DP IB Maths: AI HL

Your notes

4.13 Transition Matrices & Markov Chains

Contents

4.13.1 Markov Chains

Your notes

Markov Chains

What is meant by a "state"?

- States refer to **mutually exclusive events** with the current event **able to change over time**
- Examples of states include:
  - Daily weather conditions
    - The states could be: "sunny" and "not sunny"
  - Countries visited by an inspector each day
    - The states could be: "France", "Spain" and "Germany"
  - Store chosen for weekly grocery shop:
    - The states could be: "Foods-U-Like", "Smiley Shoppers" and "Better Buys"

What is a Markov chain?

- A **Markov chain** is a model that describes a **sequence of states** over a period of time
  - Time is measured in discrete steps
    - Such as days, months, years, etc
- The **conditions** for a Markov chain are:
  - The **probability** of a state being the **next state** in the sequence **only depends** on the **current state**
    - For example
      The 11^{th} state **only depends** on the 10^{th} state
      The first 9 states **do not affect** the 11^{th} state
    - This probability is called a **transition probability**
  - The **transition probabilities do not change** over time
    - For example
      The probability that the 11^{th} state is A given that the 10^{th} state is B is equal to the probability that the 12^{th} state is A given that the 11^{th} state is B
- A Markov chain is said to be **regular** if there is a value *k* such that in **exactly *k* steps** it is possible to reach any state regardless of the initial state
  - The chain where A can only go to B, B can only go to C and C can only go to A, is **not regular**
    - After any number of changes, A can only go to either B or C but not both
    - After 100 changes, A can end up at B but not C
    - After 500 changes, A can end up at C but not B

What is a transition state diagram?

- A **transition diagram** is a **directed graph**
  - The **vertices** are the **states**
  - The **edges** represent the **transition probabilities** between the states
- The graph can contain
  - **Loops**
    - These will be the transition probabilities of the next state being the same as the current state

- **Two edges between each pair** of vertices
  - The edges will be in opposite directions
  - Each edge will show the transition probability of the state changing in the given direction
- The **probabilities** on the **edges coming out** of a vertex **add up to 1**

Your notes

Examiner Tip

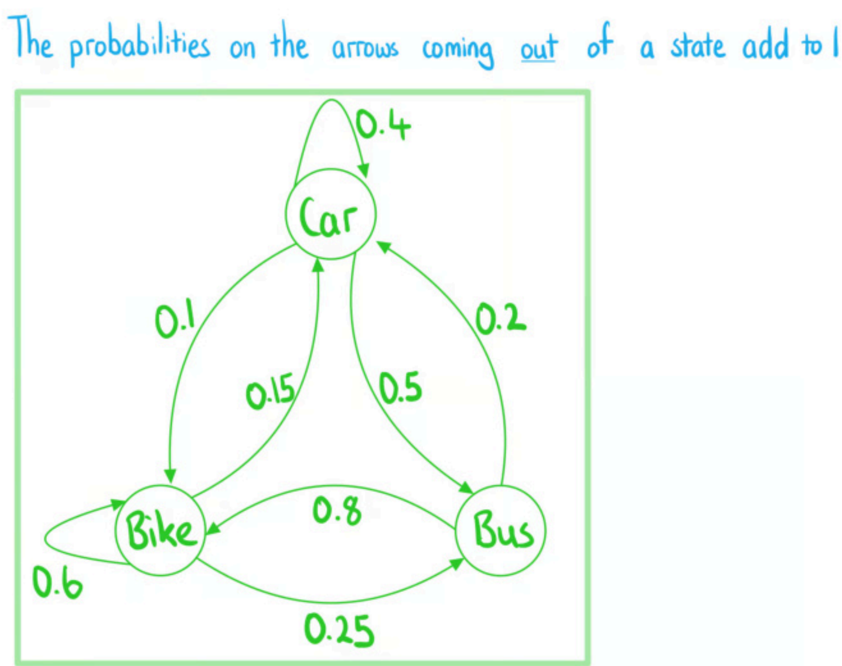
- Drawing a transition state diagram (even when the question does not ask for one) can help you visualise the problem

Worked example

Fleur travels to work by car, bike or bus. Each day she chooses her mode of transport based on the transport she chose the previous day.

- If Fleur travels by car then there is a 40% chance that she will travel by car the following day and a 10% chance that she will travel by bike.
- If Fleur travels by bike then there is a 60% chance that she will travel by bike the following day and a 25% chance that she will travel by bus.
- If Fleur travels by bus then there is an 80% chance that she will travel by bike the following day and a 20% chance that she will travel by car.

Represent this information as a transition state diagram.



4.13.2 Transition Matrices

Your notes

Transition Matrices

What is a transition matrix?

- A **transition matrix** ***T*** shows the **transition probabilities** between the current state and the next state
  - The **columns** represent the **current states**
  - The **rows** represent the **next states**
- The element of ***T*** in the i^{th} row and j^{th} column gives the transition probability t_{ij} of :
  - the **next state** being the state corresponding to **row *i***
  - **given that the current state** is the state corresponding to **column *j***
- The probabilities in each **column** must **add up to 1**
- The transition matrix depends on how you assign the states to the columns
  - Each transition matrix for a Markov chain will contain the same elements
    - The rows and columns may be in different orders though
    - E.g. Sunny (S) & Cloudy (C) could be in the order **S then C** or **C then S**

What is an initial state probability matrix?

- An **initial state probability matrix** $\mathbf{s}_0$ is a column vector which contains the **probabilities** of each state being chosen as the **initial state**
  - If you know which state was chosen as the initial state then that entry will be 1 and the others will all be zero
- You can find the **state probability matrix** $\mathbf{s}_1$ which contains the probabilities of each state being chosen after **one interval of time**
  - $\mathbf{s}_1 = T\mathbf{s}_0$

How do I find expected values after one interval of time?

- Suppose the Markov change represents a **population moving between states**
  - Examples include:
    - People in a town switching gyms each year
    - Children choosing a type of sandwich for their lunch each day
- Suppose the **total population is fixed** and equals N
- You can **multiply the state probability matrix** $\mathbf{s}_1$ by N to find the expected number of members of the population at each state

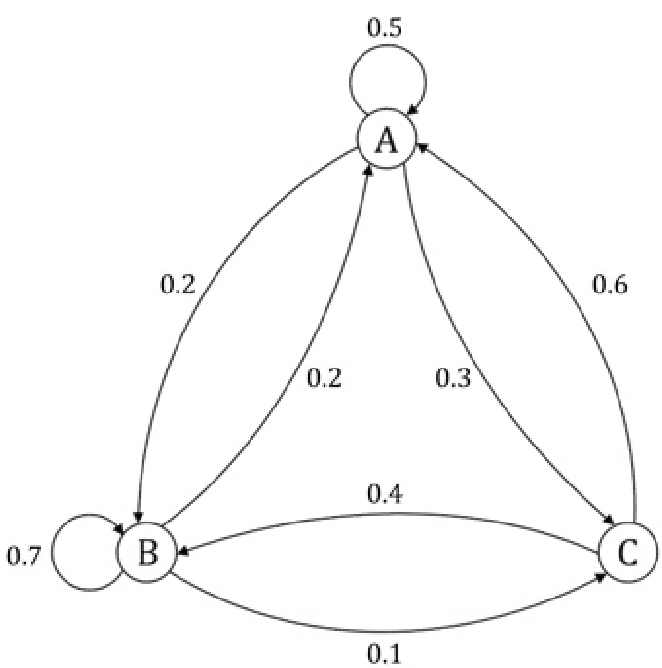
Examiner Tip

- If you are asked to find a transition matrix, check that all the probabilities within a column add up to 1
- Drawing a transition state diagram can help you to visualise the problem

Your notes

Worked example

Each year Jamie donates to one of three charities: A, B or C. At the start of each year, the probabilities of Jamie continuing donate to the same charity or changing charities are represented by the following transition state diagram:



a) Write down a transition matrix $\boldsymbol{T}$ for this system of probabilities.

Current state

$$\text{Next state}\quad \begin{matrix} & A & B & C \\ A & 0.5 & 0.2 & 0.6 \\ B & 0.2 & 0.7 & 0.4 \\ C & 0.3 & 0.1 & 0 \end{matrix}$$

$$T = \begin{pmatrix} 0.5 & 0.2 & 0.6 \\ 0.2 & 0.7 & 0.4 \\ 0.3 & 0.1 & 0 \end{pmatrix}$$

b) There is a 10% chance that charity A is the first charity that Jamie chooses, a 10% chance for charity B and an 80% chance for charity C. Find the charity which has the highest probability of being picked as the second charity after the first year.

Your notes

Write down the initial state vector $s_0 = \begin{pmatrix} 0.1 \\ 0.1 \\ 0.8 \end{pmatrix}$

$s_1 = Ts_0$ $\quad s_1 = \begin{pmatrix} 0.5 & 0.2 & 0.6 \\ 0.2 & 0.7 & 0.4 \\ 0.3 & 0.1 & 0 \end{pmatrix}\begin{pmatrix} 0.1 \\ 0.1 \\ 0.8 \end{pmatrix} = \begin{pmatrix} 0.55 \\ 0.41 \\ 0.04 \end{pmatrix}$

Charity A has the highest probability of being the second charity picked.

Your notes

Powers of Transition Matrices

How do I find powers of a transition matrix?

- You can simply use your **GDC** to find given powers of a matrix
- The power could be left in terms of an **unknown** *n*
  - In this case it would be more helpful to write the transition matrix in diagonalised form (see section **1.8.2 Applications of Matrices**) $T = PDP^{-1}$ where
    - *D* is a **diagonal matrix** of the **eigenvalues**
    - *P* is a matrix of **corresponding eigenvectors**
  - Then $T^n = PD^nP^{-1}$
    - This is given in the **formula booklet**
  - Every transition matrix always has an **eigenvalue equal to 1**

What is represented by the powers of a transition matrix?

- The powers of a transition matrix also **represent probabilities**
- The element of T^n in the i^{th} row and j^{th} column gives the **probability** t^n_{ij} of :
  - the **future state** after ***n* intervals of time** being the state corresponding to **row *i***
  - **given that** the **current state** is the state corresponding to **column *j***
- For example: Let ***T*** be a transition matrix with the element $t_{2,3}$ representing the probability that tomorrow is sunny given that it is raining today
  - The element $t^5_{2,3}$ of the matrix T^5 represents the probability that it is sunny in 5 days' time given that it is raining today
- The probabilities in **each column** must still **add up to 1**

How do I find the column state matrices?

- The column state matrix $\mathbf{s_n}$ is a column vector which contains the **probabilities** of each state being chosen after *n* intervals of time given the current state
  - $\mathbf{s_n}$ depends on $\mathbf{s_0}$
- To calculate the column state matrix you raise the transition matrix to the power *n* and multiply by the initial state matrix
  - $T^n s_0 = s_n$
    - You are given this in the **formula booklet**
- You can multiply $\mathbf{s_n}$ by the fixed population size to find the expected number of members of the population at each state after *n* intervals of time

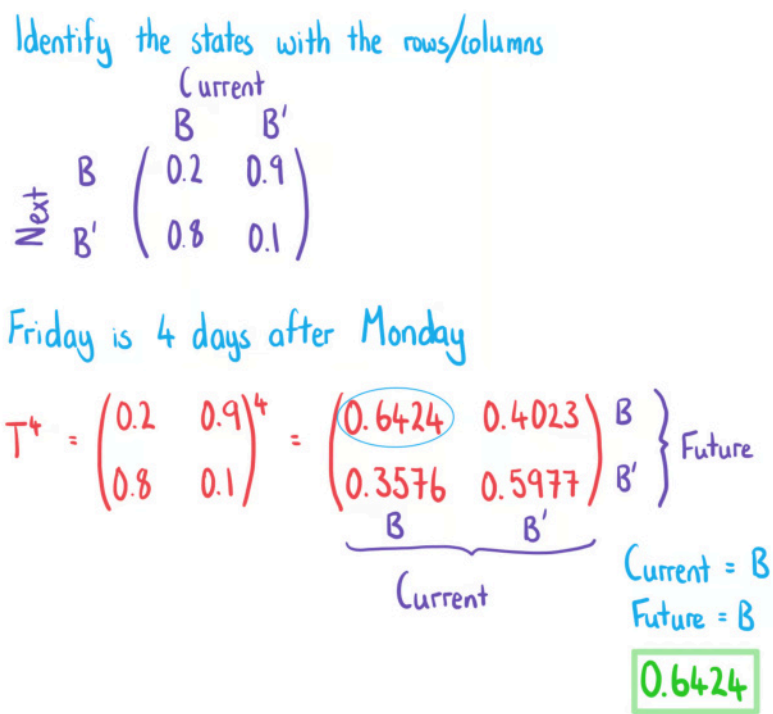
Your notes

Worked example

At a cat sanctuary there are 1000 cats. If a cat is brushed on a given day, then the probability it is brushed the following day is 0.2. If a cat is not brushed on a given day, then the probability that is will be brushed the following day is 0.9.

The transition matrix $\boldsymbol{T}$ is used to model this information with $\boldsymbol{T} = \begin{pmatrix} 0.2 & 0.9 \\ 0.8 & 0.1 \end{pmatrix}$.

a) On Monday Hippo the cat is brushed. Find the probability that Hippo will be brushed on Friday.



b) On Monday 700 cats were brushed. Find the expected number of cats that will be brushed on the following Monday.

Your notes

On Monday 700 brushed $\quad S_0 = \begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix}$

Expected numbers after 7 days

$\text{Total} \times S_7 = \text{Total} \times T^7 S_0$

$$1000 \times \begin{pmatrix} 0.2 & 0.9 \\ 0.8 & 0.1 \end{pmatrix}^7 \begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix} = \begin{pmatrix} 0.2 & 0.9 \\ 0.8 & 0.1 \end{pmatrix}^7 \begin{pmatrix} 700 \\ 300 \end{pmatrix} = \begin{pmatrix} 515.36309 \\ 484.63691 \end{pmatrix} \begin{matrix} B \\ B' \end{matrix}$$

515 cats

Your notes

Steady State & Long-term Probabilities

What is the steady state of a regular Markov chain?

- The vector **s** is said to be a **steady state** vector if it does not change when multiplied by the transition matrix
  - $\boldsymbol{T}\mathbf{s} = \mathbf{s}$
- **Regular Markov chains** have steady states
  - A Markov chain is said to be regular if there exists a **positive integer *k*** such that **none of the entries** are **equal to 0** in the matrix $\boldsymbol{T}^k$
    - For this course all Markov chains will be regular
- The transition matrix for a regular Markov chain will have **exactly one** eigenvalue equal to 1 and the **rest will all be less than 1**
- As n gets bigger $\boldsymbol{T}^n$ tends to a matrix where **each column is identical**
  - The column matrix formed by using **one of these columns** is called the steady state column matrix **s**
  - This means that the **long-term probabilities** tend to fixed probabilities
    - $\mathbf{s}_n$ tends to **s**

How do I use long-term probabilities to find the steady state?

- As $\boldsymbol{T}^n$ tends to a matrix whose columns equal the steady state vector
  - Calculate $\boldsymbol{T}^n$ for a large value of n using your GDC
  - If the columns are identical when rounded to a required degree of accuracy then the column is the steady state vector
  - If the columns are not identical then choose a higher power and repeat

How do I find the exact steady state probabilities?

- As $\boldsymbol{T}\mathbf{s} = \mathbf{s}$ the steady state vector **s** is the **eigenvector** of $\boldsymbol{T}$ corresponding to the **eigenvalue equal to 1** whose elements sum to 1:
  - Let **s** have entries $x_1, x_2, ..., x_n$
  - Use $\boldsymbol{T}\mathbf{s} = \mathbf{s}$ to form a system of linear equations
  - There will be an infinite number of solutions so choose a value for one of the unknowns
    - For example: let $x_n = 1$
  - Ignoring the last equation solve the system of linear equations to find $x_1, x_2, ..., x_{n-1}$
  - Divide each value x_i by the sum of the values
    - This makes the values add up to 1
- You might be asked to **show this result using diagonalisation**
  - Write $\boldsymbol{T} = \boldsymbol{PDP}^{-1}$ where $\boldsymbol{D}$ is the diagonal matrix of eigenvalues and $\boldsymbol{P}$ is the matrix of eigenvectors
  - Use $\boldsymbol{T}^n = \boldsymbol{PD}^n\boldsymbol{P}^{-1}$
  - As n gets large $\boldsymbol{D}^n$ tends to a matrix where all entries are 0 apart from one entry of 1 due to the eigenvalue of 1
  - Calculate the limit of $\boldsymbol{T}^n$ which will have **identical columns**
    - You can calculate this by multiplying the three matrices ($\boldsymbol{P}, \boldsymbol{D}^\infty, \boldsymbol{P}^{-1}$) together

Your notes

Examiner Tip

- If you calculate T^{∞} by hand then a quick check is to see if the columns are identical
  - It should look like $\begin{pmatrix} a & a & a \\ b & b & b \\ c & c & c \end{pmatrix}$

Your notes

Worked example

If a cat is brushed on a given day, then the probability it is brushed the following day is 0.2. If a cat is not brushed on a given day, then the probability that is will be brushed the following day is 0.9.

The transition matrix T is used to model this information with $T = \begin{pmatrix} 0.2 & 0.9 \\ 0.8 & 0.1 \end{pmatrix}$.

a) Find an eigenvector of T corresponding to the eigenvalue 1.

$\underline{v}$ is an eigenvector of T with eigenvalue 1 if $T\underline{v} = \underline{v}$

Let $\underline{v} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$T\underline{v} = \begin{pmatrix} 0.2 & 0.9 \\ 0.8 & 0.1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.2x_1 + 0.9x_2 \\ 0.8x_1 + 0.1x_2 \end{pmatrix}$

$T\underline{v} = \underline{v}$ $\quad 0.2x_1 + 0.9x_2 = x_1 \Rightarrow 0.9x_2 = 0.8x_1 \Rightarrow 9x_2 = 8x_1$

$\quad 0.8x_1 + 0.1x_2 = x_2 \Rightarrow 0.8x_1 = 0.9x_2 \Rightarrow 8x_1 = 9x_2$

Find a solution $x_1 = 9$ and $x_2 = 8$

$\begin{pmatrix} 9 \\ 8 \end{pmatrix}$ or any scalar multiple

b) Hence find the steady state vector.

Your notes

Scale the elements so that they add to 1 $\begin{pmatrix} \frac{9}{17} \\ \frac{8}{17} \end{pmatrix}$

The eigenvector corresponding to the eigenvalue 1, whose elements add to 1, is the steady state vector.

$$\begin{pmatrix} \frac{9}{17} \\ \frac{8}{17} \end{pmatrix}$$