignment S after iteration i of the algorithm.
Let n be the number of variables.

$$Pr(X_i = 1 \mid X_{i-1} = 0) = 1$$

For $1 \leq t \leq n-1$,

$$Pr(X_i = t+1 \mid X_{i-1} = t) \geq 1/2$$

$$Pr(X_i = t-1 \mid X_{i-1} = t) \leq 1/2$$

To obtain an upper bound on the expected number of steps, consider assume that we have

$$Pr(X_i = 1 \mid X_{i-1} = 0) = 1$$

for $1 \leq t \leq n-1$,

$$Pr(X_i = t+1 \mid X_{i-1} = t) = 1/2$$

$$Pr(X_i = t-1 \mid X_{i-1} = t) = 1/2$$

and

$$Pr(X_i = n | X_{i-1} = n) = 1$$

Let D_t be the expected number of steps to termination when we have t incorrect variable assignments.
$D_n = 1 + D_{n-1}$.
$D_t = 1 + \frac{1}{2} D_{t+1} + \frac{1}{2} D_{t-1}$
We "guess"

$$D_t = t(2n - t)$$

.

$$D_0 = 0.$$

Inserting our guess in the formula for D_t we obtain

$$D_t = 1 + \frac{1}{2}(t+1)(2n-t-1) + \frac{1}{2}(t-1)(2n-t+1) =$$

$$1 + \frac{1}{2}(2nt + 2n - t^2 - t - t - 1 + 2nt - 2n - t^2 + t + t - 1) =$$

$$1 + 2nt - t^2 - 1 = t(2n-t).$$

$$D_n = 1 + D_{n-1} = 1 + (n-1)(2n-n+1) = n^2 = n(2n-n).$$

Theorem

Assuming that the formula has a satisfying assignment the expected number of steps to find that assignment is $O(n^2)$*.*

Theorem

There is a one-sided error randomized algorithm for the 2-SAT problem that terminates in $O(n^2 \log n)$ *time, with high probability returns an assignment when the formula is satisfiable, and always returns "UNSATISFIABLE" when no assignment exists.*

Proof.

By Markov's inequality, the probability that the algorithm does not find a good truth assignment in at most $2n^2$ steps when one exists is bounded by $\frac{1}{2}$. Since this is independent of where we start the algorithm, the probability that it still has not found a good truth assigment after $2n^2 \log n$ steps is at most $\frac{1}{2}^{\log n} = \frac{1}{n}$. □

A randomized 3-SAT algorithm

Given a formula with exactly 3 variables per clause, find a Boolean assignment that satisfies all clauses.

Algorithm:

1. Start with an arbitrary assignment.
2. **Repeat** till all clauses are satisfied:
   1. Pick an unsatisfied clause C.
   2. Pick a random litteral in the clause C and switch its value.

NB: 3-SAT is NPC so we should not expect our algorithm to be polynomial, even if there is a valid truth assignment.
We first try to analyze in the same way as for the 2-SAT algorithm:
Assume the formula is satisfiable and that S is a fized satisfying assignment.Let A_i the assignment after i steps and let X_i denote the number of variables whose values is the same in A_i as in S.

$$
\begin{aligned}
Pr(X_{i+1} = j+1 | X_i = j) &\geq \frac{1}{3} \\
Pr(X_{i+1} = j-1 | X_i = j) &\leq \frac{2}{3}
\end{aligned}
$$

If we assume equalities above and solve as for 2-SAT, we will get (with D_t being the expected number of steps to termination when we have t incorrect variable assignments), then we get

$$D_t = 2^{n+2} - 2^{t+2} - 3(n - t)$$

$$D_0 = \Theta(2^n)$$

This is not good since there are only 2^n truth assignments to try.

The problem is that the number of variables that agree with S becomes smaller over time with high probability.

- If we start from a random truth assignment, then w.h.p. this agrees with S in $n/2$ variables (this is the expectation).
- Once we start the algorithm, we tend to move towards 0 rather than n correct variables. Hence we are better off restarting the process several times and taking only a small number of steps ($3n$ works as we shall see) before restarting.

A modified algorithm

Modified 3-SAT Algorithm:

1. **Repeat** m times (alternatively: till all clauses are satisfied):
2. Start with uniformly random assignment.
3. **Repeat** up to $3n$ times, terminating if a satisfying assignment is found:
   1. Pick an unsatisfied clause C.
   2. Pick a random litteral in the clause C and switch its value.

Analogy to a particle move on the integer line

Consider a particle moving on the integer line:
with probability $\frac{1}{3}$ it moves up by one and with probability $\frac{2}{3}$ it moves down by one. Then the probability of exatly k moves down and $j + k$ moves up is:

$$\binom{j+2k}{k}\left(\frac{2}{3}\right)^k\left(\frac{1}{3}\right)^{j+k}$$

Let q_j denote the probability that the algorithm reaches a satisfying assignment within $3n$ steps when the initial (random) assignment disagreed with S on j variables.

$$q_j \geq \max_{k=0,\ldots,j} \binom{j+2k}{k}\left(\frac{2}{3}\right)^k\left(\frac{1}{3}\right)^{j+k}$$

In particular, with $k = j$ we have:

$$q_j \geq \binom{3j}{j} (\frac{2}{3})^j (\frac{1}{3})^{2j}$$

Using Stirlings formula, one can get (when $j > 0$)

$$\binom{3j}{j} \geq \frac{c}{\sqrt{j}} (\frac{27}{4})^j$$

$$\begin{aligned} q_j &\geq \binom{3j}{j} (\frac{2}{3})^j (\frac{1}{3})^{2j} \\ &\geq \frac{c}{\sqrt{j}} (\frac{27}{4})^j (\frac{2}{3})^j (\frac{1}{3})^{2j} \\ &\geq \frac{c}{\sqrt{j}} \frac{1}{2^j} \end{aligned}$$

Let q be the probability of reaching a satisfying assignment within $3n$ steps starting from the random initial assignment (in one round):

$$
\begin{aligned}
q &\geq \sum_{j=0}^{n} \Pr(\text{start with } j \text{ mismatches with } S) q_j \\
&\geq \frac{1}{2^n} + \sum_{j=1}^{n} \binom{n}{j} \left(\frac{1}{2}\right)^n \frac{c}{\sqrt{j}} \frac{1}{2^j} \\
&\geq \frac{c}{\sqrt{n}} \left(\frac{1}{2}\right)^n \sum_{j=1}^{n} \binom{n}{j} \left(\frac{1}{2}\right)^j (1)^{n-j} \\
&= \frac{c}{\sqrt{n}} \left(\frac{1}{2}\right)^n \left(\frac{3}{2}\right)^n \\
&= \frac{c}{\sqrt{n}} \left(\frac{3}{4}\right)^n
\end{aligned}
$$

If S exists, the number of times we have to repeat the initial random assignments is a geometric random variable with parameter q. The expected number of repetitions is $\frac{1}{q}$.
Hence, the expected number of repetitions is $\frac{\sqrt{n}}{c}\left(\frac{4}{3}\right)^n$ so the expected number of steps until a solution is found is $O(n^{\frac{3}{2}}\left(\frac{4}{3}\right)^n)$
If a denotes the expected number of steps above, then, by Markov's inequality, the probability that we need more than $2a$ steps is at most $\frac{1}{2}$ so if we repeat the outer loop (picking a new random assignment) $2ab$ times, then the probability that no solution is found when one exists is at most 2^{-b}.

Classification of States

Definition

State j is *accessible* from state i if for some integer $n \geq 0$, $P^n_{i,j} > 0$. If two states i and j are accessible from each other we say that they *communicate*, and we write $i \leftrightarrow j$.

In the graph representation $i \leftrightarrow j$ if and only if there are directed paths connecting i to j and j to i.
The communicating relation defines an equivalence relation. That is, the relation is

1. Reflexive: for any state i, $i \leftrightarrow i$;
2. Symmetric: if $i \leftrightarrow j$ then $j \leftrightarrow i$; and
3. Transitive: if $i \leftrightarrow j$ and $j \leftrightarrow k$, then $i \leftrightarrow k$.

Definition

A Markov chain is *irreducible* if all states belong to one communicating class.

Lemma

A finite Markov chain is irreducible if and only if its graph representation is a strongly connected graph.

$r_{i,j}^t =$ the probability that starting at state i the first transition to state j occurred at time t, that is,

$$r_{i,j}^t = \Pr(X_t = j \text{ and for } 1 \leq s \leq t-1,\ X_s \neq j \mid X_0 = i).$$

Definition

A state is *recurrent* if $\sum_{t \geq 1} r_{i,i}^t = 1$, and it is *transient* if $\sum_{t \geq 1} r_{i,i}^t < 1$. A Markov chain is recurrent if every state in the chain is recurrent.

The expected time to return to state i when starting at state j:

$$h_{j,i} = \sum_{t \geq 1} t \cdot r_{j,i}^t$$

Definition

A recurrent state i is *positive recurrent* if $h_{i,i} < \infty$. Otherwise, it is *null recurrent*.

Example - null recurrent states

States are the positive numbers.

$$P_{i,j} = \begin{cases} \frac{i}{i+1} & j = i+1 \\ 1 - \frac{i}{i+1} & j = 1 \\ 0 & \text{otherwise} \end{cases}$$

The probability of not having returned to state 1 within the first t steps is

$$\prod_{j=1}^{t} \frac{j}{j+1} = \frac{1}{t+1}.$$

The probability of never returning to state 1 from state 1 is 0, and state 1 is recurrent.

$$r_{1,1}^t = \frac{1}{t(t+1)}.$$

$$h_{1,1} = \sum_{t=1}^{\infty} t \cdot r_{1,1}^t = \sum_{t=1}^{\infty} \frac{1}{t+1} = \infty$$

State 1 is null recurrent.

Lemma

In a finite Markov chain,

1. *At least one state is recurrent;*
2. *All recurrent states are positive recurrent.*

Definition

A state j in a discrete time Markov chain is *periodic* if there exists an integer $\Delta > 1$ such that $\Pr(X_{t+s} = j \mid X_t = j) = 0$ unless s is divisible by Δ. A discrete time Markov chain is *periodic* if any state in the chain is periodic. A state or chain that is not periodic is *aperiodic*.

Definition

An aperiodic, positive recurrent state is an *ergodic* state. A Markov chain is *ergodic* if all its states are ergodic.

Corollary

Any finite, irreducible, and aperiodic Markov chain is an ergodic chain.

Example: The Gambler's Ruin

- Consider a sequence of independent, two players, fair gambling games.
- In each round a player wins a dollar with probability $1/2$ or loses a dollar with probability $1/2$.
- $W^t =$ the number of dollars won by player 1 up to (including) step t.
- If player 1 has lost money, this number is negative.
- $W^0 = 0$. For any t, $\mathbf{E}[W^t] = 0$.
- Player 1 must end the game if she loses ℓ_1 dollars ($W^t = -\ell_1$); player 2 must terminate when she loses ℓ_2 dollars ($W^t = \ell_2$).
- Let q be the probability that the game ends with player 1 winning ℓ_2 dollars.
- If $\ell_2 = \ell_1$, then by symmetry $q = 1/2$. What is q when $\ell_2 \neq \ell_1$?

$-\ell_1$ and ℓ_2 are recurrent states. All other states are transient. Let

P_i^t be the probability that after t steps the chain is at state i.

For $-\ell_1 < i < \ell_2$, $\lim_{t\to\infty} P_i^t = 0$.

$$\lim_{t\to\infty} P_{\ell_2}^t = q.$$

$$\lim_{t\to\infty} P_{\ell_1}^t = 1 - q.$$

$$\mathbf{E}[W^t] = \sum_{i=-\ell_1}^{\ell_2} iP_i^t = 0$$

$$\lim_{t\to\infty} \mathbf{E}[W^t] = \ell_2 q - \ell_1(1-q) = 0.$$

$$q = \frac{\ell_1}{\ell_1 + \ell_2}.$$