S&DS 351 - Stochastic Processes - Lecture 3 Notes

Bryan SebaRaj

January 15, 2025

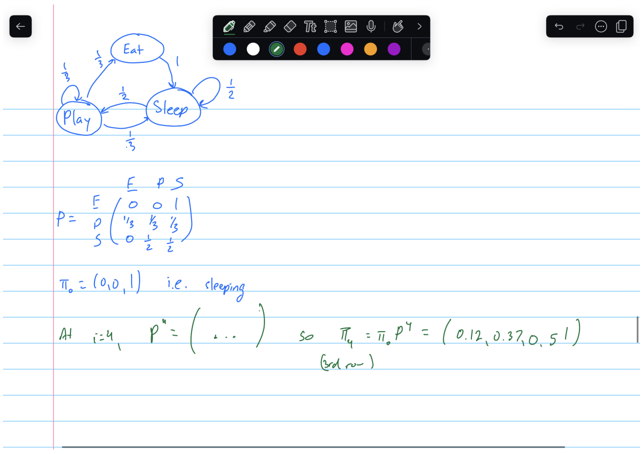
1 Recall

- Markov Chain: $\mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n, ..., X_0 = i_0) = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n)$
- Transition Matrix: $P = [p_{ij}]$ where $p_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i) = \mathbb{P}(X_1 = j | X_0 = i)$
- $\mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, ..., X_0 = i_0) = \Pi_0(i_0) P_{0,1} ... P_{n-1,n}$

2 Marginal distribution at time n

- MC: P - transition matrix
- $(\Pi_0) \in S$ - initialization distribution
- Assume $N = |S| < +\infty$
- $X_0 \sim \Pi_0$, $X_1 \sim \Pi_1$, ..., $X_n \sim \Pi_n$
- $\forall j \in S$,
$$\Pi_1(j) = \mathbb{P}_{x_0, \Pi_0}(X_1 = j) = \sum_{i \in S} \mathbb{P}(X_1 = j | X_0 = i) = \sum_{i \in S} \Pi_0(i) P_{i,j}$$
- which is the same as $\Pi_1 = \Pi_0 P$
- with identical reasoning, $\Pi_2 = \Pi_1 P = \Pi_0 p^2$, $\Pi_3 = \Pi_2 P = \Pi_0 P^3$, ...
- by induction, $\Pi_n = \Pi_0 P^n$

2.1 Example: Markov Dog



- Claim: $\forall i, j \in S$, $\forall n, (P^n)_{i,j} = \mathbb{P}(X_n = j | x_0 = i)$
- Proof: $\forall i \in S, \Pi_0 = (00...01...0)$ where $i \in S$ is 1
- Π_n is the distribution of X_n

- $\Pi_n = \Pi_0 P^n = i^{th}$ row of P^n
- $\Pi_n j = P^n_{i,j},\ \forall j \in S$
- $\mathbb{P}(X_n = j | X_0 = i)$
- Claim: $\forall i, j \in S, n \in \mathbb{N}$
- $(P^n)_{i,j} = \sum_{i_1,...,i_{n-1} \in S} P_{i_0,i_1} P_{i_1,i_2} ... P_{n-1,n}$
- i.e. all possible walks from i to j of any length less than n
- Proof:
  - We know

$$\begin{aligned}(P^n)_{i,j} = \mathbb{P}(X_n = j | X_0 = i) &= \sum_{i_1,...,i_{n-1} \in S} \mathbb{P}(X_n = j, X_{n-1} = i_{n-1}, ..., X_1 = i_1 | X_0 = i_0) \\ &= \sum_{i_1,...,i_{n-1} \in S} P_{i_0,i_1} ... P_{i_{n-1},n}\end{aligned}$$

But this is often way too computationally expensive. Note the cardinality of a matrix with a large n

2.2 Smarter way to calculate

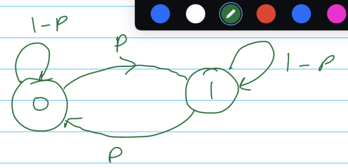
Yes, if

- P is diagonalizable
- We can quickly find its eigen decomposition
- Recall: eigenvectors and eigenvalues of matrices
- If matrix $A \in \mathbb{R}$ is symmetric, then it has n real eigenvalues and n eigenvectors that are linearly independent
- $\exists q \in \mathbb{R}^{m \times n}$ and the inv. is D-diagonal, then $A = qDq^{-1}$

2.3 Applcation to Markov Chains

- Observe, $\forall m, A^m = qDq^{-1}qDq^{-1}... = qD^m q^{-1}$
- and D^m is trivial easy to calculate (review), since its diagonal

2.4 Example:



$$P = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1-p & p \\ 1 & p & 1-p \end{array}$$

Eigenvectors: $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ w/ eigenvalues 1 and $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ w/ eigvalue $1-2p$

$$Q^{-1} = Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \qquad D = \begin{pmatrix} 1 & 0 \\ 0 & 1-2p \end{pmatrix}$$

$$P = Q D Q^{-1} \quad \text{so} \quad P^m = Q D^m Q^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1^m & 0 \\ 0 & (1-2p)^m \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$P_m = \begin{pmatrix} \frac{1}{2}(1-(1-2p)^m) & \frac{1}{2}(1+(1-2p)^m) \\ \frac{1}{2}(1+(1-2p)^m) & \frac{1}{2}(1-(1-2p)^m) \end{pmatrix}$$

- Are all MCs linear algebra?
- Not really, especially if $|S|$ is not "small enough" or P is not diagonalizable
- Example:
  - deck of 52 cards, and you want to shuffle it
  - Step 1: pick a uniformly random card
  - Step 2: put in on the top of the deck
  - Question: how many shuffles lead to a uniform "global shuffle" of the whole deck
  - S is the set of all permutations of 52 cards
  - $|S| = 52!$
  - $\Pi_n \approx (\frac{1}{52!}, ...)$. state space is too large, so need to think of it probabalistically and not just through the lens of linear algebra

2.5 Remarks

- iid process is a Markov Chain
  - $X_0, X_1, ..., X_n$ iid from μ on S
  - $\mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}, ..., X_0 = i_0) = \mathbb{P}(X_n = i_n) = \mu(i_n) = \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1})$
  - $P =$ rows of μ
- higher order markov chains
  - a stochastic process $X_0, X_1, ..., X_n$ is a Markov Chain
  - All X_n take value in a countable set S
  - $\forall n, i_0, ..., i_n \in S$,

$$\mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}, ..., X_0 = i_0) = \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}, ..., X_{n-i+1} = i_{n-i+1}$$

  - ObservationL any r^{th} order MC can be analyzed via a markov chain
  - Define $Y_n = (X_n, Y_{n+1}), n \in \mathbb{N}$
  - Takes values on $S \times S$

– $\forall n, i_0, ..., i_n, i_{n+1} \in S$,

$$\begin{aligned}
&\mathbb{P}(Y_{n+1} = (i_{n+1}, i_{n+2}|Y_n = (i_n, i_{n+1})..., Y_0 = (i_0, i_1)) \\
&= \mathbb{P}(X_{n+2} = i_{n+2}|X_{n+1} = i_{n+1}, ..., X_0 = i_0) \\
&= \mathbb{P}(X_{n+2} = i_{n+2}|X_{n+1} = i_{n+1}, X_n = i_n) \\
&= \mathbb{P}(X_{n+2} = i_{n+2}, X_{n+1} = i_{n+1}|X_{n+1} = i_{n+1}) \\
&= \mathbb{P}(Y_{n+1} = (i_{n+1}, i_{n+2})|Y_n = (i_n, i_{n+1})
\end{aligned}$$