



MVE051/MSG810 - Statistik och diskret matematik (LP4 2022-2023)

Chalmers och Göteborgs universitet

Föreläsning 6 – Markovkedjor

April 19, 2023

Dagens föreläsning

- En introduktion till Markovkedjor

Rekommenderad läsning: <https://setosa.io/ev/markov-chains/>

Exempel

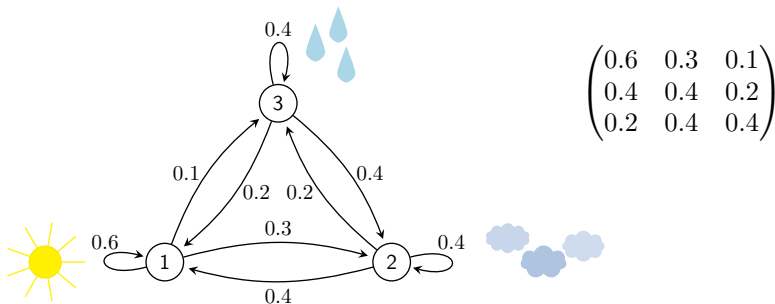
Antag att det bara finns tre olika väder på sommaren i Göteborg; (1) soligt, (2) mulet och (3) regnigt. Anta också att vädret i morgon bara beror på vädret idag, och även följande.

- (1) Om det är soligt idag så är sannolikheten för regn i morgon 0.2, sannolikheten för mulet väder i morgon 0.4 och sannolikheten för soligt väder i morgon 0.6.
- (2) Om det är mulet idag så är sannolikheten för regn i morgon 0.4, sannolikheten för mulet väder i morgon 0.4 och sannolikheten för soligt väder 0.2 i morgon.
- (3) Om det är regnigt idag så är sannolikheten för regn i morgon 0.6, sannolikheten för mulet väder 0.3 i morgon och sannolikheten för soligt väder i morgon 0.1.

Exempel

Övergångsgraf och övergångsmatrix

Vi kan sammanfatta vädret med följande graf och matris:



Markovkedja

$X_1, X_2, X_3, \dots$

Realisation

1, 2, 3, 3, 2, 1, 1, 3, 2, 1, 2, 2, 1, 3, 2, 1, 1, 2, 1, \dots

Markovkedjor

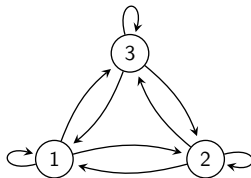
Markovkedja

En serie av stokastiska variabler $X_1, X_2, \dots$ sägs vara en Markovkedja om

$$P(X_n = k \mid X_1, X_2, \dots, X_{n-1}) = P(X_n = k \mid X_{n-1}) = P(X_1 = k \mid X_0)$$

↑
minneslös

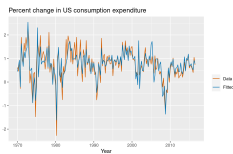
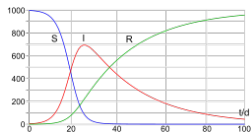
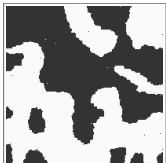
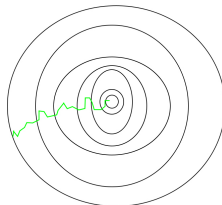
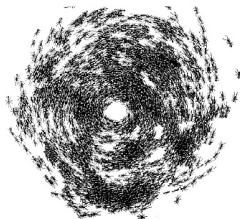
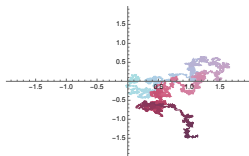
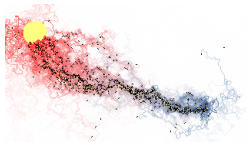
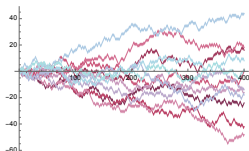
↑
tidshomogen



Frågor

- Hur ofta är Markovkedjan lika med 1?
- Hur länge behöver vi vänta innan Markovkedjan når tillstånd 2?
- Hur länge är Markovkedjan i tillstånd 3 innan den hoppar till ett annat tillstånd?

Tillämpningar



Markovkedjor

Tillstånd

Om $S = \{s_1, s_2, \dots\}$ är mängden av alla värden som X_k kan anta. Elementen i S kallas för tillstånd.

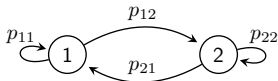
Övergång

Vid varje tidssteg sker en övergång (transition) från ett tillstånd X_k till ett annat X_k (kan vara samma).

Övergångssannolikhet

Sannolikheten för att gå från tillstånd s_i till tillstånd s_j skrivs p_{ij} , dvs.

$$p_{ij} = P(X_{k+1} = j \mid X_k = i)$$



Övergångsmatris

Övergångsmatrisen

Om S är ändlig kan vi bilda matrisen övergångsmatrisen $P = (p_{ij})$.

Övergångsmatrisen kommer alltid att uppfylla

- $p_{ij} \geq 0$.
- $\sum_j p_{ij} = P(X_j \in S \mid X_i = i) = 1$ (dvs varje radsumma är 1).

Man säger att P är en stokastisk matris.

Multiplikation med en enhetsvektor

Let $\mathbf{e}_i$ beteckna den i :e enhetsvektorn.

Då är

$$\mathbf{e}_i^T P = (p_{i1}, \dots, p_{in}) = (p_{ij})_j = (P(X_{k+1} = j \mid X_k = i))_j.$$

Egenskaper hos övergångsmatrisen

Födelningsvektorn

För $k \geq 0$ låter vi $\mathbf{u}^{(k)} = (P(X_k = j))_j$.

Vektorn $\mathbf{u}^{(0)}$ brukar kallas för startvektorn eller initialfördelningen. Ofta känner vi till $\mathbf{u}^{(0)}$ och är intresserade av $\mathbf{u}^{(k)} = (P(X_k = j))_j$ för något $k > 0$.

Multiplikation med $\mathbf{u}^{(k)}$

$$(\mathbf{u}^{(k)})^T P = \sum_i u_i^{(k)} p_{ij} = (P(X_{i+1} = j \mid X_i \sim \mathbf{u}^{(k)}))_j = \mathbf{u}^{(k+1)}.$$

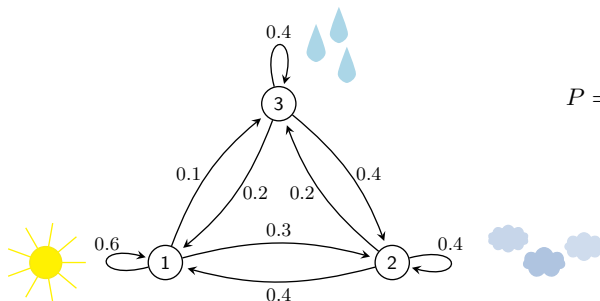
Därför gäller även

$$(\mathbf{u}^{(k)})^T P^m = (\mathbf{u}^{(k+m)})^T \quad \text{och} \quad (\mathbf{u}^{(m)})^T = (\mathbf{u}^{(0)})^T P^m.$$

Matrisen P^m

$$P^m(i, j) = P(X_{k+m} = j \mid X_k = i) = p_{ij}^{(m)}$$

Exempel



$$P = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.4 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}$$

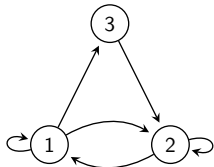
$$P^2 = \begin{pmatrix} 0.5 & 0.34 & 0.16 \\ 0.44 & 0.36 & 0.2 \\ 0.36 & 0.38 & 0.26 \end{pmatrix} \quad P^3 = \begin{pmatrix} 0.468 & 0.35 & 0.182 \\ 0.448 & 0.356 & 0.196 \\ 0.42 & 0.364 & 0.216 \end{pmatrix} \quad \dots$$

$$P(\text{regn i övermorgon om sol idag}) = P^2(1, 3) = 0.16$$

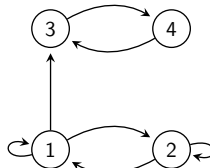
Irreducibilitet och reguljära markovkedjor

Irreducibla Markovkedjor

En Markovkedja sägs vara irreducibel om det går att gå från varje tillstånd till varje annat tillstånd i något antal steg.



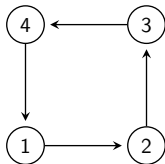
En irreducibel Markovkedja



En Markovkedja som inte är irreducibel

Reguljära Markovkedjor

En Markovkedja sägs vara reguljär om alla element i P^m är strikt större än 0 för något m . Varje reguljär Markovkedja är irreducibel.



En Markovkedja som är irreducibel men inte reguljär

Gränsfördelning och stationär fördelning

Gränsfördelning

Om en markovkedja är reguljär så kan man visa att

$$\Pi = \lim_{m \rightarrow \infty} P^m = \begin{pmatrix} \pi_1 & \pi_2 & \dots & \pi_n \\ \pi_1 & \pi_2 & \dots & \pi_n \\ \vdots & \vdots & ? & \vdots \\ \pi_1 & \pi_2 & \dots & \pi_n \end{pmatrix}$$

för någon vektor $(\pi_1, \dots, \pi_n)$, dvs.

$$P(X_{k+m} = j \mid X_k = i) \rightarrow \pi_j$$

oavsett vad i är. Vektorn π kallas för markovkedjans gränsfördelning.

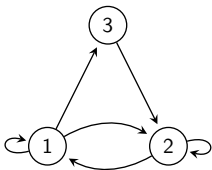
Det gäller att $\pi^T \Pi = \pi$, och

$$\pi^T P = \pi^T \Pi P = \pi^T \lim_{m \rightarrow \infty} P^m P = \pi^T \lim_{m \rightarrow \infty} P^{m+1} = \pi^T \Pi = \pi^T.$$

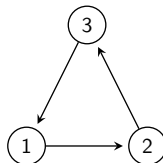
Gränsfördelning och stationär fördelning

Stationär fördelning

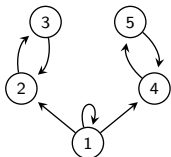
En sannolikhetsvektor π som uppfyller $\pi^T P = \pi^T$ kallas för en stationär fördelning till P .



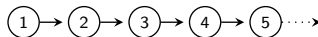
En Markovkedja som är reguljär och har en unik stationär fördelning



En Markovkedja inte är reguljär men har en unik stationär fördelning



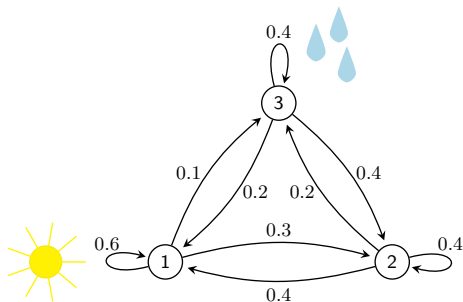
En Markovkedja har mer än en stationär fördelning.



En Markovkedja som inte har någon stationär fördelning.

Exempel

Hur ofta är det soligt väder?



$$P = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.4 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}$$

$$\Pi \approx P^{100} = \begin{pmatrix} 0.452 & 0.355 & 0.194 \\ 0.452 & 0.355 & 0.194 \\ 0.452 & 0.355 & 0.194 \end{pmatrix}, \quad \pi = (0.452, 0.355, 0.194)$$

$$P(\text{sol en slumpvist vald dag}) = \pi(1) = 0.452$$

Absorberande och transienta tillstånd

Absorberande tillstånd

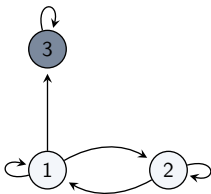
Ett tillstånd $s_i \in S$ sägs vara absorberande om man aldrig lämnar det, dvs. om $p_{ii} = 1$.

Absorberande Markovkedja

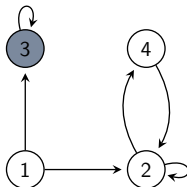
En Markovkedja sägs vara absorberande om den har minst ett absorberande tillstånd, och det går att komma till ett absorberande tillstånd från alla andra tillstånd.

Transienta tillstånd

Ett tillstånd som inte är absorberande i en absorberande markovkedja sägs vara transient.

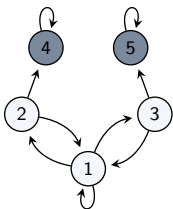


En absorberande Markovkedjor



En Markovkedja som inte är absorberande

Exempel



$$P = \left(\begin{array}{ccc|cc} 0.5 & 0.2 & 0.3 & 0 & 0 \\ 0.5 & 0 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0 & 0.5 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

Kanonisk form

Övergångsmatrisen på kanonisk form

Övergångsmatrisen $P = (p_{ij})$ till en absorberande Markovkedja med t transienta tillstånd $s_1, s_2, \dots, s_a$ och a absorberande tillstånd $s_{t+1}, s_{t+2}, \dots, s_a$ kan skrivas som

$$P = \begin{pmatrix} Q & R \\ \mathbf{0} & I_a \end{pmatrix}$$

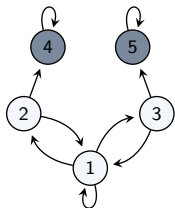
där Q är en $t \times t$ -matrix, R är en $t \times a$, $\mathbf{0}$ är en $a \times t$ -matrix med bara nollor, och I_a är en $a \times a$ enhetsmatrix.

Det följer att

$$P^n = \begin{pmatrix} Q^n & \sum_{j=0}^{n-1} Q^j R \\ \mathbf{0} & I_a \end{pmatrix}$$

där Q^n innehåller n -stegs övergångssannolikheter $p_{ij}^{(n)}$ mellan transienta tillstånd.

Exempel



$$P = \begin{pmatrix} \boxed{\begin{matrix} 0.5 & 0.2 & 0.3 \\ 0.5 & 0 & 0 \\ 0.5 & 0 & 0 \end{matrix}} & \boxed{\begin{matrix} 0 & 0 \\ 0.5 & 0 \\ 0 & 0.5 \end{matrix}} \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \boxed{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}} \end{pmatrix}$$

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

$$R = \begin{pmatrix} 0 & 0 \\ 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$$

$$P^2 = \begin{pmatrix} 0.5 & 0.1 & 0.15 & 0.1 & 0.15 \\ 0.25 & 0.1 & 0.15 & 0.5 & 0 \\ 0.25 & 0.1 & 0.15 & 0 & 0.5 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

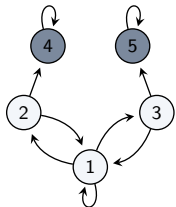
$$P^3 = \begin{pmatrix} 0.375 & 0.1 & 0.15 & 0.15 & 0.225 \\ 0.25 & 0.05 & 0.075 & 0.55 & 0.075 \\ 0.25 & 0.05 & 0.075 & 0.05 & 0.575 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Kanonisk form

En absorberande Markovkedja absorberar alltid tillslut

Om (X_i) är en absorberande Markovkedja så finns det alltid något $n \geq 0$ så att $X_k = X_n$ för alla $k \geq n$ (oavsett startfördelning). M.a.o. så är $\lim_{n \rightarrow \infty} Q^n = 0$.

Exempel



$$P = \begin{pmatrix} \boxed{\begin{matrix} 0.5 & 0.2 & 0.3 \\ 0.5 & 0 & 0 \\ 0.5 & 0 & 0 \end{matrix}} & \boxed{\begin{matrix} 0 & 0 \\ 0.5 & 0 \\ 0 & 0.5 \end{matrix}} \end{pmatrix}$$

$$Q = \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0 & 0 \\ 0.5 & 0 & 0 \end{pmatrix} \quad R = \begin{pmatrix} 0 & 0 \\ 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$$

$$Q^2 = \begin{pmatrix} 0.5 & 0.1 & 0.15 \\ 0.25 & 0.1 & 0.15 \\ 0.25 & 0.1 & 0.15 \end{pmatrix} \quad Q^4 = \begin{pmatrix} 0.313 & 0.075 & 0.113 \\ 0.188 & 0.05 & 0.075 \\ 0.188 & 0.05 & 0.075 \end{pmatrix} \quad Q^8 = \begin{pmatrix} 0.133 & 0.033 & 0.049 \\ 0.082 & 0.020 & 0.030 \\ 0.082 & 0.020 & 0.030 \end{pmatrix}$$

Kanonisk form

$$\begin{aligned}\mathbb{E}[\text{Förväntat antal besök i tillstånd } j \mid X_0 = i] &= \mathbb{E}\left[\sum_{k=0}^{\infty} \mathbf{1}(X_k = j) \mid X_0 = i\right] \\&= \sum_{k=0}^{\infty} \mathbb{E}\left[\mathbf{1}(X_k = j) \mid X_0 = i\right] = \sum_{k=0}^{\infty} P(X_k = j \mid X_0 = i) \\&= \sum_{k=0}^{\infty} Q^k(i, j) = \left(\sum_{k=0}^{\infty} Q^k\right)(i, j) = (I - Q)^{-1}(i, j).\end{aligned}$$

Kanonisk form

$$\begin{aligned}\mathbb{E}[\text{Förväntat antal besök i tillstånd } j \mid X_0 = i] &= \mathbb{E}\left[\sum_{k=0}^{\infty} \mathbf{1}(X_k = j) \mid X_0 = i\right] \\&= \sum_{k=0}^{\infty} \mathbb{E}\left[\mathbf{1}(X_k = j) \mid X_0 = i\right] = \sum_{k=0}^{\infty} P(X_k = j \mid X_0 = i) \\&= \sum_{k=0}^{\infty} Q^k(i, j) = \left(\sum_{k=0}^{\infty} Q^k\right)(i, j) = (I - Q)^{-1}(i, j).\end{aligned}$$

Den fundamentala matrisen

Man kan visa att matrisen $I - Q$ alltid är inverterbar.

Matrisen $N = (n_{ij}) = (I - Q)^{-1}$ kallas för den fundamentala matrisen.

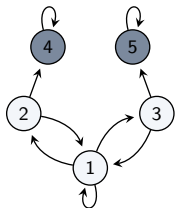
Tid i tillstånd j (innan absorption)

$$\mathbb{E}[\text{Förväntat antal besök i tillstånd } j \mid X_0 = i] = N(i, j)$$

Tid till absorption

$$\begin{aligned}\mathbb{E}[\text{Förväntad tid till absorption} \mid X_0 = i] \\&= \sum_{j \in T} \mathbb{E}[\text{Förväntat antal besök i tillstånd } j \mid X_0 = i] = \sum_{j \in T} N(i, j) = N\mathbf{1}.\end{aligned}$$

Exempel



$$P = \left(\begin{array}{ccc|cc} 0.5 & 0.2 & 0.3 & 0 & 0 \\ 0.5 & 0 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0 & 0.5 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

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

$$N = (I - Q)^{-1} = \begin{pmatrix} 4 & 0.8 & 1.2 \\ 2 & 1.4 & 0.6 \\ 2 & 0.4 & 1.6 \end{pmatrix} \quad N\mathbf{1} = \begin{pmatrix} 6 \\ 4 \\ 4 \end{pmatrix}$$

- Om vi börjar i tillstånd 1 besöker tillstånd 2 i snitt 0.8 gånger innan vi hamnar i ett av de två absorberande tillstånden (4 och 5).
- Det tar i snitt 6 tidssteg innan absorption om vi börjar i tillstånd 1.

Kanonisk form

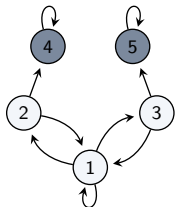
Absorptionssannolikhet

$$\begin{aligned}\lim_{n \rightarrow \infty} P^n &= \lim_{n \rightarrow \infty} \begin{pmatrix} Q^n & \sum_{j=0}^n Q^j R \\ \mathbf{0} & I_a \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \sum_{j=0}^{\infty} Q^j R \\ \mathbf{0} & I_a \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{0} & (I - Q)^{-1} R \\ \mathbf{0} & I_a \end{pmatrix} = \begin{pmatrix} \mathbf{0} & NR \\ \mathbf{0} & I_a \end{pmatrix}\end{aligned}$$

Elementen i NR beskriver absorptionssannolikheter för transienta tillstånd:

$$(NR)_{ij} = P((X_k) \text{ absorberas i } s_{t+j} \mid X_0 = s_i)$$

Exempel



$$P = \begin{pmatrix} \boxed{\begin{matrix} 0.5 & 0.2 & 0.3 \\ 0.5 & 0 & 0 \\ 0.5 & 0 & 0 \end{matrix}} & \boxed{\begin{matrix} 0 & 0 \\ 0.5 & 0 \\ 0 & 0.5 \end{matrix}} \\ \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} \end{pmatrix}$$

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

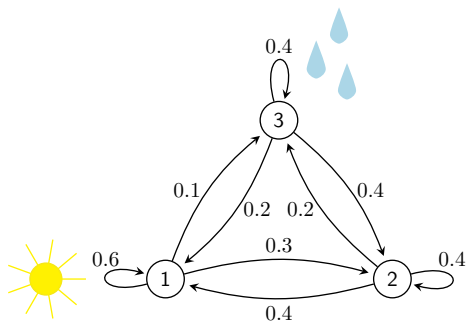
$$R = \begin{pmatrix} 0 & 0 \\ 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$$

$$N = \begin{pmatrix} 4 & 0.8 & 1.2 \\ 2 & 1.4 & 0.6 \\ 2 & 0.4 & 1.6 \end{pmatrix}$$

$$NR = \begin{pmatrix} 0.4 & 0.6 \\ 0.7 & 0.3 \\ 0.2 & 0.8 \end{pmatrix}$$

- Om vi börjar i tillstånd 2 s sannolikheten att vi absorberas i tillstånd 4 lika med 0.7.

Exempel (väder i Göteborg)



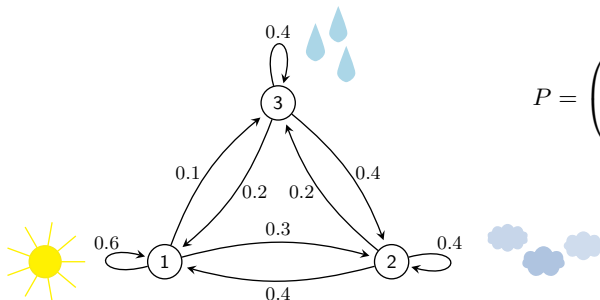
$$P = Q = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.4 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}$$

$$I - Q = \begin{pmatrix} 0.4 & -0.3 & -0.1 \\ -0.4 & 0.6 & -0.2 \\ -0.2 & -0.4 & 0.6 \end{pmatrix}$$

$$\pi^T P = \pi^T \Leftrightarrow \pi^T (P - I) = 0 \Leftrightarrow \pi = \begin{pmatrix} 0.451613 \\ 0.354839 \\ 0.193548 \end{pmatrix}$$

Exempel (väder i Göteborg)

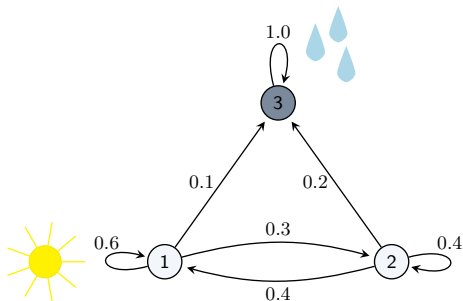
Hur länge får vi vänta i snitt innan det blir regn, om det är soligt idag?



$$P = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.4 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}$$

Exempel (väder i Göteborg)

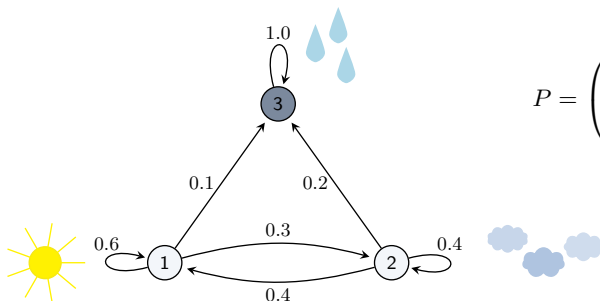
Hur länge får vi vänta i snitt innan det blir regn, om det är soligt idag?



$$P = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.4 & 0.2 \\ 0 & 0 & 1 \end{pmatrix}$$

Exempel (väder i Göteborg)

Hur länge får vi vänta i snitt innan det blir regn, om det är soligt idag?



$$P = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.4 & 0.2 \\ 0 & 0 & 1 \end{pmatrix}$$

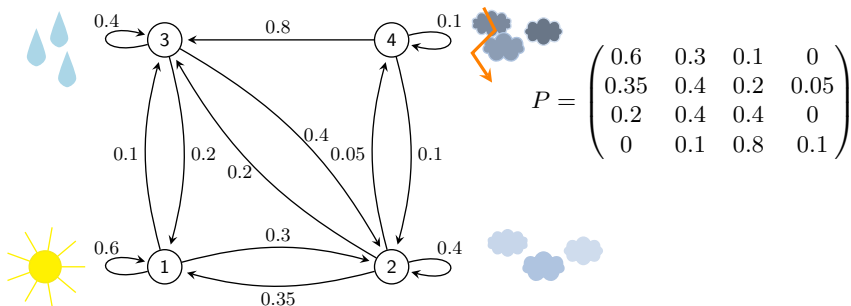
$$Q = \begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.4 \end{pmatrix}, \quad R = \begin{pmatrix} 0.1 \\ 0.2 \end{pmatrix}$$

$$I - Q = \begin{pmatrix} 0.4 & -0.3 \\ -0.4 & 0.6 \end{pmatrix}, \quad N = (I - Q)^{-1} = \begin{pmatrix} 5 & 2.5 \\ 3.333 & 3.333 \end{pmatrix}$$

$$N\mathbf{1} = \begin{pmatrix} 7.5 \\ 6.667 \end{pmatrix}, \quad NR = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

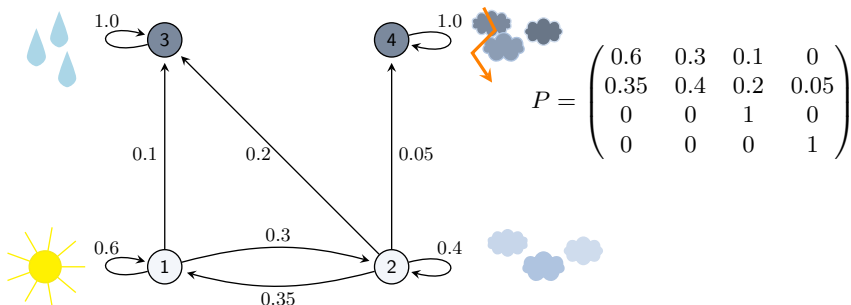
Exempel (väder i Göteborg med åska)

Om det är molnigt idag, vad annolikheten att det blir åska innan det regnar nästa gång?



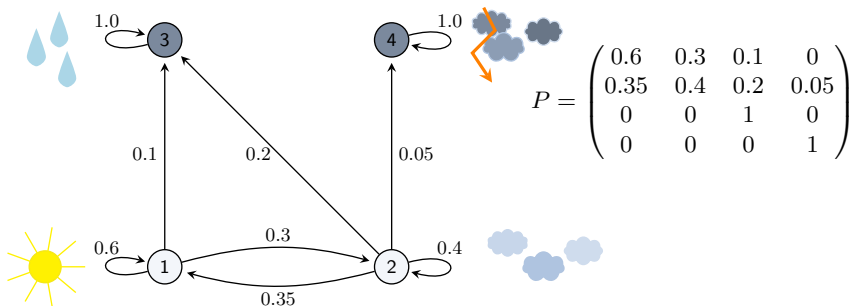
Exempel (väder i Göteborg med åska)

Om det är molnigt idag, vad annolikheten att det blir åska innan det regnar nästa gång?



Exempel (väder i Göteborg med åska)

Om det är molnigt idag, vad annolikheten att det blir åska innan det regnar nästa gång?



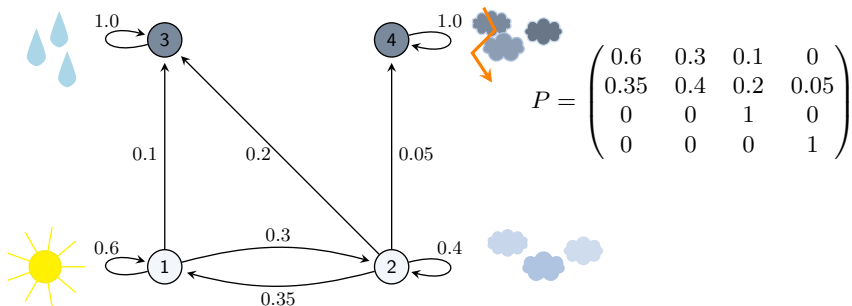
$$Q = \begin{pmatrix} 0.6 & 0.3 \\ 0.35 & 0.4 \end{pmatrix}, \quad R = \begin{pmatrix} 0.1 & 0 \\ 0.2 & 0.05 \end{pmatrix}$$

$$I - Q = \begin{pmatrix} 0.4 & -0.3 \\ -0.35 & 0.6 \end{pmatrix}, \quad N = (I - Q)^{-1} = \begin{pmatrix} 4.444 & 2.222 \\ 2.593 & 2.963 \end{pmatrix}$$

$$N\mathbf{1} = \begin{pmatrix} 6.667 \\ 5.556 \end{pmatrix}, \quad NR = \begin{pmatrix} 0.889 & 0.0111 \\ 0.852 & 0.0148 \end{pmatrix}$$

Exempel (väder i Göteborg med åska)

Om det är molnigt idag, vad annolikheten att det blir åska innan det regnar nästa gång?



$$Q = \begin{pmatrix} 0.6 & 0.3 \\ 0.35 & 0.4 \end{pmatrix}, \quad R = \begin{pmatrix} 0.1 & 0 \\ 0.2 & 0.05 \end{pmatrix}$$

$$I - Q = \begin{pmatrix} 0.4 & -0.3 \\ -0.35 & 0.6 \end{pmatrix}, \quad N = (I - Q)^{-1} = \begin{pmatrix} 4.444 & 2.222 \\ 2.593 & 2.963 \end{pmatrix}$$

$$N\mathbf{1} = \begin{pmatrix} 6.667 \\ 5.556 \end{pmatrix}, \quad NR = \begin{pmatrix} 0.889 & 0.0111 \\ 0.852 & 0.0148 \end{pmatrix}$$

$$\pi = (0.41496, 0.352627, 0.212823, 0.0195904)$$