MATH 5340 Fall '23

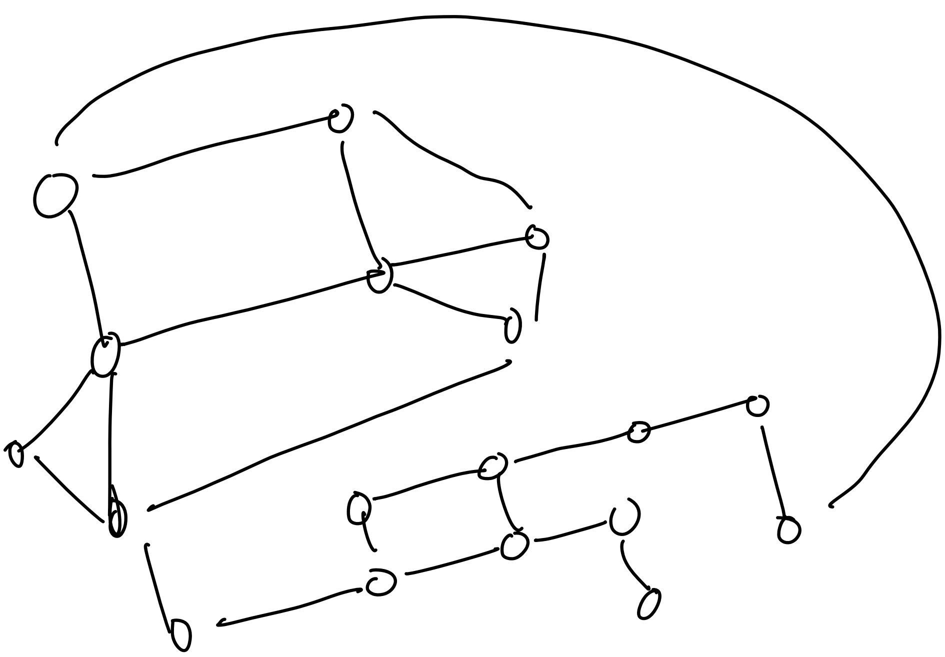
Lecture 20

Part III. Graphs, Laplacians, and Markov chains

Markov chains (stochastic processes)

Motivating example (simple random walk on a graph)

Consider a symmetric graph (G, ω_{ij}), and for simplicity assume $\omega_{ij} \in \{0, 1\}$.



in this graph, which we think of as

the board for a game, we do the following: we place a token at some initial node X_0 and then we move this token in every turn according to the following rule:

if we at the beginning of the n-th turn the token is at X_{n-1} we choose at random one of the neighbors of X_{n-1} and move our token there.

Q: Given an initial location x_0, so

$$X_0 = x_0,$$

what is the probability distribution of X_n? in other words, for any $x_j \in G$, what is

$$\text{Prob}(X_n = x_j \mid X_0 = x_0)?$$

Note (on probability distributions on a graph)

Given a finite set G, $G = \{x_1, \ldots, x_N\}$, then a probability distribution in G is a function

$$\mu : G \longrightarrow \mathbb{R}, \quad \text{and}$$

$$\mu(x_i) \geq 0 \quad \text{for } i = 1, \ldots, N,$$

$$\sum_{i=1}^{N} \mu(x_i) = 1$$

Alternatively, one can think of μ as a funct

$$\mu : 2^G \longrightarrow \mathbb{R}$$

satisfying:

- $0 \leq \mu(E) \leq 1 \quad \forall E \subset G$ $(E \in 2^G)$
- $\mu(A \cup B) = \mu(A) + \mu(B)$ if $A \cap B = \phi$.

In this case, μ defines a function $\mu : G \to \mathbb{R}$ by:

$$\mu(x_i) = \mu(\{x_i\}), \quad \text{and}$$

$$\mu(E) = \sum_{x_i \in E} \mu(x_i)$$

This is all to say that probabilities in G are described by elements of $C(G)$, which are non-negative and whose values add up to 1.

Markov chains

(finite)

A Markov chain is a sequence of random variables $X_0, X_1, X_2, \ldots$ living in a "state space" S, which is assumed finite, and satisfying the "Markov property", which says that given n, any given elements

$$\alpha_0, \alpha_1, \ldots, \alpha_n \in S$$

Then

$$\mathrm{Prob}(X_n = \alpha_n \mid X_0 = \alpha_0, \ldots, X_{n-1} = \alpha_{n-1})$$

$$= \mathrm{Prob}(X_n = \alpha_n \mid X_{n-1} = \alpha_{n-1})$$

(Here, $\text{Prob}(A|B) := \dfrac{\text{Prob}(A\cap B)}{\text{Prob}(B)}$ is the conditional probability of A given B)

For a Markov chain one defines a matrix ($S = \{x_1, x_2, \ldots, x_N\}$)

$$\pi_{ij}^{(n)} := \text{Prob}(X_{n+1} = x_j \mid X_n = x_i)$$

if these numbers don't vary with n, we say the chain is homogeneous, and $\pi_{ij} = \pi_{ij}^{(1)}$ is called the transition probability matrix

Exercise: Check that if π_{ij} is a transition probability matrix, then $\pi_{ij} \geq 0$ $\forall\, ij$ and the rows of π add up to 1.

Exercise: Given a simple graph write down π_{ij} for the Markov chain given at the beginning of the class (write your answer in terms of w_{ij})

Evolution of the probability distribution

Let's consider a homogeneous Markov chain $X_0, X_1, \dots$ with transition matrix π_{ij} and state space $S = \{x_1, \dots, x_N\}$

Problem: Given $x_* \in S$, compute the probabilities

$$u_n(x_j) = \text{Prob}(X_n = x_j \mid X_0 = x_*)$$

The key to solving this problem is using what is known as the total probability formula: (and the Markov property)

If $A = A_1 \cup \cdots \cup A_n$ (disjoint union),
and $E \subset A$, then

$$P(E) = P(E|A_1)P(A_1) + \cdots + P(E|A_n)P(A_n)$$

Exercise: If $\mu : G \to \mathbb{R}$ is a probability distribution, where G is finite, show the above holds for $P(A) := \sum_{x_i \in A} \mu(x_i)$

Lemma: For each $n \geq 1$, we have

$$u_n(x_i) = \sum_{j=1}^{N} \Pi_{ji} u_{n-1}(x_j)$$

Proof

$$u_n(x_i) = \mathrm{Prob}(X_n = x_i \mid X_0 = x_*)$$

$$= \sum_{j=1}^{N} \mathrm{Prob}\left(X_n = x_i \mid X_{n-1} = x_j, X_0 = x_*\right) \mathrm{Prob}\left(X_{n-1} = x_j \mid X_0 = x_*\right)$$

by the Markov property

$$\mathrm{Prob}\left(X_n = x_i \mid X_{n-1} = x_j, X_0 = x_*\right)$$
$$= \mathrm{Prob}\left(X_n = x_i \mid X_{n-1} = x_j\right)$$
$$\left(= \Pi_{ji} \right)$$

meanwhile $\mathrm{Prob}\left(X_{n-1} = x_j \mid X_0 = x_*\right) = u_{n-1}(x_j)$,

so we have

$$u_n(x_i) = \sum_{j=1}^{N} \Pi_{ji} \, u_{n-1}(x_j)$$

as we wanted. $\blacksquare$

Remark: If we think of u_n as a vector in $\mathbb{R}^N$

$$\left(u_n = \begin{pmatrix} u_n(x_1) \\ \vdots \\ u_n(x_N) \end{pmatrix} \text{ for each } n \in \mathbb{N} \right)$$

Then the above lemma states that

$$u_{n+1} = \Pi^t u_n \qquad \forall n$$

With this we reduce our problem to matrix multiplication, in fact

$$u_n = \left(T L^t\right)^n u_0$$