Dynamic programming operators over noncommutative spaces: an approach to optimal control of switched systems

Stéphane Gaubert* Nikolas Stott**

Stephane.Gaubert@inria.fr
*: INRIA and CMAP, Ecole polytechnique, IP Paris, CNRS
**: LocalSolver

ICODE Workshop
Jan. 8-10, 2020
Univ. Paris-Diderot

References: SG, NS arXiv:1706.04471, in CDC2017; SG, NS arXiv:1805.03284, in Math. Control Related Fields (2020); NS PhD thesis; NS arXiv:1612.05664, in Proc. AMS; X. Allamigeon, SG, E. Goubault, S. Putot, NS; A scalable algebraic method to infer quadratic invariants of switched systems, ACM Transactions on Embedded Computing Systems (TECS), Volume 15 Issue 4, August 2016

classical dynamic programming	
$\mathbb{R}^n$	
lattice order $\leqslant$	
probability measures	
Markov operator	
$P \geqslant 0$, $Pe = e$	
value function	
Bellman operator	
$[T(v)]_i = \max_j (A_{ij} + v_j)$	

classical dynamic programming	"noncommutative" dynamic programming
$\mathbb{R}^n$	S_n, symmetric matrices
lattice order $\leqslant$	Loewner order $(X \succcurlyeq 0 \iff \lambda_{\min}(X) \geqslant 0)$
probability measures	
Markov operator	
$P \geqslant 0$, $Pe = e$	
value function	
Bellman operator	
$[T(v)]_i = \max_j (A_{ij} + v_j)$	

classical dynamic programming	“noncommutative” dynamic programming
$\mathbb{R}^n$	S_n, symmetric matrices
lattice order $\leqslant$	Loewner order ($X \succcurlyeq 0 \iff \lambda_{\min}(X) \geqslant 0$)
probability measures	density matrices
Markov operator	Quantum channel
$P \geqslant 0$, $Pe = e$	$K(X) = \sum_i A_i^* X A_i$, $\sum_i A_i A_i^* = I$
value function	
Bellman operator	
$[T(v)]_i = \max_j (A_{ij} + v_j)$	

classical dynamic programming	"noncommutative" dynamic programming
$\mathbb{R}^n$	S_n, symmetric matrices
lattice order $\leqslant$	Loewner order ($X \succcurlyeq 0 \iff \lambda_{\min}(X) \geqslant 0$)
probability measures	density matrices
Markov operator	Quantum channel
$P \geqslant 0$, $Pe = e$	$K(X) = \sum_i A_i^* X A_i$, $\sum_i A_i A_i^* = I$
value function	
Bellman operator	How do we fill this box ?
$[T(v)]_i = \max_j (A_{ij} + v_j)$	what can it be used for?

The joint spectral radius

$\mathcal{A} = \{A_1, \ldots, A_m\} \subset \mathbb{R}^{n \times n}$, largest growth rate:

$$\rho(\mathcal{A}) \coloneqq \lim_{k \to \infty} \sup_{A_{i_1}, \ldots A_{i_k} \in \mathcal{A}} \|A_{i_1} \cdots A_{i_k}\|^{1/k}.$$

The joint spectral radius

$\mathcal{A} = \{A_1, \dots, A_m\} \subset \mathbb{R}^{n\times n}$, largest growth rate:

$$\rho(\mathcal{A}) := \lim_{k\to\infty} \sup_{A_{i_1},\dots A_{i_k}\in\mathcal{A}} \|A_{i_1}\cdots A_{i_k}\|^{1/k}.$$

Theorem (Blondel-Tsitsiklis - 2000)

Unless P = NP, there is no polynomial-time computable function $\hat{\rho}$ of $\mathcal{A}$ and ε satisfying

$$|\rho(\mathcal{A}) - \hat{\rho}(\mathcal{A},\varepsilon)| \leqslant \varepsilon\rho(\mathcal{A})$$

even if $\mathcal{A}$ consists of 2 matrices with entries in $\{0,1\}$.

Theorem (Barabanov, 1988)

If the set $\mathcal{A}$ is irreducible, then there is a norm ν such that

$$\max_{i\in[m]} \nu(A_i x) = \rho(\mathcal{A})\nu(x) \ , \ \forall x \, .$$

Theorem (Barabanov, 1988)

If the set $\mathcal{A}$ is irreducible, then there is a norm ν such that

$$\max_{i\in[m]} \nu(A_i x) = \rho(\mathcal{A})\nu(x) \ , \forall x \, .$$

Special case of ergodic control problem. Continuous time version: reduction to an ergodic HJ PDE (Calvez, SG, Gabriel 2014).

Theorem (Barabanov, 1988)

If the set $\mathcal{A}$ is irreducible, then there is a norm ν such that

$$\max_{i\in[m]} \nu(A_i x) = \rho(\mathcal{A})\nu(x) \; , \; \forall x \, .$$

Special case of ergodic control problem. Continuous time version: reduction to an ergodic HJ PDE (Calvez, SG, Gabriel 2014).

Certifying a upper bound of the joint spectral radius

Find a norm ν such that

$$\max_{i\in[m]} \nu(A_i x) \leqslant \rho\nu(x) \; , \; \forall x \, .$$

Then $\rho(\mathcal{A}) \leqslant \rho$.

Theorem (Barabanov, 1988)

If the set $\mathcal{A}$ is irreducible, then there is a norm ν such that

$$\max_{i \in [m]} \nu(A_i x) = \rho(\mathcal{A}) \nu(x) \ , \ \forall x \, .$$

Special case of ergodic control problem. Continuous time version: reduction to an ergodic HJ PDE (Calvez, SG, Gabriel 2014).

Certifying a upper bound of the joint spectral radius

Find a norm ν such that

$$\max_{i \in [m]} \nu(A_i x) \leqslant \rho \nu(x) \ , \ \forall x \, .$$

Then $\rho(\mathcal{A}) \leqslant \rho$.

Goal

Construct a sequence of such norms ν_k such that the corresponding upper bounds ρ_k of $\rho(\mathcal{A})$ do converge to $\rho(\mathcal{A})$.

This talk

Use ideas / techniques from:

- max-plus basis methods Fleming, McEneaney, Akian, Dower, Kaise, Qu, SG, . . .

This talk

Use ideas / techniques from:

- max-plus basis methods Fleming, McEneaney, Akian, Dower, Kaise, Qu, SG, ...
- path-complete automata Ahmadi, Parrilo, Jungers, Roozbehani

This talk

Use ideas / techniques from:

- max-plus basis methods Fleming, McEneaney, Akian, Dower, Kaise, Qu, SG, ...
- path-complete automata Ahmadi, Parrilo, Jungers, Roozbehani
- polyhedral approximation: Guglielmi, Kozyakin, Protasov ...

This talk

Use ideas / techniques from:

- max-plus basis methods Fleming, McEneaney, Akian, Dower, Kaise, Qu, SG, ...
- path-complete automata Ahmadi, Parrilo, Jungers, Roozbehani
- polyhedral approximation: Guglielmi, Kozyakin, Protasov ...
- geometry of the Loewner order

This talk

Use ideas / techniques from:

- max-plus basis methods Fleming, McEneaney, Akian, Dower, Kaise, Qu, SG, ...
- path-complete automata Ahmadi, Parrilo, Jungers, Roozbehani
- polyhedral approximation: Guglielmi, Kozyakin, Protasov ...
- geometry of the Loewner order
- non-linear Perron-Frobenius theory, nonexpansive mappings Nussbaum, Baillon, Bruck, SG, Gunawardena,

This talk

Use ideas / techniques from:

- max-plus basis methods Fleming, McEneaney, Akian, Dower, Kaise, Qu, SG, ...
- path-complete automata Ahmadi, Parrilo, Jungers, Roozbehani
- polyhedral approximation: Guglielmi, Kozyakin, Protasov ...
- geometry of the Loewner order
- non-linear Perron-Frobenius theory, nonexpansive mappings Nussbaum, Baillon, Bruck, SG, Gunawardena,
- risk-sensitive control Anantharam, Borkar

This talk

Use ideas / techniques from:

- max-plus basis methods Fleming, McEneaney, Akian, Dower, Kaise, Qu, SG, ...
- path-complete automata Ahmadi, Parrilo, Jungers, Roozbehani
- polyhedral approximation: Guglielmi, Kozyakin, Protasov ...
- geometry of the Loewner order
- non-linear Perron-Frobenius theory, nonexpansive mappings Nussbaum, Baillon, Bruck, SG, Gunawardena,
- risk-sensitive control Anantharam, Borkar

This talk

Use ideas / techniques from:

- max-plus basis methods Fleming, McEneaney, Akian, Dower, Kaise, Qu, SG, ...
- path-complete automata Ahmadi, Parrilo, Jungers, Roozbehani
- polyhedral approximation: Guglielmi, Kozyakin, Protasov ...
- geometry of the Loewner order
- non-linear Perron-Frobenius theory, nonexpansive mappings Nussbaum, Baillon, Bruck, SG, Gunawardena,
- risk-sensitive control Anantharam, Borkar

to obtain a decreasing sequence of upper approximations of the joint spectral radius.

This talk

Use ideas / techniques from:

- max-plus basis methods Fleming, McEneaney, Akian, Dower, Kaise, Qu, SG, . . .
- path-complete automata Ahmadi, Parrilo, Jungers, Roozbehani
- polyhedral approximation: Guglielmi, Kozyakin, Protasov . . .
- geometry of the Loewner order
- non-linear Perron-Frobenius theory, nonexpansive mappings Nussbaum, Baillon, Bruck, SG, Gunawardena,
- risk-sensitive control Anantharam, Borkar

to obtain a decreasing sequence of upper approximations of the joint spectral radius.

$\rightarrow$ method little sensitive to the curse of dimensionality: can deal with instances up to dimension 500 (random matrices with real entries) and even up to dimension 5000 (random matrices with nonnegative entries)

Bounds arising from piecewise quadratic norms

Look for

$$\nu(x) = \max_{v \in V} \sqrt{x^T Q_v x}$$

with V finite set, such that

$$\max_{i \in [m]} \nu(A_i x) \leqslant \rho \nu(x) \ , \forall x \, .$$

Then $\rho(\mathcal{A}) \leqslant \rho$. (Ahmadi et al), related to McEneaney's max-plus basis method)

Bounds arising from piecewise quadratic norms

Look for

$$\nu(x) = \max_{v \in V} \sqrt{x^T Q_v x}$$

with V finite set, such that

$$\max_{i \in [m]} \nu(A_i x) \leqslant \rho \nu(x) \ , \ \forall x \, .$$

Then $\rho(\mathcal{A}) \leqslant \rho$. (Ahmadi et al), related to McEneaney's max-plus basis method)

Goal: find collection of matrices $(Q_v)_v$ such that

$$\max_{i \in [n], v \in V} x^T (A_i^T Q_v A_i) x \leqslant \max_{w \in V} x^T (\rho^2 Q_w) x$$

Bounds arising from piecewise quadratic norms

Look for

$$\nu(x) = \max_{v \in V} \sqrt{x^T Q_v x}$$

with V finite set, such that

$$\max_{i \in [m]} \nu(A_i x) \leqslant \rho \nu(x) \ , \forall x \, .$$

Then $\rho(\mathcal{A}) \leqslant \rho$. (Ahmadi et al), related to McEneaney's max-plus basis method)

Goal: find collection of matrices $(Q_v)_v$ such that

$$\max_{i \in [n], v \in V} x^T (A_i^T Q_v A_i) x \leqslant \max_{w \in V} x^T (\rho^2 Q_w) x$$

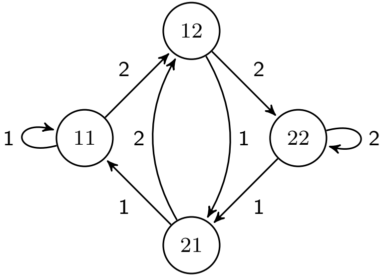
2 relaxations (Ahmadi et al.)

- For all v, i, there is w such that $A_i^T Q_v A_i \preccurlyeq \rho^2 Q_w$
- We enforce the choice of $w = \tau(v, i)$ for some transition map τ.

De Bruijn automaton, "concatenate and forget"

- Alphabet: $\Sigma := [m] = \{1, \ldots, m\}$, States: Σ^d
- Transition map τ_d:

$$\tau_d(v, i) = w \iff \begin{cases} v = i_1 i_2 \ldots i_d \\ w = i_2 \ldots i_d i \end{cases} .$$



Path-complete LMI automaton (Ahmadi et al.)

Solve family of LMIs:

$$(\mathcal{P}_\rho) \begin{cases} Q_v \succ 0 \ , \forall v \\ \rho^2 Q_w \succcurlyeq A_i^T Q_v A_i \ , \forall w = \tau_d(v, i) \end{cases}$$

Bisection:

$$\rho_d := \text{smallest } \rho \text{ such that } (\mathcal{P}_\rho) \text{ is feasible.}$$

Path-complete LMI automaton (Ahmadi et al.)

Solve family of LMIs:

$$(\mathcal{P}_\rho) \begin{cases} Q_v \succ 0 \ , \ \forall v \\ \rho^2 Q_w \succcurlyeq A_i^T Q_v A_i \ , \ \forall w = \tau_d(v, i) \end{cases}$$

Bisection:

$$\rho_d := \text{smallest } \rho \text{ such that } (\mathcal{P}_\rho) \text{ is feasible.}$$

Theorem (Ahmadi et al. - SICON 2014)

An optimal solution $(Q_v)_v$ *provides a norm*

$$\nu(x) = \max_v (x^T Q_v x)^{1/2}$$

such that

$$\rho_d \geqslant \rho(A) \geqslant \frac{1}{n^{\frac{1}{2(d+1)}}} \rho_d$$

(asymptotically exact as $d \to \infty$*).*

Path-complete LMI automaton (Ahmadi et al.)

Solve family of LMIs:

$$(\mathcal{P}_\rho) \begin{cases} Q_v \succ 0 \ , \forall v \\ \rho^2 Q_w \succcurlyeq A_i^T Q_v A_i \ , \forall w = \tau_d(v,i) \end{cases}$$

Bisection:

$$\rho_d := \text{smallest } \rho \text{ such that } (\mathcal{P}_\rho) \text{ is feasible.}$$

Theorem (Ahmadi et al. - SICON 2014)

An optimal solution $(Q_v)_v$ provides a norm

$$\nu(x) = \max_v (x^T Q_v x)^{1/2}$$

such that

$$\rho_d \geqslant \rho(A) \geqslant \frac{1}{n^{\frac{1}{2(d+1)}}} \rho_d$$

(asymptotically exact as $d \to \infty$).

Proof based on the Loewner-John theorem: the Barabanov norm can be approximated by an Euclidean norm up to a $\sqrt{n}$ multiplicative factor.

Before...

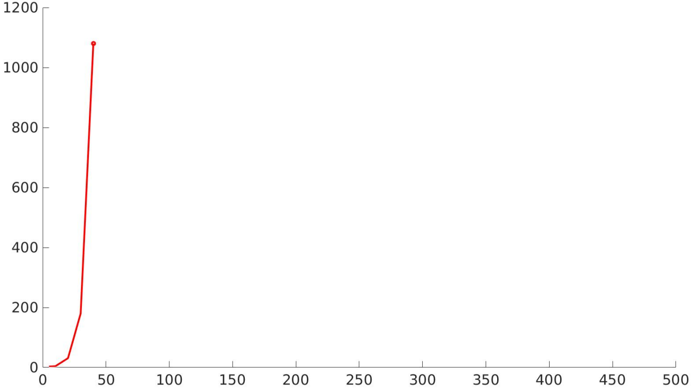


Figure: Computation time (s) vs dimension: red Ahmadi et al., ,

...Now

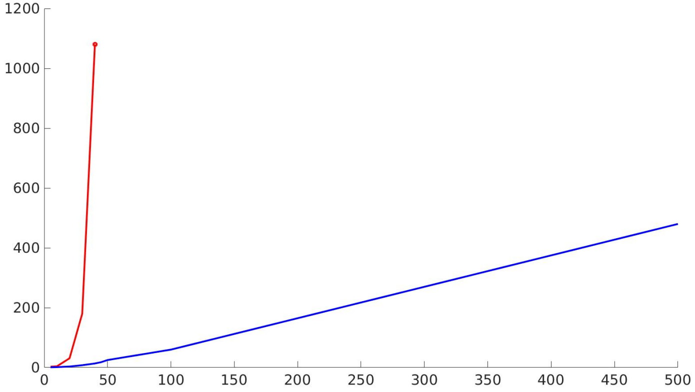


Figure: Computation time (s) vs dimension: red Ahmadi et al., blue "quantum" dynamic programming (this talk),

...Now

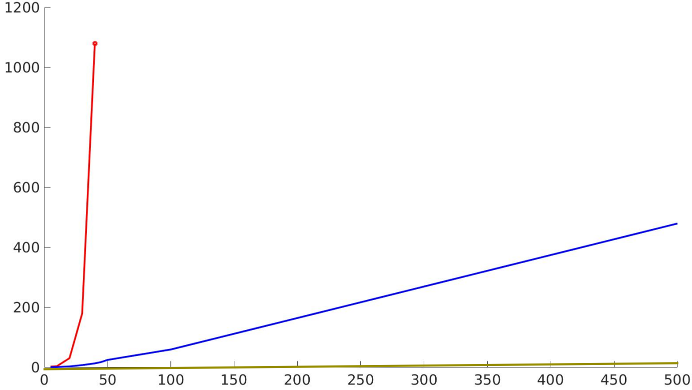


Figure: Computation time (s) vs dimension: red Ahmadi et al., blue "quantum" dynamic programming (this talk), green specialization to nonnegative matrices (this talk - MCRF, 2020)

How do we get there ?

A closer look at simplified LMIs

$$Q \succ 0 \qquad \rho^2 Q \succcurlyeq A_i^T Q A_i \;, \forall i \in [m] \, .$$

How do we get there ?

A closer look at simplified LMIs

$$Q \succ 0 \qquad \rho^2 Q \succcurlyeq A_i^T Q A_i \ , \forall i \in [m] \, .$$

Solving a wrong equation

We would like to write:

$$\text{“}\rho^2 Q \succcurlyeq \sup_{i \in [m]} A_i^T Q A_i\text{”} \, .$$

How do we get there ?

A closer look at simplified LMIs

$$Q \succ 0 \qquad \rho^2 Q \succcurlyeq A_i^T Q A_i \ , \forall i \in [m] \, .$$

Solving a wrong equation

We would like to write:

$$\text{``}\rho^2 Q \succcurlyeq \sup_{i \in [m]} A_i^T Q A_i\text{''} \, .$$

The supremum of several quadratic forms does not exist !

$\Rightarrow$ will replace supremum by a minimal upper bound

How do we get there ?

A closer look at simplified LMIs

$$Q \succ 0 \qquad \rho^2 Q \succcurlyeq A_i^T Q A_i \; , \forall i \in [m] \, .$$

Solving a wrong equation

We would like to write:

$$\text{“}\rho^2 Q \succcurlyeq \sup_{i \in [m]} A_i^T Q A_i\text{”} \, .$$

The supremum of several quadratic forms does not exist !

$\Rightarrow$ will replace supremum by a minimal upper bound

Fast computational scheme

Interior point methods are relatively slow

$\rightarrow$ Replace optimization by a fixed point approach. For nonnegative matrices, reduces to a risk-sensitive eigenproblem.

Minimal upper bounds

x is a minimal upper bound of the set $\mathcal{A}$ iff

$$\mathcal{A} \preccurlyeq x \quad \text{and} \quad \left(\mathcal{A} \preccurlyeq y \preccurlyeq x \implies y = x\right).$$

The set of minimal upper bounds: $\bigvee \mathcal{A}$.

Minimal upper bounds

x is a minimal upper bound of the set $\mathcal{A}$ iff

$$\mathcal{A} \preccurlyeq x \quad \text{and} \quad \left(\mathcal{A} \preccurlyeq y \preccurlyeq x \implies y = x\right).$$

The set of minimal upper bounds: $\bigvee \mathcal{A}$.

Theorem (Krein-Rutman - 1948)

A cone induces a lattice structure iff it is simplicial ($\cong \mathbb{R}_n^+$).

Minimal upper bounds

x is a minimal upper bound of the set $\mathcal{A}$ iff

$$\mathcal{A} \preccurlyeq x \quad \text{and} \quad \left(\mathcal{A} \preccurlyeq y \preccurlyeq x \implies y = x\right).$$

The set of minimal upper bounds: $\bigvee \mathcal{A}$.

Theorem (Krein-Rutman - 1948)

A cone induces a lattice structure iff it is simplicial ($\cong \mathbb{R}_n^+$).

Theorem (Kadison - 1951)

The Löwner order induces an anti-lattice structure: two symmetric matrices A, B have a supremum if and only if $A \preccurlyeq B$ or $B \preccurlyeq A$.

The inertia of the symmetric matrix M is the tuple (p, q, r), where

- p: number of positive eigenvalues of M,
- q: number of negative eigenvalues of M,
- r: number of zero eigenvalues of M.

Definition (Indefinite orthogonal group)

$\mathcal{O}(p, q)$ is the group of matrices S preserving the quadratic form $x_1^1 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2$:

$$S\begin{pmatrix} I_p & \\ & -I_q \end{pmatrix}S^T = \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} =: J_{p,q}$$

The inertia of the symmetric matrix M is the tuple (p,q,r), where

- p: number of positive eigenvalues of M,
- q: number of negative eigenvalues of M,
- r: number of zero eigenvalues of M.

Definition (Indefinite orthogonal group)

$\mathcal{O}(p,q)$ is the group of matrices S preserving the quadratic form $x_1^1 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2$:

$$S\begin{pmatrix} I_p & \\ & -I_q \end{pmatrix}S^T = \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} =: J_{p,q}$$

$\mathcal{O}(1,1)$ is the group of hyperbolic isometries $\begin{pmatrix} \epsilon_1 \operatorname{ch} t & \epsilon_2 \operatorname{sh} t \\ \epsilon_1 \operatorname{sh} t & \epsilon_2 \operatorname{ch} t \end{pmatrix}$, where $\epsilon_1, \epsilon_2 \in \{-1, 1\}$

The inertia of the symmetric matrix M is the tuple (p, q, r), where

- p: number of positive eigenvalues of M,
- q: number of negative eigenvalues of M,
- r: number of zero eigenvalues of M.

Definition (Indefinite orthogonal group)

$\mathcal{O}(p, q)$ is the group of matrices S preserving the quadratic form $x_1^1 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2$:

$$S\begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} S^T = \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} =: J_{p,q}$$

$\mathcal{O}(1,1)$ is the group of hyperbolic isometries $\begin{pmatrix} \epsilon_1 \operatorname{ch} t & \epsilon_2 \operatorname{sh} t \\ \epsilon_1 \operatorname{sh} t & \epsilon_2 \operatorname{ch} t \end{pmatrix}$,
where $\epsilon_1, \epsilon_2 \in \{-1, 1\}$
$\mathcal{O}(p) \times \mathcal{O}(q)$ is a maximal compact subgroup of $\mathcal{O}(p, q)$.

Theorem (Stott - Proc AMS 2018, Quantitative version of Kadison theorem)

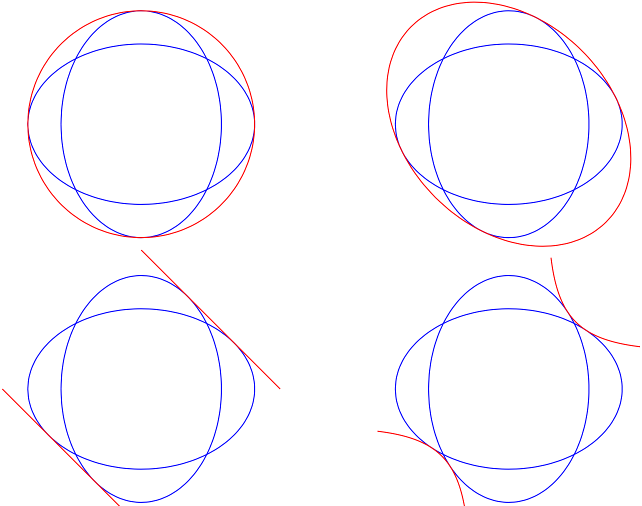
If the inertia of $A - B$ is $(p, q, 0)$, then

$$\bigvee\{A\,,B\} \;\cong\; \mathcal{O}(p,q)\Big/\mathcal{O}(p)\times\mathcal{O}(q) \;\cong\; \mathbb{R}^{pq}\,.$$

Example $p = q = 1$.

$\mathcal{O}(1,1) \big/ \mathcal{O}(1) \times \mathcal{O}(1)$ is the group of hyperbolic rotations:

$$\left\{ \begin{pmatrix} \operatorname{ch} t & \operatorname{sh} t \\ \operatorname{sh} t & \operatorname{ch} t \end{pmatrix} \mid t \in \mathbb{R} \right\}$$



Canonical selection of a minimal upper bound

Ellipsoid: $\mathcal{E}(M) = \{x \mid x^T M^{-1} x \leqslant 1\}$, where M is symmetric pos. def.

Theorem (Löwner - John)

There is a unique minimum volume ellipsoid containing a convex body $\mathcal{C}$.

Canonical selection of a minimal upper bound

Ellipsoid: $\mathcal{E}(M) = \{x \mid x^T M^{-1} x \leqslant 1\}$, where M is symmetric pos. def.

Theorem (Löwner - John)

There is a unique minimum volume ellipsoid containing a convex body $\mathcal{C}$.

Definition-Proposition (Allamigeon, SG, Goubault, Putot, NS, ACM TECS 2016)

Let $\mathcal{A} = \{A_i\}_i \subset \mathcal{S}_n^{++}$ and $\mathcal{C} = \cup_i \mathcal{E}(A_i)$. We define $\sqcup\mathcal{A}$ so that $\mathcal{E}(\sqcup\mathcal{A})$ is the Löwner ellipsoid of $\cup_{A\in\mathcal{A}}\mathcal{E}(A)$, i.e.,

$$(\sqcup\mathcal{A})^{-1} = \operatorname{argmax}_X \{\log\det X \mid X \preccurlyeq A_i^{-1}, i \in [m], \quad X \succ 0\} \ .$$

Canonical selection of a minimal upper bound

Ellipsoid: $\mathcal{E}(M) = \{x \mid x^T M^{-1} x \leqslant 1\}$, where M is symmetric pos. def.

Theorem (Löwner - John)

There is a unique minimum volume ellipsoid containing a convex body $\mathcal{C}$.

Definition-Proposition (Allamigeon, SG, Goubault, Putot, NS, ACM TECS 2016)

Let $\mathcal{A} = \{A_i\}_i \subset \mathcal{S}_n^{++}$ and $\mathcal{C} = \cup_i \mathcal{E}(A_i)$. We define $\sqcup \mathcal{A}$ so that $\mathcal{E}(\sqcup \mathcal{A})$ is the Löwner ellipsoid of $\cup_{A \in \mathcal{A}} \mathcal{E}(A)$, i.e.,

$$(\sqcup \mathcal{A})^{-1} = \mathrm{argmax}_X \{\log\det X \mid X \preccurlyeq A_i^{-1}, i \in [m], \quad X \succ 0\} \ .$$

Then, $\sqcup \mathcal{A}$ is a minimal upper bound of $\mathcal{A}$, and $\sqcup$ is the only selection that commutes with the action of invertible congruences:

$$L(\sqcup \mathcal{A})L^T = \sqcup(L \mathcal{A} L^T) \, ,$$

Theorem (Allamigeon, SG, Goubault, Putot, NS, ACM TECS 2016)

Computating $X \sqcup Y$ reduces to a square root (i.e., SDP-free!).

$$\text{Suppose } Y = I: \qquad X \sqcup I = \tfrac{1}{2}(X + I) + \tfrac{1}{2}|X - I|\,.$$

General case reduces to it by congruence: add 1 Cholesky decomposition $+\,1$ triangular inversion. Complexity: $O(n^3)$.

Theorem (Allamigeon, SG, Goubault, Putot, NS, ACM TECS 2016)

Computating $X \sqcup Y$ reduces to a square root (i.e., SDP-free!).

$$\text{Suppose } Y = I: \qquad X \sqcup I = \frac{1}{2}(X + I) + \frac{1}{2}|X - I|\,.$$

General case reduces to it by congruence: add 1 Cholesky decomposition $+\,1$ triangular inversion. Complexity: $O(n^3)$.

The Loewner selection $\sqcup$ is

Theorem (Allamigeon, SG, Goubault, Putot, NS, ACM TECS 2016)

Computating $X \sqcup Y$ reduces to a square root (i.e., SDP-free!).

$$\textit{Suppose } Y = I: \qquad X \sqcup I = \frac{1}{2}(X + I) + \frac{1}{2}|X - I| \,.$$

General case reduces to it by congruence: add 1 Cholesky decomposition $+\,1$ triangular inversion. Complexity: $O(n^3)$.

The Loewner selection $\sqcup$ is

- continuous on $\mathcal{S}_n^{++} \times \mathcal{S}_n^{++}$ but does not extend continuously to the closed cone,

Theorem (Allamigeon, SG, Goubault, Putot, NS, ACM TECS 2016)

Computating $X \sqcup Y$ reduces to a square root (i.e., SDP-free!).

Suppose $Y = I$: $$X \sqcup I = \frac{1}{2}(X + I) + \frac{1}{2}|X - I|\,.$$

General case reduces to it by congruence: add 1 Cholesky decomposition $+\,1$ triangular inversion. Complexity: $O(n^3)$.

The Loewner selection $\sqcup$ is

- continuous on $\mathcal{S}_n^{++} \times \mathcal{S}_n^{++}$ but does not extend continuously to the closed cone,
- not order-preserving,

Theorem (Allamigeon, SG, Goubault, Putot, NS, ACM TECS 2016)

Computating $X \sqcup Y$ reduces to a square root (i.e., SDP-free!).

$$\text{Suppose } Y = I: \qquad X \sqcup I = \frac{1}{2}(X+I) + \frac{1}{2}|X-I|\,.$$

General case reduces to it by congruence: add 1 Cholesky decomposition $+\,1$ triangular inversion. Complexity: $O(n^3)$.

The Loewner selection $\sqcup$ is

- continuous on $\mathcal{S}_n^{++} \times \mathcal{S}_n^{++}$ but does not extend continuously to the closed cone,
- not order-preserving,
- not associative.

Reducing the search of a joint quadratic Lyapunov function to an eigenproblem

Goal

Compute norm $\nu(x) = \sqrt{x^T Q x}$ such that $\max_{i\in[m]} \nu(A_i x) \leqslant \rho\nu(x)$.

Computation: single quadratic form

Corresponding LMI:

$$\rho^2 Q \succcurlyeq A_i^T Q A_i \ , \forall i \, .$$

Eigenvalue problem for a multivalued map

$$\rho^2 Q \in \bigvee_i A_i^T Q A_i \, .$$

Quantum dynamic programming operators

Quantum channels (0-player games)

Completely positive trace perserving operators:

$$K(X) = \sum_i A_i X A_i^* \, , \qquad \sum_i A_i^* A_i = I_n \, .$$

Quantum dynamic programming operators

Quantum channels (0-player games)

Completely positive trace perserving operators:

$$K(X) = \sum_i A_i X A_i^* \;, \qquad \sum_i A_i^* A_i = I_n \;.$$

Propagation of "non-commutative probability measures" (analogue of Fokker-Planck).

Quantum dynamic programming operator (1-player game)

$$\mathcal{T}(X) = \bigvee_i A_i^T X A_i$$

with $\bigvee$ the set of least upper bounds in Löwner order (multivalued map).

Quantum dynamic programming operators

Quantum channels (0-player games)

Completely positive trace perserving operators:

$$K(X) = \sum_i A_i X A_i^* \ , \qquad \sum_i A_i^* A_i = I_n \ .$$

Propagation of "non-commutative probability measures" (analogue of Fokker-Planck).

Quantum dynamic programming operator (1-player game)

$$\mathcal{T}(X) = \bigvee_i A_i^T X A_i$$

with $\bigvee$ the set of least upper bounds in Löwner order (multivalued map).
Propagation of norms (backward equation).

Quantum dynamic programming operator associated with an automaton

τ_d transition map of the De Bruijn automaton on d letters:

$$X \in (\mathcal{S}_n^+)^{(m^d)} \quad \text{and} \quad \mathcal{T}_w^d(X) := \bigvee_{w=\tau_d(v,i)} A_i^T X_v A_i$$

Reduces to the earlier $d = 1$ case by a block diagonal construction.

Theorem

Suppose that

$$\rho^2 X \in \mathcal{T}^d(X)$$

with $\rho > 0$ and X positive definite. Then,

$$\rho(\mathcal{A}) \leqslant \rho\,.$$

Theorem

Suppose that $\mathcal{A}$ is irreducible. Then there exists $\rho > 0$ and X such that $\sum_v X_v$ is positive definite and

$$\rho^2 X = T^d_{\sqcup}(X) \in \mathcal{T}^d(X)$$

where

$$[T^d_{\sqcup}(X)]_w := \bigsqcup_{w=\tau_d(v,i)} A_i^T X_v A_i \ .$$

Exercise: find the mistake in the following proof

We want to show that the following eigenproblem is solvable:

$$[T^d_\sqcup(X)]_w := \bigsqcup_{w=\tau_d(v,i)} A_i^T X_v A_i = \rho^2 X_w$$

1. suppose, w.l.g., $d = 0$.

Exercise: find the mistake in the following proof

We want to show that the following eigenproblem is solvable:

$$[T^d_{\sqcup}(X)]_w := \bigsqcup_{w=\tau_d(v,i)} A_i^T X_v A_i = \rho^2 X_w$$

1. suppose, w.l.g., $d = 0$.
2. Consider the noncommutative simplex,
   $\Delta := \{X \succcurlyeq 0 \colon \operatorname{trace} X = 1\}$. This set is compact and convex.

Exercise: find the mistake in the following proof

We want to show that the following eigenproblem is solvable:

$$[T^d_{\sqcup}(X)]_w := \bigsqcup_{w=\tau_d(v,i)} A_i^T X_v A_i = \rho^2 X_w$$

1. suppose, w.l.g., $d = 0$.
2. Consider the noncommutative simplex, $\Delta := \{X \succcurlyeq 0 \colon \operatorname{trace} X = 1\}$. This set is compact and convex.
3. Consider the normalized map $\tilde{T}^d_{\sqcup}(X) = (\operatorname{trace} T^d_{\sqcup}(X))^{-1} T^d_{\sqcup}(X)$. It sends Δ to Δ

Exercise: find the mistake in the following proof

We want to show that the following eigenproblem is solvable:

$$[T^d_{\sqcup}(X)]_w := \bigsqcup_{w=\tau_d(v,i)} A_i^T X_v A_i = \rho^2 X_w$$

1. suppose, w.l.g., $d = 0$.
2. Consider the noncommutative simplex,
   $\Delta := \{X \succcurlyeq 0 \colon \operatorname{trace} X = 1\}$. This set is compact and convex.
3. Consider the normalized map $\tilde{T}^d_{\sqcup}(X) = (\operatorname{trace} T^d_{\sqcup}(X))^{-1} T^d_{\sqcup}(X)$. It sends Δ to Δ
4. By Brouwer fixed point theorem, it has a fixed point

Exercise: find the mistake in the following proof

We want to show that the following eigenproblem is solvable:

$$[T^d_\sqcup(X)]_w := \bigsqcup_{w=\tau_d(v,i)} A_i^T X_v A_i = \rho^2 X_w$$

1. suppose, w.l.g., $d = 0$.
2. Consider the noncommutative simplex, $\Delta := \{X \succcurlyeq 0 \colon \operatorname{trace} X = 1\}$. This set is compact and convex.
3. Consider the normalized map $\tilde{T}^d_\sqcup(X) = (\operatorname{trace} T^d_\sqcup(X))^{-1} T^d_\sqcup(X)$. It sends Δ to Δ
4. By Brouwer fixed point theorem, it has a fixed point
5. This fixed point is an eigenvector of $\mathcal{T}^d$

Exercise: find the mistake in the following proof

We want to show that the following eigenproblem is solvable:

$$[T^d_\sqcup(X)]_w := \bigsqcup_{w=\tau_d(v,i)} A_i^T X_v A_i = \rho^2 X_w$$

1. suppose, w.l.g., $d = 0$.
2. Consider the noncommutative simplex, $\Delta := \{X \succcurlyeq 0 \colon \operatorname{trace} X = 1\}$. This set is compact and convex.
3. Consider the normalized map $\tilde{T}^d_\sqcup(X) = (\operatorname{trace} T^d_\sqcup(X))^{-1} T^d_\sqcup(X)$. It sends Δ to Δ
4. By Brouwer fixed point theorem, it has a fixed point
5. This fixed point is an eigenvector of $\mathcal{T}^d$

Exercise: find the mistake in the following proof

We want to show that the following eigenproblem is solvable:

$$[T^d_{\sqcup}(X)]_w := \bigsqcup_{w=\tau_d(v,i)} A_i^T X_v A_i = \rho^2 X_w$$

1. suppose, w.l.g., $d = 0$.
2. Consider the noncommutative simplex, $\Delta := \{X \succcurlyeq 0 \colon \operatorname{trace} X = 1\}$. This set is compact and convex.
3. Consider the normalized map $\tilde{T}^d_{\sqcup}(X) = (\operatorname{trace} T^d_{\sqcup}(X))^{-1} T^d_{\sqcup}(X)$. It sends Δ to Δ
4. By Brouwer fixed point theorem, it has a fixed point
5. This fixed point is an eigenvector of $\mathcal{T}^d$

$\sqcup$ is continuous in $\operatorname{int} \mathcal{S}_n^+ \times \operatorname{int} \mathcal{S}_n^+$, but not on its closure.

Exercise: find the mistake in the following proof

We want to show that the following eigenproblem is solvable:

$$[T_\sqcup^d(X)]_w := \bigsqcup_{w=\tau_d(v,i)} A_i^T X_v A_i = \rho^2 X_w$$

1. suppose, w.l.g., $d = 0$.
2. Consider the noncommutative simplex, $\Delta := \{X \succcurlyeq 0\colon \operatorname{trace} X = 1\}$. This set is compact and convex.
3. Consider the normalized map $\tilde{T}_\sqcup^d(X) = (\operatorname{trace} T_\sqcup^d(X))^{-1} T_\sqcup^d(X)$. It sends Δ to Δ
4. By Brouwer fixed point theorem, it has a fixed point
5. This fixed point is an eigenvector of $\mathcal{T}^d$

$\sqcup$ is continuous in $\operatorname{int}\mathcal{S}_n^+ \times \operatorname{int}\mathcal{S}_n^+$, but not on its closure.

$\rightarrow$ cannot apply naively Brouwer.

Fixing the proof of existence of eigenvectors

Lemma

For $Y_i \succ 0$, we have

$$\frac{1}{m}\sum_{i=1}^{m} Y_i \preccurlyeq \bigsqcup_{i=1}^{m} Y_i \preccurlyeq \sum_{i=1}^{m} Y_i$$

Corollary

For all $X \in \mathcal{S}_n^+$, we have

$$\frac{1}{m}K^d(X) \preccurlyeq T_{\sqcup}^d(X) \preccurlyeq K^d(X)\,,$$

with

$$K_w^d(X) = \sum_{w=\tau_d(v,i)} A_i^T X_v A_i \qquad T_{\sqcup,w}^d(X) = \bigsqcup_{w=\tau_d(v,i)} A_i^T X_v A_i\,.$$

Proof

Reduction to $K\colon X \mapsto \sum_i A_i^T X A_i$ strictly positive:

$$X \succcurlyeq 0 \implies K(X) \succ 0\,.$$

Proof

Reduction to $K\colon X \mapsto \sum_i A_i^T X A_i$ strictly positive:

$$X \succcurlyeq 0 \implies K(X) \succ 0\,.$$

Let $X \in \Delta := \{X \succcurlyeq 0\colon\ \operatorname{trace} X = 1\}$. By compactness:

$$\alpha I \preccurlyeq K(X) \preccurlyeq \beta I \text{ , with } \alpha > 0\,.$$

Proof

Reduction to $K\colon X \mapsto \sum_i A_i^T X A_i$ strictly positive:

$$X \succcurlyeq 0 \implies K(X) \succ 0\,.$$

Let $X \in \Delta \coloneqq \{X \succcurlyeq 0\colon\ \operatorname{trace} X = 1\}$. By compactness:

$$\alpha I \preccurlyeq K(X) \preccurlyeq \beta I\ \text{, with } \alpha > 0\,.$$

Then

$$\frac{\alpha}{m} I \preccurlyeq T_{\sqcup}(X) \preccurlyeq \beta I\,,$$

so $T_{\sqcup}(\Delta) \subset$ compact subset of $\operatorname{int} \Delta$.

Proof

Reduction to $K\colon X \mapsto \sum_i A_i^T X A_i$ strictly positive:

$$X \succcurlyeq 0 \implies K(X) \succ 0\,.$$

Let $X \in \Delta := \{X \succcurlyeq 0\colon \operatorname{trace} X = 1\}$. By compactness:

$$\alpha I \preccurlyeq K(X) \preccurlyeq \beta I\,, \text{ with } \alpha > 0\,.$$

Then

$$\frac{\alpha}{m} I \preccurlyeq T_{\sqcup}(X) \preccurlyeq \beta I\,,$$

so $T_{\sqcup}(\Delta) \subset$ compact subset of $\operatorname{int}\Delta$. Conclude by Brouwer's fixed point theorem. □

Computing an eigenvector

We introduce a damping parameter γ:

$$T_{\sqcup}^{\gamma}(X) = \bigsqcup_i \left(A_i^T X A_i + \gamma(\operatorname{trace} X) I_n\right).$$

Theorem

The iteration

$$X^{k+1} = \frac{T_{\sqcup}^{\gamma}(X)}{\operatorname{trace} T_{\sqcup}^{\gamma}(X)}$$

converges for a large damping: $\gamma > nm^{(3d+1)/2}$

Conjecture

The iteration converges if $\gamma > m^{1/2}n^{-1/2}$.

Experimentally: $\gamma \sim 10^{-2}$ is enough! Huge gap between conservative theoretical estimates and practice. How theoretical estimates are obtained?

Lipschitz estimations

Riemann and Thompson metrics

Two standard metrics on the cone $\mathcal{S}_n^{++}$

$$d_R(A,B) \coloneqq \| \log \operatorname{spec}(A^{-1}B)\|_2 \,.$$

$$d_T(A,B) \coloneqq \| \log \operatorname{spec}(A^{-1}B)\|_\infty \,.$$

They are invariant under the action of congruences:
$d(LAL^T, LBL^T) = d(A,B)$ for invertible L.

Lipschitz constant: $\operatorname{Lip}_M \sqcup \coloneqq \sup_{X_1,X_2,Y_1,Y_2 \succ 0} \frac{d_M(X_1 \sqcup X_2, Y_1 \sqcup Y_2)}{d_M(X_1 \oplus X_2, Y_1 \oplus Y_2)}$.

Lipschitz estimations

Riemann and Thompson metrics

Two standard metrics on the cone $\mathcal{S}_n^{++}$

$$d_R(A,B) \coloneqq \|\log \operatorname{spec}(A^{-1}B)\|_2\,.$$

$$d_T(A,B) \coloneqq \|\log \operatorname{spec}(A^{-1}B)\|_\infty\,.$$

They are invariant under the action of congruences: $d(LAL^T, LBL^T) = d(A,B)$ for invertible L.

Lipschitz constant: $\operatorname{Lip}_M \sqcup \coloneqq \sup_{X_1,X_2,Y_1,Y_2 \succ 0} \frac{d_M(X_1 \sqcup X_2, Y_1 \sqcup Y_2)}{d_M(X_1 \oplus X_2, Y_1 \oplus Y_2)}$.

Theorem

$$\operatorname{Lip}_T \sqcup = \Theta(\log n) \qquad \operatorname{Lip}_R \sqcup = 1$$

Proof.

d_T, d_R are Riemann/Finsler metrics $\to$ work locally + Schur multiplier estimation (Mathias). □

Scalability: dimension

Table: big-LMI vs Tropical Kraus

Dimension n	CPU time (tropical)	CPU time (LMI)	Error vs LMI
5	0.9 s	3.1 s	0.1 %
10	1.5 s	4.2 s	1.4 %
20	3.5 s	31 s	0.4 %
30	7.9 s	3min	0.2 %
40	13.7 s	18min	0.05 %
45	18.1 s	–	–
50	25.2 s	–	–
100	1min	–	–
500	8min	–	–

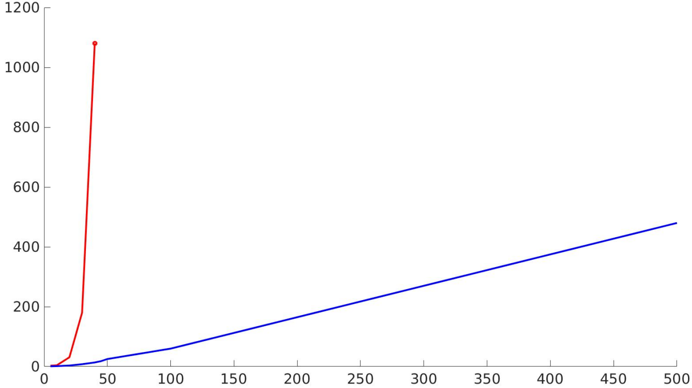


Figure: Computation time vs dimension

Scalability: graph size

$$A_1 = \begin{pmatrix} -1 & 1 & -1 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \qquad A_2 = \begin{pmatrix} -1 & 1 & -1 \\ -1 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Table: big-LMI vs Tropical Kraus: $30 - 60$ times faster.

Order d	2	4	6	8	10
Size of graph	8	32	128	512	2048
CPU time (tropical)	0.03s	0.07s	0.4s	2.0s	9.0s
CPU time (LMI)	1.9s	4.0s	24s	1min	10min
Accuracy	1.1 %	1.3 %	0.4 %	0.4 %	0.6 %

Special case of nonnegative matrices

Suppose $A_i \in \mathbb{R}_+^{n\times n}$, replace the quantum dynamic programming operator

$$X \in (\mathcal{S}_n^+)^{(m^d)} \quad \text{and} \quad \mathcal{T}_w^d(X) := \bigvee_{w=\tau_d(v,i)} A_i^T X_v A_i$$

Special case of nonnegative matrices

Suppose $A_i \in \mathbb{R}_+^{n\times n}$, replace the quantum dynamic programming operator

$$X \in (\mathcal{S}_n^+)^{(m^d)} \quad \text{and} \quad \mathcal{T}_w^d(X) := \bigvee_{w=\tau_d(v,i)} A_i^T X_v A_i$$

by the classical dynamic programming operator

$$x \in (\mathbb{R}_+^n)^{(m^d)} \quad \text{and} \quad T_w^d(x) := \sup_{w=\tau_d(v,i)} A_i^T x_v$$

Special case of nonnegative matrices

Suppose $A_i \in \mathbb{R}_+^{n\times n}$, replace the quantum dynamic programming operator

$$X \in (\mathcal{S}_n^+)^{(m^d)} \quad \text{and} \quad \mathcal{T}_w^d(X) := \bigvee_{w=\tau_d(v,i)} A_i^T X_v A_i$$

by the classical dynamic programming operator

$$x \in (\mathbb{R}_+^n)^{(m^d)} \quad \text{and} \quad T_w^d(x) := \sup_{w=\tau_d(v,i)} A_i^T x_v$$

Operators of this type arise in risk-sensitive control Anantharam, Borkar, also in games of topological entropy Asarin, Cervelle, Degorre, Dima, Horn, Kozyakin, Akian, SG, Grand-Clément, Guillaud.

Special case of nonnegative matrices

Suppose $A_i \in \mathbb{R}_+^{n\times n}$, replace the quantum dynamic programming operator

$$X \in (\mathcal{S}_n^+)^{(m^d)} \quad \text{and} \quad \mathcal{T}_w^d(X) := \bigvee_{w=\tau_d(v,i)} A_i^T X_v A_i$$

by the classical dynamic programming operator

$$x \in (\mathbb{R}_+^n)^{(m^d)} \quad \text{and} \quad T_w^d(x) := \sup_{w=\tau_d(v,i)} A_i^T x_v$$

Operators of this type arise in risk-sensitive control Anantharam, Borkar, also in games of topological entropy Asarin, Cervelle, Degorre, Dima, Horn, Kozyakin, Akian, SG, Grand-Clément, Guillaud.

Theorem

Suppose the set of nonnegative matrices $\mathcal{A}$ is positively irreducible. Then, there exists $u \in (\mathbb{R}_+)^{(m^d)} \setminus \{0\}$ such that

$$T^d(u) = \lambda_d u \ .$$

Special case of nonnegative matrices

Suppose $A_i \in \mathbb{R}_+^{n\times n}$, replace the quantum dynamic programming operator

$$X \in (\mathcal{S}_n^+)^{(m^d)} \quad \text{and} \quad \mathcal{T}_w^d(X) := \bigvee_{w=\tau_d(v,i)} A_i^T X_v A_i$$

by the classical dynamic programming operator

$$x \in (\mathbb{R}_+^n)^{(m^d)} \quad \text{and} \quad T_w^d(x) := \sup_{w=\tau_d(v,i)} A_i^T x_v$$

Operators of this type arise in risk-sensitive control Anantharam, Borkar, also in games of topological entropy Asarin, Cervelle, Degorre, Dima, Horn, Kozyakin, Akian, SG, Grand-Clément, Guillaud.

Theorem

Suppose the set of nonnegative matrices $\mathcal{A}$ is positively irreducible. Then, there exists $u \in (\mathbb{R}_+)^{(m^d)} \setminus \{0\}$ such that

$$T^d(u) = \lambda_d u \ .$$

Follows from SG and Gunawardena, TAMS 2004.

A monotone hemi-norm is a map $\nu(x) := \max_{v \in V} \langle u_v, x \rangle$ with $u_v \geqslant 0$ such that $x \mapsto \nu(x) \vee \nu(-x)$ is a norm.

Theorem (Coro. of Guglielmi and Protasov)

If $\mathcal{A} \subset \mathbb{R}_+^{n \times n}$ is positively irreducible, there is a monotone hemi-norm ν such that

$$\max_{i \in [m]} \nu(A_i x) = \rho(\mathcal{A}) \nu(x), \qquad \forall x \in \mathbb{R}_+^n$$

Theorem (Polyhedral monotone hemi-norms)

If $\mathcal{A} \subset \mathbb{R}_+^{n \times n}$ is positively irreducible, if $T^d(u) = \lambda_d u$, and $u \in (\mathbb{R}_+^n)^{(m^d)} \setminus \{0\}$, then

$$\|x\|_u := \max_{v \in [m^d]} \langle u_v, x \rangle$$

is a polyhedral monotone hemi-norm and

$$\max_{i \in [m]} \|A_i x\|_u \leqslant \lambda_d \|x\|_u \quad .$$

A monotone hemi-norm is a map $\nu(x) := \max_{v\in V}\langle u_v, x\rangle$ with $u_v \geqslant 0$ such that $x \mapsto \nu(x) \vee \nu(-x)$ is a norm.

Theorem (Coro. of Guglielmi and Protasov)

If $\mathcal{A} \subset \mathbb{R}_+^{n\times n}$ is positively irreducible, there is a monotone hemi-norm ν such that

$$\max_{i\in[m]} \nu(A_i x) = \rho(\mathcal{A})\nu(x), \qquad \forall x \in \mathbb{R}_+^n$$

Theorem (Polyhedral monotone hemi-norms)

If $\mathcal{A} \subset \mathbb{R}_+^{n\times n}$ is positively irreducible, if $T^d(u) = \lambda_d u$, and $u \in (\mathbb{R}_+^n)^{(m^d)} \setminus \{0\}$, then

$$\|x\|_u := \max_{v\in[m^d]} \langle u_v, x\rangle$$

is a polyhedral monotone hemi-norm and

$$\max_{i\in[m]} \|A_i x\|_u \leqslant \lambda_d \|x\|_u \ .$$

Moreover, $\rho(\mathcal{A}) \leqslant \lambda_d \leqslant n^{1/(d+1)}\rho(\mathcal{A})$, in particular $\lambda_d \to \lambda$ as $d \to \infty$.

How to compute λ such that $T^d(u) = \lambda u$ for some $u \neq 0$, $u \neq 0$

How to compute λ such that $T^d(u) = \lambda u$ for some $u \neq 0$, $u \neq 0$

- Policy iteration: Rothblum

How to compute λ such that $T^d(u) = \lambda u$ for some $u \neq 0$, $u \neq 0$

- Policy iteration: Rothblum
- Spectral simplex: Protasov

How to compute λ such that $T^d(u) = \lambda u$ for some $u \neq 0$, $u \neq 0$

- Policy iteration: Rothblum
- Spectral simplex: Protasov
- non-linear Collatz-Wielandt theorem + convex programming $\implies$ polytime : Akian, SG, Grand-Clément, Guillaud (ACM TOCS 2019)

How to compute λ such that $T^d(u) = \lambda u$ for some $u \neq 0$, $u \neq 0$

- Policy iteration: Rothblum
- Spectral simplex: Protasov
- non-linear Collatz-Wielandt theorem + convex programming $\implies$ polytime : Akian, SG, Grand-Clément, Guillaud (ACM TOCS 2019)

How to compute λ such that $T^d(u) = \lambda u$ for some $u \neq 0$, $u \neq 0$

- Policy iteration: Rothblum
- Spectral simplex: Protasov
- non-linear Collatz-Wielandt theorem + convex programming $\implies$ polytime : Akian, SG, Grand-Clément, Guillaud (ACM TOCS 2019)

policy iteration/spectral simplex requires computing eigenvalues (demanding), need to work with huge scale instances (dimension $N = n \times m^d$)

Krasnoselski-Mann iteration

$$x_{k+1} = \frac{1}{2}(x_k + F(x_k))$$

applies to a nonexpansive map F: $\|F(x) - F(y)\| \leqslant \|x - y\|$.

Krasnoselski-Mann iteration

$$x_{k+1} = \frac{1}{2}(x_k + F(x_k))$$

applies to a nonexpansive map F: $\|F(x) - F(y)\| \leqslant \|x - y\|$.

Theorem (Ishikawa)

Let D be a closed convex subset of a Banach space X, let F be a nonexpansive mapping sending D to a compact subset of D. Then, for any initial point $x^0 \in D$, the sequence x^k converges to a fixed point of F.

Krasnoselski-Mann iteration

$$x_{k+1} = \frac{1}{2}(x_k + F(x_k))$$

applies to a nonexpansive map F: $\|F(x) - F(y)\| \leqslant \|x - y\|$.

Theorem (Ishikawa)

Let D be a closed convex subset of a Banach space X, let F be a nonexpansive mapping sending D to a compact subset of D. Then, for any initial point $x^0 \in D$, the sequence x^k converges to a fixed point of F.

Theorem (Baillon, Bruck)

$$\|F(x^k) - x^k\| \leqslant \frac{2 \operatorname{diam}(D)}{\sqrt{\pi k}} \ ,$$

Definition (Projective Krasnoselskii-Mann iteration)

Suppose $f : \mathbb{R}_+^N \to \mathbb{R}_+^N$ is order preserving and positively homogeneous of degree 1. Choose any $v^0 \in \mathbb{R}_{>0}^N$ such that $\prod_{i\in[N]} v_i^0 = 1$,

$$v^{k+1} = \left[\frac{f(v^k)}{G\big[f(v^k)\big]} \circ v^k\right]^{1/2}, \tag{1}$$

where $x \circ y := (x_i y_i)$ and $G(x) = (x_1 \cdots x_N)^{1/N}$.

Definition (Projective Krasnoselskii-Mann iteration)

Suppose $f : \mathbb{R}_+^N \to \mathbb{R}_+^N$ is order preserving and positively homogeneous of degree 1. Choose any $v^0 \in \mathbb{R}_{>0}^N$ such that $\prod_{i\in[N]} v_i^0 = 1$,

$$v^{k+1} = \left[\frac{f(v^k)}{G\big[f(v^k)\big]} \circ v^k\right]^{1/2}, \tag{1}$$

where $x \circ y := (x_i y_i)$ and $G(x) = (x_1 \cdots x_N)^{1/N}$.

Theorem

Suppose in addition that f has a positive eigenvector. Then, the projective Krasnoselskii-Mann iteration initialized at any positive vector $v^0 \in \mathbb{R}_+^N$ such that $\prod_{i\in[N]} v_i^0 = 1$, converges towards an eigenvector of f, and $G(f(v^k))$ converges to the maximal eigenvalue of f.

Definition (Projective Krasnoselskii-Mann iteration)

Suppose $f : \mathbb{R}_+^N \to \mathbb{R}_+^N$ is order preserving and positively homogeneous of degree 1. Choose any $v^0 \in \mathbb{R}_{>0}^N$ such that $\prod_{i\in[N]} v_i^0 = 1$,

$$v^{k+1} = \left[\frac{f(v^k)}{G\big[f(v^k)\big]} \circ v^k\right]^{1/2}, \tag{1}$$

where $x \circ y := (x_i y_i)$ and $G(x) = (x_1 \cdots x_N)^{1/N}$.

Theorem

Suppose in addition that f has a positive eigenvector. Then, the projective Krasnoselskii-Mann iteration initialized at any positive vector $v^0 \in \mathbb{R}_+^N$ such that $\prod_{i\in[N]} v_i^0 = 1$, converges towards an eigenvector of f, and $G(f(v^k))$ converges to the maximal eigenvalue of f.

Proof idea. This is Krasnoselski iteration applied to $F := \log \circ f \circ \exp$ acting in the quotient of the normed space $(\mathbb{R}^n, \|\cdot\|_\infty)$ by the one dimensional subspace $\mathbb{R}\mathbf{1}_N$.

Corollary

Take $f := T^d$*, the risk-sensitive dynamic programming operator, and let*

$$\beta_k := \max_{i \in [N]} (f(v^k))_i / v_i^k \ .$$

Then,

$$\log \rho(\mathcal{A}) \leqslant \log \beta_k \leqslant \log \rho(A) + \frac{4}{\sqrt{\pi k}} d_H(v^0, u) + \frac{\log n}{d+1}$$

where d_H *is Hilbert's projective metric.*

Level d	CPU Time (s)	Eigenvalue λ_d	Relative error
1	0.01	2.165	6.8%
2	0.01	2.102	3.7%
3	0.01	2.086	2.9%
4	0.01	2.059	1.6%
5	0.02	2.041	0.7%
6	0.05	2.030	0.1%
7	0.7	2.027	0.0%
8	0.32	2.027	0.0%
9	1.12	2.027	0.0%

Table: Convergence of the hierarchy on an instance with 5×5 matrices and a maximizing cyclic product of length 6

Dimension n	Level d	Eigenvalue λ_d	CPU Time
10	2	4.287	0.01 s
	3	4.286	0.03 s
20	2	8.582	0.01 s
	3	8.576	0.03 s
50	2	22.34	0.04 s
	3	22.33	0.16 s
100	2	44.45	0.17 s
	3	44.45	0.53 s
200	2	89.77	0.71 s
	3	89.76	2.46 s
500	2	224.88	5.45 s
	3	224.88	19.7 s
1000	2	449.87	44.0 s
	3	449.87	2.7 min
2000	2	889.96	4.6 min
	3	889.96	19.2 min
5000	2	2249.69	51.9 min
	3	2249.57	3.3 h

Table: Computation time for large matrices

MEGA

The Minimal Ellipsoid Geometric Analyzer, Stott - available from http://www.cmap.polytechnique.fr/~stott/

- implements the quantum dynamic programming approach
- 1700 lines of OCaml and 800 lines of Matlab
- uses BLAS/LAPACK via LACAML for linear algebra
- uses OSDP/CSDP for some semidefinite programming
- uses Matlab for other semidefinite programming

Concluding remarks

- Reduced the approximation of the joint spectral radius to solving non-linear eigenproblems

Concluding remarks

- Reduced the approximation of the joint spectral radius to solving non-linear eigenproblems
- joint spectral radius of general matrices: "quantum" dynamic programming operator acting on the space of positive semidefinite matrices, tropical analogue of completely positive maps. "states" = bunchs of positive semidefinite matrices. yields a piecewise quadratic approximate extremal norm.

Concluding remarks

- Reduced the approximation of the joint spectral radius to solving non-linear eigenproblems
- joint spectral radius of general matrices: "quantum" dynamic programming operator acting on the space of positive semidefinite matrices, tropical analogue of completely positive maps. "states" = bunchs of positive semidefinite matrices. yields a piecewise quadratic approximate extremal norm.
- special case of nonnegative matrices: paradise of risk-sensitive eigenproblem (computationnally tractable in theory and in practice).

Concluding remarks

- Reduced the approximation of the joint spectral radius to solving non-linear eigenproblems
- joint spectral radius of general matrices: "quantum" dynamic programming operator acting on the space of positive semidefinite matrices, tropical analogue of completely positive maps. "states" = bunchs of positive semidefinite matrices. yields a piecewise quadratic approximate extremal norm.
- special case of nonnegative matrices: paradise of risk-sensitive eigenproblem (computationnally tractable in theory and in practice).
- eigenproblems solved by iterative methods, variations of Krasnoselskii-Mann, scalable.

Concluding remarks

- Reduced the approximation of the joint spectral radius to solving non-linear eigenproblems
- joint spectral radius of general matrices: "quantum" dynamic programming operator acting on the space of positive semidefinite matrices, tropical analogue of completely positive maps. "states" = bunchs of positive semidefinite matrices. yields a piecewise quadratic approximate extremal norm.
- special case of nonnegative matrices: paradise of risk-sensitive eigenproblem (computationnally tractable in theory and in practice).
- eigenproblems solved by iterative methods, variations of Krasnoselskii-Mann, scalable.
- convergence analysis considerably harder in the "quantum" case, since the dynamic programming operator is not any more nonexpansive in the natural metrics.

Concluding remarks

- Reduced the approximation of the joint spectral radius to solving non-linear eigenproblems
- joint spectral radius of general matrices: "quantum" dynamic programming operator acting on the space of positive semidefinite matrices, tropical analogue of completely positive maps. "states" = bunchs of positive semidefinite matrices. yields a piecewise quadratic approximate extremal norm.
- special case of nonnegative matrices: paradise of risk-sensitive eigenproblem (computationnally tractable in theory and in practice).
- eigenproblems solved by iterative methods, variations of Krasnoselskii-Mann, scalable.
- convergence analysis considerably harder in the "quantum" case, since the dynamic programming operator is not any more nonexpansive in the natural metrics.
- generalization to the infinitesimal / PDE case ?

Concluding remarks

- Reduced the approximation of the joint spectral radius to solving non-linear eigenproblems
- joint spectral radius of general matrices: "quantum" dynamic programming operator acting on the space of positive semidefinite matrices, tropical analogue of completely positive maps. "states" = bunchs of positive semidefinite matrices. yields a piecewise quadratic approximate extremal norm.
- special case of nonnegative matrices: paradise of risk-sensitive eigenproblem (computationnally tractable in theory and in practice).
- eigenproblems solved by iterative methods, variations of Krasnoselskii-Mann, scalable.
- convergence analysis considerably harder in the "quantum" case, since the dynamic programming operator is not any more nonexpansive in the natural metrics.
- generalization to the infinitesimal / PDE case ?

Concluding remarks

- Reduced the approximation of the joint spectral radius to solving non-linear eigenproblems
- joint spectral radius of general matrices: "quantum" dynamic programming operator acting on the space of positive semidefinite matrices, tropical analogue of completely positive maps. "states" = bunchs of positive semidefinite matrices. yields a piecewise quadratic approximate extremal norm.
- special case of nonnegative matrices: paradise of risk-sensitive eigenproblem (computationnally tractable in theory and in practice).
- eigenproblems solved by iterative methods, variations of Krasnoselskii-Mann, scalable.
- convergence analysis considerably harder in the "quantum" case, since the dynamic programming operator is not any more nonexpansive in the natural metrics.
- generalization to the infinitesimal / PDE case ?

Thank you !