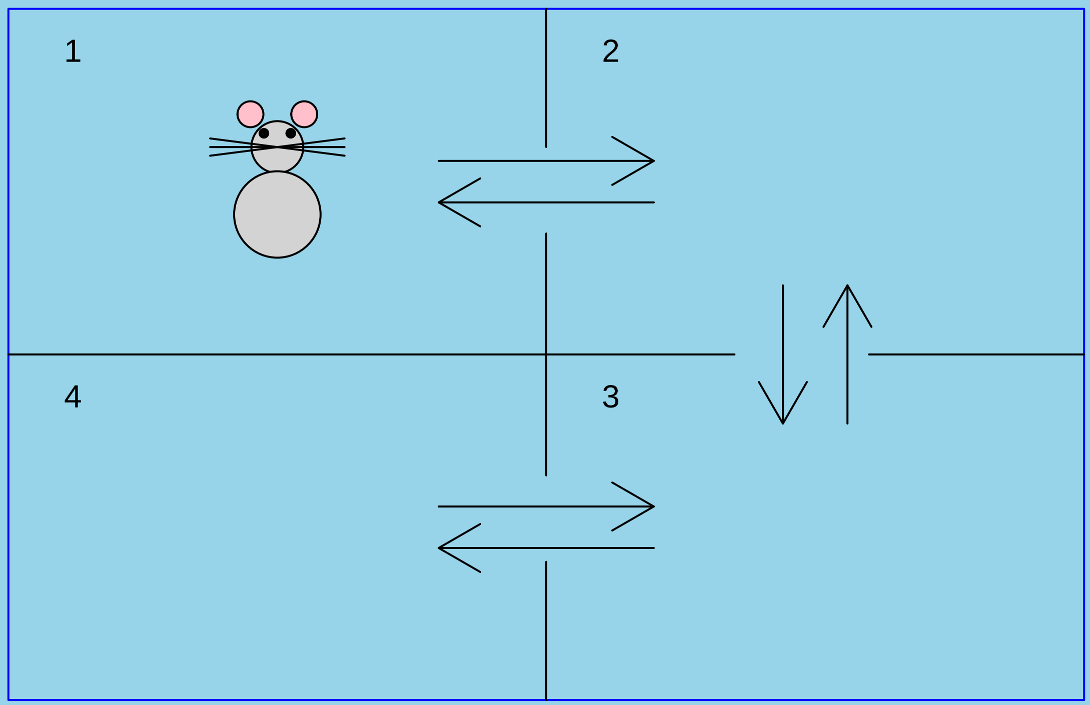
COSC/DATA 405/505

Modelling and Simulation

Outline

Example

A Mouse in a Maze



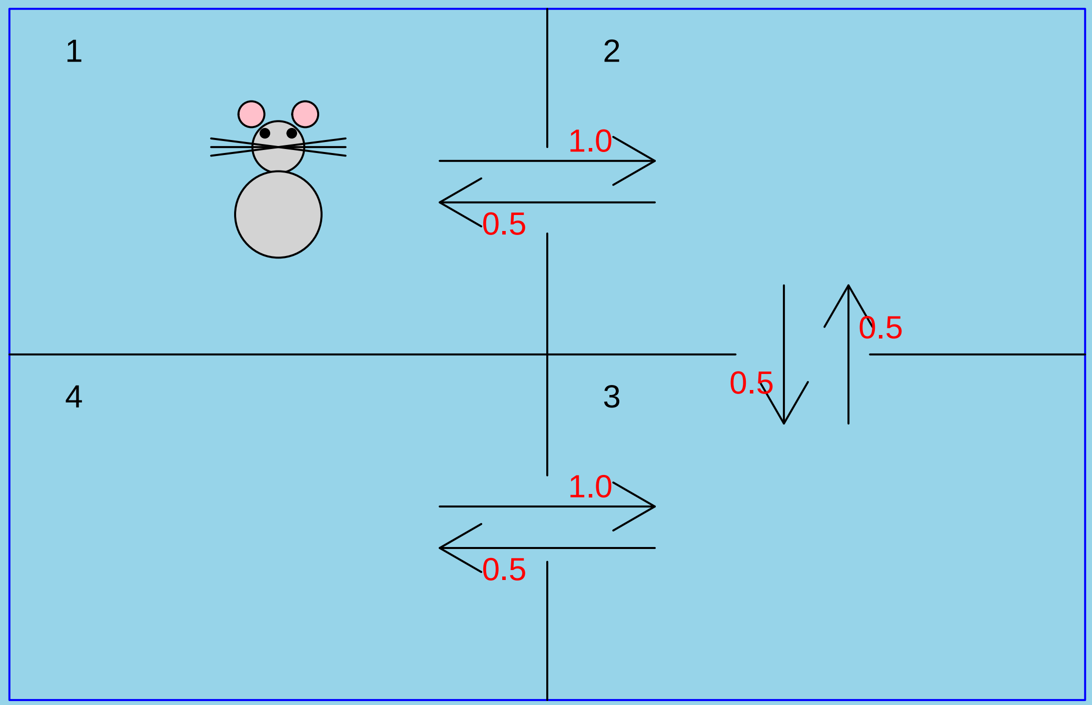
A Mouse Movement Model - Complete Randomness

If the mouse wanders aimlessly, we might attach probabilities to the *transitions* between compartments in the maze as in the following diagram.

Alternatively, the mouse might favour one of the compartments over another, perhaps because of the location of food or a particular odor.

A Mouse Movement Model - Complete Randomness

Transition Probabilities



A Mouse Movement Model - Complete Randomness

Transition Matrix

The transition probabilities can be organized systematically into a matrix, noting that impossible transitions have probability 0.

The 1st row of the matrix lists the transition probabilities from the first compartment to all other compartments. We assume that the mouse is not staying in compartment 1: 0, 1, 0, 0.

2nd row – transition probabilities out of compartment 2: 0.5, 0, 0.5, 0.

3rd row – transition probabilities from compartment 3: 0, 0.5, 0, 0.5.

4th row – transition probabilities out of compartment 4: 0, 0, 1, 0.

A Mouse Movement Model - Complete Randomness

Transition Matrix

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Entry (i, j) represents the probability of transition into compartment j, given that the mouse was in state i.

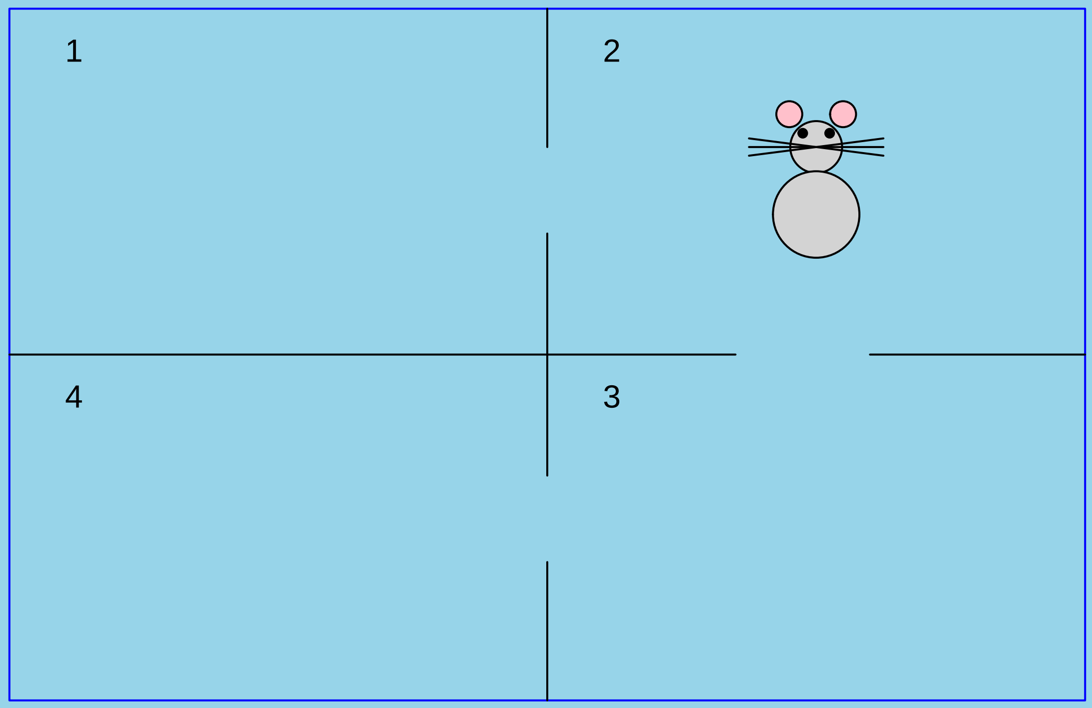
Simulating from the Mouse Movement Model

X_t = **compartment at time** t

$X_0 = 1$ **(Mouse starts in the 1st compartment.)**

Simulating from the Mouse Movement Model

At the first transition, the mouse enters the 2nd compartment: $X_1 = 2$.



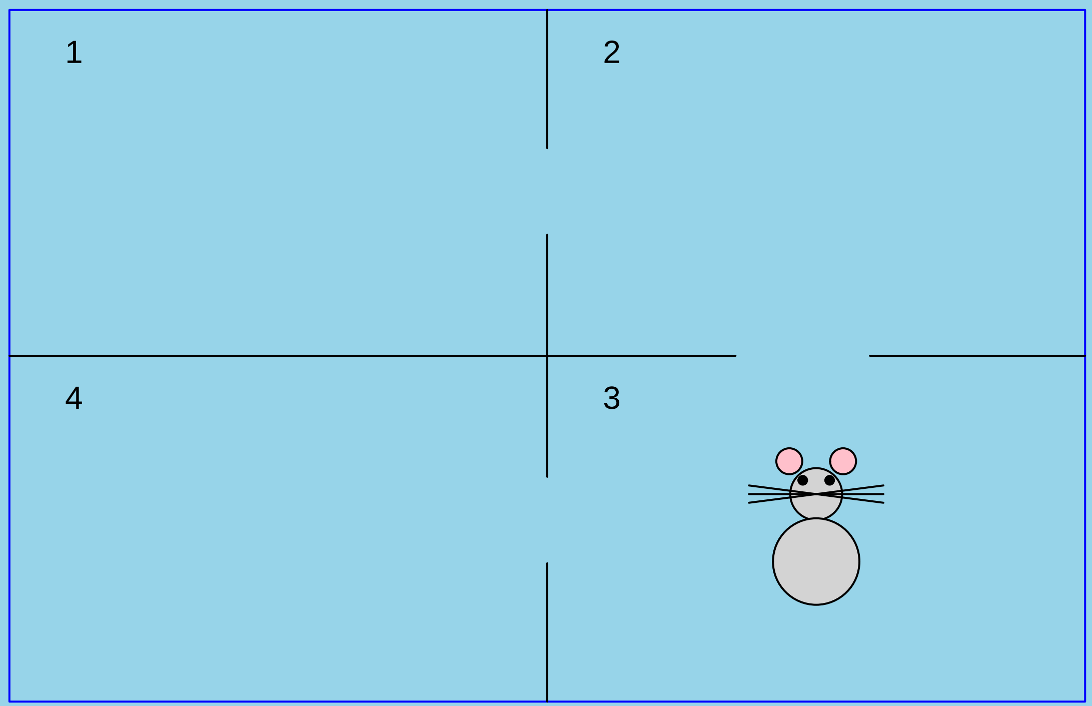
Simulating from the Mouse Movement Model

At the second transition, the mouse can enter compartment 1 with probability 0.5 or compartment 3 with probability 0.5, like a Bernoulli trial.

Generate a pseudorandom number U_2. If $U_2 < 0.5$, the mouse enters compartment 1. Otherwise, it enters compartment 3.

Simulating from the Mouse Movement Model

$U_2 = 0.61$, **so the mouse enters the 3rd compartment:** $X_2 = 3$**.**



Simulating from the Mouse Movement Model

$U_3 = 0.34$, **so the mouse enters the 2nd compartment:** $X_3 = 2$.

$U_4 = 0.88$, **so** $X_4 = 3$.

$U_5 = 0.52$, **so** $X_5 = 4$.

The mouse must re-enter compartment 3 with probability 1: $X_6 = 3$

and so on ....

$\{X_1, X_2, X_3, \ldots\}$ **is an example of a Markov chain.**

Simulating the Model in R

A simple way to simulate values of X_j given the value of X_{j-1} is to use the `sample()` function.

A simple application of `sample()`: simple random sampling from a population

```
N <- 2000000 # population size
n <- 50 # sample size
sample(1:N, size = n, replace = FALSE)
```

Simulating the Model in R

A sample from a larger population:

```
##  [1] 1527008 1122622  584376 1346146  495293
##  [6] 1467063  467262  433770  248959  104058
## [11] 1637822 1621857 1871020  772812  294664
## [16]  299495  632355  788408 1174441  517412
## [21] 1524527  774789  762264  681565 1408257
## [26]   90456  432791  862099 1995696 1716061
## [31]  549969  585086  931023  670723 1054233
## [36]  345823  908149  892966  542162 1236185
## [41] 1843070 1086765  544733  447914 1100151
## [46] 1197779  437779 1606850  294199  795967
```

Simulating the Model in R

Another example: simulating from a discrete probability distribution

$$P(X = 1) = .1, P(X = 2) = .2, P(X = 3) = .1, P(X = 4) = .4, P(X = 5) = .2$$

```
sample(1:5, size = 10, replace = TRUE,
      prob = c(.1, .2, .1, .4, .2))

##  [1] 3 5 1 4 2 2 4 5 5 4
```

Note that when simulating values in a Markov chain, we are simulating from probability distributions defined by the rows of the transition matrix.

Simulating the Model in R

```
Ntransitions <- 100000 # number of mouse moves
P <- matrix(c(0, 1, 0, 0,
              0.5, 0, 0.5, 0,
              0, 0.5, 0, 0.5,
              0, 0, 1, 0), nrow = 4,
              byrow = TRUE)  # P is the transition matrix
location <- numeric(Ntransitions)#initializing the chain
current.state <- 1 # initial compartment
for (t in 1:Ntransitions) {
   current.state <- sample(1:4,
        size = 1, prob = P[current.state, ])
   location[t] <- current.state
}
```

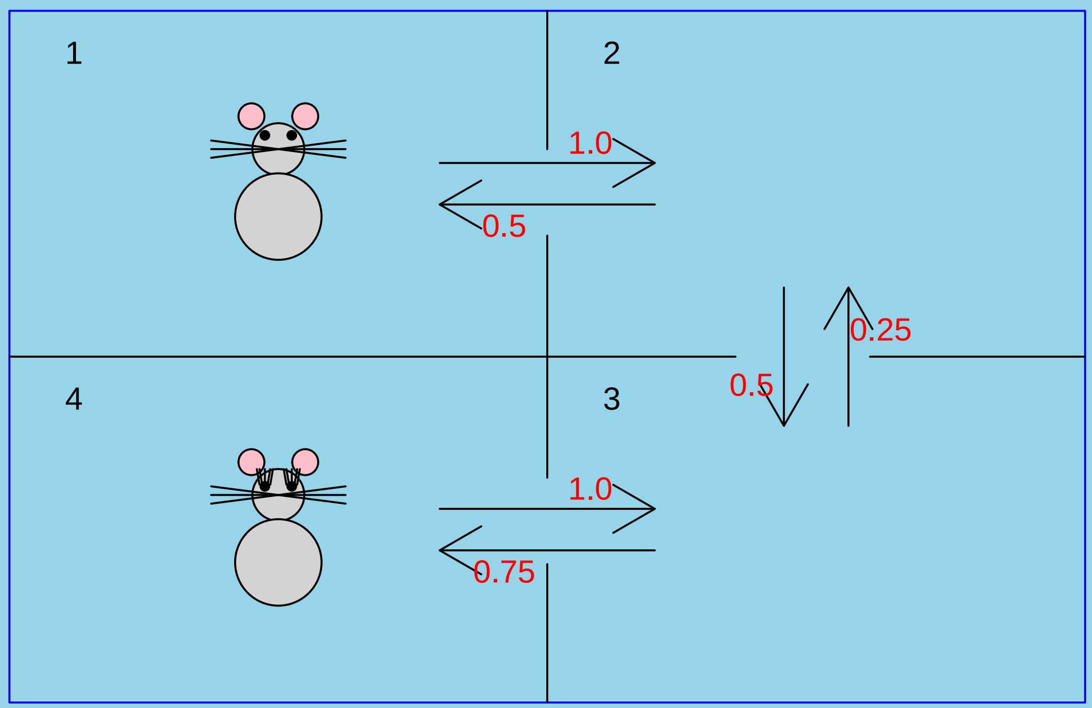
Simulating the Model in R

```
table(location) # this counts visits to each compartment

## location
##     1     2     3     4
## 16385 33104 33615 16896
```

A Mouse Movement Model - Including Some Structure

Transition Probabilities:



A Mouse Movement Model - Some Structure

Transition Matrix

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.25 & 0 & 0.75 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Simulating the More Structured Model in R

```
Ntransitions <- 100000 # number of mouse moves
P <- matrix(c(0, 1, 0, 0,
              0.5, 0, 0.5, 0,
              0, 0.25, 0, 0.75,
              0, 0, 1, 0), nrow = 4,
              byrow = TRUE)  # P is the transition matrix
location <- numeric(Ntransitions)#initializing the Markov cha
current.state <- 1 # initial compartment
for (t in 1:Ntransitions) {
   current.state <- sample(1:4,
        size = 1, prob = P[current.state, ])
   location[t] <- current.state
}
table(location) # the odor in compartment 4 is attractive
```

Simulating the More Structured Model in R

```
table(location)

## location
##     1     2     3     4
## 10215 20036 39785 29964

    # the odor in compartment 4 is attractive
```

Long Run Distribution

If we repeatedly simulate these Markov chains, we will find that the proportions of visits to each location follow specific distributions.

In the null model, this *steady state* distribution is $1/6, 1/3, 1/3, 1/6$.

In the model with the odor, the steady state distribution is different. (We will find out how to calculate it later.)

Definitions and Terminology

Definition: ***State Space*** **=** $S = \{1, 2, \ldots, m\}$**.**

e.g. For the mouse example, $S = \{1, 2, 3, 4\}$**.**

Definitions and Terminology

Definition: The sequence of random variables $X_1, X_2, X_3, \ldots,$ is called a Markov chain if

$$P(X_n = j_n | X_{n-1} = j_{n-1}, X_{n-2} = j_{n-2}, \ldots)$$

$$= P(X_n = j_n | X_{n-1} = j_{n-1})$$

where $j_n, j_{n-1}, j_{n-2}, \ldots$ are elements of S.

Definitions and Terminology

Example:

e.g. If the mouse starts in compartment 1, and enters compartment 2, then compartment 3, back to 2, back to 3, then 4, and back to 3.

$X_0 = 1, X_1 = 2, X_2 = 3, X_3 = 2, X_4 = 3, X_5 = 4, X_6 = 3, \ldots$

Definitions and Terminology

Define a matrix P with (i, j)th entry

$$p_{ij} == P(X_n = j | X_{n-1} = i)$$

p_{ij} is called the transition probability from state i to state j. P is called a transition matrix. (We are assuming that p_{ij} does not depend on n.)

Definitions and Terminology

All rows of P sum to one. That is,

$$P \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

A 3×3 Example

$$P_{33} = \begin{bmatrix} 0 & 0.4 & 0.6 \\ 0.5 & 0 & 0.5 \\ 0.25 & 0 & 0.75 \end{bmatrix}$$

```r
P33 <- matrix(c(0, 0.4, 0.6,
                0.5, 0, 0.5,
                0.25, 0, 0.75), nrow = 3,
                byrow = TRUE)
```

Probability Calculations

Theorem 0:

$P(X_n = j | X_0 = i)$ **is the** (i, j)**th element of the matrix** P^n**.**

Mouse Odor Model Example

```
P2 <- P%*%P
P2

##       [,1]  [,2]  [,3]  [,4]
## [1,] 0.500 0.000 0.500 0.000
## [2,] 0.000 0.625 0.000 0.375
## [3,] 0.125 0.000 0.875 0.000
## [4,] 0.000 0.250 0.000 0.750
```

For example, the probability of returning to compartment 1 after 2 transitions is 0.5. The probability of reaching compartment 4 in 2 transitions, if starting in compartment 2, is 0.375.

Probability Calculations

Define

$$x^{\{n\}} = [P(X_n = 1)\ P(X_n = 2)\ \ldots\ P(X_n = m)]$$

This vector is called the nth state vector of the Markov chain.

It specifies the probability distribution of X_n.

The sum of the entries of $x^{\{n\}}$ must always be one. $x^{\{0\}}$ denotes the distribution of the initial state X_0.

Probability Calculations

For the mouse example, if the mouse starts in compartment 1, $x^{\{0\}} = [1, 0, 0, 0]$.

If the mouse starts in a randomly selected compartment, $x^{\{0\}} = [.25, .25, .25, .25]$.

Probability Calculations

Theorem 1:

$$x^{\{n\}} = x^{\{0\}} P^n$$

Example - $x^{\{0\}} P_{33}$ and $x^{\{0\}} P_{33}^2$ with random starting point

```
x0 <- rep(1/3, 3)   # random starting point
x1 <- x0%*%P33  # distribution after 1 transition
x2 <- x0%*%(P33%*%P33) # distribution after 2 transitions
```

Probability Calculations - Example (cont'd)

$x^{\{0\}}$, $x^{\{1\}}$ and $x^{\{2\}}$:

```
x0

## [1] 0.3333333 0.3333333 0.3333333

x1

##      [,1]      [,2]      [,3]
## [1,] 0.25 0.1333333 0.6166667

x2

##           [,1] [,2]      [,3]
## [1,] 0.2208333  0.1 0.6791667
```

Probability Distribution of X_n When n is Large

```
P2 <- P33%*%P33
P4 <- P2%*%P2 # 4th power of P
P8 <- P4%*%P4 # 8th power of P
P16 <- P8%*%P8 # 16th power of P
x16 <- x0%*%P16 # distribution after 16 transitions
x16

##           [,1]       [,2]      [,3]
## [1,] 0.2173913 0.08695657 0.6956522

x32 <- x16%*%P16 # distribution after 32 transitions
x32

##           [,1]       [,2]      [,3]
## [1,] 0.2173913 0.08695652 0.6956522
```

The distribution of X_n no longer seems to depend on n. We have found the *long run distribution* of this Markov chain.

Does Every Markov Chain Have a Long Run Distribution?

No.

But Markov chains with *regular* transition matrices do.

Definition: P **is said to be a regular if there exists some positive integer** n **such that all entries of** P^n **are greater than zero.**

Regular Transition Matrices

Example

```
P33%*%P33%*%P33 # 3rd power of P33

##           [,1]  [,2]     [,3]
## [1,] 0.162500 0.140 0.697500
## [2,] 0.268750 0.050 0.681250
## [3,] 0.228125 0.075 0.696875
```

P_{33} is a regular transition matrix. P for the mouse odor example is not regular. (So regularity is not always necessary for a long run distribution to exist.)

Long Run Distribution

Theorem 2: If P is a regular matrix, then there exists a vector q such that

$$\lim_{n \to \infty} P^n = \begin{bmatrix} \mathbf{q} \\ \vdots \\ \mathbf{q} \end{bmatrix}$$

Long Run Distribution

Example

```
P33power <- P33%*%P33
P33power <- P33power%*%P33power
P33power <- P33power%*%P33power
P33power <- P33power%*%P33power
P33power <- P33power%*%P33power
P33power

##              [,1]       [,2]      [,3]
## [1,] 0.2173913 0.08695652 0.6956522
## [2,] 0.2173913 0.08695652 0.6956522
## [3,] 0.2173913 0.08695652 0.6956522
```

$\mathbf{q} = [0.2173, 0.08695, 0.6956]$

Long Run Distribution Example

```
P16

##             [,1]       [,2]      [,3]
## [1,] 0.2173940 0.08695419 0.6956518
## [2,] 0.2173890 0.08695851 0.6956525
## [3,] 0.2173907 0.08695700 0.6956523

P16%*%P16%*%P16

##             [,1]       [,2]      [,3]
## [1,] 0.2173913 0.08695652 0.6956522
## [2,] 0.2173913 0.08695652 0.6956522
## [3,] 0.2173913 0.08695652 0.6956522
```

Long Run Distribution

Theorem 3: If P is a regular matrix, then there exists a unique vector q such that

$$\lim_{n \to \infty} x^{\{n\}} = \mathbf{q}$$

for any initial state vector $x^{\{0\}}$. The vector q specifies the long run distribution of the Markov chain.

Long Run Distribution

Theorem 4: If P is a regular matrix, then the long run distribution vector $\mathbf{q}$ is the unique solution to the equation

$$\mathbf{q} = \mathbf{q}P$$

whose entries sum to one. The solution of the above equation is the *steady state* vector.

Long Run Distribution

Example

```
q <- c(.2173913, .08695652, 0.6956522)
q%*%P33  # test to see if equality holds here

##           [,1]       [,2]      [,3]
## [1,] 0.2173913 0.08695652 0.6956522
```

Mouse Odor Example

This is a non-regular example, but

```
q <- c(.1, .2, .4, .3)
q%*%P  # test to see if equality holds here


##        [,1] [,2] [,3] [,4]
## [1,]   0.1  0.2  0.4  0.3
```

the steady state distribution still satisfies

$$\mathbf{q} = \mathbf{q}P.$$

Law of Large Numbers for Markov Chains

Theorem 5: If $X_1, X_2, \ldots$ is a finite state Markov chain with a regular transition matrix, then for any function $f(j)$ defined on the state space, and for any initial state X_0,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(X_k) = \mathbf{q} \begin{bmatrix} f(1) \\ \vdots \\ f(m) \end{bmatrix} = \sum_{j=1}^{m} q_j f(j) \quad \textbf{with probability 1}$$

Note that

$$E[f(X)] = \sum_{j=1}^{m} q_j f(j)$$

when X has a distribution given by q. (One implication of this is Markov Chain Monte Carlo simulation, i.e. MCMC).

Law of Large Numbers for Periodic Markov Chains

The preceding result also holds for periodic Markov chains as long as the matrix of transitions within each periodic class is regular.

Law of Large Numbers for Periodic Markov Chains

Example: $f(x) = x^2$. Check result for Mouse Odor example:

```
mean(location^2) # location contains a simulated chain

## [1] 9.27848

(1:4)^2%*%c(.1, .2, .4, .3)  #  expected value of f(X^2)

##      [,1]
## [1,]  9.3
```

The two results match.

Classification of Markov Chain States

The following discussion leads to a simple way of determining whether a transition matrix is regular or not.

Definition. State i *leads* to state j if there exists $n \geq 1$ such that $P_{ij}^{(n)} > 0$.

e.g. Mouse maze: State 1 leads to state 2; State 2 leads to state 1, etc.

Classification of Markov Chain States

Definition. States i and j are said to *communicate* if i leads to j and j leads to i.

e.g. States 1 and 2 communicate.

Classification of Markov Chain States

Proposition. The ‘leads to’ relation is transitive. That is, if i leads to j and j leads to k, then i leads to k.

Mouse Maze e.g.: State 1 leads to State 2, State 2 leads to State 3, so State 1 leads to State 3. Similarly, State 3 leads to State 1. Therefore, States 1 and 3 communicate. Similarly, State 1 and State 4 communicate.

Classification of Markov Chain States

Definition. A *class* of states is defined as a subset of S in which any two members communicate.

Mouse Maze e.g.: All states communicate. They form a class.

Classification of Markov Chain States

Example:

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

States 1 and 2 communicate, but State 3 does not communicate with States 1 and 2. $\{1,2\}$ **is one class and** $\{3\}$ **is another class.**

Classification of Markov Chain States

Definition. S **is said to be** *irreducible* **if it is a class. That is, if all states in** S **communicate,** S **is said to be irreducible.**

The mouse maze state space is irreducible, but the state space for P **is not irreducible.**

Classification of Markov Chain States

Definition. **For any $i \in S$, the** ***period*** **of state i is defined to be the greatest common divisor of the set**

$$\{n > 0 : P_{ii}^{(n)} > 0\}$$

Proposition. **If i and j communicate, then the periods of i and j are the same.**

Classification of Markov Chain States

Definition. If the state space of a Markov chain is irreducible, then the period of the Markov chain is defined to be the common period of each state.

Mouse Maze e.g.: Period is 2, since if the chain starts in State 1 it can never return to State 1 in an odd number of transitions, and all states communicate.

Classification of Markov Chain States

Definition. If the period of a Markov chain is 1, the Markov chain is said to be aperiodic.

P_{33} **is aperiodic.**

Classification of Markov Chain States

Theorem. An aperiodic irreducible Markov chain with a finite state space must have a regular transition matrix.

P_{33} **is regular.** P **for the mouse maze example is not regular.**

Calculation of Steady State Vector

The steady state vector q solves

$$\mathbf{q} = \mathbf{q}P$$

which can be re-written as

$$(P^\top - I)\pi = 0$$

where $\pi = \mathbf{q}^\top$.

Calculation of Steady State Vector

Example:

```
P33 <- matrix(c(0, 0.4, 0.6,
              0.5, 0, 0.5,
              0.25, 0, 0.75), nrow = 3,
              byrow = TRUE)
A <- t(P33) - diag(rep(1,3)) # P^T - I
solve(A, rep(0,3))   # solve A pi = 0

## Error in solve.default(A, rep(0, 3)):  Lapack routine
dgesv:  system is exactly singular:  U[3,3] = 0
```

trouble! We either have too many solutions or no solutions to this problem. $P^\top - I$ is singular, so there are too many solutions. We need more equations.

Calculation of Steady State Vector

Since the steady state vector is a (discrete) probability distribution, its elements must sum to 1:

$\pi_1 + \pi_2 + \pi_3 = 1,$ **so we include this equation:**

```
A <- rbind(A, rep(1,3))
RHS <- c(rep(0,3), 1)
qr.solve(A, RHS)  # no longer a square system

## [1] 0.21739130 0.08695652 0.69565217
```

Solution of Linear Systems via QR

Every matrix A has a QR decomposition.

Suppose A is $n \times m$ with $n > m$ (as in our case, where $n = m + 1$).

$$A = QR$$

where Q is an $n \times n$ orthogonal matrix, i.e. $Q^\top Q = I$, and R is an $n \times m$ upper triangular matrix. Among other things, this means that all entries of R's bottom $n - m$ rows are 0's.

Solution of Linear Systems via QR (`qr.solve()`)

Now, solve

$$
\begin{aligned}
Ax &= y \\
\rightsquigarrow QRx &= y \\
\rightsquigarrow Q^\top QRx &= Q^\top y \\
\rightsquigarrow Rx &= Q^\top y
\end{aligned}
$$

which can be solved for x using backward substitution, starting at row m.

If the last $n - m$ elements of $Q^\top y$ are 0, then the system has a solution. If not, the system has no exact solution, but the result is the least-squares estimate.

Mouse Odor Example

Recall the mouse odor model:

```
P <- matrix(c(0, 1, 0, 0,
              0.5, 0, 0.5, 0,
              0, 0.25, 0, 0.75,
              0, 0, 1, 0), nrow = 4,
              byrow = TRUE)  # P is the transition matrix
A <- t(P) - diag(rep(1, 4))  # P^T - I
A <- rbind(A, rep(1,4))  # additional row
RHS <- c(rep(0,4), 1)
qr.solve(A, RHS)

## [1] 0.1 0.2 0.4 0.3
```

An Industrial Example

At the beginning of each day, a batch of containers arrives at a stockyard having capacity to store 6 containers.

The batch size has the discrete probability distribution $\{q_0 = .4, q_1 = 0.3, q_2 = 0.2, q_3 = 0.1\}$**.**

If the stockyard does not have sufficient space to store the whole batch, the batch as a whole is taken elsewhere.

Each day, as long as there are containers in the stockyard, exactly one container is removed from the stockyard.

An Industrial Example

1. Find the transition matrix for the Markov chain $\{X_1, X_2, \ldots\}$, where $X_t =$ the number of containers in the stockyard at the beginning of the tth day.

2. Find the long run distribution for this Markov chain.

3. Suppose a profit of \$100 is realized for each container that spends a night at the stockyard. Calculate the long-run average daily profit.

An Industrial Example - Transition Matrix

$X_t =$ **the number of containers at the start of day** t**.**

The state space is

$$S = \{0, 1, 2, 3, 4, 5\}.$$

Note that the stockyard could be empty at the beginning of a day, and because it can only hold 6 containers, it could never end a day with more than 5 containers, since 1 is always taken away.

An Industrial Example

Transition Matrix

$$P = \begin{bmatrix} 0.7 & 0.2 & 0.1 & 0 & 0 & 0 \\ 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 \\ 0 & 0.4 & 0.3 & 0.2 & 0.1 & 0 \\ 0 & 0 & 0.4 & 0.3 & 0.2 & 0.1 \\ 0 & 0 & 0 & 0.5 & 0.3 & 0.2 \\ 0 & 0 & 0 & 0 & 0.7 & 0.3 \end{bmatrix}$$

An Industrial Example

Equilibrium Distribution, i.e. Steady-State Distribution

```r
P <- matrix(c(0.7, .2, .1, 0, 0, 0, 0.4, 0.3, 0.2, 0.1,
0, 0, 0, 0.4, 0.3, 0.2, 0.1, 0, 0, 0, 0.4, 0.3, 0.2, 0.1,
0, 0, 0, 0.5, 0.3, 0.2,  0, 0, 0, 0, 0.7, 0.3),
   nrow=6, byrow = TRUE)
A <- t(P) - diag(rep(1, 6)) # P^T - I
A <- rbind(A, rep(1,6))
RHS <- c(rep(0,6), 1)
options(digits=4)
pi <- qr.solve(A, RHS)
pi


## [1] 0.23273 0.17455 0.18909 0.18545 0.14909
## [6] 0.06909
```

An Industrial Example - Expected Long-Run Daily Profit

$$\textbf{Profit} = 100 \times X$$

where X is the number of containers in the yard at the beginning of a day.

$$E[X] = \sum_{i=0}^{5} i\pi_i$$

```
sum(pi*(0:5))

## [1] 2.051
```

so

$$E[100X] = 205.10$$