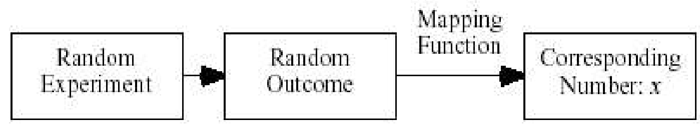
MARKOV CHAINS

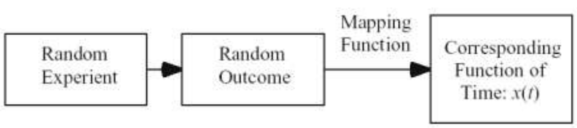
Random Variables



We mentioned before that the Random Variable is to assign a numerical value to the outcome of a random experiments.

Random Process

- A **random process** assigns a random **function of time** as the outcome of a random experiment.



- The sequence of events leading to assigning a time function X(t) to the outcome of a random experiment.

Random Process

- A random process X(t) is described by:
  - The sample space S which includes all possible outcomes of a random experiment.
  - The **sample function X(t)** which is the time function associated with an outcome s. The values of the sample function could be discrete or continuous.
  - The **ensemble** which is the set of all possible time functions produced by the random experiment.

Random Process

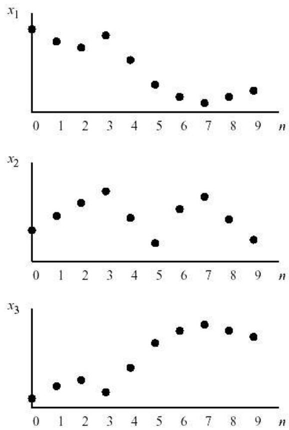
- Time could be continuous (t) or discrete (n).
- The function value could be continuous or discrete.

Random Process

- We have four possible types of random processes:
  1. Discrete time, discrete value
  2. Discrete time, continuous value
  3. Continuous time, discrete value
  4. Continuous time, continuous value

Random Process

- An example of a discrete time, discrete value random process for an observation of 10 samples where only three random functions are possible.



Markov process

- A Markov process is a random processes where the value of the random variable at instant n depends **only** on its immediate past value at instant n-1(**memoryless property**).
- In a Markov process the random variable represents the **state** of the system at a given instant n.
- How to understand it?

Markov Process

- We see examples of Markov processes in many real life situations:
  - Stock price
  - Exchange rate fluctuation
  - Customer arrivals and departures at banks
  - Checkout counters at supermarkets
  - Random walk such as Brownian motion
  - Gambler's wealth

Discrete-Time Markov Process

- In a discrete-time Markov process the hold time assumes discrete values. As a result, changes in the states occur over a finite or countable time values.

Markov Chains

- If the **state space** of a Markov process is **discrete**, the Markov process is called a **Markov chain**.
- The states are labeled by the integers 0, 1, 2, and so on.

Memoryless Property of Markov Chains

- In a discrete-time Markov chain, the value of the random variable s(n) is a function of its immediate past value, i.e., s(n) depends on s(n-1).
- This is referred to as the **Markov property** or **memoryless property** of the Markov chain.
- Example: gambler’s wealth

Markov Chains

The probability that the Markov chain is in state s_i at time n is a function of its past state s_j at time n-1 only

$p[S(n)=s_i] = f(S(n-1) = s_j)$

Transition Probability

- Assume at time n-1 we are in state j and at time n we are in state i, transition probability from j to i

$$p_{ij}(n) = p\left[S(n) = i \mid S(n-1) = j\right]$$

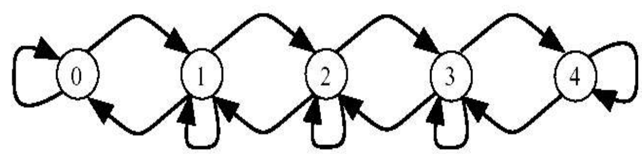
- For **Homogeneous** Markov chains, the transition probabilities do not depend on n:

$$p_{ij} = p\left[S(n) = i \mid S(n-1) = j\right]$$

Example

Consider a data buffer in a certain communication device such as a network router. Assume the buffer could accommodate at most four packets. We say the buffer size is B=4. Identify the states of this buffer and show the possible transitions between states assuming at any time step at most one packet can arrive or leave the buffer. Finally explain why the buffer could be studied using Markov chain analysis.

Solution:



The occupancy states of a buffer of size four and the possible transitions between the states.

Transition Probability

- Define the probability of finding our system in state i at the n-th step as

$$s_i(n) = p\left[X(n) = i\right]$$

Transition Probability

$s_i(n)$ can be written as

$s_i(n) = \sum_j \; p_{ij} \, s_j(n-1)$

Express this in matrix form

$s(n) = \mathbf{P}s(n-1)$

Transition Matrix

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1,m} \\ p_{21} & p_{22} & \cdots & p_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m,1} & p_{m,2} & \cdots & p_{m,m} \end{bmatrix}$$

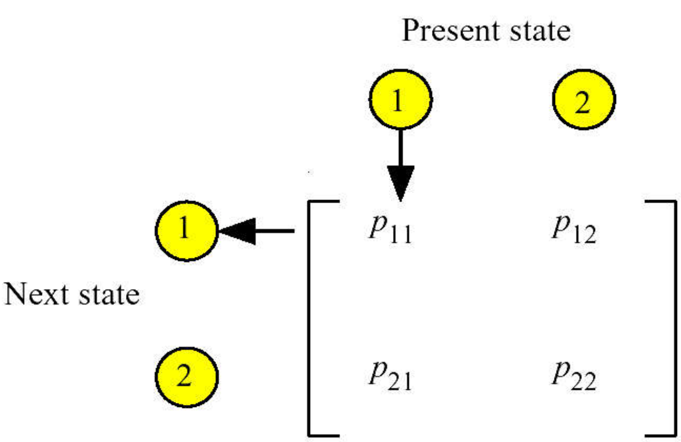
$$\sum_{i=1}^{m} p_{ij} = 1$$

Distribution Vector

$$\mathbf{s}(n) = \begin{bmatrix} s_1(n) & s_2(n) & \cdots & s_m(n) \end{bmatrix}^t$$

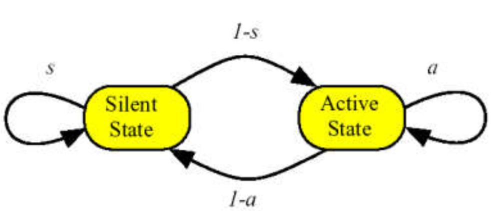
$$\sum_{i=1}^{m} s_i(n) = 1 \qquad n = 0, 1, 2, \cdots$$

Transition Matrix



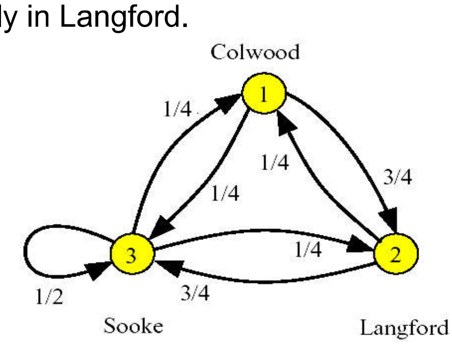
Example: An on-off source is often used in telecommunications to simulate voice traffic. Such a source has two states: the silent state s_1 where the source does not send any data packets and the active state s_2 where the source sends one packet per time step. If the source were in s_1 it has a probability *s* of staying in that state for one more time step. When it is in state s_2, it has a probability *a* of staying in that state. Obtain the transition matrix for describing the source.

- Solution:



$$\mathbf{P} = \begin{bmatrix} s & 1-a \\ 1-s & a \end{bmatrix}$$

- **Example:** Assume that the probability that a delivery truck moves between three cities at the start of each day is shown in figure below. Write down the transition matrix and the initial distribution vector assuming that the truck was initially in Langford.

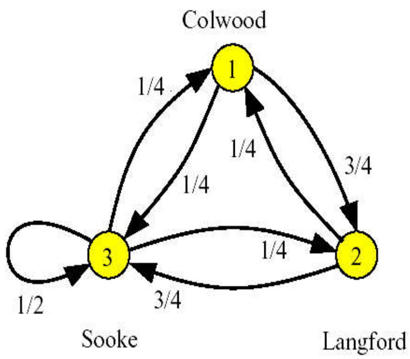


Solution:

City	State index
Colwood	1
Langford	2
Sooke	3

$$\mathbf{P} = \begin{bmatrix} 0 & 1/4 & 1/4 \\ 3/4 & 0 & 1/4 \\ 1/4 & 3/4 & 1/2 \end{bmatrix}$$

$$\mathbf{s}(0) = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^t$$

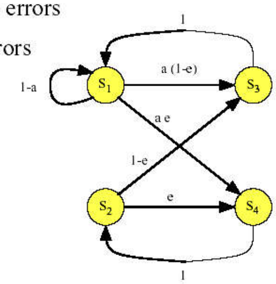


- Example: Assume an on-off data source that generates equal length packets with probability a per time step. The channel introduces errors in the transmitted packets such that the probability of a packet is received in error is e. Draw a Markov chain state transition diagram and write the equivalent state transition matrix.

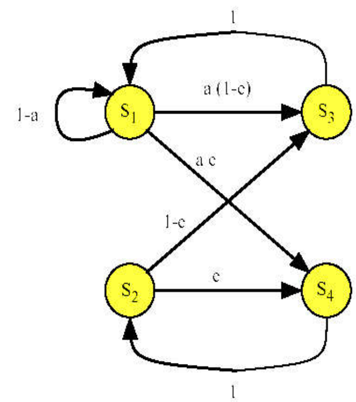
Solution:

The source is assumed to be in one of four states based on packet availability and feedback from receiver.

State	Significance
1	Source is idle
2	Source is retransmitting a frame
3	Source transmitted a frame with no errors
4	Source transmitted a frame with errors



$$\mathbf{P} = \begin{bmatrix} 1-a & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ a(1-e) & 1-e & 0 & 0 \\ a\,e & e & 0 & 0 \end{bmatrix}$$

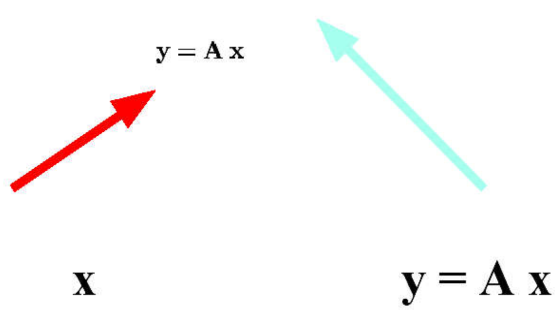


State Transition Matrix

1. $0 \le p_{ij} \le 1$ for all values of i and j. When all the elements of $\mathbf{P}$ are nonnegative, we have a **nonnegative matrix**.
2. The sum of each column is exactly 1 (i.e. $\sum_{j=1}^{m} p_{ji} = 1$).
3. The magnitude of all eigenvalues obey the condition $|\lambda_i| \le 1$.

Eigenvalue of a matrix

We can think of a matrix as a **transformation** that rotates and scales any vector it multipliers



Eigenvalue of a matrix

A vector $\mathbf{x}$ is an **eigenvector** of the matrix $\mathbf{A}$ when the vector is not rotated, just scaled by the **eigenvalue** λ

$$\mathbf{y} = \mathbf{A}\,\mathbf{x}$$
$$= \lambda\mathbf{x}$$

$\mathbf{x}$

$\mathbf{y} = \mathbf{A}\,\mathbf{x}$

$= \lambda\,\mathbf{x}$

Eigenvalue & Eigenvector

$$\mathbf{A}\mathbf{x} \;=\; \lambda\mathbf{x}, \quad where\ \mathbf{x} \neq \mathbf{0}$$

$$det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = \mathbf{0}$$

$$(\mathbf{A} - \lambda_{\mathbf{i}}\mathbf{I})\mathbf{x} = \mathbf{0}$$

Theorems

Let $\mathbf{P}$ be any $m \times m$ column stochastic matrix. Then $\mathbf{P}$ has 1 as an eigenvalue.

The sum of columns of any matrix $\mathbf{A}$ will not change when it is premultiplied by a column stochastic matrix $\mathbf{P}$.

Constructing **P**

1. All possible states of the system are identified and labeled. The labeling of the states is arbitrary although some labeling schemes would render the transition matrix easier to visualize.
2. All possible transitions between the states are drawn.
3. The probability of every transition in the state diagram is written down.

Constructing **P**

4. The transition matrix is constructed.
5. Relabeling of the states is always possible. That will change the locations of the matrix elements and make the structure of the matrix more visible. This rearrangement will still produce a column stochastic matrix and will not disturb its eigenvalues or the **directions** of its eigenvectors.

Transient Analysis of Markov Chains

$$\begin{aligned} \mathbf{s}(n) &= \mathbf{P}\mathbf{s}(n-1) \\ &= \mathbf{P}\mathbf{P}\mathbf{s}(n-2) \\ &= \cdots \\ &= \mathbf{P}^n\mathbf{s}(0) \end{aligned}$$

Given **P** and the initial conditions **s**(0), we can find **s**(n).

Example:

A computer memory system is composed of very fast on-chip cache, fast on-board RAM, and slow hard disk. When the computer is accessing a block from each memory system, the next block required could come form any of the three available memory systems. This is modeled as a Markov chain with the state of the system representing the memory from which the current block came from: state s_1 corresponds to the cache, state s_2 corresponds to the RAM, and state s_3 corresponds to the hard disk. The transition matrix is given by

$$\mathbf{P} = \begin{bmatrix} 0.7 & 0.1 & 0 \\ 0.2 & 0.7 & 0.1 \\ 0.1 & 0.2 & 0.9 \end{bmatrix}$$

Find the probability that after three consecutive block accesses the system will read a block from the cache. Assume the initial state is s_1.

Solution:

Starting distribution vector is $\mathbf{s}(0) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^t$ and we have

$$\mathbf{P}^2 = \begin{bmatrix} 0.51 & 0.14 & 0.01 \\ 0.29 & 0.53 & 0.16 \\ 0.20 & 0.33 & 0.83 \end{bmatrix}$$

$$\mathbf{P}^3 = \begin{bmatrix} 0.386 & 0.151 & 0.023 \\ 0.325 & 0.432 & 0.197 \\ 0.289 & 0.417 & 0.780 \end{bmatrix}$$

Solution(cont'd):

The distribution vector after three iterations is

$$\mathbf{s}(3) = \mathbf{P}^3\, \mathbf{s}(0) = \begin{bmatrix} 0.386 & 0.325 & 0.289 \end{bmatrix}^t$$

The probability that the system will read a block from the cache is 0.386.

Property of $\mathbf{P}^n$

- $\mathbf{P}^n$ remains a column stochastic matrix.
- A nonzero element in $\mathbf{P}$ can increase or decrease in $\mathbf{P}^n$ but can never become zero.
- A zero element in $\mathbf{P}$ could remain zero or increase in $\mathbf{P}^n$ but can never become negative.

Find **s**(n)

1. Repeated multiplication of **P**
2. Expanding the initial distribution vector **s**(0)
3. Diagonalizing the matrix **P**
4. Using the Jordan canonic form of **P**
5. Using the z-transform

Finding **s**(n) by Expanding **s**(0)

$$\mathbf{s}(0) = c_1\, \mathbf{x}_1 + c_2\, \mathbf{x}_2 + \cdots + c_m\, \mathbf{x}_m$$

In matrix form we can write:

$$\mathbf{s}(0) = \mathbf{X}\,\mathbf{c}$$

where

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_m \end{bmatrix}$$

$$\mathbf{c} = \begin{bmatrix} c_1 & c_2 & \cdots & c_m \end{bmatrix}^t$$

$\mathbf{x}_i$ is the ith eigenvector of $\mathbf{P}$

We find $\mathbf{X}$ using MATLAB

```
X = eig(P)
```

We solve for $\mathbf{c}$ using MATLAB also

```
c = X\s
```


Example

A Markov chain has the state matrix

$$\mathbf{P} = \begin{bmatrix} 0.3 & 0.2 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.4 & 0.4 \end{bmatrix}$$

Check to see if this matrix has distinct eigenvalues. If so, expand the initial distribution vector

$$\mathbf{s}(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^t$$

in terms of the eigenvectors of $\mathbf{P}$.

The eigenvalues of $\mathbf{P}$ are $\lambda_1 = 1$, $\lambda_2 = 0.1$, $\lambda_3 = 0.2$, and $\lambda_4 = 0$. The eigenvectors corresponding to these eigenvalues are

$$\mathbf{X} = \begin{bmatrix} 0.439 & -0.707 & 0.0 & 0.0 \\ 0.548 & 0.707 & 0.707 & 0.0 \\ 0.395 & 0.0 & 0.0 & 0.707 \\ 0.592 & 0.0 & -0.707 & -0.707 \end{bmatrix}$$

Therefore, we can expand $\mathbf{s}(0)$ in terms of the corresponding eigenvectors.

$$\mathbf{s}(0) = \mathbf{X}\,\mathbf{c}$$

where $\mathbf{c}$ given by

$$\mathbf{c} = \begin{bmatrix} 0.507 & -1.1 & 0.707 & 0.283 \end{bmatrix}^t$$

Finding **s**(n)

$$\begin{aligned} \mathbf{s}(1) &= \mathbf{P}\,\mathbf{s}(0) \\ &= \mathbf{P}\,(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_m\mathbf{x}_m) \\ &= c_1\lambda_1\mathbf{x}_1 + c_2\lambda_2\mathbf{x}_2 + \cdots + c_m\lambda_m\mathbf{x}_m \end{aligned}$$

We can write this in matrix form:

$$\mathbf{s}(1) = \mathbf{XDc}$$

Finding **s**(n)

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{bmatrix}$$

MATLAB helps us find $\mathbf{D}$

```
[X,D] = eig(P)
```

Finding **s**(n)

Now we can find $\mathbf{s}(2)$

$$\begin{aligned}\mathbf{s}(2) &= \mathbf{P}\,\mathbf{s}(1)\\ &= \mathbf{P}\,\mathbf{X}\,\mathbf{D}\,\mathbf{c}\end{aligned}$$

But we have

$$\mathbf{P}\mathbf{X} = \mathbf{X}\,\mathbf{D}$$

Therefore we have $\mathbf{s}(2)$:

$$\mathbf{s}(2) = \mathbf{X}\,\mathbf{D}^2\,\mathbf{c}$$

Finding **s**(n)

$$\mathbf{D}^2 = \begin{bmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m^2 \end{bmatrix}$$

Finding **s**(n)

In general $\mathbf{s}(n)$ is given by

$$\mathbf{s}(n) = \mathbf{X}\mathbf{D}^n\mathbf{c}$$

where

$$\mathbf{D}^n = \begin{bmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m^n \end{bmatrix}$$

Example

A Markov chain has the state matrix

$$\mathbf{P} = \begin{bmatrix} 0.3 & 0.2 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.4 & 0.4 \end{bmatrix}$$

Find the values of the distribution vector at time steps 2, 5, and 20.

Assume that the initial distribution is $[1 \quad 0 \quad 0 \quad 0]^t$

Solution:

For the given transition matrix, we can write

$$\mathbf{s}(n) = \mathbf{X}\mathbf{D}^n\mathbf{c}$$

where the matrices $\mathbf{X}$, $\mathbf{D}$, and $\mathbf{c}$ are given by

$$\begin{aligned}\mathbf{X} &= \begin{bmatrix} 0.439 & -0.707 & 0.0 & 0.0 \\ 0.548 & 0.707 & 0.707 & 0.0 \\ 0.395 & 0.0 & 0.0 & 0.707 \\ 0.592 & 0.0 & -0.707 & -0.707 \end{bmatrix}\\ \mathbf{D} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\\ \mathbf{c} &= \begin{bmatrix} 0.507 & -1.1 & 0.707 & 0.283 \end{bmatrix}^t\end{aligned}$$

Thus we can simply write

$$\begin{aligned} \mathbf{s}(2) &= \mathbf{X}\mathbf{D}^2\mathbf{c} \\ &= \begin{bmatrix} 0.23 & 0.29 & 0.2 & 0.28 \end{bmatrix}^t \end{aligned}$$

$$\begin{aligned} \mathbf{s}(5) &= \mathbf{X}\mathbf{D}^5\mathbf{c} \\ &= \begin{bmatrix} 0.22 & 0.28 & 0.2 & 0.3 \end{bmatrix}^t \end{aligned}$$

$$\begin{aligned} \mathbf{s}(20) &= \mathbf{X}\mathbf{D}^{20}\mathbf{c} \\ &= \begin{bmatrix} 0.22 & 0.28 & 0.2 & 0.3 \end{bmatrix}^t \end{aligned}$$

Example:
A Markov chain has the transition matrix

$$\mathbf{P} = \begin{bmatrix} 0.1 & 0.4 & 0.2 \\ 0.1 & 0.4 & 0.6 \\ 0.8 & 0.2 & 0.2 \end{bmatrix}$$

Check to see if this matrix has distinct eigenvalues. If so,
(a) Expand the initial distribution vector

$$\mathbf{s}(0) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^t$$

in terms of the eigenvectors of $\mathbf{P}$.
(b) Find the value of $\mathbf{s}(3)$.

Solution:

The eigenvalues of $\mathbf{P}$ are

$$\begin{aligned} \lambda_1 &= 1 \\ \lambda_2 &= -0.15 + j0.3122 \\ \lambda_3 &= -0.15 - j0.3122 \end{aligned}$$

We see that complex eigenvalues appear as complex conjugate pairs.

The eigenvectors corresponding to these eigenvalues are

$$\mathbf{X} = \begin{bmatrix} -0.4234 & -0.1698 + j0.3536 & -0.1698 - j0.3536 \\ -0.6726 & -0.5095 - j0.3536 & -0.5095 + j0.3536 \\ -0.6005 & 0.6794 & 0.6794 \end{bmatrix}$$

We see that the eigenvectors corresponding to the complex eigenvalues appear also as complex conjugate pairs.

Express $\mathbf{s}(0)$ in terms of the corresponding eigenvectors.

$$\mathbf{s}(0) = \mathbf{X}\,\mathbf{c}$$

where $\mathbf{c}$ given by

$$\mathbf{c} = \begin{bmatrix} -5863 & -0.2591 - j0.9312 & -0.2591 + j0.9312 \end{bmatrix}^t$$

The distribution vector at time step $n = 3$ is

$$\mathbf{s}(3) = \mathbf{X}\mathbf{D}^3\mathbf{c} = \begin{bmatrix} 0.2850 & 0.3890 & 0.3260 \end{bmatrix}^t$$

Finding **s**(n) by diagonalizing **P**

Matrix $\mathbf{P}$ is diagonalizable when it has m distinct eigenvalues and we can write

$$\mathbf{P} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1}$$

$\mathbf{X}$ is the matrix whose columns are the eigenvectors of $\mathbf{P}$ and $\mathbf{D}$ is a diagonal matrix whose diagonal elements are the eigenvalues arranged according to the ordering of the columns of $\mathbf{X}$.

MATLAB provides a very handy function for fi nding the matrices $\mathbf{X}$ and $\mathbf{D}$ using the single command

```
[X,D] = eig(P)
```

Finding **s**(n) by diagonalizing **P**

We calculate $\mathbf{P}^2$ as

$$\begin{aligned}\mathbf{P}^2 &= \left(\mathbf{XDX}^{-1}\right) \times \left(\mathbf{XDX}^{-1}\right) \\ &= \mathbf{XD}^2\mathbf{X}^{-1}\end{aligned}$$

Finding **s**(n) by diagonalizing **P**

In general we have

$$\mathbf{P}^n = \mathbf{XD}^n\mathbf{X}^{-1}$$

where

$$\mathbf{D}^n = \begin{bmatrix} \lambda_1^n & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2^n & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3^n & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_m^n \end{bmatrix}$$

Finding **s**(n) by diagonalizing **P**

It is easy therefore to find $\mathbf{s}(n)$ at any value for n by simply finding $\mathbf{D}^n$ and then evaluate the simple matrix multiplication expression

$$\begin{aligned}\mathbf{s}(n) &= \mathbf{P}^n\,\mathbf{s}(0) \\ &= \mathbf{XD}^n\mathbf{X}^{-1}\,\mathbf{s}(0)\end{aligned}$$

Expanding $\mathbf{P}^n$ in terms of its Eigenvalues

Expanding $\mathbf{P}^n$ in terms of the eigenvalues of $\mathbf{P}$ will give us additional insights into the transient behavior of Markov chains.

We have $\mathbf{P}^n = \mathbf{XD}^n\mathbf{X}^{-1}$

Expanding $\mathbf{P}^n$ in terms of its Eigenvalues

Thus we can express $\mathbf{P}^n$ in the above equation in the form

$$\mathbf{P}^n = \mathbf{A}_1 + \lambda_2^n\mathbf{A}_2 + \cdots + \lambda_m^n\mathbf{A}_m$$

Where we assumed that $\lambda_1 = 1$ and all other eigenvalues have magnitudes lesser than unity (fractions) because $\mathbf{P}$ is column stochastic. This shows that as time progresses, n becomes large and the powers of λ_i^n will quickly decrease. The main contribution to $\mathbf{P}^n$ will be due to $\mathbf{A}_1$ only.

Expanding $\mathbf{P}^n$ in terms of its Eigenvalues

The matrices $\mathbf{A}_i$ can be determined from the product

$$\mathbf{A}_i = \mathbf{X}\,\mathbf{Y}_i\,\mathbf{X}^{-1}$$

where $\mathbf{Y}_i$ is the **selection matrix** which has zeros everywhere except for element $y_{ii} = 1$.

Expanding $\mathbf{P}^n$ in terms of its Eigenvalues

For a $\mathbf{3} \times \mathbf{3}$ matrix, we can write

$$\mathbf{P}^n = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix} \begin{bmatrix} \lambda_1^n & 0 & 0 \\ 0 & \lambda_2^n & 0 \\ 0 & 0 & \lambda_3^n \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix}^{-1}$$

Expanding $\mathbf{P}^n$ in terms of its Eigenvalues

$$\mathbf{A}_1 = \mathbf{X} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{X}^{-1}$$

$$\mathbf{A}_2 = \mathbf{X} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{X}^{-1}$$

$$\mathbf{A}_3 = \mathbf{X} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{X}^{-1}$$

Example

The following is a diagonalizable state matrix.

$$\mathbf{P} = \begin{bmatrix} 0.1 & 0.3 & 1 \\ 0.2 & 0.3 & 0 \\ 0.7 & 0.4 & 0 \end{bmatrix}$$

Express $\mathbf{P}^n$ in terms of its eigenvalues.

Solution:

First thing is to check that $\mathbf{P}$ is diagonalizable by checking that it has three distinct eigenvalues. Using the MATLAB function `select` that we developed, we find that

$$\mathbf{P}^n = \mathbf{A}_1 + \lambda_2^n \mathbf{A}_2 + \lambda_3^n \mathbf{A}_3$$

where $\lambda_1 = 1$.

$$\mathbf{A}_1 = \begin{bmatrix} 0.476 & 0.476 & 0.476 \\ 0.136 & 0.136 & 0.136 \\ 0.388 & 0.388 & 0.388 \end{bmatrix}$$

$$\mathbf{A}_2 = \begin{bmatrix} 0.495 & 0.102 & -0.644 \\ -0.093 & -0.019 & 0.121 \\ -0.403 & -0.083 & 0.524 \end{bmatrix}$$

$$\mathbf{A}_3 = \begin{bmatrix} 0.028 & -0.578 & 0.168 \\ -0.043 & 0.883 & -0.257 \\ 0.015 & -0.305 & 0.089 \end{bmatrix}$$

Notice that matrix $\mathbf{A}_1$ is column stochastic and all its columns are equal. Notice also that the sum of columns for matrices $\mathbf{A}_2$ and $\mathbf{A}_3$ is zero.

Example

A Markov chain has the transition matrix

$$\mathbf{P} = \begin{bmatrix} 0.2 & 0.4 & 0.2 \\ 0.1 & 0.4 & 0.6 \\ 0.7 & 0.2 & 0.2 \end{bmatrix}$$

Check to see if this matrix has distinct eigenvalues. If so:

(a) Expand the transition matrix $\mathbf{P}^n$ in terms of its eigenvalues.

(b) Find the value of $\mathbf{s}(3)$.

Assume that the initial distribution is $[1 \ \ 0 \ \ 0]^t$.

Solution:

The eigenvalues of $\mathbf{P}$ are

$\lambda_1 = 1$,

$\lambda_2 = -0.1 + j0.3$, and

$\lambda_3 = -0.1 - j0.3$.

We see that complex eigenvalues appear as complex conjugate pairs.

The eigenvectors corresponding to these eigenvalues are

$$\mathbf{X} = \begin{bmatrix} 0.476 & -0.39 - j0.122 & -0.39 + j0.122 \\ 0.660 & 0.378 - j0.523 & 0.378 + j0.523 \\ 0.581 & 0.011 + j0.645 & 0.011 - j0.645 \end{bmatrix}$$

We see that the eigenvectors corresponding to the complex eigenvalues appear also as complex conjugate pairs.

Using the function `select(P)`, we express $\mathbf{P}^n$ according to

$$\mathbf{P}^n = \mathbf{A}_1 + \lambda_2^n \mathbf{A}_2 + \lambda_3^n \mathbf{A}_3$$

$$\begin{aligned} \mathbf{A}_1 &= \begin{bmatrix} 0.277 & 0.277 & 0.277 \\ 0.385 & 0.385 & 0.385 \\ 0.339 & 0.339 & 0.339 \end{bmatrix} \\ \mathbf{A}_2 &= \begin{bmatrix} 0.362 + j0.008 & -0.139 - j0.159 & -0.139 + j0.174 \\ -0.192 + j0.539 & 0.308 - j0.128 & -0.192 - j0.295 \\ -0.169 - j0.546 & -0.169 + j0.287 & 0.331 + j0.121 \end{bmatrix} \\ \mathbf{A}_3 &= \begin{bmatrix} 0.362 - j0.008 & -0.139 + j0.159 & -0.139 - j0.174 \\ -0.192 - j0.539 & 0.308 + j0.128 & -0.192 + j0.295 \\ -0.169 + j0.546 & -0.169 - j0.287 & 0.331 - j0.121 \end{bmatrix} \end{aligned}$$

The distribution vector at time step $n = 3$ is

$$\begin{aligned} \mathbf{s}(3) &= \mathbf{P}^3\, \mathbf{s}(0) \\ &= \left[\mathbf{A}_1 + \lambda_2^3 \mathbf{A}_2 + \lambda_3^3 \mathbf{A}_3\right] \mathbf{s}(0) \\ &= \begin{bmatrix} 0.285 & 0.389 & 0.326 \end{bmatrix}^t \end{aligned}$$

Property of Matrix **A1**

Matrix $\mathbf{A}_1$ is column stochastic and all its columns are identical.

$$\mathbf{A}_1 = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_1 & \cdots & \mathbf{x}_1 \end{bmatrix}$$

Each column of $\mathbf{A}_1$ represents the steady-state distribution vector for the Markov chain.

The steady state distribution vector $\mathbf{s}(\infty) = \mathbf{s}$ must be independent of the initial value $\mathbf{s}(0)$ and equals any column of $\mathbf{A}_1$.