Stats 318: Lecture 1

Agenda:

- Course organization
- Definition of Markov chains
- First examples

Finite Markov chains

- Discrete-time stochastic process $\{X_t\}$, $t = 0, 1, 2, \ldots$, where X_t takes values in a finite set $\mathcal{X}$
- Possible values for X_t are called the *states* of the system
- Joint probability (joint law of $X_{0:t}$) can be specified via
  - Initial prob. dist. of X_0 denoted $\mathbb{P}(X_0)$
  - The *transition probabilities* $\mathbb{P}(X_t \mid X_{0:t-1})$

  This is because

$$\mathbb{P}(X_{0:t}) = \mathbb{P}(X_0)\mathbb{P}(X_1 \mid X_0)\mathbb{P}(X_2 \mid X_{0:1}) \ldots \mathbb{P}(X_t \mid X_{0:t-1})$$

Finite Markov chains

Interested in a special class of processes, which obey the Markov property:

Definition (Markov property)

The conditional distribution of X_{t+1} given $X_{0:t}$ depends only on X_t; i.e., to make predictions of the system in the future, it suffices to consider only the present state and not the past history

Formally,

$$\mathbb{P}(X_{t+1} \mid X_{0:t}) = \mathbb{P}(X_{t+1} \mid X_t)$$

Transition probabilities may or may not depend on time

- If time independent, process is *homegeneous*
- If time dependent, process is *inhomogeneous*

Unless specified otherwise, we are concerned with homogeneous Markov processes

Transition matrices

Definition

Let P be an $n \times n$ matrix with elements $P(x, y)$. A random process $\{X_t\}$ is said to be a (homogeneous) Markov chain with transition matrix P if for all $x_0, \ldots, x_{t+1} \in \mathcal{X}$,

$$\mathbb{P}(X_{t+1} = x_{t+1} \mid X_{0:t} = x_{0:t}) = \mathbb{P}(X_{t+1} = x_{t+1} \mid X_t = x_t) = P(x_t, x_{t+1})$$

Transition probability matrix must obey

$$P(x, y) \geq 0 \quad \forall x, y \in \mathcal{X} \qquad \text{prob. are non-negative}$$

and

$$\sum_{y \in \mathcal{X}} P(x, y) = 1 \qquad \text{prob. sum to one}$$

Transition graph

Transition matrix sometimes represented by its transition graph

- Nodes are states
- Oriented edges from i to j if $P(i,j) > 0$
- Edges adorned with label $P(i,j)$

Agenda:

- Course organization
- Definition of Markov chains
- First examples

Random walks [possibly infinite state space]

- State space is $\mathbb{Z}$
- At each time step, random walker moves

Right	w.p. p
Left	w.p. $1-p$

- Transition matrix is

$$P(x, x+1) = p \qquad P(x, x-1) = 1-p \qquad P(x, y) = 0 \text{ for all other pairs}$$

- $p = 1/2$ is a symmetric random walk

Extensions

- Can add boundary points: e.g. random walk on $\{0, \ldots, n\}$ with absorbing or reflecting boundaries
- Random walk over $\mathbb{Z}^2$, $\mathbb{Z}^3$, ...
  Move to lattice neighbor selected at random

Random walk on a graph

- A graph $G = (V, E)$ is a set of nodes V and a set of edges connecting pairs of nodes
- Write $v_1 \sim v_2$ if nodes $v_1, v_2 \in V$ are connected
- State space: V
- Dynamics: at each time, the chain chooses a new state uniformly at random from the states adjacent to the current state

This is a called a simple random walk on a graph

Random walk on a directed graph

- In a directed graph, edges have an orientation (think of the www)
- Random walker chooses an outgoing link from current node at random (to be well defined, stay put if no outgoing link, say)
- This models a random web surfer:
  each time you visit a webpage, click at random on of its hyperkinks

Iventory example

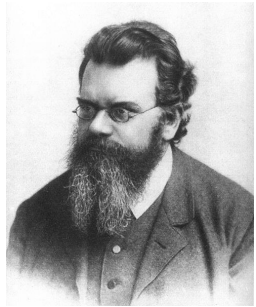
- $\{D_t\}_{t\geq 1}$ is an iid sequence of non-negatiev integers representing demands on consecutive periods $(t-1, t]$
- Commonly used inventory management policy (called an (s, s) policy) regulates the inventory position according to

$$X_{t+1} = \begin{cases} X_t - D_{t+1} & X_t - D_{t+1} \geq s \\ s & \text{otherwise} \end{cases}$$

The Boltzmann-Poincaré controversy



H. Poincaré, 1854-1912



L. Boltzmann, 1844-1906

- Poincaré: a conservative system visits its initial state infinitely often (recurrence theorem)
- Boltzmann: second principle of thermodynamics (Carnot); entropy is increasing, i.e. there is an arrow of time

Understanding this paradox





Paul and Tatiana Ehrenfest

Paul and Tatiana Ehrenfest proposed a probabilistic diffusion model to expose both aspects (reversibility and irreversibility)

The Ehrenfest model

Suppose n molecules of a gas are in a rectangular container divided into two equal parts by a permeable barrier

At each time $t = 1, 2, 3, \ldots$, a molecule selected at random changes side

Microscopic description

- State of system is $x = (x_1, \ldots, x_n)$ where $x_i = 0$ or 1 depending on wehther ith particle is in the 1st or 2nd part
- State space $\{0, 1\}^n$
- Transition matrix

$$P(x, y) = \begin{cases} \frac{1}{n} & \sum_i |x_i - y_i| = 1 \\ 0 & \text{otherwise} \end{cases}$$

The Ehrenfest model

Macroscopic description (unlabeled molecules)

- State of system is x is number of molecules in first container
- State space $\{0, 1, \ldots, n\}$
- Transition matrix

$$P(x,y) = \begin{cases} \frac{n-x}{n} & y = x+1 \text{ and } x < n \\ \frac{x}{n} & y = x-1 \text{ and } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Thermodynamic irreversibility

Eherenfest model

- This system is reversible: if all the molecules are initially on the left, we will find them all on the left infinitely often.
- But the Poincaré's time of return is about 2^n. Imagine this number when $n = 6.02 \times 10^{23}$!
- A statistical description gives the arrow of time
  The system will reach a statistical equilibrium and the entropy is increasing

The traveling salesman problem

A salesman needs to visits $n = 100$ cities

What is the shortest round-trip route that visits each city exactly once and then returns home (the starting city)?

- Routes are chains of segments which connect pairs of cities.
- There are $(n-1)! = 99! \approx 4.6 \times 10^{155}$ possible routes.
- This is an NP-complete problem
  It is impossible to find the best possible route

Randomized solution by simulated annealing

Algorithm selects two segments at random

- If segments cross each other, they are removed and replaced by two new segments which do not cross and have the same four end points
- Otherwise
  - algorithm does nothing with large probability
  - algorithm removes the segments and replace them with two new crossing segments with small probability

Note: algorithm chooses to make mistakes from time to time to avoid local minima

Theorem

If the probability of "re-crossing" decays sufficiently slowly to zero, then the algorithm converges to the optimal route

Epidemiological model

The SIR model without vital dynamic (Kermack and McKendrick)

- Population of constant size N
- Interested $(I(t), S(t), R(t))$: infected, susceptible and recovered at time t
- Ignores so-called vital dynamics (birth and death) sometimes called demography

$$\begin{aligned}
\frac{dS}{dt} &= -\frac{\beta I S}{N} \\
\frac{dI}{dt} &= \frac{\beta I S}{N} - \gamma I \\
\frac{dR}{dt} &= \gamma I
\end{aligned}$$

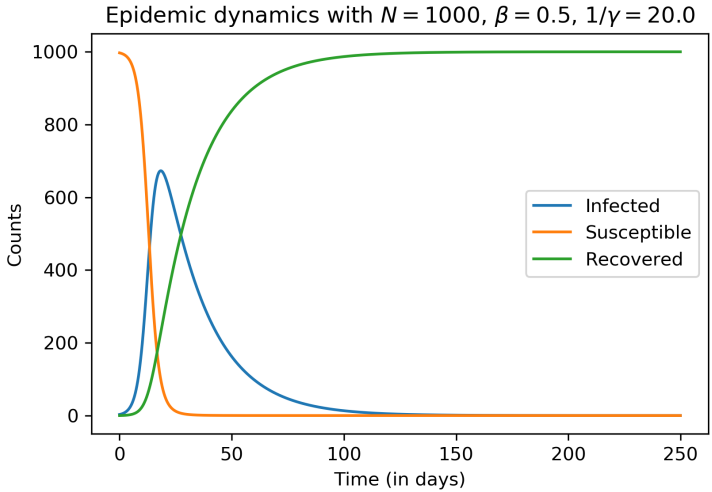
$\beta > 0$ contact rate

$\gamma > 0$ recovery rate

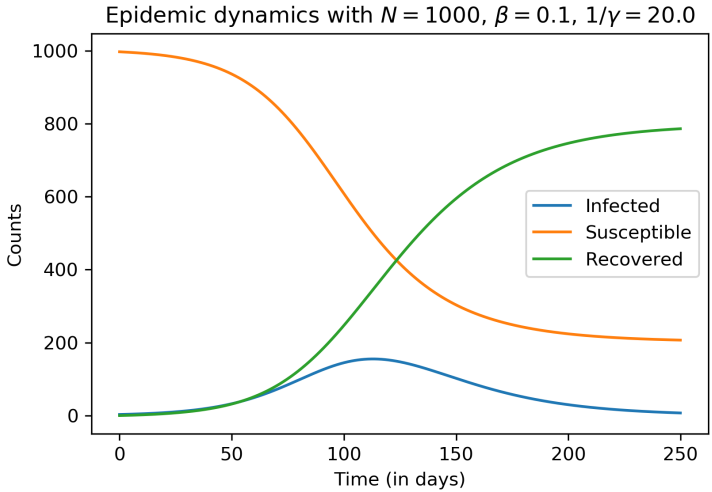
$R_0 = \frac{\beta}{\gamma}$

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$$

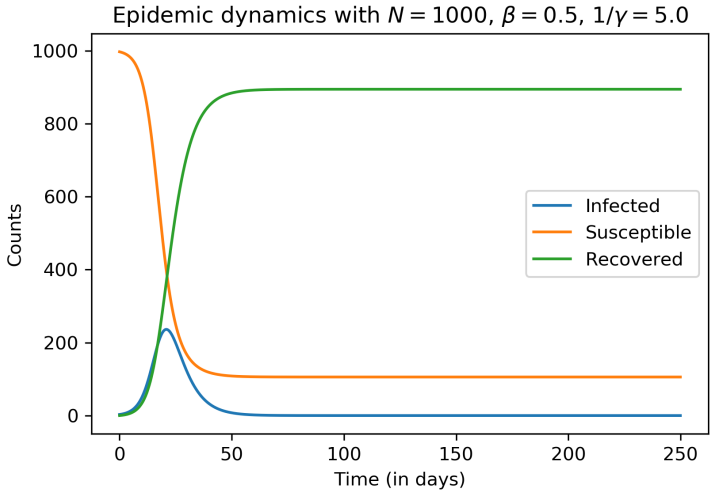
Solution with $N = 1,000$



Solution with $N = 1,000$



Solution with $N = 1,000$



Stochastic version

- State space is $\{S(t) = s, I(t) = i\}$
- Consider evolution over short time window Δt
- Transition probabilities

$$P((s,i),(s',i') = \begin{cases} (\beta/N)is\Delta t & (s',i') = (s-1,i+1) \\ \gamma i \Delta t & (s',i') = (s,i-1) \\ 1-(\beta/N)is\Delta t - \gamma i \Delta t & (s',i') = (s,i) \\ 0 & \text{otherwise} \end{cases}$$

Stochasticity of number of infected

