STATS 217: Introduction to Stochastic Processes I

Lecture 13

Evolution of distributions

- Consider a DTMC $(X_n)_{n\geq 0}$ on S with transition matrix P.
- Suppose we start the chain from a random initial state distributed according to λ. We will use the notation $X_0 \sim \lambda$. This just means that

some prob. dis. on S

$$\mathbb{P}[X_0 = i] = \lambda_i \quad \forall i \in S.$$

Evolution of distributions

- Consider a DTMC $(X_n)_{n\geq 0}$ on S with transition matrix P.
- Suppose we start the chain from a random initial state distributed according to λ. We will use the notation $X_0 \sim \lambda$. This just means that

$$\mathbb{P}[X_0 = i] = \lambda_i \quad \forall i \in S.$$

- What is the distribution of X_1?
- More generally, what is the distribution of X_n?

Evolution of distributions

Recall: $X_0 \sim \lambda$

For any $j \in S$, we have

$$\mathbb{P}[X_n = j] = \sum_{i \in S} \mathbb{P}[X_0 = i \wedge X_n = j]$$

law of total prob.

$$= \sum_{i \in S} \underbrace{\mathbb{P}[X_n = j \mid X_0 = i]}_{= p^n_{ij} = (\underline{P}^n)_{ij}} \underbrace{\mathbb{P}[X_0 = i]}_{= \lambda_i}$$

Evolution of distributions

For any $j \in S$, we have

$$
\begin{aligned}
\mathbb{P}[X_n = j] &= \sum_{i \in S} \mathbb{P}[X_0 = i \wedge X_n = j] \\
&= \sum_{i \in S} \mathbb{P}[X_n = j \mid X_0 = i]\mathbb{P}[X_0 = i] \\
&= \sum_{i \in S} \lambda_i \cdot p_{ij}^n
\end{aligned}
$$

Evolution of distributions

For any $j \in S$, we have

$$\begin{aligned}
\mathbb{P}[X_n = j] &= \sum_{i \in S} \mathbb{P}[X_0 = i \wedge X_n = j] \\
&= \sum_{i \in S} \mathbb{P}[X_n = j \mid X_0 = i]\mathbb{P}[X_0 = i] \\
&= \sum_{i \in S} \lambda_i \cdot p_{ij}^n \\
&= \sum_{i \in S} \lambda_i \cdot (P^n)_{ij}
\end{aligned}$$

$$(\lambda_1 \ldots \lambda_S) \begin{pmatrix} p_{11}^n & \cdots & p_{1S}^n \\ \vdots & & \\ p_{S1}^n & \cdots & p_{SS}^n \end{pmatrix}$$

$$= (\lambda \underline{P}^n)_j$$

sanity check: $\lambda = \delta_i \longrightarrow (\delta_i \underline{P}^n)_j = \underline{P}^n_{ij}$

Evolution of distributions

For any $j \in S$, we have

$$\begin{aligned}
\mathbb{P}[X_n = j] &= \sum_{i \in S} \mathbb{P}[X_0 = i \wedge X_n = j] \\
&= \sum_{i \in S} \mathbb{P}[X_n = j \mid X_0 = i]\mathbb{P}[X_0 = i] \\
&= \sum_{i \in S} \lambda_i \cdot p_{ij}^n \\
&= \sum_{i \in S} \lambda_i \cdot (P^n)_{ij} \\
&= (\lambda P^n)_j.
\end{aligned}$$

Stationary distributions

- So, if $X_0 \sim \lambda$, then $X_n \sim \lambda P^n$.

Stationary distributions

- So, if $X_0 \sim \lambda$, then $X_n \sim \lambda P^n$.
- **A stationary distribution** for P is a probability distribution π on S satisfying

$$\pi P = \pi.$$

Stationary distributions

- So, if $X_0 \sim \lambda$, then $X_n \sim \lambda P^n$.
- **A stationary distribution** for P is a probability distribution π on S satisfying

$$\pi P = \pi.$$

in particular, this is a left eigenvec. for P w/ eigenvalue 1.

- Therefore, if π is a stationary distribution for P, then

$$X_0 \sim \pi \implies X_n \sim \pi \quad \forall n \geq 1.$$

Existence and uniqueness

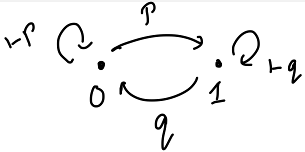
- A Markov chain $(X_n)_{n\geq 0}$ on S with transition matrix P is said to be **irreducible** if all the states for a single communicating class.
- Recall that this means that for all $i, j \in S$, there exists some t (possibly depending on i and j) such that $(P^t)_{i,j} > 0$.
- Recall also that since S is finite, this means that all states in S are recurrent.

Existence and uniqueness

- A Markov chain $(X_n)_{n\geq 0}$ on S with transition matrix P is said to be **irreducible** if all the states for a single communicating class.
- Recall that this means that for all $i, j \in S$, there exists some t (possibly depending on i and j) such that $(P^t)_{i,j} > 0$.
- Recall also that since S is finite, this means that all states in S are recurrent.
- Next time: Let P be the transition matrix of an irreducible Markov chain. Then, there exists a unique probability distribution π satisfying $\pi P = \pi$.

"Perron-Frobenius thm"

Example



Two state chain: $S = \{0, 1\}$ and for $p, q \in (0, 1]$,

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q. \end{bmatrix}$$

- Since $p, q > 0$, the chain is irreducible.
- By the theorem, there is a unique stationary distribution.

$*$ $(\pi_1 \ \ \pi_2)\begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} = (\pi_1 \ \ \pi_2)$

$\pi_1(1-p) + \pi_2 q = \pi_1$

$* \text{ s.t. } \pi_1 + \pi_2 = 1.$

$\Rightarrow \pi_2 q = \pi_1 p$

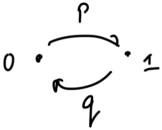
substitute here

$\pi_1 = \frac{q}{p+q}, \quad \pi_2 = \frac{p}{p+q}$

Example

Two state chain: $S = \{0, 1\}$ and for $p, q \in (0, 1]$,

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q. \end{bmatrix}$$



- Since $p, q > 0$, the chain is irreducible.
- By the theorem, there is a unique stationary distribution.
- By solving $\pi P = \pi$ and using that π is a probability distribution, we get (check!) the solution

$$\pi = \left(\frac{q}{p+q}, \frac{p}{p+q} \right).$$

Doubly-stochastic Markov chains

- Consider a DTMC on S with transition matrix P.
- We know that entries of each row of the transition matrix P sum to 1.
- Suppose also that the columns of P sum to 1. Then, P is said to be a **doubly-stochastic transition matrix**.

Doubly-stochastic Markov chains

- Consider a DTMC on S with transition matrix P.
- We know that entries of each row of the transition matrix P sum to 1.
- Suppose also that the columns of P sum to 1. Then, P is said to be a **doubly-stochastic transition matrix**.
- For instance, last time, in our study of waiting times for patterns in coin tossing, we encountered the doubly-stochastic transition matrix

$$P = \begin{bmatrix} & HH & HT & TH & TT \\ HH & 1/2 & 1/2 & 0 & 0 \\ HT & 0 & 0 & 1/2 & 1/2 \\ TH & 1/2 & 1/2 & 0 & 0 \\ TT & 0 & 0 & 1/2 & 1/2 \end{bmatrix}.$$

Doubly-stochastic Markov chains

- Consider a DTMC on S with transition matrix P.
- We know that entries of each row of the transition matrix P sum to 1.
- Suppose also that the columns of P sum to 1. Then, P is said to be a **doubly-stochastic transition matrix**.
- For instance, last time, in our study of waiting times for patterns in coin tossing, we encountered the doubly-stochastic transition matrix

e.g.

$\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$

$$P = \begin{array}{c} \\ HH \\ HT \\ TH \\ TT \end{array} \begin{array}{c} \begin{array}{cccc} HH & HT & TH & TT \end{array} \\ \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix} \end{array}.$$

- Problem 1, Homework 4: Let P be a doubly-stochastic transition matrix on the state space S. Then, the uniform distribution on S is a stationary distribution.

Detailed balance conditions

a priori, need not be stat. dis.

- Consider a DTMC on S with transition matrix P. Let μ be a probability distribution on S.
- We say that μ satisfies the **detailed balance conditions** with respect to P if

$$\mu_i P_{ij} = \mu_j P_{ji} \quad \forall i, j \in S.$$

$(\mu_1, \ldots, \mu_n)$

Detailed balance conditions

- Consider a DTMC on S with transition matrix P. Let μ be a probability distribution on S.
- We say that μ satisfies the **detailed balance conditions** with respect to P if

$$\mu_i P_{ij} = \mu_j P_{ji} \quad \forall i, j \in S.$$

- If μ satisfies the detailed balance conditions with respect to P, then μ is a stationary distribution for P.

Detailed balance conditions

- Consider a DTMC on S with transition matrix P. Let μ be a probability distribution on S.
- We say that μ satisfies the **detailed balance conditions** with respect to P if

$$\mu_i P_{ij} = \mu_j P_{ji} \quad \forall i, j \in S.$$

- If μ satisfies the detailed balance conditions with respect to P, then μ is a stationary distribution for P. wts : $\mu P = \mu \iff (\mu P)_i = \mu_i \ \forall \ i$
- Indeed, for all $i \in S$,

$$(\mu P)_i = \sum_{j \in S} \underbrace{\mu_j P_{ji}}_{= \mu_i P_{ij}} = \sum_{j \in S} \mu_i P_{ij} = \mu_i \Big(\underbrace{\sum_{j \in S} P_{ij}}_{= 1} \Big)$$

.

Detailed balance conditions

- Consider a DTMC on S with transition matrix P. Let μ be a probability distribution on S.
- We say that μ satisfies the **detailed balance conditions** with respect to P if

$$\mu_i P_{ij} = \mu_j P_{ji} \quad \forall i, j \in S.$$

- If μ satisfies the detailed balance conditions with respect to P, then μ is a stationary distribution for P.
- Indeed, for all $i \in S$,

$$(\mu P)_i = \sum_{j \in S} \mu_j P_{ji} = \sum_{j \in S} \mu_i P_{ij}$$

Detailed balance conditions

- Consider a DTMC on S with transition matrix P. Let μ be a probability distribution on S.
- We say that μ satisfies the **detailed balance conditions** with respect to P if

$$\mu_i P_{ij} = \mu_j P_{ji} \quad \forall i, j \in S.$$

- If μ satisfies the detailed balance conditions with respect to P, then μ is a stationary distribution for P.
- Indeed, for all $i \in S$,

$$(\mu P)_i = \sum_{j \in S} \mu_j P_{ji} = \sum_{j \in S} \mu_i P_{ij} = \mu_i \sum_{j \in S} P_{ij}$$

Detailed balance conditions

- Consider a DTMC on S with transition matrix P. Let μ be a probability distribution on S.
- We say that μ satisfies the **detailed balance conditions** with respect to P if

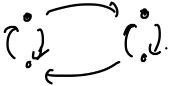
$$\mu_i P_{ij} = \mu_j P_{ji} \quad \forall i, j \in S.$$

note: μ_i is uniform
$\Rightarrow$ P is symmetric

- If μ satisfies the detailed balance conditions with respect to P, then μ is a stationary distribution for P.
- Indeed, for all $i \in S$,

$$(\mu P)_i = \sum_{j \in S} \mu_j P_{ji} = \sum_{j \in S} \mu_i P_{ij} = \mu_i \sum_{j \in S} P_{ij} = \mu_i.$$

non-example :



Detailed balance conditions (DBC)

if π satisfies DBC wrt $\underline{P}$, also say that $\underline{P}$ is "reversible".

- If π satisfies the detailed balance conditions with respect to P and $X_0 \sim \pi$, then

$$\mathbb{P}[X_0 = x_0, \ldots, X_n = x_n] = \pi_{x_0} P_{x_0,x_1} \ldots P_{x_{n-1},x_n}$$

$$\|$$

$$\underline{\mathbb{P}}[X_0 = x_n, \ldots, X_n = x_0]$$

Detailed balance conditions

- If π satisfies the detailed balance conditions with respect to P and $X_0 \sim \pi$, then

$\pi_{x_0} P_{x_0 x_1} = \pi_{x_1} P_{x_1 x_0}$

$$\begin{aligned}\mathbb{P}[X_0 = x_0, \ldots, X_n = x_n] &= \pi_{x_0} P_{x_0,x_1} \ldots P_{x_{n-1},x_n} \\ &= P_{x_1,x_0} \cdot \underbrace{\pi_{x_1} P_{x_1,x_2}}_{= \pi_{x_2} P_{x_2, x_1}} \ldots P_{x_{n-1},x_n}\end{aligned}$$

Detailed balance conditions

- If π satisfies the detailed balance conditions with respect to P and $X_0 \sim \pi$, then

$$\begin{aligned}\mathbb{P}[X_0 = x_0, \ldots, X_n = x_n] &= \pi_{x_0} P_{x_0,x_1} \ldots P_{x_{n-1},x_n} \\ &= P_{x_1,x_0} \cdot \pi_{x_1} P_{x_1,x_2} \ldots P_{x_{n-1},x_n} \\ &= P_{x_1,x_0} P_{x_2,x_1} \cdot \pi_{x_2} P_{x_2,x_3} \ldots P_{x_{n-1},x_n}\end{aligned}$$

Detailed balance conditions

- If π satisfies the detailed balance conditions with respect to P and $X_0 \sim \pi$, then

$$\begin{aligned}
\mathbb{P}[X_0 = x_0, \dots, X_n = x_n] &= \pi_{x_0} P_{x_0,x_1} \dots P_{x_{n-1},x_n} \\
&= P_{x_1,x_0} \cdot \pi_{x_1} P_{x_1,x_2} \dots P_{x_{n-1},x_n} \\
&= P_{x_1,x_0} P_{x_2,x_1} \cdot \pi_{x_2} P_{x_2,x_3} \dots P_{x_{n-1},x_n} \\
&= \dots \\
&= \pi_{x_n} P_{x_n,x_{n-1}} \dots P_{x_1,x_0}
\end{aligned}$$

Detailed balance conditions

- If π satisfies the detailed balance conditions with respect to P and $X_0 \sim \pi$, then

$$\begin{aligned}
\mathbb{P}[X_0 = x_0, \dots, X_n = x_n] &= \pi_{x_0} P_{x_0,x_1} \dots P_{x_{n-1},x_n} \\
&= P_{x_1,x_0} \cdot \pi_{x_1} P_{x_1,x_2} \dots P_{x_{n-1},x_n} \\
&= P_{x_1,x_0} P_{x_2,x_1} \cdot \pi_{x_2} P_{x_2,x_3} \dots P_{x_{n-1},x_n} \\
&= \dots \\
&= \pi_{x_n} P_{x_n,x_{n-1}} \dots P_{x_1,x_0} \\
&= \mathbb{P}[X_0 = x_n, \dots, X_n = x_0].
\end{aligned}$$

- For this reason, such chains are also called **reversible**.

if $X_0 \sim \pi$

$(X_0, \dots, X_n) \sim (X_n, \dots, X_0)$.

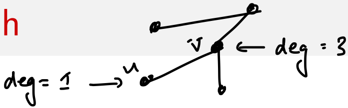
Detailed balance conditions

- If π satisfies the detailed balance conditions with respect to P and $X_0 \sim \pi$, then

$$\begin{aligned}
\mathbb{P}[X_0 = x_0, \ldots, X_n = x_n] &= \pi_{x_0} P_{x_0,x_1} \ldots P_{x_{n-1},x_n} \\
&= P_{x_1,x_0} \cdot \pi_{x_1} P_{x_1,x_2} \ldots P_{x_{n-1},x_n} \\
&= P_{x_1,x_0} P_{x_2,x_1} \cdot \pi_{x_2} P_{x_2,x_3} \ldots P_{x_{n-1},x_n} \\
&= \ldots \\
&= \pi_{x_n} P_{x_n,x_{n-1}} \ldots P_{x_1,x_0} \\
&= \mathbb{P}[X_0 = x_n, \ldots, X_n = x_0].
\end{aligned}$$

- For this reason, such chains are also called **reversible**.
- In many interesting examples, the detailed balance conditions provide an efficient way of finding the stationary distribution.

Example: Random walk on a graph



- $G = (V, E)$ is a graph, where V is the set of vertices and E is the set of edges.
- For vertices $u \neq v \in V$, we say that $u \sim v$ if and only if there is an edge between u and v.
- For a vertex $u \in V$, $\deg(u)$ denotes the degree of u i.e. the number of vertices it is connected to.

$$\sum_{u \in V} \deg(u) = 2|E|$$

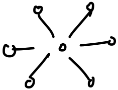
Example: Random walk on a graph

- $G = (V, E)$ is a graph, where V is the set of vertices and E is the set of edges.
- For vertices $u \neq v \in V$, we say that $u \sim v$ if and only if there is an edge between u and v.
- For a vertex $u \in V$, $\deg(u)$ denotes the degree of u i.e. the number of vertices it is connected to.
- Note that $\sum_{u \in V} \deg(u) = 2|E|$. "handshaking lemma".

Example: Random walk on a graph

- $G = (V, E)$ is a graph, where V is the set of vertices and E is the set of edges.
- For vertices $u \neq v \in V$, we say that $u \sim v$ if and only if there is an edge between u and v.
- For a vertex $u \in V$, $\deg(u)$ denotes the degree of u i.e. the number of vertices it is connected to.
- Note that $\sum_{u \in V} \deg(u) = 2|E|$.
- Recall that the transition matrix of the random walk is given by

$$P_{u,v} = \begin{cases} \frac{1}{\deg(u)} & \text{if } v \sim u \\ 0 & \text{otherwise.} \end{cases}$$



Example: Random walk on a graph

- Consider the distribution π where $\pi_u = \deg(u)/2|E|$.
- Then, π is a probability distribution.

Example: Random walk on a graph

- Consider the distribution π where $\pi_u = \deg(u)/2|E|$.
- Then, π is a probability distribution.
- We claim that π satisfies the detailed balance conditions with respect to P.

need to check: $\pi_u P_{uv} = \pi_v P_{vu} \quad \forall\, u, v.$

① case $\underline{1}$: $u \not\sim v$: both sides are 0

② case 2: $u \sim v$. $\quad \pi_u \underbrace{P_{uv}}_{= \frac{1}{\deg(u)}} = \frac{\cancel{\deg(u)}}{2|E|} \frac{1}{\cancel{\deg(u)}}$

$\pi_v P_{vu}$

Example: Random walk on a graph

- Consider the distribution π where $\pi_u = \deg(u)/2|E|$.
- Then, π is a probability distribution.
- We claim that π satisfies the detailed balance conditions with respect to P.
- There are two cases. If $u \not\sim v$, then the condition is clearly satisfied since $P_{uv} = P_{vu} = 0$.

Example: Random walk on a graph

- Consider the distribution π where $\pi_u = \deg(u)/2|E|$.
- Then, π is a probability distribution.
- We claim that π satisfies the detailed balance conditions with respect to P.
- There are two cases. If $u \not\sim v$, then the condition is clearly satisfied since $P_{uv} = P_{vu} = 0$.
- If $u \sim v$, then

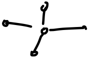
$$\pi_u P_{uv} = \frac{\deg(u)}{2|E|} \cdot \frac{1}{\deg(u)} = \frac{1}{2|E|} = \pi_v P_{vu}.$$

Example: Random walk on a graph

- Consider the distribution π where $\pi_u = \deg(u)/2|E|$.
- Then, π is a probability distribution.
- We claim that π satisfies the detailed balance conditions with respect to P.
- There are two cases. If $u \not\sim v$, then the condition is clearly satisfied since $P_{uv} = P_{vu} = 0$.
- If $u \sim v$, then

$$\pi_u P_{uv} = \frac{\deg(u)}{2|E|} \cdot \frac{1}{\deg(u)} = \frac{1}{2|E|} = \pi_v P_{vu}.$$

- Therefore, π is a stationary distribution for P. Note that P is irreducible if and only if the graph G is connected i.e., there is a path from any vertex to any other vertex, in which case, π is the unique stationary distribution.




Example: The Ehrenfest urn

The Ehrenfest urn. n balls are distributed among two urns, urn A and urn B. At each time, we select a ball uniformly at random and move it from its current urn to the other urn.

Example: The Ehrenfest urn

The Ehrenfest urn. n balls are distributed among two urns, urn A and urn B. At each time, we select a ball uniformly at random and move it from its current urn to the other urn.

- Let X_t denote the number of balls in urn A at time t. Then, $(X_t)_{t \geq 0}$ is a DTMC on $\{\overset{0}{\cancel{1}}, \ldots, n\}$ with transition matrix P given by

$$P_{jk} = \begin{cases} j/n & \text{if } k = j - 1 \\ (n - j)/n & \text{if } k = j + 1 \\ 0 & \text{otherwise.} \end{cases}$$

Example: The Ehrenfest urn

The Ehrenfest urn. n balls are distributed among two urns, urn A and urn B. At each time, we select a ball uniformly at random and move it from its current urn to the other urn.

- Let X_t denote the number of balls in urn A at time t. Then, $(X_t)_{t\geq 0}$ is a DTMC on $\{1, \ldots, n\}$ with transition matrix P given by

$$P_{jk} = \begin{cases} j/n & \text{if } k = j-1 \\ (n-j)/n & \text{if } k = j+1 \\ 0 & \text{otherwise.} \end{cases}$$

- Note that P is clearly irreducible.

$\left\{ \text{by the thm, } \exists \text{ unique stationary dis.} \right\}$

Example: The Ehrenfest urn

- Let π be the distribution on $\{0, \ldots, n\}$ given by

$$\pi_x = 2^{-n} \cdot \binom{n}{x}.$$

- By the binomial theorem, π is a probability distribution.

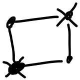
Example: The Ehrenfest urn

intuition: consider the random walk on $\{0,1\}^n$

· n neighbors, move to one of them uniformly

- Let π be the distribution on $\{0, \ldots, n\}$ given by

$$\pi_x = 2^{-n} \cdot \binom{n}{x}.$$



- By the binomial theorem, π is a probability distribution.
- **Exercise:** check that π satisfies the detailed balance condition with respect to P.
- Hence, π is the unique stationary distribution for P.

$\pi_x \quad P_{x,x+1}$ $\qquad\qquad$ $\pi_{x+1} \quad P_{x+1,x}$

$2^{-n}\binom{n}{x}\frac{n-x}{n}$ $\qquad\qquad$ $2^{-n}\binom{n}{x+1}\frac{x+1}{n}$

ex: think about the # of 1s at the current coord.