Insert Markov pun here

A talk on Probabilistic Automata and Markov Chains

Aniruddha Deb

IIT Delhi

June 10, 2021

Stochastic vectors and stochastic matrices

Stochastic vectors and stochastic matrices

A **Stochastic Vector** $\vec{a}$ is a vector in $\mathbb{R}^n$ which satisfies the following conditions:

- if $s \in \vec{a}$, then $0 < s < 1$
- the sum of all the elements of $\vec{a}$ is 1

Stochastic vectors and stochastic matrices

A **Stochastic Vector** $\vec{a}$ is a vector in $\mathbb{R}^n$ which satisfies the following conditions:

- if $s \in \vec{a}$, then $0 < s < 1$
- the sum of all the elements of $\vec{a}$ is 1

A **Stochastic Matrix** (aka Markov matrix) $\mathcal{S}$ is a matrix whose column vectors are stochastic vectors.

Stochastic vectors and stochastic matrices

A **Stochastic Vector** $\vec{a}$ is a vector in $\mathbb{R}^n$ which satisfies the following conditions:

- if $s \in \vec{a}$, then $0 < s < 1$
- the sum of all the elements of $\vec{a}$ is 1

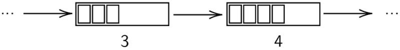
A **Stochastic Matrix** (aka Markov matrix) $\mathcal{S}$ is a matrix whose column vectors are stochastic vectors.

Stochastic vectors can be thought of as a form of **discrete probability distribution**: given a finite number of possible states, $\vec{a}$ gives the probability of transitioning to the ith state.

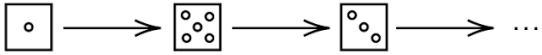
A short recap of NFA's

Deterministic and Nondeterministic Finite State Machines

Combining both these concepts, a **Deterministic Finite State Automata (DFA)** is one whose next state is uniquely determined by the current state. As an example, a loading bar at n% can only progress to a state at (n+1)%



In contrast, a **Nondeterministic Finite State Automata (NFA)** is one whose next state is not uniquely determined by it's current state. An example would be the output of a dice roll: it's a NFA, because the number of states are finite, and the next state is independent of the current state

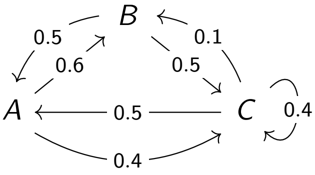


Probabilistic Automata

A **Probabilistic Automata** (PA) generalizes the concept of a NFA, so that the transition probabilities are given by the transition function (the transition function for such an automata is a stochastic matrix / markov matrix)

Probabilistic Automata

A **Probabilistic Automata** (PA) generalizes the concept of a NFA, so that the transition probabilities are given by the transition function (the transition function for such an automata is a stochastic matrix / markov matrix)

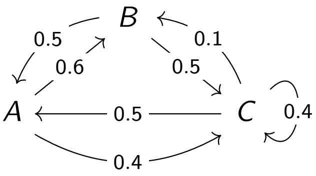


Modeling PA's: Markov Processes

PA's lend themselves to being modeled by markov matrices very well. As an example, the process in the previous slide can be modeled as follows:

Modeling PA's: Markov Processes

PA's lend themselves to being modeled by markov matrices very well. As an example, the process in the previous slide can be modeled as follows:

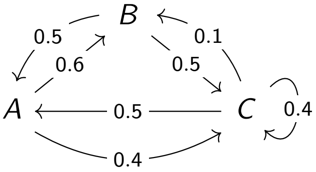


$$\vec{a}_A = \begin{bmatrix} 0 \\ 0.6 \\ 0.4 \end{bmatrix} \quad \vec{a}_B = \begin{bmatrix} 0.5 \\ 0 \\ 0.5 \end{bmatrix} \quad \vec{a}_C = \begin{bmatrix} 0.5 \\ 0.1 \\ 0.4 \end{bmatrix}$$

$$\mathcal{S} = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.6 & 0 & 0.1 \\ 0.4 & 0.5 & 0.4 \end{bmatrix}$$

Modeling PA's: Markov Processes

PA's lend themselves to being modeled by markov matrices very well. As an example, the process in the previous slide can be modeled as follows:



$$\vec{a}_A = \begin{bmatrix} 0 \\ 0.6 \\ 0.4 \end{bmatrix} \; \vec{a}_B = \begin{bmatrix} 0.5 \\ 0 \\ 0.5 \end{bmatrix} \; \vec{a}_C = \begin{bmatrix} 0.5 \\ 0.1 \\ 0.4 \end{bmatrix}$$

$$\mathcal{S} = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.6 & 0 & 0.1 \\ 0.4 & 0.5 & 0.4 \end{bmatrix}$$

Now, given an initial state a, $\mathcal{S}a$ would give the final state after one transition. By induction, $\mathcal{S}^n a$ would give the final state after n transitions.

More on Markov Matrices

- A markov matrix $\mathcal{S}$ multiplied by a stochastic vector $\vec{a}$ gives a stochastic vector

Proof: Consider the row vector $T = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}$. This acts like a map from $\mathbb{R}^n \to \mathbb{R}$ defined as $T\vec{a} = a_1 + a_2 + \dots + a_n$. For a stochastic vector $\vec{a}$, $T\vec{a} = 1$. If we apply this map on $\mathcal{S}\vec{a}$, then we get

$$\begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix} \mathcal{S}\vec{a} = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix} \vec{a} = 1$$

(Since every column vector of a stochastic matrix is a stochastic vector, $T\mathcal{S} = T$)

Hence, $\vec{a^1} = \mathcal{S}\vec{a}$ is a stochastic vector. By induction, $\vec{a^n} = \mathcal{S}^n\vec{a}$ is also a stochastic vector. $\blacksquare$

Eigenvalues and Eigenvectors of $\mathcal{S}$

Eigenvalues and Eigenvectors of $\mathcal{S}$

Can you find an eigenvalue of $\mathcal{S}$ by inspection?

More verbosely, what is an eigenvalue of a matrix, all of whose columns sum to 1?

Eigenvalues and Eigenvectors of $\mathcal{S}$

Can you find an eigenvalue of $\mathcal{S}$ by inspection?

More verbosely, what is an eigenvalue of a matrix, all of whose columns sum to 1?

- 1 is an eigenvalue of every stochastic matrix $\mathcal{S}$

Proof: we need to show $|\mathcal{S} - I| = 0$. Notice that since every column of $\mathcal{S}$ sums to 1, every column of $\mathcal{S} - I$ would sum to 0. Hence $|\mathcal{S} - I| = 0$. ■

- All other eigenvalues of $\mathcal{S}$ are less than 1

Proof: Suppose there exists $\lambda > 1$ such that $\mathcal{S}\vec{a} = \lambda\vec{a}$. let a_L be the largest element of $\vec{a}$. Every entry in $\lambda\vec{a}$ can be written as

$$a_i = \sum_{j=1}^{n} a_j \mathcal{S}_{ij} < a_L \sum_{j=1}^{n} \mathcal{S}_{ij} = a_L$$

but since $\lambda a_L \in \lambda\vec{a}$ and $\lambda > 1$, we have a contradiction. Hence, all the eigenvalues of $\mathcal{S}$ are less than 1. ■

Analysing the steady state

What is a steady state? When the PA reaches a state that does not change in the next iteration, we call it a steady state. Using a markov matrix, we can define it as:

$$\mathcal{S}\vec{\tau} = \vec{\tau}$$

It's easy to see that $\vec{\tau}$ is the eigenvector corresponding to the eigenvalue 1. The question then arises

Proof: