Ergodicity Coefficients and Associated Polytopes

Incheoul Chung

Missouri State University,
Springfield, MO

July 11, 2017

Ergodicity

- Ergodicity has to do with the long-term behaviour of dynamical systems.
- Dynamical system described by a measure-preserving transformation.
- Time average equal spatial averages
- In a probability context, time average equals average over the probability space.
- State of the process after a long time is independent of the initial state.
- Possible when dynamical system is sufficiently "mixing".

Ergodicity

- Examples of ergodic dynamical systems:
- Irreducible Markov Chains
- Irrational rotation of the circle $\mathbb{R}/\mathbb{Z}$
- Bernoulli Shift
  - Let m be a probability measure on $\mathbb{R}$. Let $(\mathbb{R}^{\mathbb{N}}, \mathcal{E}, P)$ be a probability space such that there exists a sequence Y_i of i.i.d. random variables with distribution m.
  - The shift operator $\theta : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ defined by
  - $\theta(x_1, x_2, x_3, x_4, x_5, \dots) = (x_2, x_3, x_4, x_5, x_6, \dots)$
  - θ is an ergodic transformation.
- Example of a nonergodic dynamical system:
- The double potential well in classical mechanics

Ergodicity

For time-homogeneous Markov chains, every irreducible Markov chain is ergodic.

Theorem

Ergodic Theorem: Let Ω be a finite state space, f a real-valued function defined on Ω, μ any probability distribution on Ω. Define $E_\mu = \sum_{x \in \Omega} f(x)\mu(x)$. If (X_t) is an irreducible Markov chain, π is the stationary distribution of the Markov chain, then

$$P_\mu \left\{ \lim_{t \to \infty} \frac{1}{t} \sum_{s=0}^{t-1} f(X_s) = E_\pi(f) \right\} = 1$$

Ergodicity

A formal definition of ergodicity for finite, inhomogeneous Markov chains.

Definition

Let $\{S_k\}$ be a sequence of $n \times n$, row-stochastic matrices, $k \geq 1$, and let $t_{ij}^{(p,r)}$ be the (i,j)th entry of the forward product $T^{(p,r)} = S_{p+1}S_{p+2}\ldots S_{p+r}$. The sequence $\{S_k\}$ is said to be weakly ergodic if for all $1 \leq i, j, k, \leq n, p \geq 0$,

$$\lim_{r\to\infty} t_{ik}^{(p,r)} - t_{jk}^{(p,r)} = 0$$

Ergodicity

Roughly speaking, this means that a sequence of stochastic matrices is weakly ergodic if the rows equalize as the number of products increase.

Ergodicity coefficients are used to determine if a sequence is weakly ergodic.

Definition

A coefficient of ergodicity, or ergodicity coefficient, is a continuous scalar function $\mu(\cdot)$ defined for stochastic matrices S that satisfies

$$0 \leq \mu(S) \leq 1$$

A coefficient of ergodicity is proper iff

$$\mu(S) = 0 \iff rank(S) = 1$$

Ergodicity

Theorem

Let μ be a proper coefficient of ergodicity, $\{S_k\}$ a sequence of stochastic matrices, $k \geq 1$. Let $T^{(p,r)} = S_{p+1}S_{p+2}\ldots S_{p+r}$. The sequence $\{S_k\}$ is weakly ergodic if $\forall p \geq 0$,

$$\lim_{r\to\infty} \mu(T^{(p,r)}) = 0$$

Ergodicity

- Many different kinds of ergodicity coefficients exist but I will focus on those generated from vector norms.

Definition

$$\tau_p(S) = \max_{\substack{||z||_p=1 \\ z^T\mathbb{1}=0}} ||S^T z||_p$$

In particular, we focus on τ_1 and τ_∞ because their limits of maximum are convex polytopes.

Ergodicity

Theorem

[1] If S, S_1, S_2 are stochastic matrices, then

1. $0 \leq \tau_1(S) \leq 1$
2. $|\lambda| \leq \tau_1(S)$ *for all eigenvalues* $|\lambda| < 1$ *of* S.
3. $\tau_1(S_1 S_2) \leq \tau_1(S_1)\tau_1(S_2)$
4. $|\tau_1(S_1) - \tau_1(S_2)| \leq \tau_1(S_1 - S_2)$
5. $\tau_1(S) = 0 \iff rank(S) = 1$

Ergodicity

Theorem

[1] If S, S_1, S_2 are stochastic matrices, then

1. $0 \leq \tau_\infty(S) \leq ||S||_1$
2. $|\tau_\infty(S_1) - \tau_\infty(S_2)| \leq \tau_\infty(S_1 - S_2)$
3. $\tau_\infty(S_1 S_2) \leq \tau_\infty(S_1)\tau_\infty(S_2)$
4. $\tau_\infty(S) = 0 \iff rank(S) = 1$

Because $\tau_\infty(S)$ can be greater than 1, τ_∞ is not a coefficient of ergodicity in a strict sense. But it is still useful.

Polytopes

There are two equivalent definitions of convex polytopes:

Definition

A convex polytope is a set that can be realized as the convex hull of finitely many points.

Definition

A convex polytope is a bounded set that can be realized as the intersection of halfspaces and hyperplanes.

Convex polytopes are automatically compact sets since it is the continuous image of a simplex.

Polytopes

Definition

Let $V \subseteq \mathbb{R}^d$. A supporting hyperplane of V is a hyperplane S such that one of the two halfspaces associated with S completely contains V and S contains at least one boundary point of V.

Theorem

If K is a convex set in $\mathbb{R}^d$, and $x_0 \in K$ is a point on the boundary of K, then there exists a supporting hyperplane containing x_0.

Polytopes

Definition

Let $K \subset \mathbb{R}^d$ be a convex polytope. A set $F \subseteq K$ is a face iff $F = \emptyset$ or $F = K$, or if there exists a supporting hyperplane H of K such that $F = H \cap K$. $\emptyset$ and K are called improper faces of K. Every other face is called a proper face. The faces that are exactly a single point are called vertices, and the maximal proper faces are called facets.

Convex polytopes have a natural CW complex structure where the k-skeleton consists of the union of all the faces of dimension k or less. Faces of a convex polytope are convex polytopes themselves.

Polytopes

Theorem

Let $||\cdot||$ be a seminorm on $\mathbb{R}^n$ and $\mathcal{P}$ a convex polytope embedded in $\mathbb{R}^n$. Then there exists a vertex v of $\mathcal{P}$, such that $||v|| = M = \max_{x \in \mathcal{P}} ||x||$.

Proof:
Suppose $\mathcal{P}$ is p-dimensional, $0 \leq p \leq n$. Let $S^{(q)}(\mathcal{P})$ be the q-skeleton of $\mathcal{P}$. Say that $x \in S^{(q)}$ such that $||x|| = M$. If $q = 0$, we are done. Otherwise, find $x_1, x_2 \in S^{(q-1)}(\mathcal{P})$, $t \in [0,1]$ such that $x = tx_1 + (1-t)x_2$. Such two points always exist because each face of $\mathcal{P}$ is convex.

$$M = ||x|| = ||tx_1 + (1-t)x_2|| \leq t||x_1|| + (1-t)||x_2|| \leq M$$

Thus $||x_1|| = ||x_2|| = M$. If $q - 1 = 0$, we are done. If not, simply repeat the process until we reach a vertex.

Polytopes

Define

$$U_1^{n-1} = \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^{n} x_i = 0 \right\} \cap \{ x \in \mathbb{R}^n \mid ||x||_1 \leq 1 \}$$

$$U_\infty^{n-1} = \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^{n} x_i = 0 \right\} \cap \{ x \in \mathbb{R}^n \mid ||x||_\infty \leq 1 \}$$

Then

$$\tau_1(S) = \max_{x \in U_1^{n-1}} ||S^T x||_1$$

$$\tau_\infty(S) = \max_{x \in U_\infty^{n-1}} ||S^T x||_\infty$$

Polytopes

Theorem

U_1^{n-1} is a convex polytope with vertices of form $\frac{1}{2}(e_i - e_j)$ where $i \neq j$.

Theorem

U_∞^{n-1} is a convex polytope. If n is even, the vertices are points in $\mathbb{R}^n$ such that $\frac{n}{2}$ coefficients are equal to 1 and the other $\frac{n}{2}$ are equal to -1. If n is odd, the vertices are points in $\mathbb{R}^n$ such that $\frac{\lfloor n \rfloor}{2}$ coefficients are 1, a different $\frac{\lfloor n \rfloor}{2}$ many coefficients are -1 and the last remaining coefficient is equal to 0.

Polytopes

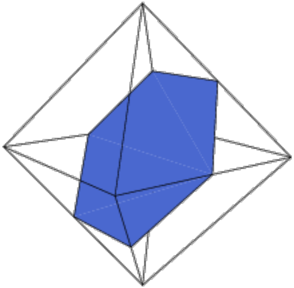
For a fixed stochastic matrix S, the functional $x \mapsto ||S^T x||_p$ is a seminorm. So by the previous theorems, we have that

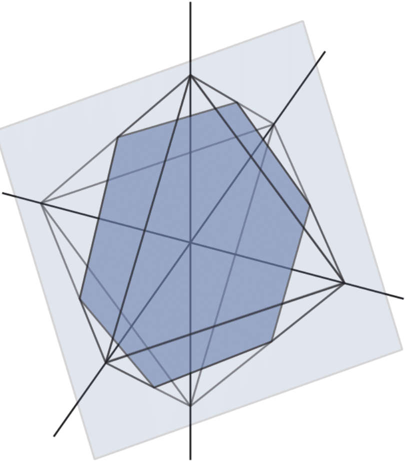
$$\tau_1(S) = \frac{1}{2} \max_{i,j} \sum_{k=1}^{n} |S_{ik} - S_{jk}|$$

$$\tau_\infty(S) = \max_{\phi \in E_n} \max_{1 \le k \le d} \left| \sum_{i=1}^{d} \phi(i) S_{ik} \right|$$

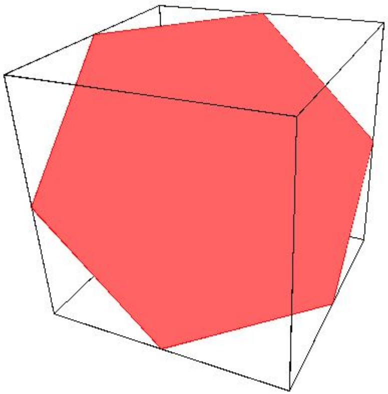
where E_n is the set of vertices of U_∞^{n-1}.
This gives an explicit form of the ergodicity coefficients that is useful for computations.

Polytopes

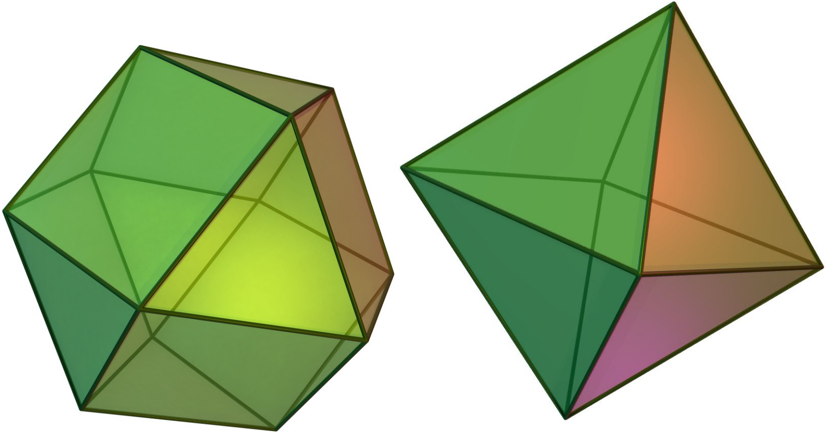




Polytopes



Polytopes



Picture References

- https://ckrao.wordpress.com/2015/02/27/cross-sections-of-a-cube/
- http://mathworld.wolfram.com/RegularOctahedron.html
- https://en.wikipedia.org/wiki/Cuboctahedron
- https://en.wikipedia.org/wiki/Octahedron
- Dr. Senger

References I

Ipsen, Ilse CF and Selee, Teresa M, *Ergodicity coefficients defined by vector norms*, SIAM Journal on Matrix Analysis and Applications (2011): 153-200