Stochastic Matrix

Definition

Stochastic Matrix: We say that a matrix, $A \in \mathbb{R}^{n \times n}$, is a stochastic matrix if and only if

Stochastic Matrix

Definition

Stochastic Matrix: We say that a matrix, $A \in \mathbb{R}^{n \times n}$, is a stochastic matrix if and only if

1. All elements of A are non-negative. IE for all $1 \leq i, j \leq n$, $a_{ij} \geq 0$

Stochastic Matrix

Definition

Stochastic Matrix: We say that a matrix, $A \in \mathbb{R}^{n \times n}$, is a stochastic matrix if and only if

1. All elements of A are non-negative. IE for all $1 \leq i, j \leq n$, $a_{ij} \geq 0$
2. All columns sum to 1.

Stochastic Matrix

Definition

Stochastic Matrix: We say that a matrix, $A \in \mathbb{R}^{n \times n}$, is a stochastic matrix if and only if

1. All elements of A are non-negative. IE for all $1 \leq i, j \leq n$, $a_{ij} \geq 0$
2. All columns sum to 1.

Example

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & 0 \\ \frac{1}{2} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 1 \end{bmatrix}$$

Positive Stochastic Matrix

Definition

Positive Stochastic Matrix: We say that a matrix, $A \in \mathbb{R}^{n \times n}$, is a positive stochastic matrix if and only if

Positive Stochastic Matrix

Definition

Positive Stochastic Matrix: We say that a matrix, $A \in \mathbb{R}^{n \times n}$, is a positive stochastic matrix if and only if

1. A is a stochastic matrix

Positive Stochastic Matrix

Definition

Positive Stochastic Matrix: We say that a matrix, $A \in \mathbb{R}^{n \times n}$, is a positive stochastic matrix if and only if

1. A is a stochastic matrix
2. All elements of A are positive. IE for all $1 \leq i, j \leq n$, $a_{ij} > 0$

Positive Stochastic Matrix

Definition

Positive Stochastic Matrix: We say that a matrix, $A \in \mathbb{R}^{n\times n}$, is a positive stochastic matrix if and only if

1. A is a stochastic matrix
2. All elements of A are positive. IE for all $1 \leq i, j \leq n$, $a_{ij} > 0$

Example

$$A = \begin{bmatrix} .1 & .7 & .3 \\ .8 & .2 & .3 \\ .1 & .1 & .4 \end{bmatrix}$$

Identifying (Positive) Stochastic Matrices

Determine which of the following matrices are positive stochastic, stochastic, or neither.

Identifying (Positive) Stochastic Matrices

Determine which of the following matrices are positive stochastic, stochastic, or neither.

$$A = \begin{bmatrix} .1 & .2 & .7 \\ .2 & .3 & .5 \\ .4 & .5 & .1 \end{bmatrix}, B = \begin{bmatrix} .3 & .8 & .15 \\ .3 & .1 & .05 \\ .4 & .1 & .8 \end{bmatrix}, C = \begin{bmatrix} .2 & .8 & 0 \\ .7 & .1 & .2 \\ .5 & .1 & .8 \end{bmatrix}, D = \begin{bmatrix} .1 & .2 & .2 \\ .3 & .1 & .2 \\ .5 & .1 & .2 \end{bmatrix}, E = \begin{bmatrix} .2 & .8 & 0 \\ .3 & .1 & .2 \\ .5 & .1 & .8 \end{bmatrix}$$

Eigenvalues of Stochastic Matrices Part 1

Theorem

Let $A \in \mathbb{R}^{n\times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Eigenvalues of Stochastic Matrices Part 1

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

First we show that 1 is an eigenvalue of A. Recall that A and $A^\top$ have the same eigenvalues. Since we know that A is stochastic, we know it's columns sum to 1. Therefore, the rows of $A^\top$ also sum to 1.

Eigenvalues of Stochastic Matrices Part 1

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

First we show that 1 is an eigenvalue of A. Recall that A and $A^\top$ have the same eigenvalues. Since we know that A is stochastic, we know it's columns sum to 1. Therefore, the rows of $A^\top$ also sum to 1. In other words, for each $1 \leq i \leq n$

$$\sum_{j=1}^{n} \left(A^\top\right)_{ij}$$

Eigenvalues of Stochastic Matrices Part 1

Theorem

Let $A \in \mathbb{R}^{n\times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

First we show that 1 is an eigenvalue of A. Recall that A and $A^\top$ have the same eigenvalues. Since we know that A is stochastic, we know it's columns sum to 1. Therefore, the rows of $A^\top$ also sum to 1. In other words, for each $1 \leq i \leq n$

$$\sum_{j=1}^{n} \left(A^\top\right)_{ij} = \sum_{j=1}^{n} A_{ji}$$

Eigenvalues of Stochastic Matrices Part 1

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

First we show that 1 is an eigenvalue of A. Recall that A and $A^\top$ have the same eigenvalues. Since we know that A is stochastic, we know it's columns sum to 1. Therefore, the rows of $A^\top$ also sum to 1. In other words, for each $1 \leq i \leq n$

$$\sum_{j=1}^{n} \left(A^\top\right)_{ij} = \sum_{j=1}^{n} A_{ji} = 1$$

Let $\mathbf{1}_n$ be the vector of all 1's, then

Eigenvalues of Stochastic Matrices Part 1

Theorem

Let $A \in \mathbb{R}^{n\times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

First we show that 1 is an eigenvalue of A. Recall that A and $A^\top$ have the same eigenvalues. Since we know that A is stochastic, we know it's columns sum to 1. Therefore, the rows of $A^\top$ also sum to 1. In other words, for each $1 \leq i \leq n$

$$\sum_{j=1}^{n} \left(A^\top\right)_{ij} = \sum_{j=1}^{n} A_{ji} = 1$$

Let $\mathbf{1}_n$ be the vector of all 1's, then

$$A^\top \mathbf{1}_n = \mathbf{1}_n$$

So, $(\lambda, \mathbf{1}_n)$ is an eigenpair of A. □

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n \times n}$ *be a stochastic matrix. Then we know that* $\lambda = 1$ *is an eigenvalue of* A *and if* λ *is an eigenvalue of* A *then* $|\lambda| \leq 1$

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

Now, we will show that if λ is an eigenvalue of A, then $|\lambda| \leq 1$.

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

Now, we will show that if λ is an eigenvalue of A, then $|\lambda| \leq 1$.
Recall as before that A and $A^\top$ share eigenvalues, so let $(\lambda, \mathbf{x})$ be an eigenpair of $A^\top$.

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

Now, we will show that if λ is an eigenvalue of A, then $|\lambda| \leq 1$.
Recall as before that A and $A^\top$ share eigenvalues, so let $(\lambda, \mathbf{x})$ be an eigenpair of $A^\top$. This means $A^\top \mathbf{x} = \lambda \mathbf{x}$.

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n\times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

Now, we will show that if λ is an eigenvalue of A, then $|\lambda| \leq 1$.
Recall as before that A and $A^\top$ share eigenvalues, so let $(\lambda, \mathbf{x})$ be an eigenpair of $A^\top$. This means $A^\top \mathbf{x} = \lambda \mathbf{x}$.
Consider the j^{th} entry of $\lambda \mathbf{x}$.

$$(\lambda \mathbf{x}) = \lambda x_j = \sum_{i=1}^{n} A_{ij} x_i$$

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n\times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

Now, we will show that if λ is an eigenvalue of A, then $|\lambda| \leq 1$.
Recall as before that A and $A^\top$ share eigenvalues, so let $(\lambda, \mathbf{x})$ be an eigenpair of $A^\top$. This means $A^\top\mathbf{x} = \lambda\mathbf{x}$.
Consider the j^{th} entry of $\lambda\mathbf{x}$.

$$(\lambda\mathbf{x}) = \lambda x_j = \sum_{i=1}^{n} A_{ij} x_i$$

If we pick the index j such that $|x_j| \leq |x_i|$ for all i, then we have that

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

Now, we will show that if λ is an eigenvalue of A, then $|\lambda| \leq 1$.
Recall as before that A and $A^\top$ share eigenvalues, so let $(\lambda, \mathbf{x})$ be an eigenpair of $A^\top$. This means $A^\top \mathbf{x} = \lambda \mathbf{x}$.
Consider the j^{th} entry of $\lambda \mathbf{x}$.

$$(\lambda \mathbf{x}) = \lambda x_j = \sum_{i=1}^{n} A_{ij} x_i$$

If we pick the index j such that $|x_j| \leq |x_i|$ for all i, then we have that

$$|\lambda x_j| = |\lambda| \, |x_j|$$

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

Now, we will show that if λ is an eigenvalue of A, then $|\lambda| \leq 1$.
Recall as before that A and $A^\top$ share eigenvalues, so let $(\lambda, \mathbf{x})$ be an eigenpair of $A^\top$. This means $A^\top \mathbf{x} = \lambda \mathbf{x}$.
Consider the j^{th} entry of $\lambda \mathbf{x}$.

$$(\lambda \mathbf{x}) = \lambda x_j = \sum_{i=1}^{n} A_{ij} x_i$$

If we pick the index j such that $|x_j| \leq |x_i|$ for all i, then we have that

$$|\lambda x_j| = |\lambda|\,|x_j| = \left| \sum_{i=1}^{n} A_{ij} x_i \right|$$

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

Now, we will show that if λ is an eigenvalue of A, then $|\lambda| \leq 1$.
Recall as before that A and $A^\top$ share eigenvalues, so let $(\lambda, \mathbf{x})$ be an eigenpair of $A^\top$. This means $A^\top \mathbf{x} = \lambda \mathbf{x}$.
Consider the j^{th} entry of $\lambda \mathbf{x}$.

$$(\lambda \mathbf{x}) = \lambda x_j = \sum_{i=1}^{n} A_{ij} x_i$$

If we pick the index j such that $|x_j| \leq |x_i|$ for all i, then we have that

$$|\lambda x_j| = |\lambda| \, |x_j| \; = \left| \sum_{i=1}^{n} A_{ij} x_i \right| \leq \sum_{i=1}^{n} |A_{ij}| \, |x_i|$$

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

Now, we will show that if λ is an eigenvalue of A, then $|\lambda| \leq 1$.
Recall as before that A and $A^\top$ share eigenvalues, so let $(\lambda, \mathbf{x})$ be an eigenpair of $A^\top$. This means $A^\top \mathbf{x} = \lambda \mathbf{x}$.
Consider the j^{th} entry of $\lambda \mathbf{x}$.

$$(\lambda \mathbf{x}) = \lambda x_j = \sum_{i=1}^{n} A_{ij} x_i$$

If we pick the index j such that $|x_j| \leq |x_i|$ for all i, then we have that

$$|\lambda x_j| = |\lambda| \, |x_j| = \left| \sum_{i=1}^{n} A_{ij} x_i \right| \leq \sum_{i=1}^{n} |A_{ij}| \, |x_i| = \sum_{i=1}^{n} A_{ij} \, |x_i|$$

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

Now, we will show that if λ is an eigenvalue of A, then $|\lambda| \leq 1$.
Recall as before that A and $A^\top$ share eigenvalues, so let $(\lambda, \mathbf{x})$ be an eigenpair of $A^\top$. This means $A^\top \mathbf{x} = \lambda \mathbf{x}$.
Consider the j^{th} entry of $\lambda \mathbf{x}$.

$$(\lambda \mathbf{x}) = \lambda x_j = \sum_{i=1}^{n} A_{ij} x_i$$

If we pick the index j such that $|x_j| \leq |x_i|$ for all i, then we have that

$$|\lambda x_j| = |\lambda|\,|x_j| = \left|\sum_{i=1}^{n} A_{ij} x_i\right| \leq \sum_{i=1}^{n} |A_{ij}|\,|x_i| = \sum_{i=1}^{n} A_{ij}\,|x_i| \leq \sum_{i=1}^{n} A_{ij}\,|x_j|$$

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

Now, we will show that if λ is an eigenvalue of A, then $|\lambda| \leq 1$.
Recall as before that A and $A^\top$ share eigenvalues, so let $(\lambda, \mathbf{x})$ be an eigenpair of $A^\top$. This means $A^\top \mathbf{x} = \lambda \mathbf{x}$.
Consider the j^{th} entry of $\lambda \mathbf{x}$.

$$(\lambda \mathbf{x}) = \lambda x_j = \sum_{i=1}^{n} A_{ij} x_i$$

If we pick the index j such that $|x_j| \leq |x_i|$ for all i, then we have that

$$|\lambda x_j| = |\lambda| \, |x_j| \; = \left| \sum_{i=1}^{n} A_{ij} x_i \right| \leq \sum_{i=1}^{n} |A_{ij}| \, |x_i| \; = \sum_{i=1}^{n} A_{ij} \, |x_i| \; \leq \sum_{i=1}^{n} A_{ij} \, |x_j| \; = |x_j| \sum_{i=1}^{n} A_{ij}$$

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n\times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

Now, we will show that if λ is an eigenvalue of A, then $|\lambda| \leq 1$.
Recall as before that A and $A^\top$ share eigenvalues, so let $(\lambda, \mathbf{x})$ be an eigenpair of $A^\top$. This means $A^\top \mathbf{x} = \lambda \mathbf{x}$.
Consider the j^{th} entry of $\lambda \mathbf{x}$.

$$(\lambda \mathbf{x}) = \lambda x_j = \sum_{i=1}^{n} A_{ij} x_i$$

If we pick the index j such that $|x_j| \leq |x_i|$ for all i, then we have that

$$|\lambda x_j| = |\lambda|\,|x_j| = \left|\sum_{i=1}^{n} A_{ij} x_i\right| \leq \sum_{i=1}^{n} |A_{ij}|\,|x_i| = \sum_{i=1}^{n} A_{ij}\,|x_i| \leq \sum_{i=1}^{n} A_{ij}\,|x_j| = |x_j| \sum_{i=1}^{n} A_{ij} = |x_j|$$

Eigenvalues of Stochastic Matrices Part 2

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then we know that $\lambda = 1$ is an eigenvalue of A and if λ is an eigenvalue of A then $|\lambda| \leq 1$

Proof.

Now, we will show that if λ is an eigenvalue of A, then $|\lambda| \leq 1$.
Recall as before that A and $A^\top$ share eigenvalues, so let $(\lambda, \mathbf{x})$ be an eigenpair of $A^\top$. This means $A^\top \mathbf{x} = \lambda \mathbf{x}$.
Consider the j^{th} entry of $\lambda \mathbf{x}$.

$$(\lambda \mathbf{x}) = \lambda x_j = \sum_{i=1}^{n} A_{ij} x_i$$

If we pick the index j such that $|x_j| \leq |x_i|$ for all i, then we have that

$$|\lambda x_j| = |\lambda| \, |x_j| = \left| \sum_{i=1}^{n} A_{ij} x_i \right| \leq \sum_{i=1}^{n} |A_{ij}| \, |x_i| = \sum_{i=1}^{n} A_{ij} \, |x_i| \leq \sum_{i=1}^{n} A_{ij} \, |x_j| = |x_j| \sum_{i=1}^{n} A_{ij} = |x_j|$$

Dividing by $|x_j|$ gives us that $|\lambda| \leq 1$ $\square$

Eigenvalues of Positive Stochastic Matrices

Theorem

If A is a positive stochastic matrix, then $\dim\left(E(A,1)\right) = 1$.

A (Very) Simple Weather Model

We will model weather as follows

A (Very) Simple Weather Model

We will model weather as follows

a) There are two states: (1) Sunny and (2) Rainy

A (Very) Simple Weather Model

We will model weather as follows

a) There are two states: (1) Sunny and (2) Rainy

b) The current weather can always be observed

A (Very) Simple Weather Model

We will model weather as follows

a) There are two states: (1) Sunny and (2) Rainy
b) The current weather can always be observed
c) The current weather informs the weather for tomorrow, and we know the probability of tomorrow's weather given today's weather

A (Very) Simple Weather Model

We will model weather as follows

a) There are two states: (1) Sunny and (2) Rainy
b) The current weather can always be observed
c) The current weather informs the weather for tomorrow, and we know the probability of tomorrow's weather given today's weather

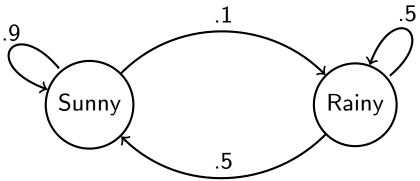
We can build a model to predict the weather. An example is given below

A (Very) Simple Weather Model

We will model weather as follows

a) There are two states: (1) Sunny and (2) Rainy
b) The current weather can always be observed
c) The current weather informs the weather for tomorrow, and we know the probability of tomorrow's weather given today's weather

We can build a model to predict the weather. An example is given below

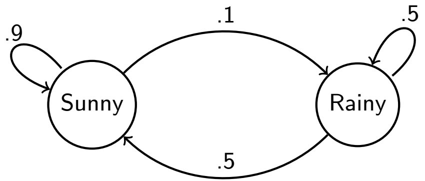


A (Very) Simple Weather Model

We will model weather as follows

a) There are two states: (1) Sunny and (2) Rainy
b) The current weather can always be observed
c) The current weather informs the weather for tomorrow, and we know the probability of tomorrow's weather given today's weather

We can build a model to predict the weather. An example is given below



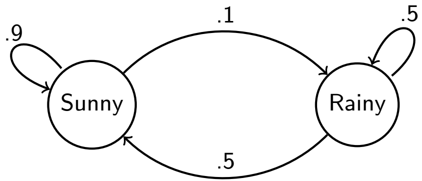
1. If we assign each node a number from 1 to n where n is the number of states we have we can collect these in a transition matrix P.

A (Very) Simple Weather Model

We will model weather as follows

a) There are two states: (1) Sunny and (2) Rainy
b) The current weather can always be observed
c) The current weather informs the weather for tomorrow, and we know the probability of tomorrow's weather given today's weather

We can build a model to predict the weather. An example is given below



1. If we assign each node a number from 1 to n where n is the number of states we have we can collect these in a transition matrix P.

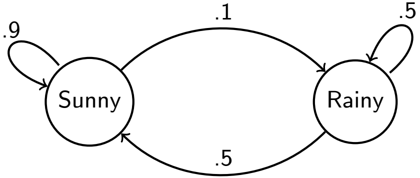
$$P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$$

A (Very) Simple Weather Model

We will model weather as follows

a) There are two states: (1) Sunny and (2) Rainy
b) The current weather can always be observed
c) The current weather informs the weather for tomorrow, and we know the probability of tomorrow's weather given today's weather

We can build a model to predict the weather. An example is given below



1. If we assign each node a number from 1 to n where n is the number of states we have we can collect these in a transition matrix P.

$$P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$$

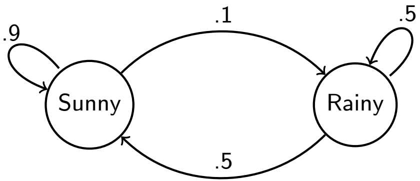
2. P_{ij} is the probability of going from state i to state j.

A (Very) Simple Weather Model

We will model weather as follows

a) There are two states: (1) Sunny and (2) Rainy
b) The current weather can always be observed
c) The current weather informs the weather for tomorrow, and we know the probability of tomorrow's weather given today's weather

We can build a model to predict the weather. An example is given below



1. If we assign each node a number from 1 to n where n is the number of states we have we can collect these in a transition matrix P.

$$P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$$

2. P_{ij} is the probability of going from state i to state j.
3. P is also a stochastic matrix!

Markov Chains in a Nutshell

An intuitive view of a Markov Chain is a math model describing some kind of “experiment” a large number of times under the same condition

Markov Chains in a Nutshell

An intuitive view of a Markov Chain is a math model describing some kind of "experiment" a large number of times under the same condition

- Each experiment has the same n possible outcomes

Markov Chains in a Nutshell

An intuitive view of a Markov Chain is a math model describing some kind of "experiment" a large number of times under the same condition

- Each experiment has the same n possible outcomes
- There is some set of probabilistic rules that determine how the system moves from one state to another

Markov Chains in a Nutshell

An intuitive view of a Markov Chain is a math model describing some kind of "experiment" a large number of times under the same condition

- Each experiment has the same n possible outcomes
- There is some set of probabilistic rules that determine how the system moves from one state to another
- These probabilities only depend on the current state

Markov Chains in a Nutshell

An intuitive view of a Markov Chain is a math model describing some kind of “experiment” a large number of times under the same condition

- Each experiment has the same n possible outcomes
- There is some set of probabilistic rules that determine how the system moves from one state to another
- These probabilities only depend on the current state

This can be applied to many fields including: Epidemiology, Finance, Statistics, etc.

Markov Chains in Terms of Linear Algebra

A more formal definition with the tools that we have is as follows:
A Markov chain is a sequence of state vectors $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$ in $\mathbb{R}^n$ and a stochastic (transition) matrix P such that for $k = 0, 1, 2, \ldots$

$$\mathbf{x}_{k+1} = P\mathbf{x}_k$$

Markov Chains in Terms of Linear Algebra

A more formal definition with the tools that we have is as follows:
A Markov chain is a sequence of state vectors $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$ in $\mathbb{R}^n$ and a stochastic (transition) matrix P such that for $k = 0, 1, 2, \ldots$

$$\mathbf{x}_{k+1} = P\mathbf{x}_k$$

Where $\mathbf{x}_k = \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix}$ is a vector of probabilities we are in each state $1, \ldots n$.

Revisiting our (Very) Simple Weather Model

Recall that our simple weather model has transition matrix: $P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$.

Revisiting our (Very) Simple Weather Model

Recall that our simple weather model has transition matrix: $P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$. Lets say that it is sunny today, so our initial state vector is $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and we want to figure out how likely it is to be each kind of weather 3 days from today.

Revisiting our (Very) Simple Weather Model

Recall that our simple weather model has transition matrix: $P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$. Lets say that it is sunny today, so our initial state vector is $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and we want to figure out how likely it is to be each kind of weather 3 days from today. We want to find $\mathbf{x}_3$.

Revisiting our (Very) Simple Weather Model

Recall that our simple weather model has transition matrix: $P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$. Lets say that it is sunny today, so our initial state vector is $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and we want to figure out how likely it is to be each kind of weather 3 days from today. We want to find $\mathbf{x}_3$.

$$\mathbf{x}_3 = P\mathbf{x}_2$$

Revisiting our (Very) Simple Weather Model

Recall that our simple weather model has transition matrix: $P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$. Lets say that it is sunny today, so our initial state vector is $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and we want to figure out how likely it is to be each kind of weather 3 days from today. We want to find $\mathbf{x}_3$.

$$\mathbf{x}_3 = P\mathbf{x}_2 = P(P\mathbf{x}_1)$$

Revisiting our (Very) Simple Weather Model

Recall that our simple weather model has transition matrix: $P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$. Lets say that it is sunny today, so our initial state vector is $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and we want to figure out how likely it is to be each kind of weather 3 days from today. We want to find $\mathbf{x}_3$.

$$\mathbf{x}_3 = P\mathbf{x}_2 = P(P\mathbf{x}_1) = P(P(P\mathbf{x}_0))$$

Revisiting our (Very) Simple Weather Model

Recall that our simple weather model has transition matrix: $P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$. Lets say that it is sunny today, so our initial state vector is $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and we want to figure out how likely it is to be each kind of weather 3 days from today. We want to find $\mathbf{x}_3$.

$$\mathbf{x}_3 = P\mathbf{x}_2 = P(P\mathbf{x}_1) = P(P(P\mathbf{x}_0)) = P^3\mathbf{x}_0$$

Revisiting our (Very) Simple Weather Model

Recall that our simple weather model has transition matrix: $P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$. Lets say that it is sunny today, so our initial state vector is $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and we want to figure out how likely it is to be each kind of weather 3 days from today. We want to find $\mathbf{x}_3$.

$$\mathbf{x}_3 = P\mathbf{x}_2 = P(P\mathbf{x}_1) = P(P(P\mathbf{x}_0)) = P^3\mathbf{x}_0$$

We just compute

$$P^3 = \begin{bmatrix} .844 & .78 \\ .156 & .22 \end{bmatrix}$$

Revisiting our (Very) Simple Weather Model

Recall that our simple weather model has transition matrix: $P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$. Lets say that it is sunny today, so our initial state vector is $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and we want to figure out how likely it is to be each kind of weather 3 days from today. We want to find $\mathbf{x}_3$.

$$\mathbf{x}_3 = P\mathbf{x}_2 = P(P\mathbf{x}_1) = P(P(P\mathbf{x}_0)) = P^3\mathbf{x}_0$$

We just compute

$$P^3 = \begin{bmatrix} .844 & .78 \\ .156 & .22 \end{bmatrix}$$

So,

$$\mathbf{x}_3 = P^3\mathbf{x}_0$$

Revisiting our (Very) Simple Weather Model

Recall that our simple weather model has transition matrix: $P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$. Lets say that it is sunny today, so our initial state vector is $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and we want to figure out how likely it is to be each kind of weather 3 days from today. We want to find $\mathbf{x}_3$.

$$\mathbf{x}_3 = P\mathbf{x}_2 = P(P\mathbf{x}_1) = P(P(P\mathbf{x}_0)) = P^3\mathbf{x}_0$$

We just compute

$$P^3 = \begin{bmatrix} .844 & .78 \\ .156 & .22 \end{bmatrix}$$

So,

$$\mathbf{x}_3 = P^3\mathbf{x}_0 = \begin{bmatrix} .844 & .78 \\ .156 & .22 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Revisiting our (Very) Simple Weather Model

Recall that our simple weather model has transition matrix: $P = \begin{bmatrix} .9 & .5 \\ .1 & .5 \end{bmatrix}$. Lets say that it is sunny today, so our initial state vector is $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and we want to figure out how likely it is to be each kind of weather 3 days from today. We want to find $\mathbf{x}_3$.

$$\mathbf{x}_3 = P\mathbf{x}_2 = P(P\mathbf{x}_1) = P(P(P\mathbf{x}_0)) = P^3\mathbf{x}_0$$

We just compute

$$P^3 = \begin{bmatrix} .844 & .78 \\ .156 & .22 \end{bmatrix}$$

So,

$$\mathbf{x}_3 = P^3\mathbf{x}_0 = \begin{bmatrix} .844 & .78 \\ .156 & .22 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} .844 \\ .156 \end{bmatrix}$$

Property of Powers of Positive Stochastic Matrices

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a diagonalizable positive stochastic matrix. Then, for some large n, we have that

Property of Powers of Positive Stochastic Matrices

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a diagonalizable positive stochastic matrix. Then, for some large n, we have that

$$A^n = \left(VDV^{-1}\right)^n$$

Property of Powers of Positive Stochastic Matrices

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a diagonalizable positive stochastic matrix. Then, for some large n, we have that

$$A^n = \left(VDV^{-1}\right)^n = VD^nV^{-1}$$

Property of Powers of Positive Stochastic Matrices

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a diagonalizable positive stochastic matrix. Then, for some large n, we have that

$$A^n = \left(VDV^{-1}\right)^n = VD^nV^{-1} \approx V \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} V^{-1}$$

Property of Powers of Positive Stochastic Matrices

Theorem

Let $A \in \mathbb{R}^{n\times n}$ be a diagonalizable positive stochastic matrix. Then, for some large n, we have that

$$A^n = \left(VDV^{-1}\right)^n = VD^nV^{-1} \approx V \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} V^{-1}$$

Note: We can say something very similar for non-diagonalizable matrices, but it's harder to phrase

Steady State (Equilibrium)

Definition

Steady State: The steady state of a stochastic matrix A is an eigenvector $\mathbf{v}$ associated with $\lambda = 1$ such that all the entries are *positive* and sum to 1.

Steady State (Equilibrium)

Definition

Steady State: The steady state of a stochastic matrix A is an eigenvector $\mathbf{v}$ associated with $\lambda = 1$ such that all the entries are *positive* and sum to 1.

This solution is a representation of what happens in the long run of a Markov chain!

Steady State (Equilibrium)

Definition

Steady State: The steady state of a stochastic matrix A is an eigenvector $\mathbf{v}$ associated with $\lambda = 1$ such that all the entries are *positive* and sum to 1.

This solution is a representation of what happens in the long run of a Markov chain! Computing this can be challenging in general if our Eigenspace associated with $\lambda = 1$ has more than one basis vector.

Steady State of a Positive Stochastic Matrix

However, if $A \in \mathbb{R}^{n \times n}$ is a *positive* stochastic matrix, then we know $\dim\left(E(A,1)\right) = 1$.

Steady State of a Positive Stochastic Matrix

However, if $A \in \mathbb{R}^{n \times n}$ is a *positive* stochastic matrix, then we know $\dim\left(E(A, 1)\right) = 1$. Meaning, it's a lot easier! We can just:

Steady State of a Positive Stochastic Matrix

However, if $A \in \mathbb{R}^{n\times n}$ is a *positive* stochastic matrix, then we know $\dim\left(E(A,1)\right) = 1$. Meaning, it's a lot easier! We can just:

1. Solve $(A - I_n)\mathbf{v} = \mathbf{0}$

Steady State of a Positive Stochastic Matrix

However, if $A \in \mathbb{R}^{n \times n}$ is a *positive* stochastic matrix, then we know $\dim(E(A, 1)) = 1$. Meaning, it's a lot easier! We can just:

1. Solve $(A - I_n)\mathbf{v} = \mathbf{0}$
2. Divide $\mathbf{v}$ by the sum of its elements.

Steady State of a Positive Stochastic Matrix

However, if $A \in \mathbb{R}^{n \times n}$ is a *positive* stochastic matrix, then we know $\dim(E(A,1)) = 1$. Meaning, it's a lot easier! We can just:

1. Solve $(A - I_n)\mathbf{v} = \mathbf{0}$
2. Divide $\mathbf{v}$ by the sum of its elements.
3. This new vector is the steady state vector!

Steady State Example

Let's find the steady state of a different stochastic matrix

$$P = \begin{bmatrix} .3 & .4 & .5 \\ .3 & .4 & .3 \\ .4 & .2 & .2 \end{bmatrix}$$

See Jupyter Notebook

A Practical Application: PageRank

The original algorithm that Google used in it's search engine was based on PageRank, which ranks websites based on the number of links to and from it.

The textbook discusses it in the end of chapter 5.6:
`https://textbooks.math.gatech.edu/ila/stochastic-matrices.html`.

A more comprehensive history can be found at the wikipedia page:
`https://en.wikipedia.org/wiki/PageRank#History`