Le comportement asymptotique des chaînes de Markov

Ludovick Bouthat

Université Laval

SUMM; January 2023

Acknowledgement

This research is a collaborate effort with Pr. Nicolas Doyon, Pr. Javad Mashreghi and Pr. Frédéric Morneau-Guérin.

It was done with the financial help of the Vanier Scholarship.

Bourses d'études supérieures du Canada
Vanier
Canada Graduate Scholarships

Stochastic processes and Markov chains

Definition

A discrete-time Markov process is a sequence $X_1, X_2, X_3, \ldots$ of random variables with values in some state space $\mathcal{S}$ which respects the Markov property : *information useful for the prediction of the future does not depend on the past*.

Stochastic processes and Markov chains

Definition

A discrete-time Markov process is a sequence $X_1, X_2, X_3, \ldots$ of random variables with values in some state space $\mathcal{S}$ which respects the Markov property : *information useful for the prediction of the future does not depend on the past*.

Example (Coupon collector problem)

Context : A person collects cards of n players from a sports team, which he finds inside candy bars ; in each tablet, there is a $1/n$ chance of obtaining the card $\#n$.

Definition : $\mathcal{S} := \{\text{"Having } k \text{ distinct cards"} \mid k = 0, 1, \ldots, n\}$.

Transition matrices

A transition matrix, or *stochastic matrix*, is a matrix

$$P = \begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,j} & \cdots & P_{1,n} \\ P_{2,1} & P_{2,2} & \cdots & P_{2,j} & \cdots & P_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ P_{i,1} & P_{i,2} & \cdots & P_{i,j} & \cdots & P_{i,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ P_{n,1} & P_{n,2} & \cdots & P_{n,j} & \cdots & P_{n,n} \end{bmatrix},$$

where $P_{i,j}$ is the probability of transitioning from state i to state j in a Markov chain in one time step.

Transition matrices

A transition matrix, or *stochastic matrix*, is a matrix

$$P = \begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,j} & \cdots & P_{1,n} \\ P_{2,1} & P_{2,2} & \cdots & P_{2,j} & \cdots & P_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ P_{i,1} & P_{i,2} & \cdots & P_{i,j} & \cdots & P_{i,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ P_{n,1} & P_{n,2} & \cdots & P_{n,j} & \cdots & P_{n,n} \end{bmatrix},$$

where $P_{i,j}$ is the probability of transitioning from state i to state j in a Markov chain in one time step.

The probability of going from state i to state j in k steps is given by the coefficient (i,j) of the matrix P^k.

Coupon collector problem

Part 2

- $S_1 :=$ "0 card" ;
- $S_2 :=$ "1 card" ;
- $S_3 :=$ "2 distinct cards" ;
- $S_4 :=$ "3 distinct card".

The associated transition matrix is :

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Coupon collector problem

Part 2

- S_1 := "0 card" ;
- S_2 := "1 card" ;
- S_3 := "2 distinct cards" ;
- S_4 := "3 distinct card".

The associated transition matrix is :

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Hence,

$$P^2 = \begin{bmatrix} 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & \frac{1}{9} & \frac{2}{3} & \frac{2}{9} \\ 0 & 0 & \frac{4}{9} & \frac{5}{9} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad P^5 = \begin{bmatrix} 0 & \frac{1}{81} & \frac{10}{27} & \frac{50}{81} \\ 0 & \frac{1}{243} & \frac{62}{243} & \frac{20}{27} \\ 0 & 0 & \frac{32}{243} & \frac{211}{243} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Doubly stochastic matrices

Definition

A square matrix is *doubly stochastic* if :

- nonnegative coefficients ;
- row sums $= 1$;
- column sums $= 1$.

The set of doubly stochastic matrices of order n is denoted by Ω_n.

Example

$$D = \begin{bmatrix} 0.1 & 0.3 & 0.6 \\ 0.4 & 0.2 & 0.4 \\ 0.5 & 0.5 & 0 \end{bmatrix}.$$

Important properties

- Ω_n is closed under classical matrix multiplication, i.e., $D_1 D_2 \in \Omega_n$ if $D_1, D_2 \in \Omega_n$.

Important properties

- Ω_n is closed under classical matrix multiplication, i.e., $D_1 D_2 \in \Omega_n$ if $D_1, D_2 \in \Omega_n$.
- If $D \in \Omega_n$, $De = e$ and 1 is an eigenvalue of D.

Important properties

- Ω_n is closed under classical matrix multiplication, i.e., $D_1 D_2 \in \Omega_n$ if $D_1, D_2 \in \Omega_n$.
- If $D \in \Omega_n$, $De = e$ and 1 is an eigenvalue of D.
- If $D \in \Omega_n$, every eigenvalues of D lie in the unit disk.

Important cases

Example

The *uniform* matrix and the *circular shift matrix*

$$J_n = \frac{1}{n}\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix} \quad \& \quad K_n = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Important cases

Properties

1. If $D \in \Omega_n$, $DJ_n = J_nD = J_n$;

Important cases

Properties

1. If $D \in \Omega_n$, $DJ_n = J_nD = J_n$;

2. $$K_n^2 = \begin{bmatrix} 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \end{bmatrix}, \ \cdots \ , \ K_n^{n-1} = I_n, \ K_n^n = K_n$$

Main question

- Let $(D_n) \subseteq \Omega_n$ be such that D_k is the transition matrix of a doubly stochastic Markov chain at step k.

Main question

• Let $(D_n) \subseteq \Omega_n$ be such that D_k is the transition matrix of a doubly stochastic Markov chain at step k.

Question : What is the long-term behavior of a Markov chain formed from doubly stochastic matrices, i.e., what can we say about $D_1 D_2 D_2 \cdots$?

Main question

- Let $(D_n) \subseteq \Omega_n$ be such that D_k is the transition matrix of a doubly stochastic Markov chain at step k.

Question : What is the long-term behavior of a Markov chain formed from doubly stochastic matrices, i.e., what can we say about $D_1 D_2 D_2 \cdots$?

Definition

A Markov chain is *homogeneous* if $D_n = D$ for every $n \geq 1$. Otherwise, it is *non-homogeneous*.

Useful facts

Definition

A doubly stochastic matrix is ***irreducible*** if its eigenvalue 1 its eigenvalue 1 is simple. It is ***primitive*** if its eigenvalue 1 is the only one on the unit circle.

Proposition

For any doubly stochastic matrix D, there exist a permutation matrix P such that

$$P^{tr} DP = D_1 \oplus D_2 \oplus \cdots \oplus D_r,$$

where D_k are irreducible doubly stochastic matrices or smaller order.

Useful facts

Proposition

If $D \in \Omega_n$ is irreducible and $h \geq 1$ denotes the number of unimodular eigenvalue of D, then there exist a permutation matrix P such that

$$P^{tr} DP = \begin{bmatrix} 0 & D_1 & 0 & \cdots & 0 \\ 0 & 0 & D_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & D_{h-1} \\ D_h & 0 & 0 & \cdots & 0 \end{bmatrix} = (D_1 \oplus \cdots \oplus D_h) \left(K_h \otimes I_{n/h}\right),$$

where $D_j \in \Omega_{n/h}$ for $1 \leq j \leq h$.

Useful facts

Corollary

If $D \in \Omega_n$, then there exist a permutation matrix P such that

$$P^{tr} DP = \bigoplus_{i=1}^{r} D_i = \bigoplus_{i=1}^{r} (D_{i,1} \oplus \cdots \oplus D_{i,h_i}) \left(K_{h_i} \otimes I_{n_i/h_i}\right),$$

where each $D_j \in \Omega_{n_j}$ is irreducible and each $D_{i,j} \in \Omega_{n_i/h_i}$.

Convergent irreducible d.s. matrices

Theorem

Let D be an $n \times n$ irreducible doubly stochastic matrix. Then D^m converges as $m \to \infty$ if and only if D is primitive. Moreover, if it converges, then $\lim_{m\to\infty} D^m = J_n$.

Convergent irreducible d.s. matrices

Theorem

Let D be an $n \times n$ irreducible doubly stochastic matrix. Then D^m converges as $m \to \infty$ if and only if D is primitive. Moreover, if it converges, then $\lim_{m\to\infty} D^m = J_n$.

Example

$$D = \begin{bmatrix} 0.1 & 0.3 & 0.6 \\ 0.4 & 0.2 & 0.4 \\ 0.5 & 0.5 & 0 \end{bmatrix} \rightsquigarrow D^8 = \begin{bmatrix} 0.334636 & 0.334634 & 0.33073 \\ 0.333332 & 0.333335 & 0.333332 \\ 0.332031 & 0.332031 & 0.335938 \end{bmatrix}$$

Convergent d.s. matrices

Corollary

For almost all doubly stochastic matrices D, $D^k \to J_n$ as $k \to \infty$.

Convergent d.s. matrices

Corollary

For almost all doubly stochastic matrices D, $D^k \to J_n$ as $k \to \infty$.

Corollary

Let D be an $n \times n$ doubly stochastic matrix and let h be the number of eigenvalues of unit modulus of D. Let P be a permutation matrix such that

$$P^{tr} DP \; = \; D_1 \oplus D_2 \oplus \cdots \oplus D_r,$$

where each $D_i \in \Omega_{n_i}$ is irreducible. Then D^m converges as $m \to \infty$ if and only if D has precisely h eigenvalues equal to 1. Moreover, if it converges, then $r = h$ and $\lim_{m\to\infty} D^m = P(J_{n_1} \oplus J_{n_2} \oplus \cdots \oplus J_{n_r})P^{tr}$.

Cyclic irreducible d.s. matrices

Theorem

Let D be an $n \times n$ irreducible doubly stochastic matrix. Then D is cyclic of order p if and only if there exists a permutation matrix P such that

$$P^{tr} D P = K_p \otimes J_{n/p} = \begin{bmatrix} 0 & J_{n/p} & 0 & \cdots & 0 \\ 0 & 0 & J_{n/p} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & J_{n/p} \\ J_{n/p} & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Moreover, p is the number of eigenvalues of D on the unit circle.

Cyclic d.s. matrices

Theorem

Let D be an $n \times n$ doubly stochastic matrix. Then D is cyclic of order p if and only if there exists a permutation matrix P and positive integers r, h_i, k_i $(1 \leq i \leq r)$ satisfying $p = \mathrm{LCM}(h_1, h_2, \ldots, h_r)$ and $h_1 k_1 + h_2 k_2 + \cdots + h_r k_r = n$ such that

$$P^{tr} D P = \bigoplus_{j=1}^{r} \left(K_{h_j} \otimes J_{k_j} \right).$$

In that case, D has r eigenvalues equal to 1 and $h = h_1 + h_2 + \cdots + h_r$ eigenvalues on the unit circle.

Diverging irreducible d.s. matrices

Theorem

Let $D \in \Omega_n$ be an irreducible doubly stochastic matrix, and let P be a permutation matrix such that

$$P^{tr} DP = \left(\oplus_{i=1}^{h} D_i\right)\left(K_h \otimes I_{n/h}\right),$$

where $h := h(D)$ and $D_i \in \Omega_{n/h}$ $(1 \leq i \leq h)$. Then for every $0 \leq r \leq h-1$,

$$D^{mh+r} \xrightarrow{m\to\infty} P\left(K_h^r \otimes J_{n/h}\right) P^{tr}.$$

Diverging d.s. matrices

Theorem

Let $D \in \Omega_n$ be an irreducible doubly stochastic matrix, and let P be a permutation matrix such that

$$P^{tr} DP = \left(\oplus_{i=1}^{h} D_i\right)\left(K_h \otimes I_{n/h}\right),$$

where $h := h(D)$ and $D_i \in \Omega_{n/h}$ $(1 \leq i \leq h)$. Then for every $0 \leq r \leq h-1$,

$$D^{mh+r} \xrightarrow{m\to\infty} P\left(K_h^r \otimes J_{n/h}\right) P^{tr}.$$

A sufficient result

Question : What can we say about the infinite product $D_1 D_2 D_2 \cdots$?

A sufficient result

Question : What can we say about the infinite product $D_1 D_2 D_2 \cdots$?

Theorem

Let $A_1, A_2, \ldots$ be a sequence of $n \times n$ doubly stochastic matrices and let $\sigma_2(A)$ be the second largest singular value of A. If $\sum_{k=1}^{\infty}(1 - \sigma_2(A_k)) = \infty$, then $\lim_{m\to\infty} A_1 A_2 \cdots A_m = J$.

A sufficient result

Question : What can we say about the infinite product $D_1 D_2 D_2 \cdots$?

Theorem

Let $A_1, A_2, \ldots$ be a sequence of $n \times n$ doubly stochastic matrices and let $\sigma_2(A)$ be the second largest singular value of A. If $\sum_{k=1}^{\infty}(1 - \sigma_2(A_k)) = \infty$, then $\lim_{m\to\infty} A_1 A_2 \cdots A_m = J$.

Theorem (Schwarz, 1980)

Let $A_1, A_2, \ldots$ be a sequence of $n \times n$ doubly stochastic matrices and let $\nu(A) := \min_{i,j} a_{ij}$. If $\sum_{k=1}^{\infty} \nu(A_k) = \infty$, then $\lim_{m\to\infty} A_1 A_2 \cdots A_m = J$.

Improving a result of Schwarz

- $n\nu(D) \leq 1 - \sigma_2(D)$ for any $D \in \Omega_n$.

Improving a result of Schwarz

- $n\nu(D) \leq 1 - \sigma_2(D)$ for any $D \in \Omega_n$.

$$\sum_{k=1}^{\infty} \nu(A_k) = \infty \quad \implies \quad \sum_{k=1}^{\infty} (1 - \sigma_2(A_k)) = \infty.$$

Improving a result of Schwarz

- $n\nu(D) \leq 1 - \sigma_2(D)$ for any $D \in \Omega_n$.

$$\sum_{k=1}^{\infty} \nu(A_k) = \infty \implies \sum_{k=1}^{\infty} (1 - \sigma_2(A_k)) = \infty.$$

Example

Consider the case $A_k = A$ for each k, where

$$A = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}.$$

Then $A \in \Omega_n$, $\nu(A) = 0$, and the singular values of A are $1, 1/2$ and 0 so that $1 - \sigma_2(A) = 1/2$.

References

- Štefan Schwarz (1980) Infinite products of doubly stochastic matrices. *Acta Math. Univ. Comenian.*, 39 : 131–150.
- Ludovick Bouthat, Apoorva Khare, Javad Mashreghi & Frédéric Morneau-Guérin (2021) The *p*-norm of circulant matrices, *Linear and Multilinear Algebra*, DOI : 10.1080/03081087.2021.1983513
- K. R. Sahasranand (2022) The *p*-norm of circulant matrices via Fourier analysis. arXiv preprint arXiv :2111.11389.