Tractable Robust Markov Decision Processes

Julien Grand-Clément (HEC Paris, ISOM Department)
Nian Si (HKUST, IEDA)
Shengbo Wang (Stanford, MS& E $\rightarrow$ USC, ISE)

ICSP 2025

This talk in one slide

Research question:

Which models of uncertainty sets lead to *tractable* robust MDPs?

Why it's interesting?

Many models: s-rec., sa-rec., r-rec., d-rec., k-rec., (ξ, η)-rec., etc.

with different properties/proofs, no principled way to check tractability

This talk in one slide

Research question:

Which models of uncertainty sets lead to *tractable* robust MDPs?

Why it's interesting?

Many models: s-rec., sa-rec., r-rec., d-rec., k-rec., (ξ, η)-rec., etc.

with different properties/proofs, no principled way to check tractability

Main novelty: necessary and sufficient condition for tractability

This talk in one slide

Research question:

Which models of uncertainty sets lead to *tractable* robust MDPs?

Why it's interesting?

Many models: s-rec., sa-rec., r-rec., d-rec., k-rec., (ξ, η)-rec., etc.

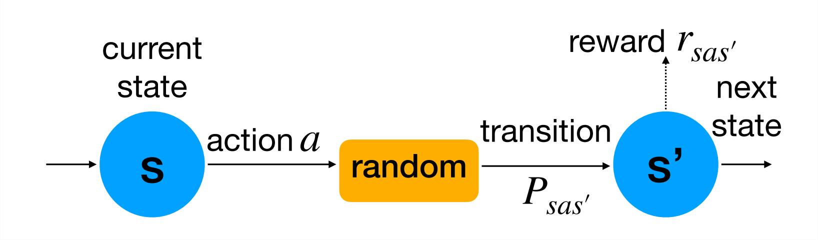
with different properties/proofs, no principled way to check tractability

Main novelty: necessary and sufficient condition for tractability

Main results:

1. Only s-rectangular models are tractable in all generality!
2. We uncover many *weakly tractable* models, "by design"
3. Unified analysis of "tractability" for different models of uncertainty

Setup for robust Markov decision process



- Finite set of states and actions: $\mathcal{S}, \mathcal{A}$
- Initial distribution over the states $\boldsymbol{\mu} \in \Delta(\mathcal{S})$
- Rewards $r_{sas'}$ for current state-action (s, a) and next state s'
- Transition proba. $\boldsymbol{P} = (P_{sas'})$, unknown: $\boldsymbol{P} \in \mathcal{P}$, convex compact
- History-dependent policy $\pi \in \Pi_{\mathrm{H}}$: maps all finite histories to $\Delta(\mathcal{A})$

Objective for robust MDPs

Discounted value function: $\boldsymbol{v}^{\pi,\boldsymbol{P}} \in \mathbb{R}^{\mathcal{S}}$ defined as

$$v_s^{\pi,\boldsymbol{P}} = \mathbb{E}^{\pi,\boldsymbol{P}}\left[\sum_{t=0}^{\infty} \gamma^t r_{s_t a_t s_{t+1}} \mid s_0 = s\right], \forall\, s \in \mathcal{S}.$$

Main objective of RMDPs: Solve

$$\sup_{\pi \in \Pi_{\mathrm{H}}} \inf_{\boldsymbol{P} \in \mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}}$$

Objective for robust MDPs

Discounted value function: $\boldsymbol{v}^{\pi,\boldsymbol{P}} \in \mathbb{R}^{\mathcal{S}}$ defined as

$$v_s^{\pi,\boldsymbol{P}} = \mathbb{E}^{\pi,\boldsymbol{P}}\left[\sum_{t=0}^{\infty} \gamma^t r_{s_t a_t s_{t+1}} \mid s_0 = s\right], \forall\, s \in \mathcal{S}.$$

Main objective of RMDPs: Solve

$$\sup_{\pi \in \Pi_{\mathrm{H}}} \inf_{\boldsymbol{P} \in \mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}}$$

Theorem [LT07, WKR13, GBZ$^+$18]

In all generality:

- Deciding $\min_{\boldsymbol{P} \in \mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}} \geq \alpha$ is NP-hard.
- Optimal policies may need to be history-dependent

Objective for robust MDPs

Discounted value function: $\boldsymbol{v}^{\pi,\boldsymbol{P}} \in \mathbb{R}^{\mathcal{S}}$ defined as

$$v_s^{\pi,\boldsymbol{P}} = \mathbb{E}^{\pi,\boldsymbol{P}}\left[\sum_{t=0}^{\infty} \gamma^t r_{s_t a_t s_{t+1}} \mid s_0 = s\right], \forall\, s \in \mathcal{S}.$$

Main objective of RMDPs: Solve

$$\sup_{\pi \in \Pi_{\mathrm{H}}} \inf_{\boldsymbol{P} \in \mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}}$$

Theorem [LT07, WKR13, GBZ+18]

In all generality:

- Deciding $\min_{\boldsymbol{P}\in\mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}} \geq \alpha$ is NP-hard.
- Optimal policies may need to be history-dependent

When are RMDPs "tractable"?

Stationary/deterministic policies, algos for $\min_{\boldsymbol{P}\in\mathcal{P}}$ and $\sup_{\pi \in \Pi_{\mathrm{H}}} \inf_{\boldsymbol{P} \in \mathcal{P}}$ problems...

Rectangularity

s-rectangularity [WKR13]

"The adversary chooses $P_{sas'}$ independently across different s":

$$\mathcal{P} = \times_{s \in \mathcal{S}} \mathcal{P}_s, \quad \mathcal{P}_s = (P_{sas'})_{as'} \subset \Delta(\mathcal{S})^{\mathcal{A}}$$

Rectangularity

s-rectangularity [WKR13]

"The adversary chooses $P_{sas'}$ independently across different s":

$$\mathcal{P} = \times_{s \in \mathcal{S}} \mathcal{P}_s, \quad \mathcal{P}_s = (P_{sas'})_{as'} \subset \Delta(\mathcal{S})^{\mathcal{A}}$$

Example 1: $\mathcal{P}$ s-rectangular, based on ℓ_∞-distance from nominal $\hat{\boldsymbol{P}}$:

$$\mathcal{P} = \{(\boldsymbol{P}_{sa}) \mid |P_{sas'} - \hat{P}_{sas'}| \leq \epsilon, \forall\, (s, a, s')\}$$

Rectangularity

s-rectangularity [WKR13]

"The adversary chooses $P_{sas'}$ independently across different s":

$$\mathcal{P} = \times_{s \in \mathcal{S}} \mathcal{P}_s, \quad \mathcal{P}_s = (P_{sas'})_{as'} \subset \Delta(\mathcal{S})^{\mathcal{A}}$$

Example 2: $\mathcal{P}$ non rectangular, based on ℓ_1-distance from nominal $\hat{\boldsymbol{P}}$:

$$\mathcal{P} = \{(\boldsymbol{P}_{sa}) \mid \sum_{(s,a,s')} |P_{sas'} - \hat{P}_{sas'}| \leq \epsilon\}$$

Rectangularity

s-rectangularity [WKR13]

"The adversary chooses $P_{sas'}$ independently across different s":

$$\mathcal{P} = \times_{s \in \mathcal{S}} \mathcal{P}_s, \quad \mathcal{P}_s = (P_{sas'})_{as'} \subset \Delta(\mathcal{S})^{\mathcal{A}}$$

Example 3: $\mathcal{P}$ non rectangular, based on underlying factors:

$$\begin{aligned} \mathcal{P} = \{(\boldsymbol{P}_{sa}) \,|\boldsymbol{P}_{sa} = \frac{1}{2}\boldsymbol{w}^1 + \frac{1}{2}\boldsymbol{w}^2, \\ (\boldsymbol{w}^1, \boldsymbol{w}^2) \in \mathcal{W}^1 \times \mathcal{W}^2\}, \\ \mathcal{W}^1, \mathcal{W}^2 \subset \Delta(\mathcal{S}) \end{aligned}$$

Properties of optimal policies [WKR13]

For $\mathcal{P}$ s-rec. convex compact, there exists stationary optimal policies.

Properties of optimal policies [WKR13]

For $\mathcal{P}$ s-rec. convex compact, there exists stationary optimal policies.

An appealing property of s-rectangular models:

the policy evaluation problem can be solved

Properties of optimal policies [WKR13]

For $\mathcal{P}$ s-rec. convex compact, there exists stationary optimal policies.

An appealing property of s-rectangular models:

the policy evaluation problem can be solved by dynamic programming:

$$\min_{\boldsymbol{P} \in \mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi, \boldsymbol{P}} = \boldsymbol{\mu}^\top \boldsymbol{u}^{\pi}$$

Properties of optimal policies [WKR13]

For $\mathcal{P}$ s-rec. convex compact, there exists stationary optimal policies.

An appealing property of s-rectangular models:

the policy evaluation problem can be solved by dynamic programming:

$$\min_{\boldsymbol{P} \in \mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi, \boldsymbol{P}} = \boldsymbol{\mu}^\top \boldsymbol{u}^\pi$$

with $\boldsymbol{u}^\pi$ the unique fixed-point of the worst-case Bellman operator:

$$u_s^\pi = \min_{\boldsymbol{P} \in \mathcal{P}} \sum_{a \in \mathcal{A}} \pi_{sa} \boldsymbol{P}_{sa}^\top \left(\boldsymbol{r}_{sa} + \gamma \boldsymbol{u}^\pi \right), \forall\, s \in \mathcal{S}.$$

But there are other models in the literature!

But there are other models in the literature!

- r-rec. [GBZ$^+$18, GGC22]:

$$\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa}, \boldsymbol{W} = (\boldsymbol{w}^1, ..., \boldsymbol{w}^r) \in \times_{i \in [r]} \mathcal{W}^i, \boldsymbol{u}_{sa} \text{ fixed}$$

But there are other models in the literature!

- r-rec. [GBZ+18, GGC22]:

$$\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa}, \boldsymbol{W} = (\boldsymbol{w}^1, ..., \boldsymbol{w}^r) \in \times_{i\in[r]}\mathcal{W}^i, \boldsymbol{u}_{sa} \text{ fixed}$$

- (ξ, η)-rec. [HMDL24]: r-rec. + $\boldsymbol{u}$'s can vary too

But there are other models in the literature!

- r-rec. [GBZ$^+$18, GGC22]:

$$\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa}, \boldsymbol{W} = (\boldsymbol{w}^1, ..., \boldsymbol{w}^r) \in \times_{i\in[r]}\mathcal{W}^i, \boldsymbol{u}_{sa} \text{ fixed}$$

- (ξ, η)-rec. [HMDL24]: r-rec. + $\boldsymbol{u}$'s can vary too
- d-rec. [MLB$^+$22]: r-rec. with finite horizon + copies of states

But there are other models in the literature!

- r-rec. [GBZ$^+$18, GGC22]:

$$\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa}, \boldsymbol{W} = (\boldsymbol{w}^1, ..., \boldsymbol{w}^r) \in \times_{i\in[r]}\mathcal{W}^i, \boldsymbol{u}_{sa} \text{ fixed}$$

- (ξ, η)-rec. [HMDL24]: r-rec. + $\boldsymbol{u}$'s can vary too
- d-rec. [MLB$^+$22]: r-rec. with finite horizon + copies of states
- k-rec. [MMX16]: finite horizon, transitions can change max. k times

But there are other models in the literature!

- r-rec. [GBZ+18, GGC22]:

$$\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa}, \boldsymbol{W} = (\boldsymbol{w}^1, ..., \boldsymbol{w}^r) \in \times_{i\in[r]}\mathcal{W}^i, \boldsymbol{u}_{sa} \text{ fixed}$$

- (ξ, η)-rec. [HMDL24]: r-rec. + $\boldsymbol{u}$'s can vary too
- d-rec. [MLB+22]: r-rec. with finite horizon + copies of states
- k-rec. [MMX16]: finite horizon, transitions can change max. k times

Some of these are "tractable", some are not, for some we don't know...

But there are other models in the literature!

- r-rec. [GBZ$^+$18, GGC22]:

$$\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa}, \boldsymbol{W} = (\boldsymbol{w}^1, ..., \boldsymbol{w}^r) \in \times_{i\in[r]}\mathcal{W}^i, \boldsymbol{u}_{sa} \text{ fixed}$$

- (ξ, η)-rec. [HMDL24]: r-rec. + $\boldsymbol{u}$'s can vary too
- d-rec. [MLB$^+$22]: r-rec. with finite horizon + copies of states
- k-rec. [MMX16]: finite horizon, transitions can change max. k times

Some of these are “tractable”, some are not, for some we don't know...

Can we find a <u>necessary</u> and sufficient condition for <u>tractability?</u>

Definition: s-tractability

An uncertainty set $\mathcal{P}$ is s-tractable if, for any parameters:

any rewards $(r_{sas'})$, discount factor $\gamma \in [0, 1)$, initial distribution $\boldsymbol{\mu}$,

the policy evaluation problem can be solved by dynamic programming:

$$\min_{\boldsymbol{P} \in \mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi, \boldsymbol{P}} = \boldsymbol{\mu}^\top \boldsymbol{u}^\pi \quad (1)$$

Definition: s-tractability

An uncertainty set $\mathcal{P}$ is s-tractable if, for any parameters:

any rewards $(r_{sas'})$, discount factor $\gamma \in [0,1)$, initial distribution $\boldsymbol{\mu}$,

the policy evaluation problem can be solved by dynamic programming:

$$\min_{\boldsymbol{P} \in \mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}} = \boldsymbol{\mu}^\top \boldsymbol{u}^\pi \tag{1}$$

with $\boldsymbol{u}^\pi$ the unique fixed-point of the worst-case Bellman operator:

$$u_s^\pi = \min_{\boldsymbol{P} \in \mathcal{P}} \sum_{a \in \mathcal{A}} \pi_{sa} \boldsymbol{P}_{sa}^\top \left(\boldsymbol{r}_{sa} + \gamma \boldsymbol{u}^\pi \right), \forall\, s \in \mathcal{S}.$$

Definition: s-tractability

An uncertainty set $\mathcal{P}$ is s-tractable if, for any parameters:

any rewards $(r_{sas'})$, discount factor $\gamma \in [0, 1)$, initial distribution $\boldsymbol{\mu}$,

the policy evaluation problem can be solved by dynamic programming:

$$\min_{\boldsymbol{P} \in \mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}} = \boldsymbol{\mu}^\top \boldsymbol{u}^\pi \tag{1}$$

with $\boldsymbol{u}^\pi$ the unique fixed-point of the worst-case Bellman operator:

$$u_s^\pi = \min_{\boldsymbol{P} \in \mathcal{P}} \sum_{a \in \mathcal{A}} \pi_{sa} \boldsymbol{P}_{sa}^\top \left(\boldsymbol{r}_{sa} + \gamma \boldsymbol{u}^\pi \right), \forall\, s \in \mathcal{S}.$$

Note in (1): LHS = global min. over $\mathcal{P}$ + non-linear objective

RHS = state-wise minimization, linear objectives

Definition: s-tractability

An uncertainty set $\mathcal{P}$ is s-tractable if, for any parameters:

any rewards $(r_{sas'})$, discount factor $\gamma \in [0,1)$, initial distribution $\boldsymbol{\mu}$,

the policy evaluation problem can be solved by dynamic programming:

$$\min_{\boldsymbol{P} \in \mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}} = \boldsymbol{\mu}^\top \boldsymbol{u}^\pi \tag{1}$$

with $\boldsymbol{u}^\pi$ the unique fixed-point of the worst-case Bellman operator:

$$u_s^\pi = \min_{\boldsymbol{P} \in \mathcal{P}} \sum_{a \in \mathcal{A}} \pi_{sa} \boldsymbol{P}_{sa}^\top \left(\boldsymbol{r}_{sa} + \gamma \boldsymbol{u}^\pi \right), \forall\, s \in \mathcal{S}.$$

Note in (1): LHS = global min. over $\mathcal{P}$ + non-linear objective

RHS = state-wise minimization, linear objectives

$\Rightarrow \boldsymbol{u}^\pi$ is worst-case value function for s-rectangular extension of $\mathcal{P}$

Definition: s-tractability

An uncertainty set $\mathcal{P}$ is s-tractable if, for any parameters:

any rewards $(r_{sas'})$, discount factor $\gamma \in [0,1)$, initial distribution $\boldsymbol{\mu}$,

the policy evaluation problem can be solved by dynamic programming:

$$\min_{\boldsymbol{P} \in \mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}} = \boldsymbol{\mu}^\top \boldsymbol{u}^\pi \tag{1}$$

with $\boldsymbol{u}^\pi$ the unique fixed-point of the worst-case Bellman operator:

$$u_s^\pi = \min_{\boldsymbol{P} \in \mathcal{P}} \sum_{a \in \mathcal{A}} \pi_{sa} \boldsymbol{P}_{sa}^\top \left(\boldsymbol{r}_{sa} + \gamma \boldsymbol{u}^\pi \right), \forall\, s \in \mathcal{S}.$$

Note in (1): LHS = global min. over $\mathcal{P}$ + non-linear objective

RHS = state-wise minimization, linear objectives

$\Rightarrow \boldsymbol{u}^\pi$ is worst-case value function for s-rectangular extension of $\mathcal{P}$

Necessary and sufficient condition for s-tractability?

Theorem

Assume that $\mathcal{P}$ is convex compact. Then

$$\mathcal{P} \text{ is s-tractable if and only } \mathcal{P} \text{ is s-rectangular.}$$

Theorem

Assume that $\mathcal{P}$ is convex compact. Then

$$\mathcal{P} \text{ is s-tractable if and only } \mathcal{P} \text{ is s-rectangular.}$$

Proof sketch:

1. Goal: $\mathcal{P}$ is a Cartesian product over s

Theorem

Assume that $\mathcal{P}$ is convex compact. Then

$$\mathcal{P} \text{ is s-tractable if and only } \mathcal{P} \text{ is s-rectangular.}$$

Proof sketch:

1. Goal: $\mathcal{P}$ is a Cartesian product over s
2. $\mathcal{P}$ convex $\Rightarrow \mathcal{P} = \mathrm{conv}(\mathrm{ext}(\mathcal{P}))$

Theorem

Assume that $\mathcal{P}$ is convex compact. Then

$$\mathcal{P} \text{ is s-tractable if and only } \mathcal{P} \text{ is s-rectangular.}$$

Proof sketch:

1. Goal: $\mathcal{P}$ is a Cartesian product over s
2. $\mathcal{P}$ convex $\Rightarrow \mathcal{P} = \text{conv}(\text{ext}(\mathcal{P}))$
3. If we show $\text{ext}(\mathcal{P})$ is s-rectangular, then $\mathcal{P}$ is s-rectangular

Theorem

Assume that $\mathcal{P}$ is convex compact. Then

$$\mathcal{P} \text{ is s-tractable if and only } \mathcal{P} \text{ is s-rectangular.}$$

Proof sketch:

1. Goal: $\mathcal{P}$ is a Cartesian product over s
2. $\mathcal{P}$ convex $\Rightarrow \mathcal{P} = \text{conv}(\text{ext}(\mathcal{P}))$
3. If we show $\text{ext}(\mathcal{P})$ is s-rectangular, then $\mathcal{P}$ is s-rectangular
4. $\text{ext}(\mathcal{P})$ is recovered by minimizing linear forms over $\mathcal{P}$

Theorem

Assume that $\mathcal{P}$ is convex compact. Then

$$\mathcal{P} \text{ is s-tractable if and only } \mathcal{P} \text{ is s-rectangular.}$$

Proof sketch:

1. Goal: $\mathcal{P}$ is a Cartesian product over s
2. $\mathcal{P}$ convex $\Rightarrow \mathcal{P} = \mathrm{conv}(\mathrm{ext}(\mathcal{P}))$
3. If we show $\mathrm{ext}(\mathcal{P})$ is s-rectangular, then $\mathcal{P}$ is s-rectangular
4. $\mathrm{ext}(\mathcal{P})$ is recovered by minimizing linear forms over $\mathcal{P}$
5. s-tractable $\Rightarrow$ minimizing linear forms over $\mathcal{P}_s$ for each $s \in \mathcal{S}$ recovers a kernel in $\mathcal{P}$

Weakly tractable models

Note that r-rec. models are not s-tractable *in all generality*!

But [GBZ+18, GGC22] show "tractability" of r-rectangular models...

$\Rightarrow$ some additional assumptions in the rewards!

Weakly tractable models

Note that r-rec. models are not s-tractable *in all generality*!

But [GBZ[+]18, GGC22] show "tractability" of r-rectangular models...

$\Rightarrow$ some additional assumptions in the rewards!

Weak tractability

$\mathcal{P}$ is weakly s-tractable if it is s-tractable, under the additional assumption that the rewards do not depend on the next states:

$$r_{sas'} = r_{sa}, \forall\, s, a, s'.$$

Weakly tractable models

Note that r-rec. models are not s-tractable *in all generality*!

But [GBZ$^+$18, GGC22] show "tractability" of r-rectangular models...

$\Rightarrow$ some additional assumptions in the rewards!

Weak tractability

$\mathcal{P}$ is weakly s-tractable if it is s-tractable, under the additional assumption that the rewards do not depend on the next states:

$$r_{sas'} = r_{sa}, \forall\ s, a, s'.$$

1. What are the implications of weak tractability?
2. Necessary and sufficient condition for weak tractability?
3. Other weakly tractable models than r-rec.?

Implications of weak s-tractability:

1. We can efficiently solve $\min_{\boldsymbol{P} \in \mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi, \boldsymbol{P}}$ (DP + convex program)
2. We can efficiently solve $\max_{\pi \in \Pi_\mathrm{S}} \min_{\boldsymbol{P} \in \mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi, \boldsymbol{P}}$

Implications of weak s-tractability:

1. We can efficiently solve $\min_{\boldsymbol{P}\in\mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}}$ (DP + convex program)
2. We can efficiently solve $\max_{\pi\in\Pi_{\mathrm{S}}} \min_{\boldsymbol{P}\in\mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}}$

Additional implications for $\mathcal{P}$ convex:

1. Optimality of stationary policies
2. We can efficiently solve $\sup_{\pi\in\Pi_{\mathrm{H}}} \min_{\boldsymbol{P}\in\mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}}$
3. Equivalence between stationary and non-stationary adversaries

Implications of weak s-tractability:

1. We can efficiently solve $\min_{\boldsymbol{P}\in\mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}}$ (DP + convex program)
2. We can efficiently solve $\max_{\pi\in\Pi_\mathrm{S}} \min_{\boldsymbol{P}\in\mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}}$

Additional implications for $\mathcal{P}$ convex:

1. Optimality of stationary policies
2. We can efficiently solve $\sup_{\pi\in\Pi_\mathrm{H}} \min_{\boldsymbol{P}\in\mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}}$
3. Equivalence between stationary and non-stationary adversaries

An important point: $\mathcal{P}$ weakly tractable $\Rightarrow \min_{\boldsymbol{P}\in\mathcal{P}} \boldsymbol{\mu}^\top \boldsymbol{v}^{\pi,\boldsymbol{P}} = \boldsymbol{\mu}^\top \boldsymbol{u}^{\pi}$...

... with $\boldsymbol{u}^\pi$ = fixed-point for the s-rectangular extension!

So non-rectangularity is "useless" if $\mathcal{P}$ is weakly tractable

Necessary and sufficient condition for weak tractability

The following statements are equivalent:

1. $\mathcal{P}$ is weakly s-tractable
2. the Weak Simultaneous Solvability Property (Weak SSP) holds:

$$\forall\,(\pi, \boldsymbol{V}) \in \Pi_S \times \mathbb{R}^{\mathcal{S}}, \cap_{s \in \mathcal{S}} \arg\min_{\boldsymbol{P} \in \mathcal{P}} \langle \boldsymbol{P}_s, \boldsymbol{\pi}_s \boldsymbol{V}^\top \rangle \neq \emptyset \quad \text{(Weak SSP)}$$

Weak SSP $\iff$ we can optimize linear forms over $\mathcal{P}$ independently over s, but objective = rank one matrices

Necessary and sufficient condition for weak tractability

The following statements are equivalent:

1. $\mathcal{P}$ is weakly s-tractable
2. the Weak Simultaneous Solvability Property (Weak SSP) holds:

$$\forall\,(\pi, \boldsymbol{V}) \in \Pi_{\mathsf{S}} \times \mathbb{R}^{\mathcal{S}}, \cap_{s\in\mathcal{S}} \arg\min_{\boldsymbol{P}\in\mathcal{P}} \langle \boldsymbol{P}_s, \boldsymbol{\pi}_s \boldsymbol{V}^\top \rangle \neq \emptyset \quad \text{(Weak SSP)}$$

Weak SSP $\iff$ we can optimize linear forms over $\mathcal{P}$ independently over s, but objective = rank one matrices

Which models of uncertainty are weakly s-tractable?

$$\forall\,(\pi, \boldsymbol{V}) \in \Pi_{\mathsf{S}} \times \mathbb{R}^{\mathcal{S}}, \cap_{s\in\mathcal{S}} \arg\min_{\boldsymbol{P}\in\mathcal{P}} \sum_{s,a} \pi_{sa} \boldsymbol{P}_{sa}^{\top} \boldsymbol{V} \neq \emptyset. \qquad \text{(Weak SSP)}$$

Many models of uncertainty are weakly s-tractable:

$$\forall\,(\pi, \boldsymbol{V}) \in \Pi_{\mathsf{S}} \times \mathbb{R}^{\mathcal{S}}, \cap_{s\in\mathcal{S}} \arg\min_{\boldsymbol{P}\in\mathcal{P}} \sum_{s,a} \pi_{sa} \boldsymbol{P}_{sa}^{\top} \boldsymbol{V} \neq \emptyset. \qquad \text{(Weak SSP)}$$

Many models of uncertainty are weakly s-tractable:

1. s-rec. works

$$\forall\,(\pi, \boldsymbol{V}) \in \Pi_{\mathsf{S}} \times \mathbb{R}^{\mathcal{S}}, \cap_{s\in\mathcal{S}} \arg\min_{\boldsymbol{P}\in\mathcal{P}} \sum_{s,a} \pi_{sa} \boldsymbol{P}_{sa}^{\top} \boldsymbol{V} \neq \emptyset. \qquad \text{(Weak SSP)}$$

Many models of uncertainty are weakly s-tractable:

1. s-rec. works
2. r-rec. works:

$$\forall\,(\pi, \boldsymbol{V}) \in \Pi_{\mathsf{S}} \times \mathbb{R}^{\mathcal{S}}, \cap_{s\in\mathcal{S}} \arg\min_{\boldsymbol{P}\in\mathcal{P}} \sum_{s,a} \pi_{sa} \boldsymbol{P}_{sa}^{\top} \boldsymbol{V} \neq \emptyset. \qquad \text{(Weak SSP)}$$

Many models of uncertainty are weakly s-tractable:

1. s-rec. works
2. r-rec. works: $\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa} \Rightarrow \boldsymbol{P}_{sa}^{\top}\boldsymbol{V} = \boldsymbol{u}_{sa}^{\top}\boldsymbol{W}^{\top}\boldsymbol{V}$

$$\forall\,(\pi, \boldsymbol{V}) \in \Pi_{\mathsf{S}} \times \mathbb{R}^{\mathcal{S}}, \cap_{s\in\mathcal{S}} \arg\min_{\boldsymbol{P}\in\mathcal{P}} \sum_{s,a} \pi_{sa} \boldsymbol{P}_{sa}^{\top} \boldsymbol{V} \neq \emptyset. \qquad \text{(Weak SSP)}$$

Many models of uncertainty are weakly s-tractable:

1. s-rec. works
2. r-rec. works: $\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa} \Rightarrow \boldsymbol{P}_{sa}^{\top}\boldsymbol{V} = \boldsymbol{u}_{sa}^{\top}\boldsymbol{W}^{\top}\boldsymbol{V}$
3. $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2$ such that $\mathcal{P}_1$ is r-rec. and $\mathcal{P}_2$ is s-rec.

$$\forall\,(\pi, \boldsymbol{V}) \in \Pi_{\mathsf{S}} \times \mathbb{R}^{\mathcal{S}}, \cap_{s \in \mathcal{S}} \arg\min_{\boldsymbol{P} \in \mathcal{P}} \sum_{s,a} \pi_{sa} \boldsymbol{P}_{sa}^\top \boldsymbol{V} \neq \emptyset. \qquad \text{(Weak SSP)}$$

Many models of uncertainty are weakly s-tractable:

1. s-rec. works
2. r-rec. works: $\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa} \Rightarrow \boldsymbol{P}_{sa}^\top \boldsymbol{V} = \boldsymbol{u}_{sa}^\top \boldsymbol{W}^\top \boldsymbol{V}$
3. $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2$ such that $\mathcal{P}_1$ is r-rec. and $\mathcal{P}_2$ is s-rec.
4. (ξ, η)-rec. works: $\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa}$, $\boldsymbol{W}$ and $\boldsymbol{u}$ in Cartesian product set

$$\forall\,(\pi, \boldsymbol{V}) \in \Pi_{\mathsf{S}} \times \mathbb{R}^{\mathcal{S}}, \cap_{s\in\mathcal{S}} \arg\min_{\boldsymbol{P}\in\mathcal{P}} \sum_{s,a} \pi_{sa} \boldsymbol{P}_{sa}^{\top} \boldsymbol{V} \neq \emptyset. \qquad \text{(Weak SSP)}$$

Many models of uncertainty are weakly s-tractable:

1. s-rec. works
2. r-rec. works: $\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa} \Rightarrow \boldsymbol{P}_{sa}^{\top}\boldsymbol{V} = \boldsymbol{u}_{sa}^{\top}\boldsymbol{W}^{\top}\boldsymbol{V}$
3. $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2$ such that $\mathcal{P}_1$ is r-rec. and $\mathcal{P}_2$ is s-rec.
4. (ξ, η)-rec. works: $\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa}$, $\boldsymbol{W}$ and $\boldsymbol{u}$ in Cartesian product set
5. Other models in the paper;

$$\forall\,(\pi, \boldsymbol{V}) \in \Pi_{\mathsf{S}} \times \mathbb{R}^{\mathcal{S}}, \cap_{s\in\mathcal{S}} \arg\min_{\boldsymbol{P}\in\mathcal{P}} \sum_{s,a} \pi_{sa} \boldsymbol{P}_{sa}^{\top} \boldsymbol{V} \neq \emptyset. \qquad \text{(Weak SSP)}$$

Many models of uncertainty are weakly s-tractable:

1. s-rec. works
2. r-rec. works: $\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa} \Rightarrow \boldsymbol{P}_{sa}^{\top}\boldsymbol{V} = \boldsymbol{u}_{sa}^{\top}\boldsymbol{W}^{\top}\boldsymbol{V}$
3. $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2$ such that $\mathcal{P}_1$ is r-rec. and $\mathcal{P}_2$ is s-rec.
4. (ξ, η)-rec. works: $\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa}$, $\boldsymbol{W}$ and $\boldsymbol{u}$ in Cartesian product set
5. Other models in the paper; what's important is the unified analysis

$$\forall\,(\pi, \boldsymbol{V}) \in \Pi_{\mathsf{S}} \times \mathbb{R}^{\mathcal{S}}, \cap_{s\in\mathcal{S}} \arg\min_{\boldsymbol{P}\in\mathcal{P}} \sum_{s,a} \pi_{sa} \boldsymbol{P}_{sa}^{\top} \boldsymbol{V} \neq \emptyset. \qquad \text{(Weak SSP)}$$

Many models of uncertainty are weakly s-tractable:

1. s-rec. works
2. r-rec. works: $\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa} \Rightarrow \boldsymbol{P}_{sa}^{\top}\boldsymbol{V} = \boldsymbol{u}_{sa}^{\top}\boldsymbol{W}^{\top}\boldsymbol{V}$
3. $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2$ such that $\mathcal{P}_1$ is r-rec. and $\mathcal{P}_2$ is s-rec.
4. (ξ, η)-rec. works: $\boldsymbol{P}_{sa} = \boldsymbol{W}\boldsymbol{u}_{sa}$, $\boldsymbol{W}$ and $\boldsymbol{u}$ in Cartesian product set
5. Other models in the paper; what's important is the unified analysis

But non-rectangularity appears useless!

Discussion and conclusion

Take-aways:

1. The case for s-rectangular uncertainty sets:
   - Other models are not *always* tractable
   - Nonrectangular models "bring down" to rectangular extensions

Discussion and conclusion

Take-aways:

1. The case for s-rectangular uncertainty sets:
   - Other models are not *always* tractable
   - Nonrectangular models "bring down" to rectangular extensions
2. Change of paradigm for the design of uncertainty sets:
   Weak SSP "by design" $\Rightarrow$ tractability "for free"

Discussion and conclusion

Take-aways:

1. The case for s-rectangular uncertainty sets:
   - Other models are not *always* tractable
   - Nonrectangular models "bring down" to rectangular extensions
2. Change of paradigm for the design of uncertainty sets:
   Weak SSP "by design" $\Rightarrow$ tractability "for free"
3. "Hardness" of $\sup_{\pi \in \Pi_{\mathrm{H}}} \min_{\boldsymbol{P} \in \mathcal{P}}$ lies in $\min_{\boldsymbol{P} \in \mathcal{P}}$
   e.g. r-rec.: [GBZ+18] then [GGC22]

Discussion and conclusion

Take-aways:

1. The case for s-rectangular uncertainty sets:
   - Other models are not *always* tractable
   - Nonrectangular models "bring down" to rectangular extensions
2. Change of paradigm for the design of uncertainty sets:
   Weak SSP "by design" $\Rightarrow$ tractability "for free"
3. "Hardness" of $\sup_{\pi \in \Pi_{\mathrm{H}}} \min_{\boldsymbol{P} \in \mathcal{P}}$ lies in $\min_{\boldsymbol{P} \in \mathcal{P}}$
   e.g. r-rec.: [GBZ+18] then [GGC22]

Next steps:

- Beyond dynamic programming: gradient based-methods?

More in the paper + if you are interested in this topic:

Yan Li and Alexander Shapiro: *Rectangularity and duality of DRMDPs*

Discussion and conclusion

Take-aways:

1. The case for s-rectangular uncertainty sets:
   - Other models are not *always* tractable
   - Nonrectangular models "bring down" to rectangular extensions
2. Change of paradigm for the design of uncertainty sets:
   Weak SSP "by design" $\Rightarrow$ tractability "for free"
3. "Hardness" of $\sup_{\pi \in \Pi_{\mathrm{H}}} \min_{\boldsymbol{P} \in \mathcal{P}}$ lies in $\min_{\boldsymbol{P} \in \mathcal{P}}$
   e.g. r-rec.: [GBZ+18] then [GGC22]

Next steps:

- Beyond dynamic programming: gradient based-methods?

More in the paper + if you are interested in this topic:

Yan Li and Alexander Shapiro: *Rectangularity and duality of DRMDPs*

Thank you!

References

Joel Goh, Mohsen Bayati, Stefanos A Zenios, Sundeep Singh, and David Moore.
Data uncertainty in Markov chains: Application to cost-effectiveness analyses of medical innovations.
Operations Research, 66(3):697–715, 2018.

Vineet Goyal and Julien Grand-Clément.
Robust Markov decision processes: Beyond rectangularity.
Mathematics of Operations Research, 2022.

Yang Hu, Haitong Ma, Bo Dai, and Na Li.
Efficient duple perturbation robustness in low-rank mdps.
arXiv preprint arXiv:2404.08089, 2024.

Yann Le Tallec.
Robust, risk-sensitive, and data-driven control of Markov decision processes.
PhD thesis, Massachusetts Institute of Technology, 2007.

- Xiaoteng Ma, Zhipeng Liang, Jose Blanchet, Mingwen Liu, Li Xia, Jiheng Zhang, Qianchuan Zhao, and Zhengyuan Zhou.
  Distributionally robust offline reinforcement learning with linear function approximation.
  arXiv preprint arXiv:2209.06620, 2022.
- S. Mannor, O. Mebel, and H. Xu.
  Robust MDPs with k-rectangular uncertainty.
  Mathematics of Operations Research, 41(4):1484–1509, 2016.
- Wolfram Wiesemann, Daniel Kuhn, and Berç Rustem.
  Robust Markov decision processes.
  Mathematics of Operations Research, 38(1):153–183, 2013.