Convergence of Markov Chains

Probability and Statistics for Data Science

Carlos Fernandez-Granda

These slides are based on the book Probability and Statistics for Data Science by Carlos Fernandez-Granda, available for purchase here. A free preprint, videos, code, slides and solutions to exercises are available at https://www.ps4ds.net

Markov property

$\tilde{a}_1$, $\tilde{a}_2$, ..., $\tilde{a}_n$ satisfy the Markov property if:

$\tilde{a}_{i+1}$ is conditionally independent of $\tilde{a}_1, \ldots, \tilde{a}_{i-1}$ given $\tilde{a}_i$

$$p_{\tilde{a}_{i+1} \,|\, \tilde{a}_1, \ldots, \tilde{a}_i}\left(a_{i+1} \,|\, a_1, a_2, \ldots, a_i\right) = p_{\tilde{a}_{i+1} \,|\, \tilde{a}_i}\left(a_{i+1} \,|\, a_i\right)$$

$$p_{\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_n}\left(a_1, a_2, \ldots, a_n\right) = p_{\tilde{a}_1}\left(a_1\right) p_{\tilde{a}_2 \,|\, \tilde{a}_1}\left(a_2 \,|\, a_1\right) p_{\tilde{a}_3 \,|\, \tilde{a}_2}\left(a_3 \,|\, a_2\right) \ldots$$

Finite state Markov chain

Each entry takes value in finite set of *states* $\{s_1, \ldots, s_m\}$

Marginal pmf represented by state vector:

$$\pi_i := \begin{bmatrix} p_{\tilde{a}_i}(s_1) \\ p_{\tilde{a}_i}(s_2) \\ \cdots \\ p_{\tilde{a}_i}(s_m) \end{bmatrix}$$

Time homogeneous finite state Markov chain

All transition probabilities are the same

$$p_{\tilde{a}_{i+1} \mid \tilde{a}_i}(a_{i+1} \mid a_i) = p_{\text{cond}}(a_{i+1} \mid a_i) \qquad 1 \le i \le n-1$$

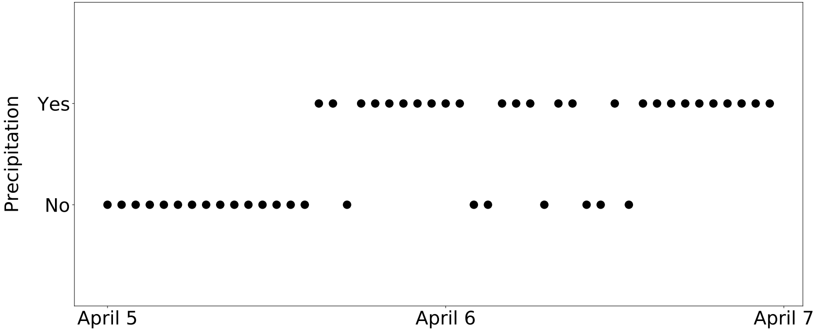
Transition matrix

$$T := \begin{bmatrix} p_{\text{cond}}(s_1 \mid s_1) & p_{\text{cond}}(s_1 \mid s_2) & \cdots & p_{\text{cond}}(s_1 \mid s_m) \\ p_{\text{cond}}(s_2 \mid s_1) & p_{\text{cond}}(s_2 \mid s_2) & \cdots & p_{\text{cond}}(s_1 \mid s_m) \\ \cdots & \cdots & \cdots & \cdots \\ p_{\text{cond}}(s_m \mid s_1) & p_{\text{cond}}(s_m \mid s_2) & \cdots & p_{\text{cond}}(s_m \mid s_m) \end{bmatrix}$$

State vector

$$\pi_i = T\pi_{i-1} = T^{i-1}\pi_1$$

Precipitation data



Markov-chain model

Marginal probabilities

No	88.7
Yes	11.3

1-step conditional probabilities

Hour h + 1 \ *Hour h*	**No**	**Yes**
No	96.0	31.2
Yes	4.0	68.8

State vector

$$\pi_1 := \begin{bmatrix} 0.887 \\ 0.113 \end{bmatrix}$$

Transition matrix

$$T := \begin{bmatrix} 0.960 & 0.312 \\ 0.040 & 0.688 \end{bmatrix}$$

$$\pi_2 = T\,\pi_1 = \pi_1!$$

$$\pi_i = T^{i-1}\,\pi_1 = \pi_1$$

Car rental

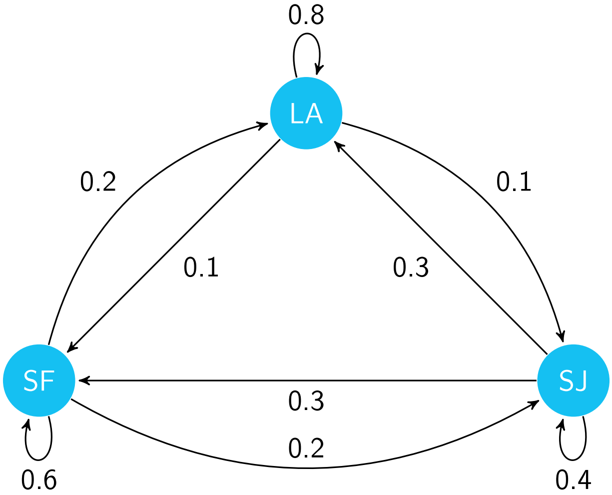
Goal: Model location of cars

3 locations (states): Los Angeles, San Francisco, San Jose

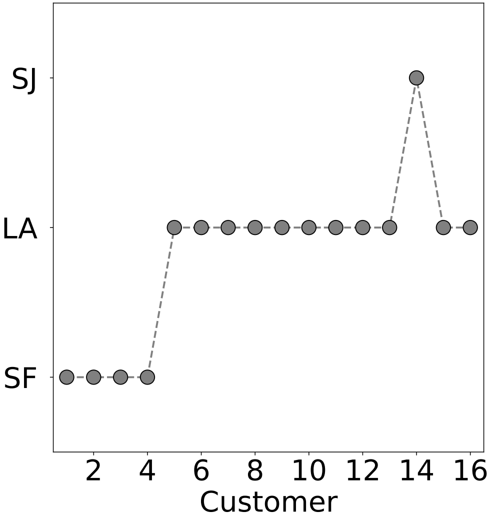
Transition probabilities:

$$\begin{array}{cc} \begin{array}{ccc} \text{San Francisco} & \text{Los Angeles} & \text{San Jose} \end{array} & \\ \begin{pmatrix} 0.6 & 0.1 & 0.3 \\ 0.2 & 0.8 & 0.3 \\ 0.2 & 0.1 & 0.4 \end{pmatrix} & \begin{array}{c} \text{San Francisco} \\ \text{Los Angeles} \\ \text{San Jose} \end{array} \end{array}$$

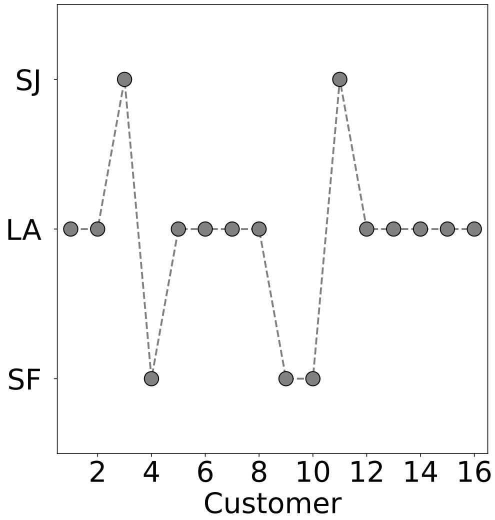
Car rental



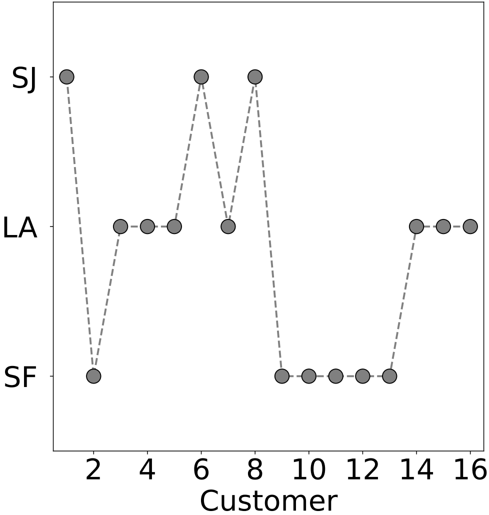
Realization



Realization



Realization



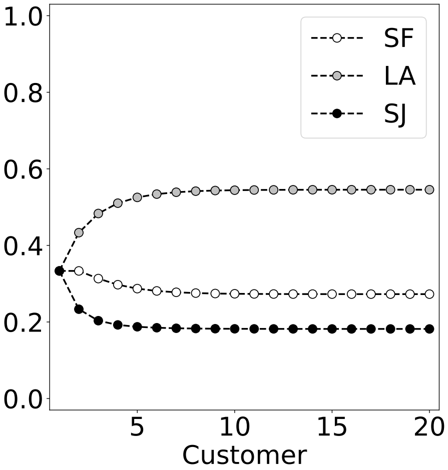
State vector and transition matrix

Cars are initially allocated to each location with same probability

$$\pi_1 := \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} \qquad T := \begin{bmatrix} 0.6 & 0.1 & 0.3 \\ 0.2 & 0.8 & 0.3 \\ 0.2 & 0.1 & 0.4 \end{bmatrix}$$

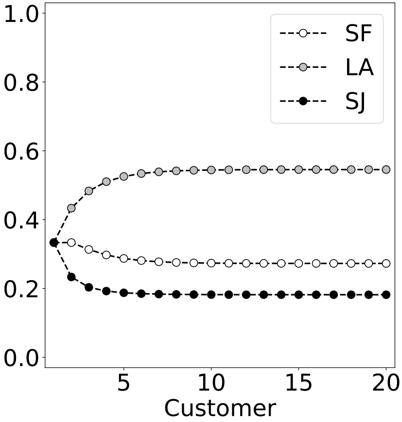
$$\pi_i = T^{i-1}\,\pi_1$$

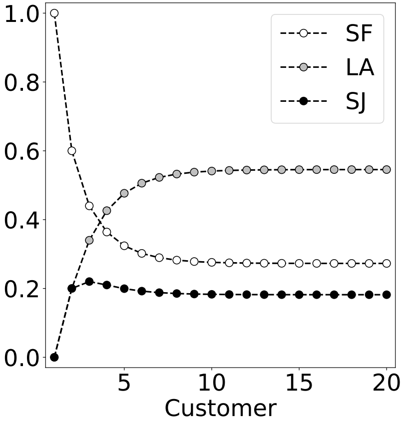
Evolution of the state vector



Different initial state vectors

$$\pi_1 := \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} \qquad \pi_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

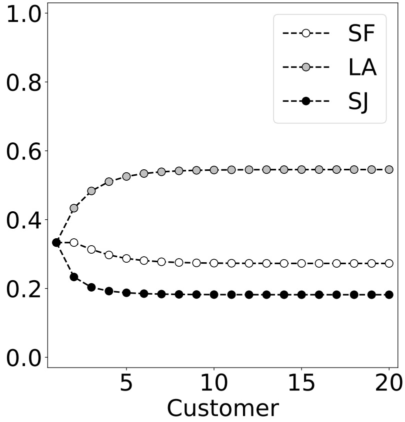


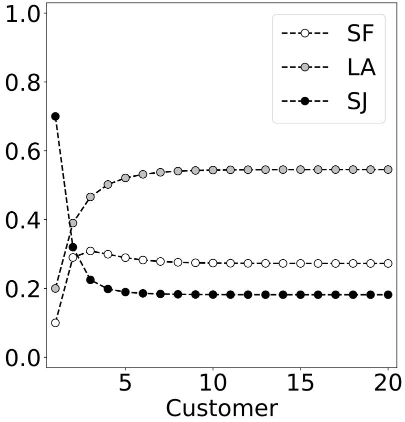


Different initial state vectors

$$\pi_1 := \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

$$\pi_1 := \begin{bmatrix} 0.1 \\ 0.2 \\ 0.7 \end{bmatrix}$$





Asymptotic analysis

$$\lim_{i \to \infty} \pi_i = \lim_{i \to \infty} T^{i-1} \pi_1$$

Eigendecomposition of the transition matrix

$$T = \underbrace{\begin{bmatrix} 0.27 & 0.37 & 0.37 \\ 0.55 & -0.50 & 0.13 \\ 0.18 & 0.13 & -0.50 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.57 & 0 \\ 0 & 0 & 0.23 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 1.28 & -0.87 & 0.70 \\ 0.70 & 0.13 & -1.44 \end{bmatrix}}_{Q^{-\mathbf{1}}}$$

$$\begin{aligned} T^{i-1} &= (Q\Lambda Q^{-1})^{i-1} = Q\Lambda Q^{-1} Q\Lambda Q^{-1} \cdots Q\Lambda Q^{-1} \\ &= Q\Lambda^{i-1} Q^{-1} \end{aligned}$$

Asymptotic analysis

$$
\begin{aligned}
&\lim_{i\to\infty} \pi_i \\
&= \lim_{i\to\infty} T^{i-1}\pi_1 = \lim_{i\to\infty} Q\Lambda^{i-1}Q^{-1}\pi_1 \\
&= \lim_{i\to\infty} Q \begin{bmatrix} 1^{i-1} & 0 & 0 \\ 0 & 0.57^{i-1} & 0 \\ 0 & 0 & 0.23^{i-1} \end{bmatrix} Q^{-1}\pi_1 \\
&= Q \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lim_{i\to\infty} 0.57^{i-1} & 0 \\ 0 & 0 & \lim_{i\to\infty} 0.23^{i-1} \end{bmatrix} Q^{-1}\pi_1 \\
&= \begin{bmatrix} 0.27 & 0.37 & 0.37 \\ 0.55 & -0.50 & 0.13 \\ 0.18 & 0.13 & -0.50 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1.28 & -0.87 & 0.70 \\ 0.70 & 0.13 & -1.44 \end{bmatrix} \pi_1 \\
&= \begin{bmatrix} 0.27 \\ 0.55 \\ 0.18 \end{bmatrix} \sum_{j=1}^{3} \pi_1[j] = \begin{bmatrix} 0.27 \\ 0.55 \\ 0.18 \end{bmatrix}
\end{aligned}
$$

Stationary distribution

π_* is a stationary distribution of a finite-state time-homogeneous Markov chain with transition matrix T if

$$T\pi_* = \pi_*$$

Precipitation

Marginal probabilities

No	88.7
Yes	11.3

1-step conditional probabilities

Hour h (columns); *Hour h + 1* (rows)

	No	**Yes**
No	96.0	31.2
Yes	4.0	68.8

State vector

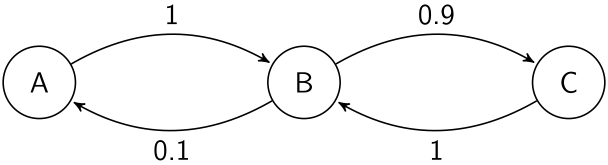
$$\pi_1 := \begin{bmatrix} 0.887 \\ 0.113 \end{bmatrix}$$

Transition matrix

$$T := \begin{bmatrix} 0.960 & 0.312 \\ 0.040 & 0.688 \end{bmatrix}$$

$$T = \underbrace{\begin{bmatrix} 0.887 & -0.632 \\ 0.113 & -0.632 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0.648 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} 1 & 1 \\ -0.179 & 1.40 \end{bmatrix}}_{Q^{-1}}$$

Another example



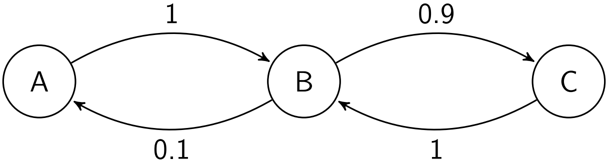
$$T = \begin{bmatrix} 0 & 0.1 & 0 \\ 1 & 0 & 1 \\ 0 & 0.9 & 0 \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 0.05 & 0.05 & 0.477 \\ 0.5 & -0.5 & 0 \\ 0.45 & 0.45 & -0.477 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1.89 & 0 & -0.21 \end{bmatrix}}_{Q^{-1}}$$

$\pi[2] = 0$

$$
\begin{aligned}
\pi_i &= T^{i-1}\pi_1 = Q\Lambda^{i-1}Q^{-1}\pi_1 \\
&= Q \begin{bmatrix} 1^{i-1} & 0 & 0 \\ 0 & (-1)^{i-1} & 0 \\ 0 & 0 & 0 \end{bmatrix} Q^{-1}\pi_1 \\
&= \left(\begin{bmatrix} 0.05 \\ 0.5 \\ 0.45 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} + (-1)^{i-1} \begin{bmatrix} 0.05 \\ -0.5 \\ 0.45 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \end{bmatrix} \right) \pi_1 \\
&= \begin{bmatrix} 0.05 \\ 0.5 \\ 0.45 \end{bmatrix} + (-1)^{i-1} \begin{bmatrix} 0.05 \\ -0.5 \\ 0.45 \end{bmatrix} = \begin{cases} \begin{bmatrix} 0.1 \\ 0 \\ 0.9 \end{bmatrix} & \text{if } i \text{ is even} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \text{if } i \text{ is odd} \end{cases}
\end{aligned}
$$

Periodic Markov chain



What we have learned

To analyze the convergence of Markov chains

Definition of stationary distribution