Discrete Mathematical Models

Lecture 20

Kane Townsend

Semester 2, 2024

Applications of Bayes' Theorem

Theorem (Bayes' Theorem)
For any probability experiment with sample space S, for any $n \in \mathbb{N}$, for any partition $\{B_1, B_2, \ldots, B_n\}$ of S and for any event $A \subseteq S$, if $\mathbb{P}(A) \neq 0$ and for all $i \in \{1, 2, \ldots, n\}$ we have $\mathbb{P}(B_i) \neq 0$, then for all $k \in \{1, 2, \ldots, n\}$ we have

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)}$$

Applications of Bayes' Theorem

Theorem (Bayes' Theorem)
For any probability experiment with sample space S, for any $n \in \mathbb{N}$, for any partition $\{B_1, B_2, \ldots, B_n\}$ of S and for any event $A \subseteq S$, if $\mathbb{P}(A) \neq 0$ and for all $i \in \{1, 2, \ldots, n\}$ we have $\mathbb{P}(B_i) \neq 0$, then for all $k \in \{1, 2, \ldots, n\}$ we have

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)}$$

- Solving Monty Hall problem
- Drug testing
- Disease testing
- Defective item rates

Applications of Bayes' Theorem

Theorem (Bayes' Theorem)
For any probability experiment with sample space S, for any $n \in \mathbb{N}$, for any partition $\{B_1, B_2, \ldots, B_n\}$ of S and for any event $A \subseteq S$, if $\mathbb{P}(A) \neq 0$ and for all $i \in \{1, 2, \ldots, n\}$ we have $\mathbb{P}(B_i) \neq 0$, then for all $k \in \{1, 2, \ldots, n\}$ we have

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)}$$

- Solving Monty Hall problem
- Drug testing
- Disease testing
- Defective item rates

A **false positive** means that a patient gets a positive test of having the disease when they do not have the disease.

A **false negative** means that a patient gets a negative test of having the disease when they do have the disease.

Example 9.9.3 from Epp.

Consider a medical test that screens for a disease found in 5 people in 1,000. Suppose that the false positive rate is 3% and the false negative rate is 1%. Then 99% of the time a person who has the condition tests positive for it, and 97% of the time a person who does not have the condition tests negative for it.

Example 9.9.3 from Epp.

Consider a medical test that screens for a disease found in 5 people in 1,000. Suppose that the false positive rate is 3% and the false negative rate is 1%. Then 99% of the time a person who has the condition tests positive for it, and 97% of the time a person who does not have the condition tests negative for it.

a. What is the probability that a randomly chosen person who tests positive for the disease actually has the disease?
b. What is the probability that a randomly chosen person who tests negative for the disease does not in fact have the disease?

Solution: Consider a random person from those screened. Let A be the event they test positive, B_1 the event they have the disease, B_2 the event they do not have the disease. Then:

Example 9.9.3 from Epp.

Consider a medical test that screens for a disease found in 5 people in 1,000. Suppose that the false positive rate is 3% and the false negative rate is 1%. Then 99% of the time a person who has the condition tests positive for it, and 97% of the time a person who does not have the condition tests negative for it.

a. What is the probability that a randomly chosen person who tests positive for the disease actually has the disease?
b. What is the probability that a randomly chosen person who tests negative for the disease does not in fact have the disease?

Solution: Consider a random person from those screened. Let A be the event they test positive, B_1 the event they have the disease, B_2 the event they do not have the disease. Then:

$$\mathbb{P}(A|B_1) = 0.99, \quad \mathbb{P}(A^c|B_1) = 0.01, \quad \mathbb{P}(A^c|B_2) = 0.97, \quad \mathbb{P}(A|B_2) = 0.03.$$

Also because 5 people in 1,000 have the disease,

$$\mathbb{P}(B_1) = 0.005 \text{ and } \mathbb{P}(B_2) = 0.995.$$

Example 9.9.3 from Epp. (Cont.)

A person tests positive, B_1 person has disease, B_2 the event they do not have the disease. Then: $\mathbb{P}(A|B_1) = 0.99, \mathbb{P}(A^c|B_1) = 0.01, \mathbb{P}(A^c|B_2) = 0.97, \mathbb{P}(A|B_2) = 0.03, \mathbb{P}(B_1) = 0.005$ and $\mathbb{P}(B_2) = 0.995$.

a. By Bayes' Theorem

$$\begin{aligned}
\mathbb{P}(B_1|A) &= \frac{\mathbb{P}(A|B_1)\mathbb{P}(B_1)}{\mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2)} \\
&= \frac{(0.99)(0.005)}{(0.99)(0.005) + (0.03)(0.995)} \\
&\approx 0.1422 \approx 14.2\%.
\end{aligned}$$

Example 9.9.3 from Epp. (Cont.)

A person tests positive, B_1 person has disease, B_2 the event they do not have the disease. Then: $\mathbb{P}(A|B_1) = 0.99, \mathbb{P}(A^c|B_1) = 0.01, \mathbb{P}(A^c|B_2) = 0.97, \mathbb{P}(A|B_2) = 0.03, \mathbb{P}(B_1) = 0.005$ and $\mathbb{P}(B_2) = 0.995$.

a. By Bayes' Theorem

$$\begin{aligned}
\mathbb{P}(B_1|A) &= \frac{\mathbb{P}(A|B_1)\mathbb{P}(B_1)}{\mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2)} \\
&= \frac{(0.99)(0.005)}{(0.99)(0.005) + (0.03)(0.995)} \\
&\approx 0.1422 \approx 14.2\%.
\end{aligned}$$

b. By Bayes' Theorem

$$\begin{aligned}
\mathbb{P}(B_2|A^c) &= \frac{\mathbb{P}(A^c|B_2)\mathbb{P}(B_2)}{\mathbb{P}(A^c|B_1)\mathbb{P}(B_1) + \mathbb{P}(A^c|B_2)\mathbb{P}(B_2)} \\
&= \frac{(0.97)(0.995)}{(0.01)(0.005) + (0.97)(0.995)} \\
&\approx 0.999948 \approx 99.995\%.
\end{aligned}$$

C3: Markov Processes (not covered in textbook)

Introduction

Markov processes

Markov processes are about probabilities. We consider

- the **state** of a system, amongst a finite number of possibilities
- the **probability** of moving between states in one time-step,
- and the probable state after **many** time-steps.
- We often don't make a sharp distinction between **proportions** and **probabilities** as you will see in the examples.

Markov processes

Markov processes are about probabilities. We consider

- the **state** of a system, amongst a finite number of possibilities
- the **probability** of moving between states in one time-step,
- and the probable state after **many** time-steps.
- We often don't make a sharp distinction between **proportions** and **probabilities** as you will see in the examples.

  This works well for large samples but you may need to be careful with small samples.

Example 1

Introductory Example

Adapted from ‘Finite Mathematics’, Maki & Thompson:

A freelance computer network consultant, let’s call her Cathy, is employed in weekly contracts. Each week she is either:

- employed (E)

Introductory Example

Adapted from 'Finite Mathematics', Maki & Thompson:

A freelance computer network consultant, let's call her Cathy, is employed in weekly contracts. Each week she is either:

- employed (E) or
- unemployed (U).

Her records support the following assumptions:

Introductory Example

Adapted from ‘Finite Mathematics’, Maki & Thompson:

A freelance computer network consultant, let’s call her Cathy, is employed in weekly contracts. Each week she is either:

- employed (E) or
- unemployed (U).

Her records support the following assumptions:

(a) If she’s employed this week, then next week she’ll be employed with probability 0.8 and unemployed with probability 0.2.

Introductory Example

Adapted from 'Finite Mathematics', Maki & Thompson:

A freelance computer network consultant, let's call her Cathy, is employed in weekly contracts. Each week she is either:

- employed (E) or
- unemployed (U).

Her records support the following assumptions:

(a) If she's employed this week, then next week she'll be employed with probability 0.8 and unemployed with probability 0.2.

(b) If she's unemployed this week, then next week she'll be employed with probability 0.6 and unemployed with probability 0.4.

System, States and Transitions

We can model Cathy's situation by a Markov process:

System, States and Transitions

We can model Cathy's situation by a Markov process:

- The **system** is Cathy herself.

System, States and Transitions

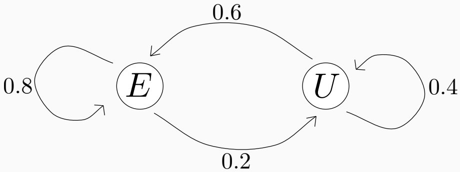
We can model Cathy's situation by a Markov process:

- The **system** is Cathy herself.
- The system can be in one of two **states**:
  - E: Employed
  - U: Unemployed

System, States and Transitions

We can model Cathy's situation by a Markov process:

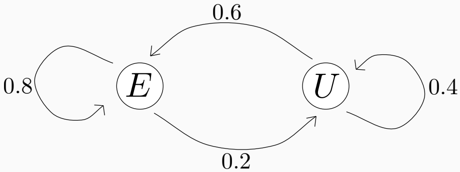
- The **system** is Cathy herself.
- The system can be in one of two **states**:
  - E: Employed
  - U: Unemployed
- A **Transition Diagram** encodes the transition probabilities:



System, States and Transitions

We can model Cathy's situation by a Markov process:

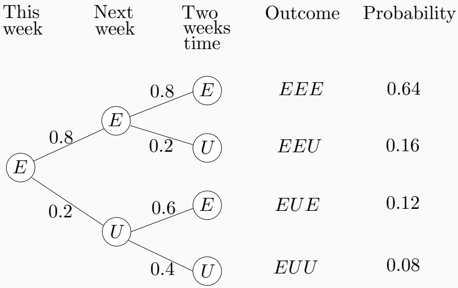
- The **system** is Cathy herself.
- The system can be in one of two **states**:
  - E: Employed
  - U: Unemployed
- A **Transition Diagram** encodes the transition probabilities:



It is a property of a Markov Process that the probability of stepping from one state to another *only depends on the current state*.

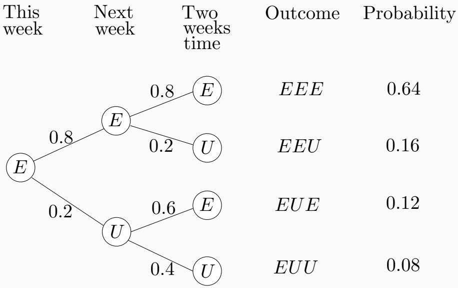
Two time-steps

If Cathy is employed this week, what is the probability that she will be employed two weeks from now? We can use a tree:



Two time-steps

If Cathy is employed this week, what is the probability that she will be employed two weeks from now? We can use a tree:



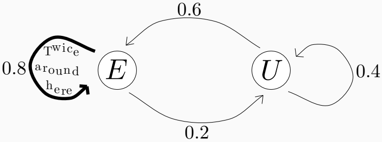
From the tree diagram, the probability that Cathy will be employed two weeks from now is

$\Pr(\text{EEE or EUE}) = \Pr(\text{EEE}) + \Pr(\text{EUE}) = 0.64 + 0.12 = 0.76.$

Two time-steps on the transition diagram

Starting employed, then employment after two weeks can be shown on the transition diagram as

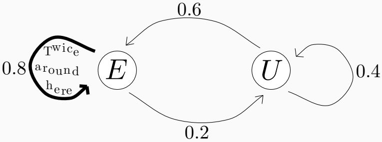
either



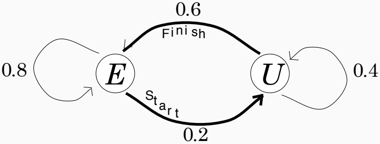
Two time-steps on the transition diagram

Starting employed, then employment after two weeks can be shown on the transition diagram as

either

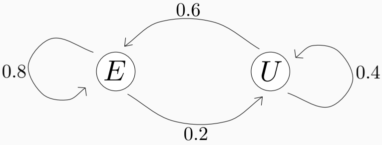


or



Transition Matrix

The information in Cathy's transition diagram



can be encoded in the transitition matrix

$$\begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix}$$

State Vectors

The **state vector** x_n shows probabilities of being in each state after n time-steps.

State Vectors

The **state vector** x_n shows probabilities of being in each state after n time-steps.

In Cathy's case, the state vector has two entries.

State Vectors

The **state vector** x_n shows probabilities of being in each state after n time-steps.

In Cathy's case, the state vector has two entries.

- The first entry records probability of employment.
- The second entry records probability of unemployment.

State Vectors

The **state vector** x_n shows probabilities of being in each state after n time-steps.

In Cathy's case, the state vector has two entries.

- The first entry records probability of employment.
- The second entry records probability of unemployment.

Initially ($n = 0$) the only possible values for the state vector are

- $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ if Cathy is employed,

State Vectors

The **state vector** x_n shows probabilities of being in each state after n time-steps.

In Cathy's case, the state vector has two entries.

- The first entry records probability of employment.
- The second entry records probability of unemployment.

Initially ($n = 0$) the only possible values for the state vector are

- $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ if Cathy is employed, and
- $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ if Cathy is unemployed,

since we assumed that, in any week, Cathy was either:
100% employed and 0% unemployed,

State Vectors

The **state vector** x_n shows probabilities of being in each state after n time-steps.

In Cathy's case, the state vector has two entries.

- The first entry records probability of employment.
- The second entry records probability of unemployment.

Initially ($n = 0$) the only possible values for the state vector are

- $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ if Cathy is employed, and
- $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ if Cathy is unemployed,

since we assumed that, in any week, Cathy was either:
100% employed and 0% unemployed, or
0% employed and 100% unemployed.

Using Matrices and State Vectors

Suppose Cathy is employed in Week 0.

Using Matrices and State Vectors

Suppose Cathy is employed in Week 0.

The initial state vector is $\mathsf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Using Matrices and State Vectors

Suppose Cathy is employed in Week 0.

The initial state vector is $x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

After one time-step, *i.e.* next week (Week 1), her probabilities of being employed or not are given by the state vector x_1.

Using Matrices and State Vectors

Suppose Cathy is employed in Week 0.

The initial state vector is $\mathrm{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

After one time-step, *i.e.* next week (Week 1), her probabilities of being employed or not are given by the state vector x_1.

This can be expressed as:

$$\begin{aligned} x_1 &= T\mathrm{x}_0 \\ &= \begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 0.8 \\ 0.2 \end{bmatrix} \end{aligned}$$

Two time-steps

In Week 2, *i.e.* after two time-steps, Cathy's chances of work are given by the state vector x_2.

Two time-steps

In Week 2, *i.e.* after two time-steps, Cathy's chances of work are given by the state vector x_2.

This can be calculated by: This can be expressed as:

$$\begin{aligned} x_2 &= T\mathrm{x}_1 \\ &= \begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix} \begin{bmatrix} 0.8 \\ 0.2 \end{bmatrix} \\ &= \begin{bmatrix} 0.64 + 0.12 \\ 0.16 + 0.08 \end{bmatrix} \\ &= \begin{bmatrix} 0.76 \\ 0.24 \end{bmatrix} \end{aligned}$$

n time-steps

$$\text{Continuing:}\quad \begin{aligned} \mathrm{x}_3 &= T\mathrm{x}_2 = \begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix} \begin{bmatrix} 0.76 \\ 0.24 \end{bmatrix} = \begin{bmatrix} 0.752 \\ 0.248 \end{bmatrix} \\ \mathrm{x}_4 &= T\mathrm{x}_3 = \begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix} \begin{bmatrix} 0.752 \\ 0.248 \end{bmatrix} = \begin{bmatrix} 0.7504 \\ 0.2496 \end{bmatrix} \\ \mathrm{x}_5 &= T\mathrm{x}_4 = \begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix} \begin{bmatrix} 0.7504 \\ 0.2496 \end{bmatrix} = \begin{bmatrix} 0.75008 \\ 0.24992 \end{bmatrix} \\ \vdots \quad & \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \end{aligned}$$

n time-steps

$$\text{Continuing:} \quad \begin{aligned} \mathrm{x}_3 &= T\mathrm{x}_2 = \begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix} \begin{bmatrix} 0.76 \\ 0.24 \end{bmatrix} = \begin{bmatrix} 0.752 \\ 0.248 \end{bmatrix} \\ \mathrm{x}_4 &= T\mathrm{x}_3 = \begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix} \begin{bmatrix} 0.752 \\ 0.248 \end{bmatrix} = \begin{bmatrix} 0.7504 \\ 0.2496 \end{bmatrix} \\ \mathrm{x}_5 &= T\mathrm{x}_4 = \begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix} \begin{bmatrix} 0.7504 \\ 0.2496 \end{bmatrix} = \begin{bmatrix} 0.75008 \\ 0.24992 \end{bmatrix} \\ \vdots \quad & \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \end{aligned}$$

$$\text{Thus:} \quad \begin{aligned} \mathrm{x}_1 &= T\mathrm{x}_0 \\ \mathrm{x}_2 &= T\mathrm{x}_1 = TT\mathrm{x}_0 = T^2\mathrm{x}_0 \\ \mathrm{x}_3 &= T\mathrm{x}_2 = TTT\mathrm{x}_0 = T^3\mathrm{x}_0 \\ \vdots \quad & \quad \vdots \qquad\qquad \vdots \\ \mathrm{x}_n &= T^n\mathrm{x}_0 \end{aligned}$$

The n^{th} power of T

Successive powers of $T = \begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix}$ are:

The n^{th} power of T

Successive powers of $T = \begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix}$ are:

$$
\begin{aligned}
T^2 = TT &= \begin{bmatrix} 0.76 & 0.72 \\ 0.24 & 0.28 \end{bmatrix} \\
T^3 = TT^2 &= \begin{bmatrix} 0.752 & 0.744 \\ 0.248 & 0.256 \end{bmatrix} \\
T4 = TT^3 &= \begin{bmatrix} 0.7504 & 0.7488 \\ 0.2496 & 0.2512 \end{bmatrix} \\
T^5 = TT^4 &= \begin{bmatrix} 0.75008 & 0.74976 \\ 0.24992 & 0.25024 \end{bmatrix} \\
\vdots \quad \vdots & \qquad\quad \vdots
\end{aligned}
$$

The n^{th} power of T

Successive powers of $T = \begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix}$ are:

$$T^2 = TT = \begin{bmatrix} 0.76 & 0.72 \\ 0.24 & 0.28 \end{bmatrix}$$

$$T^3 = TT^2 = \begin{bmatrix} 0.752 & 0.744 \\ 0.248 & 0.256 \end{bmatrix}$$

$$T4 = TT^3 = \begin{bmatrix} 0.7504 & 0.7488 \\ 0.2496 & 0.2512 \end{bmatrix}$$

$$T^5 = TT^4 = \begin{bmatrix} 0.75008 & 0.74976 \\ 0.24992 & 0.25024 \end{bmatrix}$$

$$\vdots \qquad \vdots \qquad \qquad \vdots$$

So we can guess that: $T^n \approx \begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix}$ for large values of n.

The significance of T^n for large n

We have seen that, for Cathy, $T^n \approx \begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix}$ for large values of n.

The significance of T^n for large n

We have seen that, for Cathy, $T^n \approx \begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix}$ for large values of n.

Notice that the columns of this matrix are equal, and that

$$\begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}.$$

The significance of T^n for large n

We have seen that, for Cathy, $T^n \approx \begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix}$ for large values of n.

Notice that the columns of this matrix are equal, and that

$$\begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}.$$

So, irrespective of the initial state, in the long term the state vector becomes approximately $\begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}$.

The significance of T^n for large n

We have seen that, for Cathy, $T^n \approx \begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix}$ for large values of n.

Notice that the columns of this matrix are equal, and that

$$\begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}.$$

So, irrespective of the initial state, in the long term the state vector becomes approximately $\begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}$. This means

$\boxed{\text{No matter what, eventually Cathy will be employed 75\% of the time.}}$

Definitions, Terminology, Results

Stochastic matrices and steady state vectors

A **probability vector** is a vector with non-negative entries that sum to 1.

Stochastic matrices and steady state vectors

A **probability vector** is a vector with non-negative entries that sum to 1.

A **stochastic matrix** T is a square matrix such that all its columns are probability vectors. A stochastic matrix is **positive** if all its entries are positive.

Stochastic matrices and steady state vectors

A **probability vector** is a vector with non-negative entries that sum to 1.

A **stochastic matrix** T is a square matrix such that all its columns are probability vectors. A stochastic matrix is **positive** if all its entries are positive.

A **steady state vector** v, with respect to a square matrix T has the property that $Tv = v$. A steady state vector is **normalised** if its entries sum to 1.

Stochastic matrices and steady state vectors

A **probability vector** is a vector with non-negative entries that sum to 1.

A **stochastic matrix** T is a square matrix such that all its columns are probability vectors. A stochastic matrix is **positive** if all its entries are positive.

A **steady state vector** v, with respect to a square matrix T has the property that $Tv = v$. A steady state vector is **normalised** if its entries sum to 1.

Theorem (Perron-Frobenius)
Let T be a positive stochastic matrix. Then there is a unique normalised steady state vector v with respect to T. Moreover, T^n converges to $[v \; v]$ as $n \to \infty$ and so for any initial vector v_0 with entries summing to c, $v_n = T^n v_0$ converges to cv.

More definitions

A (discrete) **Markov process** is a system that has:

- a finite number k of states,
- a sequence of time steps $n \in \mathbb{N}^*$,
- probabilities of moving from a state to another state (including itself) that depends only on your current state.

Hence, probabilities of being in a particular state at time $n \geq 1$ depend on

(i) its state at the $(n-1)$-th time step, and

(ii) a fixed stochastic matrix $T \in M_k(\mathbb{Q}_+)$ called the **transition matrix** of the process.

More definitions

A (discrete) **Markov process** is a system that has:

- a finite number k of states,
- a sequence of time steps $n \in \mathbb{N}^*$,
- probabilities of moving from a state to another state (including itself) that depends only on your current state.

Hence, probabilities of being in a particular state at time $n \geq 1$ depend on

(i) its state at the $(n-1)$-th time step, and
(ii) a fixed stochastic matrix $T \in M_k(\mathbb{Q}_+)$ called the **transition matrix** of the process.

The (i, j)-entry T_{ij} of the transition matrix T specifies the probability that the system will be in the i-th state at any time step $n \geq 1$, given that it was in the j-th state at time step $n-1$.

More definitions

A (discrete) **Markov process** is a system that has:

- a finite number k of states,
- a sequence of time steps $n \in \mathbb{N}^*$,
- probabilities of moving from a state to another state (including itself) that depends only on your current state.

Hence, probabilities of being in a particular state at time $n \geq 1$ depend on

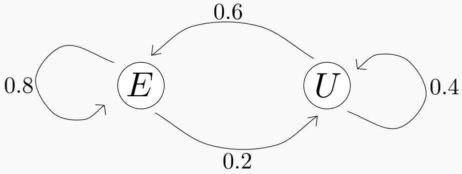
(i) its state at the $(n-1)$-th time step, and
(ii) a fixed stochastic matrix $T \in M_k(\mathbb{Q}_+)$ called the **transition matrix** of the process.

The (i, j)-entry T_{ij} of the transition matrix T specifies the probability that the system will be in the i-th state at any time step $n \geq 1$, given that it was in the j-th state at time step $n-1$.

A **transition diagram** is a complete weighted directed graph with k vertices representing the states of the system and the edge from the j-th vertex to the i-th vertex labelled with the probability T_{ij}.

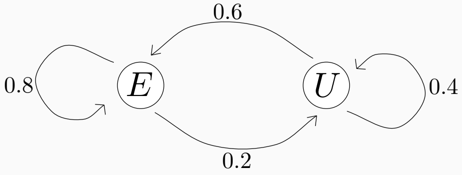
Cathy's Example is a Markov Process

Transition diagram:



Cathy's Example is a Markov Process

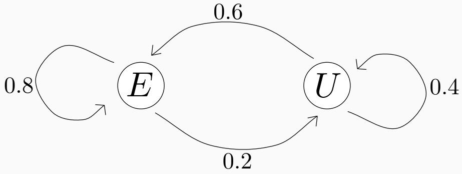
Transition diagram:



The transition matrix is a positive stochastic matrix given by $\begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix}$

Cathy's Example is a Markov Process

Transition diagram:

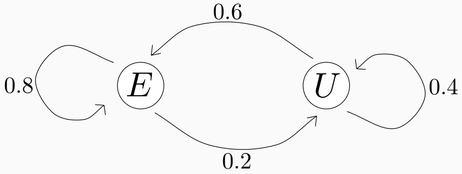


The transition matrix is a positive stochastic matrix given by $\begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix}$

Perron-Frobenius applies and T^n converges to $\begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix}$ as $n \to \infty$

Cathy's Example is a Markov Process

Transition diagram:



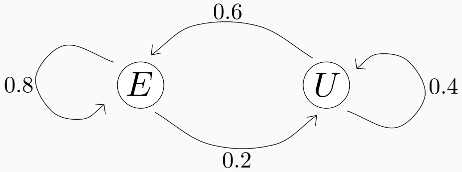
The transition matrix is a positive stochastic matrix given by $\begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix}$

Perron-Frobenius applies and T^n converges to $\begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix}$ as $n \to \infty$

Hence, $v_n = T^n v_0$ converges to $\begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}$ as $n \to \infty$

Cathy's Example is a Markov Process

Transition diagram:



The transition matrix is a positive stochastic matrix given by $\begin{bmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{bmatrix}$

Perron-Frobenius applies and T^n converges to $\begin{bmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{bmatrix}$ as $n \to \infty$

Hence, $v_n = T^n v_0$ converges to $\begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}$ as $n \to \infty$

Furthermore, the normalised steady state vector is $\begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}$.

Markov Processes Have No Memory

- The state of a Markov Process at time n only depends on fixed transition probabilities and its state at time $n-1$.

Markov Processes Have No Memory

- The state of a Markov Process at time n only depends on fixed transition probabilities and its state at time $n - 1$.
- It does not depend on the state at any earlier time. In other words, it is a *first-order* (matrix) recurrence.

Markov Processes Have No Memory

- The state of a Markov Process at time n only depends on fixed transition probabilities and its state at time $n-1$.
- It does not depend on the state at any earlier time. In other words, it is a *first-order* (matrix) recurrence.
- Because of this, Markov processes are said to “have no memory".

Finding steady state vectors

- One way to find a steady state vector of a Markov process is to do as we did in the example - namely multiply together enough copies of T to see the higher powers tending to a limit.

Finding steady state vectors

- One way to find a steady state vector of a Markov process is to do as we did in the example - namely multiply together enough copies of T to see the higher powers tending to a limit.
- There are more direct methods of finding steady state vectors that use linear algebra. We will cover this in the next lecture using a larger example.