Lecture 8 notes

Scribed by Ryan Rossi

Adjacency Matrix Facts

A is an adjacency matrix for an unweighted graph without self-loops. Also known as a simple graph in this class, but other sources may have conflicting definitions.

- $A\mathbf{e}$ gives the *out-degree* of all vertices

- $A^T\mathbf{e}$ gives the *in-degree* of all vertices

Theorem. $[A^k]_{i,j} = $ # of paths between i and j of length k.
Proof.
If $k = 1$, then the paths of length 1 are exactly the edges.
If $k = 2$,

$$[A^2]_{ij} = \sum_r \left[A_{ir} A_{rj} = \begin{cases} 1 & i \to r \to j \\ 0 & \text{else} \end{cases} \right].$$

This equation counts the number of vertices r such that there is a path from i to r to j. This is a length 2 path.
For k=3, the case is similar

$$[A^3]_{ij} = \sum_{r,s} \left[A_{ir} A_{rs} A_{sj} = \begin{cases} 1 & i \to r \to s \to j \\ 0 & \text{else} \end{cases} \right].$$

This equation counts vertices r and s such that a path i to r to s to j exists. This is a path of length 3.
For general k, we proceed inductively. Assume it works for $k = \ell$,

$$[A^{(\ell+1)}]_{ij} = [AA^\ell]_{ij} = \sum_r \left[A_{i,r}[A^\ell]_{r,j} = \begin{cases} [A^\ell]_{r,j} & i \to r \to \cdots \to j \\ 0 & \text{otherwise} \end{cases} \right].$$

Now, we are counting the vertices r that prepend a length ℓ path from i to r, and then r to j in ℓ steps. Again, this is exactly the number of paths from i to j in $\ell + 1$ steps.
End proof.
Note that the above formulation counts the number of paths, which can repeat both vertices and edges.

Counting Triangles

Suppose A is an adjacency matrix for an undirected simple graph, then

$$\text{diag}(A^3) = \mathbf{t}$$

gives twice the number of triangles around the ith vertex in the ith entry of $\mathbf{t}$:

$$t_i = [A^3]_{i,i} = \text{ twice the triangles around } i .$$

This occurs because a triangle is exactly a path of length 3.

The count of all triangles is given by,

$$\text{trace}(\boldsymbol{A}^3) = \sum_{i=1}^{n} [\boldsymbol{A}^3]_{ii}$$

which overcounts the number of triangles in the graph by a factor of 6.

Estimating the trace via eigenvalues

The trace can be computed by summing the eigenvalues, therefore the number of triangles is given by,

$$\text{trace}(\boldsymbol{A}^3) = \sum_{i=1}^{n} \lambda_i^3.$$

Assuming rapid decay in the eigenvalues such that

$$\lambda_1^3 \geq \lambda_2^3 \geq \cdots \geq \lambda_s^3 \gg \lambda_{s+1}^3 \geq \cdots \lambda_n^3$$

then a reasonable estimate of the trace is

$$\text{trace}(\boldsymbol{A}^3) \approx \sum_{i=1}^{s} \lambda_i^3.$$

Tsourakakis worked on this idea in "Fast Counting of Triangles in Large Real Networks without Counting: Algorithms and Laws"

Randomized trace estimators

The numberof triangles can also be estimated,

$$\text{trace}(\boldsymbol{A}^3) = \frac{1}{n} E\left[\frac{\mathbf{x}^T A^3 \mathbf{x}}{\mathbf{x}^T \mathbf{x}}\right]$$

where E is the expectation and $\mathbf{x} \sim N(0,1)$ is a random vector with random normals. Haim Avron and Sivan Toledo worked on some new analysis for these estimators in a recent paper: Randomized algorithms for estimating the trace of an implicit symmetric positive semi-definite matrix. Journal of the ACM, 58:8:1-8:34, April 2011.

Markov Chains

The motivation/introduction to Markov chains was presented as slides which can be found at: http://www.cs.purdue.edu/homes/dgleich/nmcomp/slides/lecture-7.pdf

A Markov chain is an instance of a stochastic process. A stochastic process is a sequence of random variables,

$$X_0, X_1, X_2, \ldots \Leftrightarrow (X_n \geq 0)$$

This is a discrete time stochastic process (in contrast to continuous). The possible values of $X_i \in \mathcal{S}$ denotes the state space of the chain.

Example Consider a sequence of coin flips drawn from a random variable,

$$X_n \sim \begin{cases} H & \text{w/ prob } 0.55 \\ T & \text{w/ prob } 0.45 \end{cases}$$

This is an instance of an *i.i.d* chain (independent identically distributed). The state space, $\mathcal{S} = \{H, T\}$

Example

$$X_{n+1} = \begin{cases} X_{n+1} & \text{w/ prob } 0.5 \\ X_{n-1} & \text{w/ prob } 0.5 \end{cases}$$

The above is known as a random-walk process on integers, $\mathcal{S} = \mathbb{Z}$

Formal Definition

More formally, a Markov chain is a stochastic process where

$$\mathbb{P}\{X_{n+1} = S_i \mid X_0, X_1, ..., X_n\} = \mathbb{P}\{X_{n+1} = \mathcal{S}_i \mid X_n\}$$

We focus mostly on Time homogeneous Markov chains (or stationary Markov chains).

$$\mathbb{P}\{\boldsymbol{X}_{n+1} = S_j \mid \boldsymbol{X}_n = \mathcal{S}_i\}$$

where the probability of the transition is independent of n, the current "time-step" of the chain.

Random Walks

A traditional definition of random-walks from a well-known probability textbook:

$$S_n = \sum_{i=1}^{n} \boldsymbol{X}_i, \text{ where } (X_n, n \geq 0) \text{ is iid.}$$

This definition seems hard to adapt for things like random walks on graphs.

For this class, we will use the terms "random walk" and "Markov chain" almost interchangeably. The distinction between them will be mainly one of semantics. A Markov chain is a probabilistic construct, a random walk is a topological construct. For instance, for each Markov chain, we can associate a directed graph:

$$V = \mathcal{S}$$

$$E = \{(S_i, S_j) \mid \mathbb{P}\{X_{n+1} = S_j \mid X_n = S_i\} > 0\}.$$

For this reason, writing down the transition graph of a random walk is a convenient way to describe the non-zero transition probabilities of a Markov chain.

`Need a figure here illustrating the relationship`

Uniform random walk Given a graph, a uniform random walk is a Markov chain where the probability of making any transition is uniform over all possible choices. That is, when the chain/walk is at a state/vertex, then the next state /vertex is picked uniformly among all edges from the vertex.

Stochastic Matrices

$(X_n, n \geq 0)$ is a Markov chain; index the states in $\mathcal{S}$ (assumed to be a finite set in this class),

$$1, 2, ..., |\mathcal{S}|$$

Then the transition matrix for X is:

$$\boldsymbol{P}_{ij} = \mathbb{P}\{\boldsymbol{X}_{n+1} = \mathcal{S}_j \mid \boldsymbol{X}_n = \mathcal{S}_i\}$$

The following properties must also hold for stochastic matrices (from probability):

$$\sum_j \boldsymbol{P}_{ij} = 1$$

Formally, a stochastic matrix is non-negative, with rows that sum to 1.

$$\boldsymbol{P}_{ij} \geq 0,$$
$$\boldsymbol{P}\mathbf{e} = \mathbf{e} \iff \sum_j P_{ij} = 1$$

The above, $\boldsymbol{P}\mathbf{e}$, sums the probabilities in each row, and therefore $\mathbf{e}$ should be a vector of all 1's.

Besides row stochastic matrices, there are also column stochastic matrices,

$$\boldsymbol{P}_{ij} \geq 0, \boldsymbol{P}^T\mathbf{e} = \mathbf{e} \iff \sum_i \boldsymbol{P}_{ij} = 1$$

There are also doubly stochastic matrices (very special type of mathematical object). They are characterized by the Birkhoff-von Neumann theorem.

State Properties

A state in a Markov chain can have a few different properties.

Absorbing State.

A state is called absorbing if it is impossible to leave this state. Therefore, the state i is absorbing if $\boldsymbol{P}_{ii} = 1$ and $\boldsymbol{P}_{ij} = 0$ for $i \neq j$.

For instance, consider the two state diagram where $\mathcal{S}_1 \rightarrow \mathcal{S}_2$ given by the transition matrix (stochastic row matrix),

$$\boldsymbol{P} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Thus, a random-walk on this state diagram would always end up at $\mathcal{S}_2$ (acting as a sink; absorbing) with 0 probability of transitioning to $\mathcal{S}_1$.

Transient State

In the previous example, $\mathcal{S}_1$ is a perfect example of a transient state. Informally, a transient state is one for which the Markov chain may not return. That is, given that the chain was at state i, then the probability that the chain is ever at state i again is less than 1. This is a little tricky to formalize in terms of probability. We'll formalize it shortly in terms of the strongly connected component structure of the Markov chain as a random walk on a graph.

Recurrent state

The opposite of a transient state is a recurrent state! This is a state that the chain will always revisit. A very simple example is the absorbing state. However, an ergodic state is simply one that the chain will always revisit in the future. Again, this concept is easiest to formalize in terms of the strong component structure, so we'll delay the formality.

Periodic state

A periodic state is a special type of recurrent state which can only be revisited on a specific period. Consider a simple directed cycle between three vertices:

Insert a picture here

Then a periodic state will be

An erdodic state

A recurrent state that isn't periodic is called ergodic. These are also called "aperiodic".

Relationship to strongly connected components

In the second on random walks, we mentioned that every Markov chain can be considered as a directed graph where the states are vertices and the non-zero transition probabilities are the edges. By analyzing the structure of the strongly connected components of this graph, we can easily formalize the definition of the types of states.

Recall that a strongly connected component of a directed graph is a set of vertices where there are directed paths between all pairs of vertices. That is, if i and j are in a strong component, then there is a directed path from i to j and j to i.

```
Insert a picture of a graph and its strong components
```

We can also define a component graph, which is a new graph where each strong component in a single vertex, and the edges reflect ways of moving between the strong components. Note that this graph *must be acyclic*. Any cycle would have produced a larger strong component. Consequently, the component graph is a directed acyclic graph!

The terminal nodes in this dag (those at the end of the dag) are recurrent states. The other nodes in the dag correspond to transient components, and thus, identify the transient states. To see why this is the case, consider that for any of these transient components, there is a non-zero probability of leaving the component. Once a walk leaves the component, the walk cannot return. Hence, the probability that the walk will ever visit "that transient vertex" again is less than 1.

The absorbing states are those terminal components with exactly one state. The periodic states are those where the greatest common divisor of the

A stochastic matrix permutation

The strongly connected component structure of the the Markov chain means that we can permute the stochastic matrix for the Markov chain into a particular form.

Let $T_1, \ldots, T_k$ be the transient components. Let $R_1, \ldots, R_m$ be the recurrent components that aren't absorbing and let A represent the set of absorbing vertices.

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{P}_{T_1} & \boldsymbol{P}_{T_1,T_2} & \cdots & \boldsymbol{P}_{T_1,T_k} & \boldsymbol{P}_{T_1,R_1} & \boldsymbol{P}_{T_1,R_2} & \cdots & \boldsymbol{P}_{T_1,R_m} & \boldsymbol{P}_{T_1,A} \\ & \boldsymbol{P}_{T_2} & \cdots & \boldsymbol{P}_{T_2,T_k} & \boldsymbol{P}_{T_2,R_1} & \boldsymbol{P}_{T_2,R_2} & \cdots & \boldsymbol{P}_{T_k,R_m} & \boldsymbol{P}_{T_1,A} \\ & & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ & & & \boldsymbol{P}_{T_k} & \boldsymbol{P}_{T_k,R_1} & \boldsymbol{P}_{T_k,R_2} & \cdots & \boldsymbol{P}_{T_k,R_m} & \boldsymbol{P}_{T_k,A} \\ & & & & \boldsymbol{P}_{R_1} & 0 & \cdots & 0 & 0 \\ & & & & & \boldsymbol{P}_{R_2} & \ddots & 0 & 0 \\ & & & & & & \ddots & & \vdots \\ & & & & & & & \boldsymbol{P}_{R_m} & \\ & & & & & & & & \boldsymbol{I} \end{bmatrix}$$

Note that there are no transitions out of the recurrent states components.