Coefficients of Ergodicity
An Introduction

Ilse Ipsen

Joint work with: Teresa Selee
Thanks to: Carl Meyer
Ergodicity

Long term behavior of dynamical systems

Homogeneous & inhomogeneous Markov chains:

Coefficient of ergodicity:
If $S$ is stochastic matrix then

$$\tau(S) = 1 - \min_{i,j} \sum_{k} \min\{s_{ik}, s_{jk}\}$$

Convergence rate of products

to a matrix with equal rows

(Doeblin 1937, Paz 1971, Iosifescu 1980)
Coefficients of ergodicity =
Bounds for eigenvalues and singular values

- Stochastic matrices and ergodicity
- Coefficient of ergodicity for stochastic matrices
- Applications
- Extension to general matrices
- Bounds for eigenvalues and singular values
- Eigenvalue inclusion regions
Square matrix $S$ is **stochastic** if

- Elements are non-negative: $S \geq 0$
- Rows sum to one: $S\mathbf{1} = \mathbf{1}$

**Power method:** $S^k \rightarrow ?$ as $k \rightarrow \infty$

If $S$ is **irreducible** then $S^k \rightarrow \mathbf{1}v^T$ all rows the same

$S^k \rightarrow \text{rank one}$

Markov chain is **ergodic**

**Ergodicity:** powers converge to **rank-one** matrix
Rate of Convergence

Eigenvalues of stochastic $S$: $1 \geq |\lambda_2| \geq |\lambda_3| \geq \ldots$

If $S$ irreducible then $\lambda_2 \neq 1$

Dominant right eigenvector: $S \mathbf{1} = \mathbf{1}$

Dominant left eigenvector: $\pi^T S = \pi^T$, $\pi > 0$, $\|\pi\|_1 = 1$

If $|\lambda_2| < 1$ then $S^k \to \mathbf{1}\pi^T$ as $k \to \infty$

$|\lambda_2(S^k)| = |\lambda_2|^k \to 0$

$|\lambda_2|$: asymptotic convergence rate to rank one matrix

Subdominant eigenvalue is measure of ergodicity

(Rothblum & Tan 1985, Gross & Rothblum 1993)
Products of Different Matrices

Different stochastic matrices $S_j$

How fast $\prod_{j=1}^k S_j \rightarrow \text{rank one}$ as $k \rightarrow \infty$?

Cannot use eigenvalues: $\lambda_2 \left( \prod_j S_j \right) \neq \prod_j \lambda_2(S_j)$

Need a replacement for $|\lambda_2|$ with some kind of multiplicative property recognizes when matrix has rank one

$\Rightarrow$ Coefficient of ergodicity
Coefficients of Ergodicity

General Definition (Seneta 1973)

Continuous scalar function $\mu(\cdot)$
Defined for stochastic matrices $S$
$0 \leq \mu(S) \leq 1$

Proper: $\mu(S) = 0 \iff \text{rank}(S) = 1$

Example (Seneta 1979)

$\tau(S) = \max_{z^T \mathbf{1} = 0, \|z\|_1 = 1} \|S^T z\|_1$

where $\|z\|_1 = \sum_i |z_i|$
2 × 2 Stochastic Matrices

\[ \tau(S) = \max_z \|S^T z\|_1 \]

where

\[
z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \quad z_1 + z_2 = 0 \quad |z_1| + |z_2| = 1 \quad z^T \mathbf{1} = 0
\]

therefore

\[ z = \pm \begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix} \]

\[ S = \begin{pmatrix} s_{11} & 1 - s_{11} \\ s_{21} & 1 - s_{21} \end{pmatrix} \]

\[ \tau(S) = \frac{1}{2} (|s_{11} - s_{21}| + |(1 - s_{11}) - (1 - s_{21})|) \]

\[ = |s_{11} - s_{21}| \]
Properties

\[ \tau(S) = \max_{z^T \mathbf{1} = 0, \|z\|_1 = 1} \|S^T z\|_1 \]

- \(0 \leq \tau(S) \leq 1\) \hspace{1cm} (bounded)

\[ \tau(S) \leq \|S^T\|_1 = \|S\|_\infty = 1 \]

- \(|\tau(S_1) - \tau(S_2)| \leq \|S_1 - S_2\|_\infty\) \hspace{1cm} (continuous)

Same proof as for norms

- \(\tau(S) = 0\) if and only \(\text{rank}(S) = 1\) \hspace{1cm} (proper)

If \(\tau(S) = 0\) then

\[ S^T z = 0 \text{ for } z = \frac{1}{2}(e_i - e_j) \]

All columns of \(S^T\) equal \(\text{rank}(S) = 1\)
More Properties

\[ \tau(S) = \max_{z^T \mathbf{1} = 0, \|z\|_1 = 1} \|S^T z\|_1 \]

- If \( \lambda \neq 1 \) eigenvalue of \( S \) then \( |\lambda| \leq \tau(S) \)

**Bound for subdominant eigenvalues**

Left eigenvectors of \( \lambda \perp \) right eigenvectors of \( 1 \)

\[ y^T S = \lambda y^T \quad S \mathbf{1} = \mathbf{1} \quad \implies \quad y^T \mathbf{1} = 0 \]

- \( \tau(S_1 S_2) \leq \tau(S_1) \tau(S_2) \) \hspace{1cm} (submultiplicative)

**Same proof as for norms**

- **Explicit expression**

\[ \tau(S) = \frac{1}{2} \max_{i,j} \|S^T (e_i - e_j)\|_1 \]

\[ = 1 - \min_{i,j} \sum \min\{s_{ik}, s_{jk}\} \]
2 × 2 Stochastic Matrices

\[ S = \begin{pmatrix} s_{11} & 1 - s_{11} \\ s_{21} & 1 - s_{21} \end{pmatrix} \]

\[ \tau(S) = \frac{1}{2} \left( |s_{11} - s_{21}| + |(1 - s_{11}) - (1 - s_{21})| \right) \]

\[ = |s_{11} - s_{21}| \]

\[ = |\lambda_2(S)| \]

If \( S \) is irreducible then \( 0 < s_{11} < 1, \ 0 < s_{21} < 1 \)

\[ \tau(S) = |\lambda_2(S)| < 1 \]

In general: when is \( \tau(S) < 1? \)
When is $\tau(S)$ useful?

**Scrambling matrices:** stochastic $S$ with $\tau(S) < 1$

- **Markov matrices:** at least one positive column
  
  $$
  S = \begin{pmatrix}
  0 & \frac{1}{2} & \frac{1}{2} \\
  0 & 0 & 1 \\
  0 & 0 & 1
  \end{pmatrix}
  $$

  Eigenvalues: 1, 0, 0
  
  $\tau(S) = \frac{1}{2} \left| \text{row 1} - \text{row 2} \right| = \frac{1}{2}$

- **Positive matrices:** all elements positive
Scrambling and Ergodicity

If stochastic matrices $S_j$ are scrambling

$$\tau(S_j) \leq \gamma < 1 \quad \text{for all } j$$

then (submultiplicativity)

$$\tau\left( \prod_{j=1}^{k} S_j \right) \leq \prod_{j=1}^{k} \tau(S_j) \leq \gamma^k$$

$$\prod_{j=1}^{k} S_j \rightarrow \text{rank one} \quad \text{with rate } \gamma \quad \text{as } k \rightarrow \infty$$

Markov chain (weakly) ergodic
Applications for $\tau$

$$\tau(S) = \max_{z^T \mathbb{1} = 0, \|z\|_1 = 1} \|S^T z\|_1$$

- Convergence rate of stochastic products
  $$\prod_{j=1}^{k} S_j \text{ as } k \rightarrow \infty$$

- PageRank computation
  $$G = \alpha S + (1 - \alpha) \mathbb{1} v^T, \quad 0 < \alpha < 1$$

  Convergence rate of power method bounded by
  $$\tau(G) \leq \alpha \tau(S) \leq \alpha$$

- Condition number for stationary distribution $\pi$
  $$\pi^T S = \pi^T \quad \pi \geq 0 \quad \|\pi\|_1 = 1$$
Condition Number for Stationary Distribution

Stochastic and irreducible matrices $S$ and $S + E$

$$\pi^T S = \pi^T \quad \pi > 0 \quad \|\pi\|_1 = 1$$

$$\omega^T (S + E) = \omega^T \quad \omega > 0 \quad \|\omega\|_1 = 1$$

If $\tau(S) < 1$ then

$$\frac{\|\omega - \pi\|_1}{\|\omega\|_1} \leq \frac{\|E\|_\infty}{1 - \tau(S)}$$

Condition number bound: $1/(1 - \tau(S))$ (Seneta 1988)

Eigenvector sensitivity: eigenvalue gap $1/(1 - |\lambda_2|)$
Possible Extensions

\[ \tau(S) = \max_{z^T \mathbf{1} = 0, \|z\|_1 = 1} \|S^T z\|_1 \]

- Different norms
- Larger class of matrices
- Vectors different from \( \mathbf{1} \)

Change norm to \( \| \cdot \|_\infty \)
Infinity-Norm Coefficient

\[ \tau_\infty(S) = \max_{z^T \mathbf{1} = 0, \|z\|_\infty = 1} \|S^T z\|_\infty \]

- **Continuity and perfect conditioning**
  \[ |\tau_\infty(S_1) - \tau_\infty(S_2)| \leq \|S_1 - S_2\|_1 \]

- **Proper**
  \[ \tau_\infty(S) = 0 \text{ if and only if } \text{rank}(S) = 1 \]

- **Eigenvalue bound**
  \[ |\lambda| \leq \tau_\infty(S) \text{ for any eigenvalue } \lambda \neq 1 \text{ of } S \]

- **Submultiplicativity**
  \[ \tau_\infty(S_1 S_2) \leq \tau_\infty(S_1) \tau_\infty(S_2) \]

But . . .
$\tau_\infty$ is unbounded

$$S = \begin{pmatrix}
1 & 0 & \ldots & 0 & 0 \\
0 & 0 & \ldots & 0 & 1 \\
1 & 0 & \ldots & 0 & 0 \\
0 & 0 & \ldots & 0 & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & 0 & 1
\end{pmatrix}$$

$\tau_\infty(S) = \|S\|_1 = n/2$

$\tau_\infty(S) > 1$ possible \hspace{1cm} (Seneta 79, Tan 82, Rhodius 88)

$\tau_\infty$ is not coefficient of ergodicity in the strict sense
General Definition (Seneta 1984)

Given:

- $m \times n$ complex matrix $A$
- $m \times 1$ complex vector $v$

\[
\tau_p(v, A) = \max_{z^*v=0, \|z\|_p=1} \|A^*z\|_p
\]

Conjugate transposes: $A^*$, $z^*$

- $S$ is stochastic, $v = \mathbb{1}$

\[
\tau_1(\mathbb{1}, S) = \tau(S) \quad \tau_\infty(\mathbb{1}, S) = \tau_\infty(S)
\]

- $A$ is non-negative with constant row sums

\[
\tau_1(\mathbb{1}, A) = \frac{1}{2} \max_{i,j} \|A^T(e_i - e_j)\|_1
\]
Eigenvalue Bounds (p-Norm)

\[ \tau_p(v, A) = \max_{z^* v = 0, \|z\|_p = 1} \|A^* z\|_p \]

If A is square

Eigenvalues: \( |\lambda_1| \geq |\lambda_2| \geq \ldots \)

Dominant eigenvector: \( A w = \lambda_1 w \)

Subdominant eigenvalue: \( \lambda \neq \lambda_1 \)

\[ |\lambda| \leq \tau_p(w, A) \]

\( \tau_p(w, A) \) is bound on subdominant eigenvalues
Singular Value Bounds (2-Norm)

\[ \tau_2(v, A) = \max_{z^*v=0, \|z\|_2=1} \|A^*z\|_2 \]

Singular values of A: \[ \sigma_1(A) \geq \sigma_2(A) \geq \ldots \]

For any \( v \)

\[ \sigma_2(A) \leq \tau_2(v, A) \leq \sigma_1(A) \]

\( \tau_2(v, A) \) is bound on subdominant singular values

Why?

\[ \sigma_2(A) = \sigma_2(A^*) = \min_w \max_{\|z\|_2=1, z^*w=0} \|A^*z\|_2 \]

\[ \sigma_2(A) = \min_w \tau_2(w, A) \]
Eigenvalue Bounds for Normal Matrices (2-Norm)

$$\tau_2(v, A) = \max_{z^* v = 0, \|z\|_2 = 1} \|A^* z\|_2$$

Eigenvalues of $A$: $|\lambda_1(A)| \geq |\lambda_2(A)| \geq \ldots$

If $A$ is normal (e.g. Hermitian, real symmetric) then for any $v$

$$|\lambda_2(A)| \leq \tau_2(v, A) \leq |\lambda_1(A)|$$

$\tau_2(v, A)$ is bound on subdominant eigenvalues

$$|\lambda_2(A)| = \min_w \tau_2(w, A)$$
Extension to Subspaces

Given:

- \( m \times n \) complex matrix \( A \)
- \( m \times k \) complex matrix \( V \)

\[
\tau_p(V, A) = \max_{z^*V = 0, \|z\|_p = 1} \|A^*z\|_p
\]

(Rothblum & Tan 1985, Hartfiel & Rothblum 1998)

\( z^*V = 0 \iff z \in \text{Null}(V^*) \)

\( \tau_p(V, A) = p\text{-norm of “A restricted to Null}(V^*)” \)
Bounds for Smaller Singular Values (2-Norm)

\[ \tau_2(V, A) = \max_{z^*V=0, \|z\|_2=1} \|A^*z\|_2 \]

Singular values of \( A \): \( \sigma_1(A) \geq \sigma_2(A) \geq \ldots \)

For any matrix \( V \) with \( k \) columns

\[ \sigma_{k+1}(A) \leq \tau_2(V, A) \leq \sigma_1(A) \]

In particular:

\[ \sigma_{k+1}(A) = \min_W \tau_2(W, A) \]

minimum ranges over all matrices \( W \) with \( k \) columns
Eigenvalue Inclusion Regions

\[ \tau_2(V, A) = \max_{z^*v = 0, \|z\|_2 = 1} \|A^*z\|_2 \]

\( n \times n \) normal matrix \( A \)

If \( V \) has \( k \) orthonormal columns then the disk

\[ |\lambda - \rho| \leq \tau_2(V, A - \rho I) \]

contains at least \( n - k \) eigenvalues of \( A \)

This is a special case of a **Lehmann bound**

(Lehmann 1963, Beattie & Ipsen 2003)
Summary

\[ \tau_p(v, A) = \max_{z^*v = 0, \|z\|_p = 1} \|A^*z\|_p \]

- Bound on subdominant singular values of A \((\text{if } p = 2)\)
- Bound on subdominant eigenvalues of A
  \(\text{if } v \text{ is dominant eigenvector}\)
- Measure of ergodicity
  Convergence rate of power method to rank-one
  \(\text{if } A \text{ stochastic, } p = 1, v = \mathbb{1}\)
Summary

\[ \tau_p(V, A) = \max_{z^* v = 0, \|z\|_p = 1} \|A^* z\|_p \]

\( V \) has \( k \) columns

- Bound on \( n - k \) smallest singular values of \( A \) (if \( p = 2 \))

- Eigenvalue inclusion region: Lehmann bound (if \( p = 2 \), \( A \) is normal, \( V \) has orthonormal columns)
The End