Lyapunov Exponent of Rank One Matrices: Ergodic Formula and Inapproximability of the Optimal Distribution

Jason Altschuler

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology

Based on joint work with **Pablo Parrilo** (MIT)

CDC - December 2019

Stochastic linear systems

Given a set $\mathcal{A} = \{A_1, \ldots, A_n\}$ of square matrices and a probability distribution p over $\{1, \ldots, n\}$, consider

$$x_{k+1} = A_{\sigma_k} x_k$$

where $\sigma_k \in \{1, \ldots, n\}$ are i.i.d. from p.

Stochastic linear systems

Given a set $\mathcal{A} = \{A_1, \ldots, A_n\}$ of square matrices and a probability distribution p over $\{1, \ldots, n\}$, consider

$$x_{k+1} = A_{\sigma_k} x_k$$

where $\sigma_k \in \{1, \ldots, n\}$ are i.i.d. from p.

Appears in several different contexts: control theory, dynamical systems, ergodic theory, computer image generation of fractals, etc.

Stochastic linear systems

Given a set $\mathcal{A} = \{A_1, \ldots, A_n\}$ of square matrices and a probability distribution p over $\{1, \ldots, n\}$, consider

$$x_{k+1} = A_{\sigma_k} x_k$$

where $\sigma_k \in \{1, \ldots, n\}$ are i.i.d. from p.

Appears in several different contexts: control theory, dynamical systems, ergodic theory, computer image generation of fractals, etc.

Two key problems:

- **Analysis problem:** Given $(\mathcal{A}, p)$, compute convergence rate.
- **Design problem:** Given $\mathcal{A}$, optimize convergence rate (by designing p).

Connection to the Lyapunov exponent

What is the "convergence rate" of the stochastic linear system $x_{k+1} = A_{\sigma_k} x_k$?

$$\underbrace{R_p(\mathcal{A})}_{\textbf{Lyapunov spectal radius}} := \lim_{k\to\infty} \|x_k\|^{1/k} = \lim_{k\to\infty} \|A_{\sigma_k} \cdots A_{\sigma_2} A_{\sigma_1}\|^{1/k}$$

Connection to the Lyapunov exponent

What is the "convergence rate" of the stochastic linear system $x_{k+1} = A_{\sigma_k} x_k$?

$$\underbrace{R_p(\mathcal{A})}_{\textbf{Lyapunov spectal radius}} := \lim_{k\to\infty} \|x_k\|^{1/k} = \lim_{k\to\infty} \|A_{\sigma_k} \cdots A_{\sigma_2} A_{\sigma_1}\|^{1/k}$$

- (Furstenberg-Kesten 1960)[1] $R_p(\mathcal{A}) = e^{\lambda_p(\mathcal{A})}$ a.s., where

$$\underbrace{\lambda_p(\mathcal{A})}_{\textbf{Lyapunov exponent}} := \lim_{k\to\infty} \frac{1}{k} \mathbb{E}\left[\log \|A_{\sigma_k} \cdots A_{\sigma_2} A_{\sigma_1}\|\right]$$

[1] Under mild conditions.

Connection to the Lyapunov exponent

What is the "convergence rate" of the stochastic linear system $x_{k+1} = A_{\sigma_k} x_k$?

$$\underbrace{R_p(\mathcal{A})}_{\textbf{Lyapunov spectal radius}} := \lim_{k\to\infty} \|x_k\|^{1/k} = \lim_{k\to\infty} \|A_{\sigma_k} \cdots A_{\sigma_2} A_{\sigma_1}\|^{1/k}$$

- (Furstenberg-Kesten 1960)[1] $R_p(\mathcal{A}) = e^{\lambda_p(\mathcal{A})}$ a.s., where

$$\underbrace{\lambda_p(\mathcal{A})}_{\textbf{Lyapunov exponent}} := \lim_{k\to\infty} \frac{1}{k}\mathbb{E}\left[\log \|A_{\sigma_k} \cdots A_{\sigma_2} A_{\sigma_1}\|\right]$$

- Stability characterization: converges iff $e^{\lambda_p(\mathcal{A})} < 1$, i.e.

$$x_{k+1} = A_{\sigma_k} x_k \text{ is stable} \iff \lambda_p(\mathcal{A}) < 0$$

[1] Under mild conditions.

Stochastic linear systems are hard...

Analysis problem: Given $(\mathcal{A}, p)$, **compute** convergence rate $e^{\lambda_p(\mathcal{A})}$.

Design problem: Given $\mathcal{A}$, **optimize** convergence rate $\min_{p \in \Delta_n} \lambda_p(\mathcal{A})$.

Stochastic linear systems are hard...

Analysis problem: Given $(\mathcal{A}, p)$, **compute** convergence rate $e^{\lambda_p(\mathcal{A})}$.

- Deciding stability (i.e. if $\lambda_p(\mathcal{A}) < 0$) is undecidable [Tsitsiklis-Blondel 1997].
- Still: how to compute/approximate? Special cases?

Design problem: Given $\mathcal{A}$, **optimize** convergence rate $\min_{p\in\Delta_n} \lambda_p(\mathcal{A})$.

Stochastic linear systems are hard...

Analysis problem: Given $(\mathcal{A}, p)$, **compute** convergence rate $e^{\lambda_p(\mathcal{A})}$.

- Deciding stability (i.e. if $\lambda_p(\mathcal{A}) < 0$) is undecidable [Tsitsiklis-Blondel 1997].
- Still: how to compute/approximate? Special cases?

Design problem: Given $\mathcal{A}$, **optimize** convergence rate $\min_{p \in \Delta_n} \lambda_p(\mathcal{A})$.

- Deciding stabilizability (i.e. if $\min_{p \in \Delta_n} \lambda_p(\mathcal{A}) < 0$) is NP-hard [A.-Parrilo 2019].
- Hard even for "simple" case of rank one matrices, in contrast to analogous optimization for the Joint Spectral Radius!

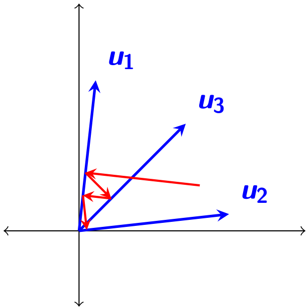
Rank one setting

Consider simple setting: symmetric, rank one matrices $\mathcal{A} = \{A_i = u_i u_i^T\}_{i=1}^n$.

Rank one setting

Consider simple setting: symmetric, rank one matrices $\mathcal{A} = \{A_i = u_i u_i^T\}_{i=1}^n$.

Geometrically, the stochastic linear system $x_{k+1} = A_{\sigma_k} x_k = u_{\sigma_k}(u_{\sigma_k}^T x_k)$ corresponds to projecting state x_k on random lines u_{σ_k}.[2]



[2] Assuming w.l.o.g. that each $\|u_i\| = 1$, since it is easy to compute the effect of re-normalizing the matrices on the Lyapunov exponent.

Rank one: analysis problem

Theorem (Lyapunov exponent of rank one matrices). Let $\mathcal{A} = \{A_i = u_i u_i^T\}_{i=1}^n$. Then

$$\lambda_p(\mathcal{A}) = \sum_{ij=1}^n p_i p_j \log |u_i^T u_j|$$

Rank one: analysis problem

Theorem (Lyapunov exponent of rank one matrices). Let $\mathcal{A} = \{A_i = u_i u_i^T\}_{i=1}^n$. Then

$$\lambda_p(\mathcal{A}) = \sum_{ij=1}^n p_i p_j \log |u_i^T u_j|$$

- Ergodic formula: quadratic form in p.
  - In contrast, general case has no simple formula.

Rank one: analysis problem

Theorem (Lyapunov exponent of rank one matrices). Let $\mathcal{A} = \{A_i = u_i u_i^T\}_{i=1}^n$. Then

$$\lambda_p(\mathcal{A}) = \sum_{ij=1}^n p_i p_j \log |u_i^T u_j|$$

- Ergodic formula: quadratic form in p.
  - In contrast, general case has no simple formula.
- Computable, and in polynomial time.
  - In contrast, general case is undecidably hard to approximate.

Rank one: analysis problem

Theorem (Lyapunov exponent of rank one matrices). Let $\mathcal{A} = \{A_i = u_i u_i^T\}_{i=1}^n$. Then

$$\lambda_p(\mathcal{A}) = \sum_{ij=1}^n p_i p_j \log |u_i^T u_j|$$

- Ergodic formula: quadratic form in p.
  - In contrast, general case has no simple formula.
- Computable, and in polynomial time.
  - In contrast, general case is undecidably hard to approximate.
- Depends only on products $A_i A_j$ of length two.
  - In contrast, JSR of rank one matrices depends on products of length n.

Rank one: analysis problem

Theorem (Lyapunov exponent of rank one matrices). Let $\mathcal{A} = \{A_i = u_i u_i^T\}_{i=1}^n$. Then

$$\lambda_p(\mathcal{A}) = \sum_{ij=1}^n p_i p_j \log |u_i^T u_j|$$

- Ergodic formula: quadratic form in p.
  - In contrast, general case has no simple formula.
- Computable, and in polynomial time.
  - In contrast, general case is undecidably hard to approximate.
- Depends only on products $A_i A_j$ of length two.
  - In contrast, JSR of rank one matrices depends on products of length n.
- Proof idea: average time spent on edges of weighted graph.

Rank one: design problem

Theorem (Hardness of optimizing the Lyapunov exponent). Given a set $\mathcal{A}$ of n symmetric rank-one matrices, it is NP-hard to decide if

$$\min_{p \in \Delta_n} \lambda_p(\mathcal{A}) \overset{?}{<} 0$$

Rank one: design problem

Theorem (Hardness of optimizing the Lyapunov exponent). Given a set $\mathcal{A}$ of n symmetric rank-one matrices, it is NP-hard to decide if

$$\min_{p \in \Delta_n} \lambda_p(\mathcal{A}) \overset{?}{<} 0$$

- Thus, NP-hard to decide stabilizability of $x_{k+1} = A_{\sigma_k} x_k$.

Rank one: design problem

Theorem (Hardness of optimizing the Lyapunov exponent). Given a set $\mathcal{A}$ of n symmetric rank-one matrices, it is NP-hard to decide if

$$\min_{p \in \Delta_n} \lambda_p(\mathcal{A}) \overset{?}{<} 0$$

- Thus, NP-hard to decide stabilizability of $x_{k+1} = A_{\sigma_k} x_k$.
- Also, NP-hard even to approximate $\min_{p \in \Delta_n} \lambda_p(\mathcal{A})$.

Rank one: design problem

Theorem (Hardness of optimizing the Lyapunov exponent). Given a set $\mathcal{A}$ of n symmetric rank-one matrices, it is NP-hard to decide if

$$\min_{p \in \Delta_n} \lambda_p(\mathcal{A}) \overset{?}{<} 0$$

- Thus, NP-hard to decide stabilizability of $x_{k+1} = A_{\sigma_k} x_k$.
- Also, NP-hard even to approximate $\min_{p \in \Delta_n} \lambda_p(\mathcal{A})$.
- Proof idea: minimizing quadratic form over simplex. Reduce from Motzkin-Straus formulation of Independent Set.

Rank one: design problem

Theorem (Hardness of optimizing the Lyapunov exponent). Given a set $\mathcal{A}$ of n symmetric rank-one matrices, it is NP-hard to decide if

$$\min_{p \in \Delta_n} \lambda_p(\mathcal{A}) \overset{?}{<} 0$$

- Thus, NP-hard to decide stabilizability of $x_{k+1} = A_{\sigma_k} x_k$.
- Also, NP-hard even to approximate $\min_{p \in \Delta_n} \lambda_p(\mathcal{A})$.
- Proof idea: minimizing quadratic form over simplex. Reduce from Motzkin-Straus formulation of Independent Set.
- $\lambda_p(\mathcal{A})$ neither convex/concave in p. (Connections to non-metrizability of the Martin distance on $(1, d)$ Grassmanian...)

Summary

- **Background.** Lyapunov exponent $\lambda_p(\mathcal{A})$ dictates convergence rate of stochastic linear system $x_{k+1} = A_{\sigma_k} x_k$.
  - Analysis problem of *computing* convergence rate — known to be hard.
  - Design problem of *optimizing* convergence — open problem.

Summary

- **Background.** Lyapunov exponent $\lambda_p(\mathcal{A})$ dictates convergence rate of stochastic linear system $x_{k+1} = A_{\sigma_k} x_k$.
  - Analysis problem of *computing* convergence rate — known to be hard.
  - Design problem of *optimizing* convergence — open problem.

- **This paper.** Focusing on rank-one matrices reveals fundamental properties.
  - Analysis problem: simple ergodic formula.
  - Design problem: NP-hard.
  - Along the way, uncover other properties: convexity/concavity, special cases when computable, surprising differences with deterministic analogue, etc.

Summary

- **Background.** Lyapunov exponent $\lambda_p(\mathcal{A})$ dictates convergence rate of stochastic linear system $x_{k+1} = A_{\sigma_k} x_k$.
  - Analysis problem of *computing* convergence rate — known to be hard.
  - Design problem of *optimizing* convergence — open problem.

- **This paper.** Focusing on rank-one matrices reveals fundamental properties.
  - Analysis problem: simple ergodic formula.
  - Design problem: NP-hard.
  - Along the way, uncover other properties: convexity/concavity, special cases when computable, surprising differences with deterministic analogue, etc.

- **Extensions.** Techniques extend to more exotic settings. Optimizing convergence still NP-hard for exchangeable processes, but poly-time for Markov processes.