Discrete-time Markov chains

Balakrishna Prabhu

Outline

2.1 Discrete-time Markov chains

Stochastic processes

Definition

A stochastic process is a sequence of random variables indexed by time.

- Examples: waiting time in a supermarket, daily temperature, number in queue in the cantine,. . .

Stochastic processes

Definition

A stochastic process is a sequence of random variables indexed by time.

- Examples: waiting time in a supermarket, daily temperature, number in queue in the cantine,. . .
- Types:

Time \ State space		
	D,D	D,C
	C,D	C,C

D=Discrete
C=Continuous

- In this course, only discrete state-space processes

Discrete-time Markov chain

Definition

A stochastic process $\{X_n\}_{n\geq 0}$ is a discrete-time Markov chain (DTMC) if

1. X_n takes values in a discrete set, $\mathcal{S}$; and
2. it has the Markov property: $\forall\, i, j \in \mathcal{S}$,

$$\mathbb{P}(X_{n+1} = j | X_n = i, X_{n-1}, \ldots, X_0) = \mathbb{P}(X_{n+1} = j | X_n = i).$$

Discrete-time Markov chain

Definition

A stochastic process $\{X_n\}_{n\geq 0}$ is a discrete-time Markov chain (DTMC) if

1. X_n takes values in a discrete set, $\mathcal{S}$; and
2. it has the Markov property: $\forall\, i, j \in \mathcal{S}$,

$$\mathbb{P}(X_{n+1} = j | X_n = i, X_{n-1}, \ldots, X_0) = \mathbb{P}(X_{n+1} = j | X_n = i).$$

- Markov property: the distribution of the future conditioned on the present is independent of the past.
- At time n to predict the future, we do not need the values from the past. Only the current value X_n of the process is sufficient.
- Models correlation between variables.

Discrete-time Markov chain

Definition

A stochastic process $\{X_n\}_{n\geq 0}$ is a discrete-time Markov chain (DTMC) if

1. X_n takes values in a discrete set, $\mathcal{S}$; and
2. it has the Markov property: $\forall\, i, j \in \mathcal{S}$,

$$\mathbb{P}(X_{n+1} = j | X_n = i, X_{n-1}, \ldots, X_0) = \mathbb{P}(X_{n+1} = j | X_n = i).$$

- Markov property: the distribution of the future conditioned on the present is independent of the past.
- At time n to predict the future, we do not need the values from the past. Only the current value X_n of the process is sufficient.
- Models correlation between variables.

Definition (Time homogeneous)

A DTMC is time homegeneous if, $\forall\, i, j \in \mathcal{S}$,

$$\mathbb{P}(X_{n+1} = j | X_n = i) = \mathbb{P}(X_1 = j | X_0 = i).$$

2.2 Transition matrix

Transition matrix

- $\mathbb{P}(X_1 = j | X_0 = i)$ are also known as the one-step transition probabilities. Denote by
$$p_{i,j} := \mathbb{P}(X_1 = j | X_0 = i)$$
- One-step probabilities predict the state of the process in the next step.

Definition (Transition matrix)

Assume $\mathcal{S}_X = \{0, 1, 2, \ldots\}$. The one-step probabilities can be written in form matrix

$$\text{Current state}\quad \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \end{array} \overset{\text{Next state}}{\begin{array}{cccc} 0 & 1 & 2 & \cdots \\ \hline \end{array}}$$

$$\begin{array}{c|cccc} & 0 & 1 & 2 & \cdots \\ \hline 0 & p_{0,0} & p_{0,1} & p_{0,2} & \cdots \\ 1 & p_{1,0} & p_{1,1} & p_{1,2} & \cdots \\ 2 & p_{2,0} & p_{2,1} & p_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Denote the transition matrix by P.

Transition matrix

- $\mathbb{P}(X_1 = j | X_0 = i)$ are also known as the one-step transition probabilities. Denote by
$$p_{i,j} := \mathbb{P}(X_1 = j | X_0 = i)$$
- One-step probabilities predict the state of the process in the next step.

Definition (Transition matrix)

Assume $\mathcal{S}_X = \{0, 1, 2, \ldots\}$. The one-step probabilities can be written in form matrix

$$\text{Current state} \quad \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \end{array} \begin{array}{c} \text{Next state} \\ \begin{array}{cccc} 0 & 1 & 2 & \cdots \end{array} \\ \begin{bmatrix} p_{0,0} & p_{0,1} & p_{0,2} & \cdots \\ p_{1,0} & p_{1,1} & p_{1,2} & \cdots \\ p_{2,0} & p_{2,1} & p_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \end{array}.$$

Denote the transition matrix by P.

- Row i is the conditional distribution of X_1 given i is observed in step 0

Transition matrix

- $\mathbb{P}(X_1 = j | X_0 = i)$ are also known as the one-step transition probabilities. Denote by
$$p_{i,j} := \mathbb{P}(X_1 = j | X_0 = i)$$
- One-step probabilities predict the state of the process in the next step.

Definition (Transition matrix)

Assume $\mathcal{S}_X = \{0, 1, 2, \ldots\}$. The one-step probabilities can be written in form matrix

$$\text{Current state} \quad \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \end{array} \overset{\text{Next state}}{\begin{array}{cccc} 0 & 1 & 2 & \cdots \\ \end{array}} \begin{bmatrix} p_{0,0} & p_{0,1} & p_{0,2} & \cdots \\ p_{1,0} & p_{1,1} & p_{1,2} & \cdots \\ p_{2,0} & p_{2,1} & p_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Denote the transition matrix by P.

- Row i is the conditional distribution of X_1 given i is observed in step 0
- Sum of each row is 1

Transition matrix

- $\mathbb{P}(X_1 = j | X_0 = i)$ are also known as the one-step transition probabilities. Denote by
$$p_{i,j} := \mathbb{P}(X_1 = j | X_0 = i)$$
- One-step probabilities predict the state of the process in the next step.

Definition (Transition matrix)

Assume $\mathcal{S}_X = \{0, 1, 2, \ldots\}$. The one-step probabilities can be written in form matrix

$$\text{Current state}\quad \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \end{array} \begin{array}{c} \text{Next state} \\ \begin{array}{cccc} 0 & 1 & 2 & \cdots \end{array} \\ \begin{bmatrix} p_{0,0} & p_{0,1} & p_{0,2} & \cdots \\ p_{1,0} & p_{1,1} & p_{1,2} & \cdots \\ p_{2,0} & p_{2,1} & p_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \end{array}.$$

Denote the transition matrix by P.

- Row i is the conditional distribution of X_1 given i is observed in step 0
- Sum of each row is 1

A DTMC is completely characterized by its transition matrix

Transition matrix

Example

A simple Markov model for the daily weather in Toulouse.

- Assume with state-space $\mathcal{S} = \{\text{Sunny (0), Rainy (1)}\}$
- Transition matrix

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & \quad 1 \end{array} \\ \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix} \end{array}$$

2.3 Transition diagram

Transition diagram

- An alternate representation of a DTMC is by using a graph called the <u>transition diagram</u>

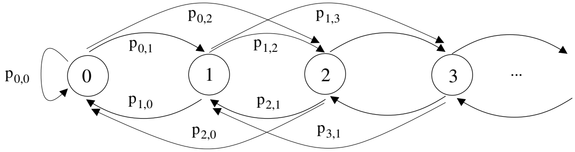


Figure: Transition diagram

- Vertices of the transition diagram are the states
- Directed edges show the transition between states
- Weight of an edge is the transition probability between its two vertices

Transition diagram

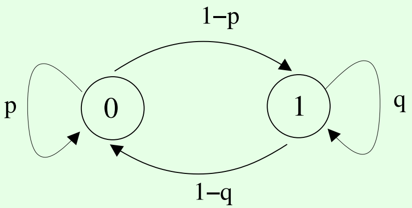
Example (Weather)

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix} \end{array}$$

Transition diagram

Example (Weather)

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \left[\begin{array}{cc} p & 1-p \\ 1-q & q \end{array} \right] \end{array}$$



2.4 n-step transition matrix

n-step transition matrix

- How to make predictions over a longer time-period?
- For $n \in \mathbb{N}$, define

$$p_{i,j}^{(n)} := \mathbb{P}(X_n = j | X_0 = i)$$

Theorem (Chapman-Kolmogorov)

For $m \leq n$,

$$p_{i,j}^{(n)} = \sum_{k \in \mathcal{S}} p_{i,k}^{(m)} \cdot p_{k,j}^{(n-m)}$$

n-step transition matrix

- How to make predictions over a longer time-period?
- For $n \in \mathbb{N}$, define

$$p_{i,j}^{(n)} := \mathbb{P}(X_n = j | X_0 = i)$$

Theorem (Chapman-Kolmogorov)

For $m \leq n$,

$$p_{i,j}^{(n)} = \sum_{k \in \mathcal{S}} p_{i,k}^{(m)} \cdot p_{k,j}^{(n-m)}$$

In matrix form,

$$P^{(n)} = P^{(m)} \cdot P^{(n-m)}$$

Proof.

$$p_{i,j}^{(n)} = \mathbb{P}(X_n = j | X_0 = i)$$

(law of total probability)

$$= \sum_{k \in \mathcal{S}} \mathbb{P}(X_n = j | X_m = k, X_0 = i) \mathbb{P}(X_m = k | X_0 = i)$$

(Markov property)

$$= \sum_{k \in \mathcal{S}} \mathbb{P}(X_n = j | X_m = k) \mathbb{P}(X_m = k | X_0 = i)$$

n-step transition matrix

Proof continued.

$$
\begin{aligned}
&\text{(time homogeneity)} \\
&= \sum_{k \in \mathcal{S}} \mathbb{P}(X_{n-m} = j | X_0 = k)\mathbb{P}(X_m = k | X_0 = i) \\
&= \sum_{k \in \mathcal{S}} p_{i,k}^{(m)} \cdot p_{k,j}^{(n-m)}
\end{aligned}
$$

□

- Predictions over a longer time-period can be deduced from predictions over two shorter time-periods

n-step transition matrix

Proof continued.

$$
\begin{aligned}
&\text{(time homogeneity)}\\
&= \sum_{k \in \mathcal{S}} \mathbb{P}(X_{n-m} = j | X_0 = k) \mathbb{P}(X_m = k | X_0 = i)\\
&= \sum_{k \in \mathcal{S}} p_{i,k}^{(m)} \cdot p_{k,j}^{(n-m)}
\end{aligned}
$$

□

- Predictions over a longer time-period can be deduced from predictions over two shorter time-periods

Corollary

$$P^{(n)} = P^n$$

That is, the n-step transition matrix is the product of n one-step transition matrices.

- The one-step transition matrix P is sufficient to make predictions over a longer time-period.

n-step transition matrix

Example (Weather)

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix} \end{array}$$

Q. Given it is raining today, what is the probability it will be sunny the day after tomorrow?

n-step transition matrix

Example (Weather)

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \left[\begin{array}{cc} p & 1-p \\ 1-q & q \end{array} \right] \end{array}$$

Q. Given it is raining today, what is the probability it will be sunny the day after tomorrow?

A. We are asked to compute $p_{1,0}^{(2)}$. From the Chapman-Kolmogorov eqns.:

$$\begin{aligned} p_{1,0}^{(2)} &= \sum_{k \in \{0,1\}} p_{1,k}^{(1)} \cdot p_{k,0}^{(1)} \\ &= p_{1,0}^{(1)} \cdot p_{0,0}^{(1)} + p_{1,1}^{(1)} \cdot p_{1,0}^{(1)} \\ &= (1-q)p + q(1-q) \end{aligned}$$

n-step transition matrix

Example (Weather)

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} \quad 0 \quad & \quad 1 \quad \end{array} \\ \left[\begin{array}{cc} p & 1-p \\ 1-q & q \end{array} \right] \end{array}$$

Q. Given it is raining today, what is the probability it will be sunny the day after tomorrow?
A. We are asked to compute $p_{1,0}^{(2)}$. From the Chapman-Kolmogorov eqns.:

$$\begin{aligned} p_{1,0}^{(2)} &= \sum_{k \in \{0,1\}} p_{1,k}^{(1)} \cdot p_{k,0}^{(1)} \\ &= p_{1,0}^{(1)} \cdot p_{0,0}^{(1)} + p_{1,1}^{(1)} \cdot p_{1,0}^{(1)} \\ &= (1-q)p + q(1-q) \end{aligned}$$

2.5 State probabilities

State probabilities

- State probabilities are the unconditional probabilities after n-steps

$$\pi_i^{(n)} := \mathbb{P}(X_n = i)$$

To compute the state probabilities, use the law of total probability

$$\begin{aligned}\pi_i^{(n)} &= \sum_{k \in \mathcal{S}} \mathbb{P}(X_) = k)\mathbb{P}(X_n = i | X_0 = k) \\ &= p_{k,i}^{(n)} \sum_{k \in \mathcal{S}} \pi_k^{(0)} p_{k,i}^{(n)}\end{aligned}$$

where $\pi_k^{(0)}$ is the belief of the current state being k.

State probabilities

- State probabilities are the unconditional probabilities after n-steps

$$\pi_i^{(n)} := \mathbb{P}(X_n = i)$$

To compute the state probabilities, use the law of total probability

$$\begin{aligned}\pi_i^{(n)} &= \sum_{k \in \mathcal{S}} \mathbb{P}(X_) = k)\mathbb{P}(X_n = i | X_0 = k) \\ &= p_{k,i}^{(n)} \sum_{k \in \mathcal{S}} \pi_k^{(0)} p_{k,i}^{(n)}\end{aligned}$$

where $\pi_k^{(0)}$ is the belief of the current state being k.

- In matrix form,

$$\pi^{(n)} = \pi^{(0)} P^n$$

where $\pi^{(n)} = [\pi_0^{(n)}, \pi_1^{(n)}, \ldots]$ is the state probability vector.

Example (Weather)

Q. Given $\pi_0^{(n)} = 0.3$, compute $\pi_0^{(2)}$.

State probabilities

- State probabilities are the unconditional probabilities after n-steps

$$\pi_i^{(n)} := \mathbb{P}(X_n = i)$$

To compute the state probabilities, use the law of total probability

$$\pi_i^{(n)} = \sum_{k \in \mathcal{S}} \mathbb{P}(X_) = k)\mathbb{P}(X_n = i | X_0 = k)$$
$$= p_{k,i}^{(n)} \sum_{k \in \mathcal{S}} \pi_k^{(0)} p_{k,i}^{(n)}$$

where $\pi_k^{(0)}$ is the belief of the current state being k.

- In matrix form,

$$\pi^{(n)} = \pi^{(0)} P^n$$

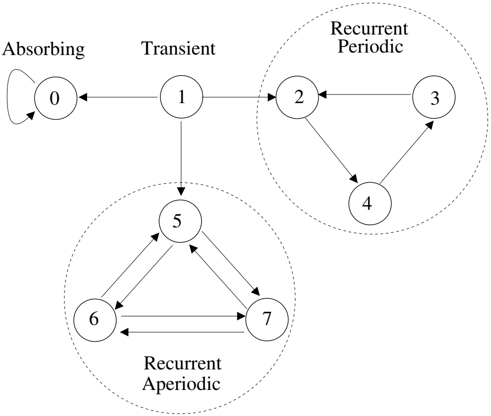
where $\pi^{(n)} = [\pi_0^{(n)}, \pi_1^{(n)}, \ldots]$ is the state probability vector.

Example (Weather)

Q. Given $\pi_0^{(n)} = 0.3$, compute $\pi_0^{(2)}$.
A. $0.3(p^2 + (1-p)(1-q)) + 0.7((1-q)p + q(1-q))$

2.6 Classification of states

Classification of states



- Transient: finite number of visits
- Recurrent: infinite number of visits
  - Positive recurrent: expected return time is finite
    - Periodic: return times are deterministic
    - Aperiodic: not periodic
  - Null recurrent: expected return time is infinite

2.7 Stationary distribution

Stationary or steady-state distribution

- How do state probabilities behave when the number of steps goes to ∞?
- Can multiply P a large number of times. Is there another way?

Theorem (Kolmogorov)

For an aperiodic and positive recurrent DTMC with transition matrix P

$$\lim_{n\to\infty} \pi^{(n)} =: \pi$$

is the solution of

$$\pi = \pi P$$
$$\sum_{i\in\mathcal{S}} \pi_i = 1 \tag{KOL}$$

Stationary or steady-state distribution

- How do state probabilities behave when the number of steps goes to ∞?
- Can multiply P a large number of times. Is there another way?

Theorem (Kolmogorov)

For an aperiodic and positive recurrent DTMC with transition matrix P

$$\lim_{n \to \infty} \pi^{(n)} =: \pi$$

is the solution of

$$\begin{aligned} \pi &= \pi P \\ \sum_{i \in \mathcal{S}} \pi_i &= 1 \end{aligned} \tag{KOL}$$

After a long time, $\pi^{(n)}$ become independent of time. It is said to be stationary or in steady-state.

Stationary or steady-state distribution

- How do state probabilities behave when the number of steps goes to ∞?
- Can multiply P a large number of times. Is there another way?

Theorem (Kolmogorov)

For an aperiodic and positive recurrent DTMC with transition matrix P

$$\lim_{n \to \infty} \pi^{(n)} =: \pi$$

is the solution of

$$\pi = \pi P$$
$$\sum_{i \in \mathcal{S}} \pi_i = 1 \quad \text{(KOL)}$$

After a long time, $\pi^{(n)}$ become independent of time. It is said to be stationary or in steady-state.

π_i is also the fraction of time the chain spends in state i

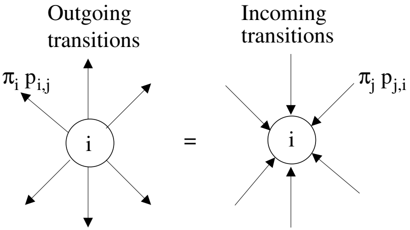
Stationary or steady-state distribution

- $\pi = \pi P$ are also called the global balance equations. To see this, the equation for π_i is

Stationary or steady-state distribution

- $\pi = \pi P$ are also called the global balance equations. To see this, the equation for π_i is

$$\begin{array}{rcl} \pi_i & = & \sum_{j \in \mathcal{S}} \pi_j p_{j,i} \\ & \Leftrightarrow & \\ \sum_{j \in \mathcal{S}} \pi_i p_{i,j} & = & \sum_{j \in \mathcal{S}} \pi_j p_{j,i} \\ \text{outgoing rate} & = & \text{incoming rate} \end{array}$$



Stationary or steady-state distribution

- If $|\mathcal{S}| = K$, then (KOL) is a linear system of $K + 1$ equations.
  - The system of the first K equations

    $$\pi = \pi P$$

    is homogeneous and has infinite number of solutions if it has at least one non-trivial solution
  - The last equation

    $$\sum_{i \in \mathcal{S}} \pi_i = 1$$

    is called the normalization equation. It is needed for determine the solution that makes π a probability vector.

Stationary or steady-state distribution

- If $|\mathcal{S}| = K$, then (KOL) is a linear system of $K + 1$ equations.
  - The system of the first K equations

    $$\pi = \pi P$$

    is homogeneous and has infinite number of solutions if it has at least one non-trivial solution
  - The last equation

    $$\sum_{i \in \mathcal{S}} \pi_i = 1$$

    is called the normalization equation. It is needed for determine the solution that makes π a probability vector.
- Observe that there are K unknown quantities (π_is) that we want to compute, and there are $K + 1$ equations.

Stationary or steady-state distribution

- If $|\mathcal{S}| = K$, then (KOL) is a linear system of $K+1$ equations.
  - The system of the first K equations

$$\pi = \pi P$$

    is homogeneous and has infinite number of solutions if it has at least one non-trivial solution
  - The last equation

$$\sum_{i \in \mathcal{S}} \pi_i = 1$$

    is called the normalization equation. It is needed for determine the solution that makes π a probability vector.
- Observe that there are K unknown quantities (π_is) that we want to compute, and there are $K+1$ equations.

Solving (KOL)

1. Remove any one equation from $\pi = \pi P$
2. Solve the remaining K equations to determine the K unknowns

Stationary or steady-state distribution

Example (Weather)

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \left[\begin{array}{cc} p & 1-p \\ 1-q & q \end{array} \right] \end{array}$$

Q. Determine the fraction of days it is sunny in this city.

Stationary or steady-state distribution

Example (Weather)

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix} \end{array}$$

Q. Determine the fraction of days it is sunny in this city.
A. We are asked to compute π_0. From (KOL), π is the solution of

Stationary or steady-state distribution

Example (Weather)

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \left[\begin{array}{cc} p & 1-p \\ 1-q & q \end{array} \right] \end{array}$$

Q. Determine the fraction of days it is sunny in this city.
A. We are asked to compute π_0. From (KOL), π is the solution of

$$\pi_0 = \pi_0 p + \pi_1 (1-q)$$

Stationary or steady-state distribution

Example (Weather)

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \left[\begin{array}{cc} p & 1-p \\ 1-q & q \end{array} \right] \end{array}$$

Q. Determine the fraction of days it is sunny in this city.
A. We are asked to compute π_0. From (KOL), π is the solution of

$$\begin{aligned} \pi_0 &= \pi_0 p + \pi_1 (1-q) \\ \pi_1 &= \pi_0 (1-p) + \pi_1 q \end{aligned}$$

Stationary or steady-state distribution

Example (Weather)

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \left[\begin{array}{cc} p & 1-p \\ 1-q & q \end{array} \right] \end{array}$$

Q. Determine the fraction of days it is sunny in this city.
A. We are asked to compute π_0. From (KOL), π is the solution of

$$\begin{aligned} \pi_0 &= \pi_0 p + \pi_1(1-q) \\ \pi_1 &= \pi_0(1-p) + \pi_1 q \\ \pi_0 + \pi_1 &= 1 \end{aligned}$$

Stationary or steady-state distribution

Example (Weather)

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \left[\begin{array}{cc} p & 1-p \\ 1-q & q \end{array} \right] \end{array}$$

Q. Determine the fraction of days it is sunny in this city.
A. We are asked to compute π_0. From (KOL), π is the solution of

$$\begin{aligned} \pi_0 &= \pi_0 p + \pi_1 (1-q) \\ \pi_1 &= \pi_0 (1-p) + \pi_1 q \\ \pi_0 + \pi_1 &= 1 \end{aligned}$$

1. Remove one equation from $\pi = \pi P$:

$$\begin{aligned} \pi_0 &= \pi_0 p + \pi_1 (1-q) \\ \pi_0 + \pi_1 &= 1 \end{aligned}$$

Stationary or steady-state distribution

Example (Weather)

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix} \end{array}$$

Q. Determine the fraction of days it is sunny in this city.
A. We are asked to compute π_0. From (KOL), π is the solution of

$$\begin{aligned} \pi_0 &= \pi_0 p + \pi_1(1-q) \\ \pi_1 &= \pi_0(1-p) + \pi_1 q \\ \pi_0 + \pi_1 &= 1 \end{aligned}$$

1. Remove one equation from $\pi = \pi P$:

$$\begin{aligned} \pi_0 &= \pi_0 p + \pi_1(1-q) \\ \pi_0 + \pi_1 &= 1 \end{aligned}$$

2. Solve these two equations to get:

$$\pi_0 = \frac{1-q}{(1-p)+(1-q)}, \quad \pi_1 = \frac{1-p}{(1-p)+(1-q)}$$

2.8 Application: Page rank

Ranking webpages

- **Objective**: rank a list of webpages containing the search terms

Ranking webpages

- **Objective**: rank a list of webpages containing the search terms
- Rank pages according to their popularity. Similar to other recommendation systems, for example, when ranking a restaurant from its reviews)

Ranking webpages

- **Objective**: rank a list of webpages containing the search terms
- Rank pages according to their popularity. Similar to other recommendation systems, for example, when ranking a restaurant from its reviews)
- For web pages, a link from another page can be seen as positive review
  - measure of popularity = the number of incoming links

Ranking webpages

- **Objective**: rank a list of webpages containing the search terms
- Rank pages according to their popularity. Similar to other recommendation systems, for example, when ranking a restaurant from its reviews)
- For web pages, a link from another page can be seen as positive review
  - measure of popularity = the number of incoming links
    ✗ Does not account for the weight of the recommender

Ranking webpages

- **Objective**: rank a list of webpages containing the search terms
- Rank pages according to their popularity. Similar to other recommendation systems, for example, when ranking a restaurant from its reviews)
- For web pages, a link from another page can be seen as positive review
  - measure of popularity = the number of incoming links
    ✗ Does not account for the weight of the recommender

Markov model for websurfing

1. The internet forms a directed graph with the webpages as the vertices and links as edges

Ranking webpages

- **Objective**: rank a list of webpages containing the search terms
- Rank pages according to their popularity. Similar to other recommendation systems, for example, when ranking a restaurant from its reviews)
- For web pages, a link from another page can be seen as positive review
  - measure of popularity = the number of incoming links
    ✗ Does not account for the weight of the recommender

Markov model for websurfing

1. The internet forms a directed graph with the webpages as the vertices and links as edges
2. Assume a websurfer follows an outgoing link chosen with uniform probability

Ranking webpages

- **Objective**: rank a list of webpages containing the search terms
- Rank pages according to their popularity. Similar to other recommendation systems, for example, when ranking a restaurant from its reviews)
- For web pages, a link from another page can be seen as positive review
  - measure of popularity = the number of incoming links
    ✗ Does not account for the weight of the recommender

Markov model for websurfing

1. The internet forms a directed graph with the webpages as the vertices and links as edges
2. Assume a websurfer follows an outgoing link chosen with uniform probability
3. Let X_n be the webpage that the websurfer finds itself on after n clicks. Show that X_n is a DTMC

Ranking webpages

- **Objective**: rank a list of webpages containing the search terms
- Rank pages according to their popularity. Similar to other recommendation systems, for example, when ranking a restaurant from its reviews)
- For web pages, a link from another page can be seen as positive review
  - measure of popularity = the number of incoming links
    ✗ Does not account for the weight of the recommender

Markov model for websurfing

1. The internet forms a directed graph with the webpages as the vertices and links as edges
2. Assume a websurfer follows an outgoing link chosen with uniform probability
3. Let X_n be the webpage that the websurfer finds itself on after n clicks. Show that X_n is a DTMC
4. Use the stationary probability of page i, π_i, as the measure of popularity

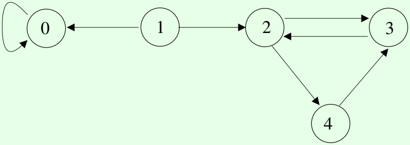
✓ Recall π_i is the fraction of visits of the surfer to page i. More visits ⇒ higher popularity.

✓ Takes into account the popularity of the recommender

$$\pi_i = \sum_j \pi_j p_{j,i}$$

Ranking webpages

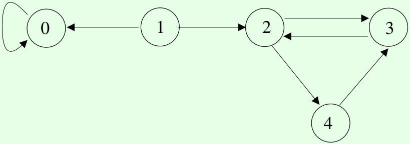
Example



$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \left[\begin{array}{ccccc} \ & \ & \ & \ & \ \\ \ & \ & \ & \ & \ \\ \ & \ & \ & \ & \ \\ \ & \ & \ & \ & \ \\ \ & \ & \ & \ & \ \end{array} \right] \end{array}.$$

Ranking webpages

Example

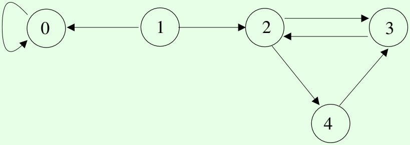


$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{array} \right] \end{array}.$$

- ✗ Internet graph is a not positive recurrent
- ✓ Make the graph positive recurrent by adding transitions

Ranking webpages

Example

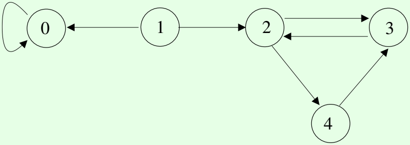


$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ & & & & \\ & & & & \\ & & & & \end{bmatrix} \end{array}.$$

- ✗ Internet graph is a not positive recurrent
- ✓ Make the graph positive recurrent by adding transitions

Ranking webpages

Example

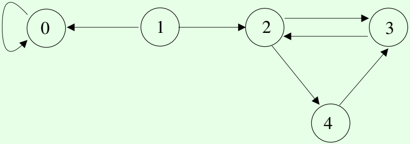


$$P = \begin{array}{c|ccccc} & 0 & 1 & 2 & 3 & 4 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0.5 & 0 & 0.5 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0.5 & 0.5 \\ 3 & & & & & \\ 4 & & & & & \end{array}.$$

- ✗ Internet graph is a not positive recurrent
- ✓ Make the graph positive recurrent by adding transitions

Ranking webpages

Example

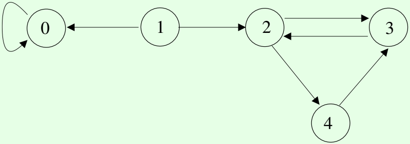


$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 1 & 0 & 0 \\ & & & & \end{bmatrix} \end{array}.$$

- ✗ Internet graph is a not positive recurrent
- ✓ Make the graph positive recurrent by adding transitions

Ranking webpages

Example



$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}.$$

- ✗ Internet graph is a not positive recurrent
- ✓ Make the graph positive recurrent by adding transitions

Ranking webpages

Modified Markov model for websurfing

1. The internet forms a directed graph with the webpages as the vertices and links as edges
2. Behaviour of a random websurfer

Ranking webpages

Modified Markov model for websurfing

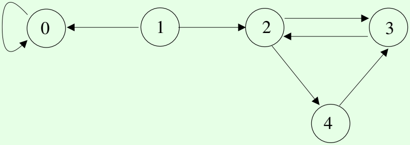
1. The internet forms a directed graph with the webpages as the vertices and links as edges
2. Behaviour of a random websurfer
    2.1 with probability α follows an outgoing link chosen with uniform probability
    2.2 with probability $(1-\alpha)$ goes to a page chosen uniformly at randomly
3. Let X_n be the webpage that the websurfer finds itself on after n clicks. Show that X_n is a DTMC
4. Use the stationary probability of page i, π_i, as the measure of popularity

Ranking webpages

Modified Markov model for websurfing

1. The internet forms a directed graph with the webpages as the vertices and links as edges
2. Behaviour of a random websurfer
   2.1 with probability α follows an outgoing link chosen with uniform probability
   2.2 with probability $(1-\alpha)$ goes to a page chosen uniformly at randomly
3. Let X_n be the webpage that the websurfer finds itself on after n clicks. Show that X_n is a DTMC
4. Use the stationary probability of page i, π_i, as the measure of popularity

Example (Modified websurfing)

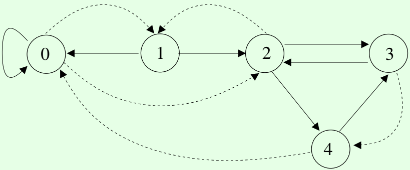


Ranking webpages

Modified Markov model for websurfing

1. The internet forms a directed graph with the webpages as the vertices and links as edges
2. Behaviour of a random websurfer
   2.1 with probability α follows an outgoing link chosen with uniform probability
   2.2 with probability $(1-\alpha)$ goes to a page chosen uniformly at randomly
3. Let X_n be the webpage that the websurfer finds itself on after n clicks. Show that X_n is a DTMC
4. Use the stationary probability of page i, π_i, as the measure of popularity

Example (Modified websurfing)



Ranking webpages

Example (Modified websurfing)

$$P = \alpha \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 1 & 0 & 0. \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} + (1-\alpha) \begin{bmatrix} & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{bmatrix}$$

Ranking webpages

Example (Modified websurfing)

$$P = \alpha \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 1 & 0 & 0. \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} + (1-\alpha) \begin{bmatrix} 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \end{bmatrix}$$

Ranking webpages

Example (Modified websurfing)

$$P = \alpha \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 1 & 0 & 0. \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} + (1-\alpha) \begin{bmatrix} 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \end{bmatrix}$$

- α can be seen as a weight given to the original graph.
  If $\alpha \approx 0$, the recommendation of the original graph has little weight