Lecture 14: Joint Spectral Radius
Stability of LTI Systems

Discrete-time linear time invariant system $x(t + 1) = Ax(t)$, $t \in \mathbb{N}$

- Growth of system solutions bounded by exponential rate $\rho$ if
  \[
  \exists K > 0 \text{ such that } \|x(t)\| \leq K\rho^t\|x(0)\|, \quad \forall t \in \mathbb{N}, \forall x(0)
  \]

- Tightest bound $\rho$ (if achievable) is the spectral radius of $A$:
  \[
  \rho(A) := \max_i |\lambda_i(A)| < 1
  \]

- $\rho(A)$ provides a quantitative stability metric
  - Qualitatively, system is stable iff $\rho(A) < 1$
Example

\[ x(t + 1) = \begin{bmatrix} 0.5 & 1 \\ 0 & 0.5 \end{bmatrix} x(t), \ x(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \] A is defective with \( \rho(A) = 0.5 \)

Bounding the exponential growth (in this case, decay) of \( \|x(t)\| \)

Exponential growth rate bounded by \( \rho(A) + \varepsilon, \ \forall \varepsilon > 0 \), but not by \( \rho(A) \)

\[ x(t) = \begin{bmatrix} t(0.5)^{t-1} \\ (0.5)^t \end{bmatrix} \quad \Rightarrow \quad \|x(t)\| = (0.5)^t \sqrt{1 + t^2/4} \leq K(0.5 + \varepsilon)^t \]
Stability of Switched Linear Systems

Discrete-time switched linear system \( x(t + 1) = A_{\sigma(t)}x(t), \ t \in \mathbb{N} \)

- **Arbitrary witching** among subsystems \( A_1, A_2, \ldots, A_m \)
- Solution under switching sequence \( \sigma(\cdot) \in \{1, \ldots, m\} \):
  \[
x_{\sigma}(t) = A_{\sigma(t-1)} \cdots A_{\sigma(0)}x(0), \quad t \in \mathbb{N}
  \]
- Growth of system solutions bounded by exponential rate \( \rho \) if
  \[
  \exists K > 0 \text{ such that } \|x_{\sigma}(t)\| \leq K \rho^t \|x(0)\|, \ \forall t, \forall x(0), \forall \sigma
  \]
- Tightest bound \( \rho^* \) (if achievable)?
  - \( \rho^* \) provides a quantitative metric of SLS’s stability
  - SLS is stable under arbitrary switching iff \( \rho^* < 1 \)
Example

SLS with two stable subsystems $A_1 = \begin{bmatrix} 0.5 & 1 \\ 0 & 0.5 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0.5 & 0 \\ 1 & 0.5 \end{bmatrix}$

- $\rho(A_1) = \rho(A_2) = 0.5$, which is the growth rate without switching

- $\rho(A_1A_2) = \rho \left( \begin{bmatrix} 1.25 & 0.5 \\ 0.5 & 0.25 \end{bmatrix} \right) = 1.4571$; hence average per-step growth rate under 2-periodic switching $\sigma = (2, 1, 2, 1, \ldots)$ is $\sqrt{1.4571} = 1.2071$

- $\rho(A_1A_2A_1) = 1.2374$; hence under 3-periodic $\sigma = (1, 2, 1, 1, 2, 1, \ldots)$, average per-step growth rate is $\sqrt[3]{1.2374} = 1.0736$

$\rho^* = 1.2071$, which is achieved by periodic switching $\sigma = (2, 1, 2, 1, \ldots)$
Growth Rate under Periodic Switching

For the matrix set \( \mathcal{A} = \{A_1, \ldots, A_m\} \), define

\[
\rho_t := \max_{A_{i_1}, \ldots, A_{i_t} \in \mathcal{A}} [\rho(A_{i_1} \cdots A_{i_t})]^{1/t}
\]

- Largest avg. per-step growth rate under \( t \)-periodic switching
- Generally increasing in \( k \) (but not monotonically)

Plot of \( \rho_t \) vs. \( t \) for the previous example
Another Example

\[ A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix} \]

<table>
<thead>
<tr>
<th>t</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho_t )</td>
<td>1.0000</td>
<td>1.1892</td>
<td>1.3415</td>
<td>1.3401</td>
<td>1.3190</td>
<td>1.3415</td>
<td>1.3428</td>
<td>1.3401</td>
</tr>
</tbody>
</table>

Plot of \( \rho_t \) vs. \( t \)
Joint Spectral Radius

Joint spectral radius* of the matrix set $\mathcal{A}$ is

$$\rho^*(\mathcal{A}) := \limsup_{t \to \infty} \rho_t$$

- $\limsup$ can be replaced by $\sup$
- Maximum (in limit) growth rate achieved by periodic switching

**Finiteness conjecture:** Can $\rho^*(\mathcal{A})$ always be achieved by a finite periodic switching?
- Answer is **no**†

---


Implications of JSR

Stability of SLS $x(t + 1) = A_{\sigma(k)}x(t)$

- Exponential stable under arbitrary switching iff $\rho^*(A) < 1$
- For any $\varepsilon > 0$, $\exists K > 0$ such that
  $$\|x_\sigma(t)\| \leq K (\rho^*(A) + \varepsilon)^t \|x(0)\|, \quad \forall t, \forall x(0), \forall \sigma$$

Other applications:

- Capacity of codes in coding theory
- Complexity of languages via counting overlap-free words
- Trackable graphs in graph theory
- Wavelets functions

---


Induced Matrix Norm

**Norm** of a matrix $A \in \mathbb{R}^{m \times n}$ **induced** by a vector norm $\| \cdot \|_{\mathbb{R}^n}$:

$$\|A\| := \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}}$$

- A norm on the matrix space $\mathbb{R}^{m \times n}$
- $\|Ax\| \leq \|A\| \cdot \|x\|$ for all $x \in \mathbb{R}^n$
- $\|AB\| \leq \|A\| \cdot \|B\|$ for all matrices $A, B$
- Example: $L_2$ (spectral) norm, $L_p$ norm, $L_\infty$ norm, ···

Induced matrix norm provides an over-estimate of spectral radius:

$$\|A\| \geq \rho(A)$$
Growth Estimate of LTI Solutions

Estimating the growth of solution to LTI system $x(t + 1) = Ax(t)$:

$$\|x(t)\| = \|A \cdots A \cdot x(0)\| \leq \|A\|^t \cdot \|x(0)\|, \forall t \in \mathbb{N}$$

- $\|A\|$ provides an upper bound of exponential growth rate of $\|x(t)\|$.
- Could be a terrible upper bound, e.g., $A = \begin{bmatrix} 0.5 & 100 \\ 0 & 0.5 \end{bmatrix}$.

Better estimate: sample LTI solution every $t$ steps:

$$\|x(kt)\| = \|(A^t)^k x(0)\| \leq \left(\|A^t\|^{1/t}\right)^{kt} \cdot \|x(0)\|, \forall k \in \mathbb{N}$$

- $\|A^t\|^{1/t}$ provides asymptotically tight upper bound:

$$\lim_{t \to \infty} \|A^t\|^{1/t} = \rho(A) \quad \text{(Gelfand’s Formula)}$$
Growth Estimate of SLS Solutions

For SLS $x(t + 1) = A_{\sigma(t)}x(t)$, its 1-step solution growth is at most

$$\|x(t + 1)\| \leq \|A_{\sigma(t)}\| \cdot \|x(t)\| \leq \left( \max_{A_i \in \mathcal{A}} \|A_i\| \right) \|x(t)\|$$

- $\hat{\rho}_1 := \max_{A_i \in \mathcal{A}} \|A_i\|$ provides an upper bound of JSR $\rho^*(\mathcal{A})$

Average per-step growth rate over $t$ steps is at most

$$\hat{\rho}_t := \left( \max_{A_{i_1}, \ldots, A_{i_t} \in \mathcal{A}} \|A_{i_1} \cdots A_{i_t}\| \right)^{1/t}$$

- $\|x(kt)\| \leq (\hat{\rho}_t)^{kt} \|x(0)\|$, $\forall k \in \mathbb{N}$
- $\hat{\rho}_t$ provides an (asymptotically tight) upper bound of JSR $\rho^*(\mathcal{A})$
Example

\[ A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix}. \]

\[ \hat{\rho}_{18} = 1.3654 \text{ (using spectral norm)}, \quad \rho_{18} = 1.3423 \]

Plot of \( \hat{\rho}_t \) (stars) and \( \rho_t \) (circles) vs. \( t \)
**Alternative Definition of JSR**

**Joint spectral radius** of the matrix set $\mathcal{A} = \{A_1, \ldots, A_m\}$ is

$$
\rho^*(\mathcal{A}) = \limsup_{t \to \infty} \hat{\rho}_t = \limsup_{t \to \infty} \max_{A_{i_1}, \ldots, A_{i_t} \in \mathcal{A}} \|A_{i_1} \cdots A_{i_t}\|^{1/t}
$$

- Original definition of JSR in [Rota & Strang’60]
- Proved to be equivalent in [Berger & Wang’92]
- $\limsup$ can be replaced with either $\lim$ or $\inf$
- Choice of induced matrix norm $\| \cdot \|$ does not matter
Computation of JSR

Brute force method: compute $\rho_t$ and $\hat{\rho}_t$ for large enough $t$

- Convergence may be slow (hence a large $t$ is needed)
- Requires enumerating $m^t$ matrix products
- Computation time grows exponentially in $t$

**Negative Result:** The problem of approximating the JSR of a pair of binary matrices in $\mathbb{R}^{n \times n}$ with an error less than $2^{-k}$ is NP hard.

- Unless $P = NP$, no algorithm with complexity polynomial in $(n, k)$
- Proved by reducing to the NP-complete Boolean Satisfaction (SAT) problem
- Does not rule out fast computation for special cases

---

Upper Bound of JSR

Recall that $\hat{\rho}_1$ is an upper bound of JSR:

$$\rho^* \leq \max_{A_i \in \mathcal{A}} \| A_i \|$$

- Different choices of $\| \cdot \|$ result in different upper bounds

Theorem (Kozyakin)

$$\rho^* = \inf_{\| \cdot \|} \max_{i \in \mathcal{M}} \| A_i \|$$

- Infimum ranges over all induced matrix norms
- Infimum not always achievable
- Norm achieving the infimum is called an extremal norm

\[ \dagger \text{“Structure of Extremal Trajectories of Discrete Linear Systems and the Finiteness Conjecture,” , V. Kozyakin, Automat Remote Control, 68 (2007), no. 1, 174-209} \]
Upper Bound via Ellipsoid Norms

Ellipsoid norm $\|x\|_P = \sqrt{x^T P x}$ with $P \succ 0$ induces a matrix norm

$$\|A\|_P = \sup_{x \neq 0} \frac{\|Ax\|_P}{\|x\|_P}$$

- Condition $\gamma \geq \|A\|_P$ is equivalent to the LMI $\gamma^2 P - A_i^T P A_i \succeq 0$

Upper bound of JSR given by $\| \cdot \|_P$:

$$\rho^* \leq \rho_P := \max_{A_i \in \mathcal{A}} \|A_i\|_P$$

To compute tightest upper bound, solve the SDP (with $\gamma$ given):

$$\text{find } P \succ 0$$

such that $\gamma^2 P - A_i^T P A_i \succeq 0, \forall i \in \mathcal{M}$

- Solved for decreasing $\gamma$ until the problem becomes infeasible
Quality of Ellipsoid Upper Bounds

Guaranteed quality: the tightest upper bound $\rho^*_P$ produced by ellipsoid norms satisfies
\[
\frac{1}{\sqrt{n}} \rho^*_P \leq \rho^* \leq \rho^*_P
\]
- This follows from Fritz John’s Theorem: a symmetric convex body in $\mathbb{R}^n$ can be sandwiched by two ellipsoids differing by a scaling of at most $\sqrt{n}$

Optimality: $\rho^*_P = \rho^*$ in the following cases‡
- if all $A_i$’s are symmetric: $\rho(A) = \max_i \rho(A_i)$ with $P = I$
- if all $A_i$’s are upper triangular: $\rho(A) = \max_i \rho(A_i)$ with $P$ diagonal
- if $A = \{A, A^T\}$: $\rho(A) = \sqrt{\lambda_{\max}(AA^T)}$

Example:
\[
A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix}, \quad P = \begin{bmatrix} 6.958 & 5.347 \\ 5.347 & 14.409 \end{bmatrix}, \quad \rho_P = 1.52
\]

Other (richer) families of norms

- **Composite ellipsoid norms**: Given $P_1, \ldots, P_\ell \succ 0$
  \[
  \|x\|_P := \max_{j=1,\ldots,\ell} \sqrt{x^T P_j x}
  \]
  whose unit ball is an intersection of ellipsoids

- **Polytope norm**: given $c_1, \ldots, c_\ell \in \mathbb{R}^n$
  \[
  \|x\|_c := \max_{j=1,\ldots,\ell} |c_j^T x|
  \]
  whose unit ball is a polytope

In theory, upper bounds of JSR using the above norms can be arbitrarily tight, though at the price of increased computational complexity
JSR via SOS

Idea: for a positive homogeneous polynomial $p(x)$ of degree $2d$

$$p(A_ix) \leq \gamma^{2d} p(x), \; \forall x, \; \forall A_i \in A \; \Rightarrow \; \rho^*(A) \leq \gamma$$

SOS Formulation: for fixed $d$

minimize $\gamma$

such that $p(x)$ is SOS

$$\gamma^{2d} p(x) - p(A_i x) \text{ is SOS, } i = 1, \ldots, m$$

- Quasiconvex problem: for fixed $\gamma$, constraint is convex

Quality: Solution $\rho^*_{\text{sos}}$ satisfies

$$\left(\frac{n+d-1}{d}\right)^{-\frac{1}{2d}} \rho^*_{\text{sos}} \leq \rho^* \leq \rho^*_{\text{sos}}$$

Example

\[ A_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}. \]

- \( \rho^* = 1 \)
- Using ellipsoid norms, \( \rho_P^* = \sqrt{2} \)

Using (nonconvex) SOS \( V(x) = (x_1^2 - x_2^2)^2 + \varepsilon(x_1^2 + x_2^2)^2 \), we have

\[
V(A_1x) - (1 + \varepsilon)V(x) = -(x_2^2 - x_1^2 + \varepsilon(x_1^2 + x_2^2))^2 \leq 0 \\
V(A_2x) - (1 + \varepsilon)V(x) = -(x_1^2 - x_2^2 + \varepsilon(x_1^2 + x_2^2))^2 \leq 0
\]

Thus, \( \rho^* \leq (1 + \varepsilon)^{\frac{1}{4}} \) for any \( \varepsilon > 0 \), i.e., \( \rho^* \leq 1 \)