# Regular Markov chains

## MATH 2740 – Mathematics of Data Science – Lecture 13

**Julien Arino**

julien.arino@umanitoba.ca

**Department of Mathematics @ University of Manitoba**

**Fall 202X**

The University of Manitoba campuses are located on original lands of Anishinaabeg, Ininew, Anisininew, Dakota and Dene peoples, and on the National Homeland of the Red River Métis. We respect the Treaties that were made on these territories, we acknowledge the harms and mistakes of the past, and we dedicate ourselves to move forward in partnership with Indigenous communities in a spirit of Reconciliation and collaboration.

# Outline

Markov chains

Running example – Mendelian inheritance

Repetition of the process

Regular Markov chains

## Markov chains

## Running example – Mendelian inheritance

## Repetition of the process

## Regular Markov chains

# Markov chain

A Markov chain is a *stochastic process* in which the evolution through time depends only on the current state of the system (we say the process is *memoryless*)

Markov chains are an interesting combination of matrix theory and graph theory

They form the theoretical foundation for Hidden Markov processes or Markov Chain Monte Carlo (MCMC) methods, are used in ML

Conduct an experiment with a set of $n$ possible outcomes

$$S = \{S_1, \ldots, S_n\}$$

Experiment repeated $t$ times (with $t$ large, potentially infinite)

System has *no memory*: the next state depends only on the present state

Probability of $S_i$ occurring on the next step, given that $S_j$ occurred on the last step, is

$$p_{ij} = p(S_i | S_j)$$

Suppose that $S_i$ is the current state, then one of $S_1, \ldots, S_n$ must be the next state; so

$$p_{1i} + p_{2i} + \cdots + p_{ni} = 1, \quad 1 \le i \le n$$

(Some of the $p_{ij}$ can be zero, all that is needed is that $\sum_{j=1}^{n} p_{ij} = 1$ for all $i$)

## Definition 1

An experiment with finite number of possible outcomes $S_1, \ldots, S_n$ is repeated. The sequence of outcomes is a **Markov chain** if there is a set of $n^2$ numbers $\{p_{ij}\}$ such that the conditional probability of outcome $S_i$ on any experiment given outcome $S_j$ on the previous experiment is $p_{ij}$, i.e., for $1 \le i, j \le n$, $t = 1, \ldots$,

$$p_{ij} = \mathbb{P}(S_i \text{ on experiment } t+1 \mid S_j \text{ on experiment } t)$$

Outcomes $S_1, \ldots, S_n$ are **states** and $p_{ij}$ are **transition probabilities**. $P = [p_{ij}]$ the **transition matrix**

The matrix

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{pmatrix}$$

has

- nonnegative entries, $p_{ij} \geq 0$
- entries less than 1, $p_{ij} \leq 1$
- column sum 1, which we write

$$\sum_{i=1}^{n} p_{ij} = 1, \quad j = 1, \ldots, n$$

or, using the notation $\mathbb{1}^T = (1, \ldots, 1)$,

$$\mathbb{1}^T P = \mathbb{1}^T$$

Markov chains

**Running example – Mendelian inheritance**

Repetition of the process

Regular Markov chains

# The “orange” gene

A cat’s coat color is determined by many genes. The "orange" trait comes from a specific gene called the **orange locus**

It has two *alleles* (versions):

- **O** $\Longrightarrow$ Produces *phaeomelanin* (red/orange pigment)
- **o** $\Longrightarrow$ Produces *eumelanin* (black/brown pigment)

This gene is *sex-linked*. It is located on the **X chromosome**

- This changes the rules of inheritance!

# How sex-linked genes work

Because the gene is on the X chromosome, males and females inherit it differently

**Females have two X chromosomes (XX)**

- They get two alleles for this gene (one from each parent)
- Possible genotypes: $X^O X^O$, $X^o X^o$, or $X^O X^o$

**Males have one X and one Y chromosome (XY)**

- They get *only one* allele for this gene (always from the mother)
- Possible genotypes: $X^O Y$ or $X^o Y$

# Genotype vs. phenotype

**Males (simple):**

- $X^O Y \implies$ **orange cat**
- $X^o Y \implies$ **non-orange cat** (e.g., black)

**Females (the special case):**

- $X^O X^O \implies$ **orange cat**
- $X^o X^o \implies$ **non-orange cat** (e.g., black)
- $X^O X^o \implies$ **tortoiseshell cat**

A “tortie” isn’t a simple hybrid. Both alleles ($O$ and $o$) are active in different patches of skin, creating the orange and black mottled pattern

# Example 1: Orange dad + black Mom

Let's cross an **orange male ($X^O Y$)** with a **black female ($X^o X^o$)**

| | | **Father** | |
|---|---|---|---|
| | | $X^O$ | $Y$ |
| **Mother** | $X^o$ | $X^O X^o$ | $X^o Y$ |
| | $X^o$ | $X^O X^o$ | $X^o Y$ |

**Results for their offspring:**

- **All females** ($X^O X^o$) will be **tortoiseshell**
- **All males** ($X^o Y$) will be **black** (non-orange)

# Example 2: black dad + tortoiseshell mom

Let's cross a **black male ($X^oY$)** with a **tortoiseshell female ($X^OX^o$)**

| | | **Father** | |
|---|---|---|---|
| | | $X^o$ | $Y$ |
| **Mother** | $X^O$ | $X^OX^o$ | $X^OY$ |
| | $X^o$ | $X^oX^o$ | $X^oY$ |

**Results for their offspring (1/4 chance for each):**

- $X^OX^o \implies$ **Tortoiseshell Female**
- $X^oX^o \implies$ **Black Female**
- $X^OY \implies$ **Orange Male**
- $X^oY \implies$ **Black Male**

# Fun fact: what about male tortoiseshells?

- As we saw, a male is $XY$. He can only get $X^O$ or $X^o$ from his mother, not both
- A male tortoiseshell is possible, but *extremely rare*
- It's a genetic anomaly where the cat has an extra X chromosome: **XXY**
- This genotype (e.g., $X^O X^o Y$) allows the cat to be male ($Y$) but also express both orange and non-orange alleles ($X^O X^o$), just like a female

Markov chains

Running example – Mendelian inheritance

**Repetition of the process**

Regular Markov chains

# General case

$p_i(t)$: probability that state $S_i$ occurs on the $t^{th}$ repetition of the experiment, $1 \leq i \leq n$

Since one the states $S_i$ must occur on the $t^{th}$ repetition

$$p_1(t) + p_2(t) + \cdots + p_n(t) = 1$$

$p_i(t+1)$: probability that state $S_i$, $1 \leq i \leq r$, occurs on $(t+1)^{th}$ repetition of the experiment

$n$ ways to be in state $S_i$ at step $t+1$:

1. Step $t$ is $S_1$. Probability of getting $S_1$ on $t^{th}$ step is $p_1(t)$, and probability of having $S_i$ after $S_1$ is $p_{i1}$. Therefore $P(S_i|S_1) = p_{i1}p_1(t)$
2. We get $S_2$ on step $t$ and $S_i$ on step $(t+1)$. Then $P(S_i|S_2) = p_{2i}p_2(t)$

..

n. Probability of occurrence of $S_i$ at step $t+1$ if $S_n$ at step $t$ is $P(S_i|S_n) = p_{in}p_n(t)$

$$\implies p_i(t+1) = P(S_i|S_1) + \cdots + P(S_i|S_n)$$
$$= p_{i1}p_1(t) + \cdots + p_{in}p_n(t)$$

Therefore,

$$p_1(t+1) = p_{11}p_1(t) + p_{12}p_2(t) + \cdots + p_{1n}p_n(t)$$
$$\vdots$$
$$p_n(t+1) = p_{n1}p_1(t) + p_{n2}p_2(t) + \cdots + p_{nn}p_n(t)$$

In matrix form

$$p(t+1) = Pp(t), \quad n = 1, 2, 3, \ldots$$

where $p(t) = (p_1(t), p_2(t), \ldots, p_n(t))^T$ is a probability vector and $P = (p_{ij})$ is an $n \times n$ *transition matrix*,

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ & & & \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{pmatrix}$$

So

$$\begin{pmatrix} p_1(t+1) \\ \vdots \\ p_n(t+1) \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ & & & \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{pmatrix} \begin{pmatrix} p_1(t) \\ \vdots \\ p_n(t) \end{pmatrix}$$

Easy to check that this gives the same expression as before

# Stochastic matrices

**Definition 2 (Stochastic matrix)**

The nonnegative $n \times n$ matrix $M$ is **row-stochastic** (resp. **column-stochastic**) if $\sum_{j=1}^{n} a_{ij} = 1$ for all $i = 1, \ldots, n$ (resp. $\sum_{i=1}^{n} a_{ij} = 1$ for all $j = 1, \ldots, n$)

We often say **stochastic** and let the context determine whether we mean row- or column-stochastic

If it is both row- and column-stochastic, the matrix is **doubly stochastic**

**Theorem 3**

*Let $M$ be a stochastic matrix. Then all eigenvalues $\lambda$ of $M$ are such that $|\lambda| \leq 1$. Furthermore, $\lambda = 1$ is an eigenvalue of $M$*

# Long time behaviour

Let $p(0)$ be the initial distribution vector. Then

$$\begin{aligned} p(1) &= Pp(0) \\ p(2) &= Pp(1) \\ &= P(Pp(0)) \\ &= P^2 p(0) \end{aligned}$$

Iterating, we get, for any $t$,

$$p(t) = P^t p(0)$$

Therefore,

$$\begin{aligned} \lim_{t\to+\infty} p(t) &= \lim_{t\to+\infty} P^t p(0) \\ &= \left(\lim_{t\to+\infty} P^t\right) p(0) \end{aligned}$$

if this limit exists

$$\lim_{n \to +\infty} p(t) = \left( \lim_{t \to +\infty} P^t \right) p(0)$$

Does the limit exist?

### Theorem 4

*If $M, N$ are nonsingular stochastic matrices, then $MN$ is a stochastic matrix*

### Corollary 5

*If $M$ is a nonsingular stochastic matrix, then for any $k \in \mathbb{N}$, $M^k$ is a stochastic matrix*

So $P^t$ above is stochastic

Markov chains

Running example – Mendelian inheritance

Repetition of the process

**Regular Markov chains**

# Regular Markov chains

**Definition 6 (Regular Markov chain)**

A **regular** Markov chain has $P^k$ (entry-wise) positive for some integer $k > 0$, i.e., $P^k$ has only positive entries

**Definition 7 (Primitive matrix)**

A nonnegative matrix $M$ is **primitive** if, and only if, there is an integer $k > 0$ such that $M^k$ is positive.

**Theorem 8**

*Markov chain regular* $\iff$ *transition matrix* $P$ *primitive*

# Computing $W$

If $p(t)$ converges, then $p(t+1) = Pp(t)$ at the limit, so $w = \lim_{t\to\infty} p(t)$ is a **fixed point** of the system. Write

$$w = Pw$$

and solve for $w$, i.e., find $w$ as a (right) eigenvector corresponding to the eigenvalue 1

$w$ might have to be normalized (you want a probability vector). Check that the norm $\|w\|_1$ defined by

$$\|w\|_1 = |w_1| + \cdots + |w_n| = w_1 + \cdots + w_n$$

(since $w \geq 0$) is equal to one. If not, use

$$\tilde{w} = \frac{w}{\|w\|_1}$$

# Behaviour of a regular MC

## Theorem 9

*If $P$ is the transition matrix of a regular Markov chain, then*

1. *the powers $P^t$ approach a stochastic matrix $W$*
2. *each column of $W$ is the same (column) vector $w = (w_1, \ldots, w_n)^T$*
3. *the components of $w$ are positive*

So if the Markov chain is regular

$$\lim_{t \to +\infty} p(t) = \lim_{t \to +\infty} P^t p(0) = Wp(0)$$

# Back to orange cats

Create a chain by tracking the 3 female genotypes:

- $S_1$: $X^O X^O$ (orange)
- $S_2$: $X^o X^o$ (black)
- $S_3$: $X^O X^o$ (tortoiseshell)

To make the chain regular, we mate our female with a male chosen randomly from a **fixed population** that is:

- 50% orange males ($X^O Y$)
- 50% black males ($X^o Y$)

# State 1: orange female ($X^O X^O$)

The mother is $X^O X^O$. We pick a father with 50/50 probability.

**Case 1: Father is $X^O Y$**

| Mother | $X^O$ | $Y$ |
|---|---|---|
| $X^O$ | $X^O X^O$ ($S_1$) | Male |
| $X^O$ | $X^O X^O$ ($S_1$) | Male |

Daughters: 100% $S_1$

**Case 2: Father is $X^o Y$**

| Mother | $X^o$ | $Y$ |
|---|---|---|
| $X^O$ | $X^O X^o$ ($S_3$) | Male |
| $X^O$ | $X^O X^o$ ($S_3$) | Male |

Daughters: 100% $S_3$

**Transitions from $S_1$:**

- $\mathbb{P}(S_1 \to S_1) = 0.5 \times 1.0 = \mathbf{0.5}$
- $\mathbb{P}(S_1 \to S_2) = 0$
- $\mathbb{P}(S_1 \to S_3) = 0.5 \times 1.0 = \mathbf{0.5}$

# State 2: black female ($X^oX^o$)

The mother is $X^oX^o$. We pick a father with 50/50 probability.

**Case 1: Father is $X^OY$**

| Mother | | $X^O$ | $Y$ |
|---|---|---|---|
| | $X^o$ | $X^OX^o$ ($S_3$) | Male |
| | $X^o$ | $X^OX^o$ ($S_3$) | Male |

Daughters: 100% $S_3$

**Case 2: Father is $X^oY$**

| Mother | | $X^o$ | $Y$ |
|---|---|---|---|
| | $X^o$ | $X^oX^o$ ($S_2$) | Male |
| | $X^o$ | $X^oX^o$ ($S_2$) | Male |

Daughters: 100% $S_2$

**Transitions from $S_2$:**

- $\mathbb{P}(S_2 \to S_1) = 0$
- $\mathbb{P}(S_2 \to S_2) = 0.5 \times 1.0 = \mathbf{0.5}$
- $\mathbb{P}(S_2 \to S_3) = 0.5 \times 1.0 = \mathbf{0.5}$

# State 3: tortoiseshell female ($X^O X^o$)

The mother is $X^O X^o$. We pick a father with 50/50 probability.

**Case 1: Father is $X^O Y$**

| Mother | $X^O$ | $Y$ |
|---|---|---|
| $X^O$ | $X^O X^O$ ($S_1$) | Male |
| $X^o$ | $X^O X^o$ ($S_3$) | Male |

Daughters: 50% $S_1$, 50% $S_3$

**Case 2: Father is $X^o Y$**

| Mother | $X^o$ | $Y$ |
|---|---|---|
| $X^O$ | $X^O X^o$ ($S_3$) | Male |
| $X^o$ | $X^o X^o$ ($S_2$) | Male |

Daughters: 50% $S_2$, 50% $S_3$

**Transitions from $S_3$:**

- $\mathbb{P}(S_3 \to S_1) = 0.5 \times 0.5 = \mathbf{0.25}$
- $\mathbb{P}(S_3 \to S_2) = 0.5 \times 0.5 = \mathbf{0.25}$
- $\mathbb{P}(S_3 \to S_3) = (0.5 \times 0.5) + (0.5 \times 0.5) = \mathbf{0.5}$

# Summary of the 3-state chain

The transition matrix $P$ for states $\{S_1, S_2, S_3\}$ is:

$$P = \begin{pmatrix} 0.5 & 0 & 0.5 \\ 0 & 0.5 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{pmatrix}$$

**Is this chain regular?**

- **Irreducible? Yes.** All states communicate.
  - $S_1 \to S_3 \to S_2$ (Path from $S_1$ to $S_2$)
  - $S_2 \to S_3 \to S_1$ (Path from $S_2$ to $S_1$)
  - All other paths are direct ($S_1 \to S_3$, $S_3 \to S_1$, etc.)
- **Aperiodic? Yes.** All states have self-loops ($p_{11}, p_{22}, p_{33} > 0$).

Since the chain is irreducible and aperiodic, it is **regular**.

Compute (right) eigenvector associated to 1

$$\begin{pmatrix} 1/2 & 1/4 & 0 \\ 1/2 & 1/2 & 1/2 \\ 0 & 1/4 & 1/2 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

$$\begin{aligned} \frac{1}{2}w_1 + \frac{1}{4}w_2 &= w_1 \\ \frac{1}{2}w_1 + \frac{1}{2}w_2 + \frac{1}{2}w_3 &= w_2 \\ \frac{1}{4}w_2 + \frac{1}{2}w_3 &= w_3 \end{aligned}$$

So $w_1 = w_2/2$, $w_3 = w_2/2$ and thus

$$\frac{1}{4}w_2 + \frac{1}{2}w_2 + \frac{1}{4}w_2 = w_2,$$

that is, $w_2 = w_2$, i.e., $w_2$ can take any value

$$\implies \quad w = \left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$$

```r
# Total population
nb_states = 10
# Parameters
proba_left = 0.5
proba_right = 0.5
proba_stay = 1-(proba_left+proba_right)
# Make the transition matrix
T = mat.or.vec(nr = nb_states, nc = nb_states)
for (row in 2:(nb_states-1)) {
  T[row,(row-1)] = proba_left
  T[row,(row+1)] = proba_right
  T[row, row] = proba_stay
}
T[1,2] = 1 # First row only moves right
T[nb_states, (nb_states-1)] = 1 # Last row only moves left
```

# Create the MC object

```r
# Library: markovchain
mcRW <- new("markovchain",
            states = sprintf("S_%d", 1:nb_states),
            transitionMatrix = T,
            name = "RW_reg")
```

# Show some information about the chain

```r
summary(mcRW)

## RW_reg  Markov chain that is composed by:
## Closed classes:
## S_1 S_2 S_3 S_4 S_5 S_6 S_7 S_8 S_9 S_10
## Recurrent classes:
## {S_1,S_2,S_3,S_4,S_5,S_6,S_7,S_8,S_9,S_10}
## Transient classes:
## NONE
## The Markov chain is irreducible
## The absorbing states are: NONE
```

# The equilibrium distribution

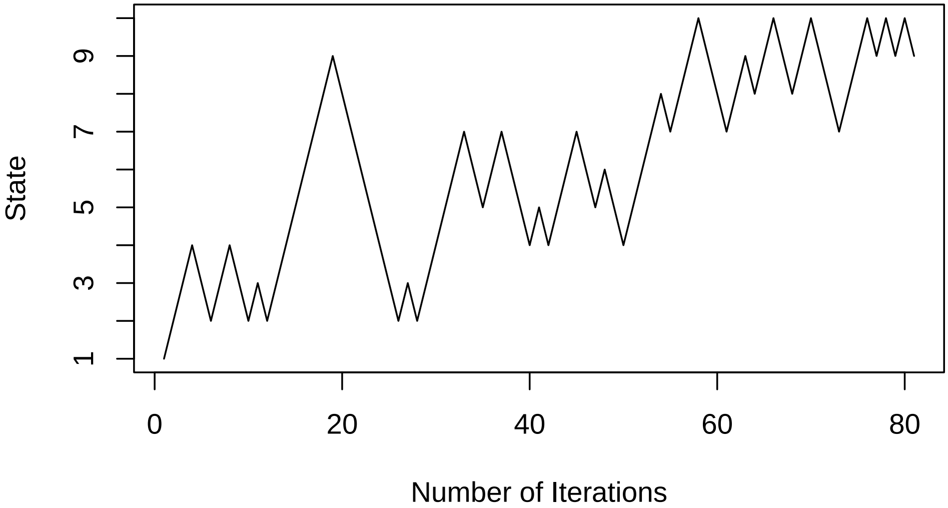
```r
steadyStates(mcRW)
```

```
##             S_1       S_2       S_3       S_4       S_5       S_6       S
## [1,] 0.05555556 0.1111111 0.1111111 0.1111111 0.1111111 0.1111111 0.11111
##            S_8       S_9       S_10
## [1,] 0.1111111 0.1111111 0.05555556
```

# Showing a realisation

```r
# Library: DTMCPack
IC = rep(0, nb_states)
IC[1] = 1
sol = DTMC(T, IC, 81, trace=TRUE)

## NULL
```



State
1
3
5
7
9
0
20
40
60
80
Number of Iterations