From infinite to finite programs: explicit error bounds with an application to approximate dynamic programming

Peyman Mohajerin Esfahani, Tobias Sutter, Daniel Kuhn, and John Lygeros

Automatic Control Laboratory

Cambridge & Oxford, February 2017

ETH zürich

Constrained Markov Decision Processes (MDPs)

- Markov control model $(S, A, \{A(s) | s \in S\}, Q, \psi, d, \ell)$
- long-run average cost

$$J_\psi(\pi, \nu) := \limsup_{n \to \infty} \frac{1}{n} \mathbb{E}_\nu^\pi \left(\sum_{t=0}^{n-1} \psi(s_t, a_t) \right)$$

- constrained MDP

$$\rho_{\min} = \begin{cases} \inf_{\nu, \pi} & J_\psi(\pi, \nu) \\ \text{s.t.} & J_d(\pi, \nu) \leq \ell \end{cases}$$

  - multi-objective problems
  - MDP under limited information constraints

- simplifying assumptions
  - $S, A \subset \mathbb{R}^n$ compact
  - $\{A(s) | s \in S\} = A$

Linear Programming Formulation

$$\text{LP}: \quad J^{\text{ac}} := \begin{cases} \min\limits_{\mu} & \langle \mu, \psi \rangle \\ \text{s.t.} & \langle \mathrm{L}\mu, u \rangle = 0 \quad \forall u \in \mathcal{C}(X) \\ & \langle \mu, d \rangle \leq \ell \\ & \mu \in \mathcal{P}(X \times A) \end{cases}$$

- Linear, continuous operator $\mathrm{L} : \mathbb{M}(X \times A) \to \mathbb{M}(X)$

$$\mathrm{L}\mu(B) := \mu(B \times A) - \int_{X \times A} Q(B|x, a)\mu(\mathrm{d}(x, a)), \quad B \in \mathcal{B}(X)$$

Theorem (LP equivalence)

Under certain technical assumptions, (LP) is solvable and $\rho_{\min} = J^{\text{ac}}$

How to approximate (LP)?

- finite LP approximation [Hernández-Lerma & Lasserre'98], [Dufour & Prieto-Rumeau'13]
- ADP [De Farias & van Roy'03]
- many more: Bertsekas, Yüksel, Boyd, ...

Infinite dimensional LPs

$$
\begin{array}{cc}
(\mathbb{X}, & \mathbb{C}) \\
\mathcal{A}\downarrow & \uparrow \mathcal{A}^* \\
(\mathbb{B}, & \mathbb{Y}) \\
\cup & \cup \\
\mathbb{K} & \mathbb{K}^*
\end{array}
$$

- dual pair normed vector spaces
- linear operator
- closed convex cone
- dual cone

Primal LP ($\mathcal{P}$)

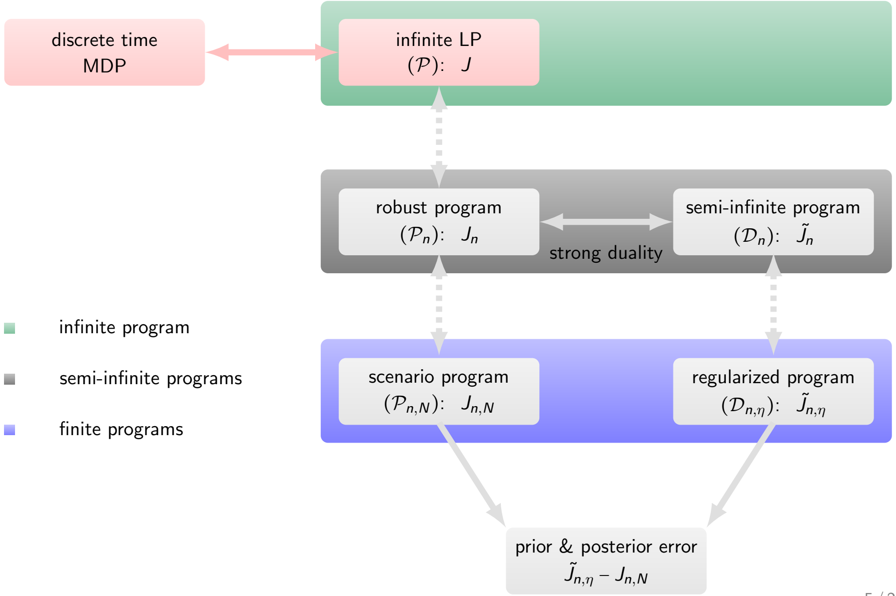
$$
J := \begin{cases} \inf_x & \langle x, c \rangle \\ \text{s.t.} & \mathcal{A}x \geq_{\mathbb{K}} b \\ & x \in \mathbb{X} \end{cases}
$$

Dual LP ($\mathcal{D}$)

$$
\tilde{J} := \begin{cases} \sup_y & \langle b, y \rangle \\ \text{s.t.} & \mathcal{A}^* y = c \\ & y \in \mathbb{K}^* \end{cases}
$$

- weak duality $\quad J \geq \tilde{J}$
- zero duality gap $\quad J = \tilde{J}$

Outline



discrete time MDP
infinite LP
$(\mathcal{P})$: J
robust program
$(\mathcal{P}_n)$: J_n
strong duality
semi-infinite program
$(\mathcal{D}_n)$: $\tilde{J}_n$
scenario program
$(\mathcal{P}_{n,N})$: $J_{n,N}$
regularized program
$(\mathcal{D}_{n,\eta})$: $\tilde{J}_{n,\eta}$
prior & posterior error
$\tilde{J}_{n,\eta} - J_{n,N}$
infinite program
semi-infinite programs
finite programs

Semi-infinite approximation

- linearly independent elements $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{X}$
- $\mathbb{X}_n := \operatorname{span}\{x_1, \cdots, x_n\}$, $\quad \alpha \in \mathbb{R}^n$, $\quad \boldsymbol{c} := [\langle x_1, c\rangle, \cdots, \langle x_n, c\rangle]$
- unit "ball" $\mathrm{B}_n := \{\sum_{i=1}^n \alpha_i x_i \in \mathbb{X}_n : \|\alpha\|_{\mathcal{R}} \le \theta_{\mathcal{P}}\}$
- operator $\mathcal{A}_n : \mathbb{R}^n \to \mathbb{B}$, $\mathcal{A}_n\alpha := \sum_{i=1}^n \alpha_i \mathcal{A} x_i$

Primal semi-inf. program $(\mathcal{P}_n)$

$$J_n := \begin{cases} \inf\limits_{\alpha} & \alpha^\top \boldsymbol{c} \\ \text{s.t.} & \mathcal{A}_n\alpha \succeq_{\mathbb{K}} b \\ & \|\alpha\|_{\mathcal{R}} \le \theta_{\mathcal{P}} \end{cases}$$

Dual semi-inf. program $(\mathcal{D}_n)$

$$\tilde{J}_n := \begin{cases} \sup\limits_{y\in\mathbb{Y}} & \langle b, y\rangle - \theta_{\mathcal{P}}\|\mathcal{A}_n^* y - \boldsymbol{c}\|_{\mathcal{R}^*} \\ \text{s.t.} & y \in \mathbb{K}^*, \end{cases}$$

Assumptions

1. there exists $\gamma > 0$ so that $\|\mathcal{A}_n^* y\|_{\mathcal{R}^*} \ge \gamma\|y\|_*$ for every $y \in \mathbb{K}^*$
2. $(\mathcal{P}_n)$ is feasible and $\theta_{\mathcal{P}} > \|b\|/\gamma$

Theorem (Semi-infinite approximation)

Let $r_n := x^\star - \Pi_{\mathrm{B}_n}(x^\star)$, then

$$0 \le J_n - J \le \|r_n\| \left(\|c\|_* + \frac{2\theta_{\mathcal{P}}\|\mathcal{A}\|\|\boldsymbol{c}\|_{\mathcal{R}^*}}{\gamma\theta_{\mathcal{P}} - \|b\|} \right)$$

Semi-infinite approximation — Proof sketch

Theorem (Semi-infinite approximation)

Let $r_n := x^\star - \Pi_{\mathrm{B}_n}(x^\star)$, then

$$0 \le J_n - J \le \|r_n\| \left(\|c\|_* + \frac{2\theta_{\mathcal{P}} \|\mathcal{A}\| \|\boldsymbol{c}\|_{\mathcal{R}^*}}{\gamma\theta_{\mathcal{P}} - \|b\|} \right)$$

1. zero duality gap between $(\mathcal{P}_n)$ and $(\mathcal{D}_n)$
2. Sion's minimax theorem
3. bounded dual optimizer

$$\|y_n^\star\|_* \le \frac{2\theta_{\mathcal{P}} \|\boldsymbol{c}\|_{\mathcal{R}^*}}{\gamma\theta_{\mathcal{P}} - \|b\|}$$

MDP setting

$$J := \begin{cases} \inf_x & \langle x, c \rangle \\ \text{s.t.} & \mathcal{A}x \geq_{\mathbb{K}} b \\ & x \in \mathbb{X} \end{cases}$$

- $$\begin{cases} \mathbb{X} := \mathbb{R} \times \mathcal{L}(S), & \mathbb{C} := \mathbb{R} \times \mathcal{M}(S) \\ \|x\| := \max\{|x_1|, \|x_2\|_\infty\}, & \|c\|_* = |c_1| + \|c_2\|_{\mathrm{W}} \\ \mathbb{B} := \mathcal{L}(S \times A), & \mathbb{Y} := \mathcal{M}(S \times A) \\ \|b\| := \|b\|_{\mathrm{L}}, & \|y\|_* = \|y\|_{\mathrm{W}} \\ \mathbb{K} := \mathcal{L}_+(S \times A), & \mathbb{K}^* := \mathcal{M}_+(S \times A) \end{cases}$$
- $$\begin{cases} b := -\psi \\ c := (-1, 0) \\ \mathcal{A}(\rho, u)(s, a) := -\rho - u(s) + Qu(s, a) \end{cases}$$
- bounded dual optimizer in $(\mathcal{D}_n)$

$$\|y_n^\star\|_{\mathrm{W}} \leq 1$$

- inf-sup condition paramter $\gamma = 1$ (average cost, min-max structure)

MDP setting (cont'd)

- $\mathbb{U}_n := \{\sum_{i=1}^n \alpha_i u_i \ : \ \|\alpha\|_{\mathcal{R}} \leq \theta_{\mathcal{P}}\}$
- L_Q Lipschitz constant of the transition kernel Q

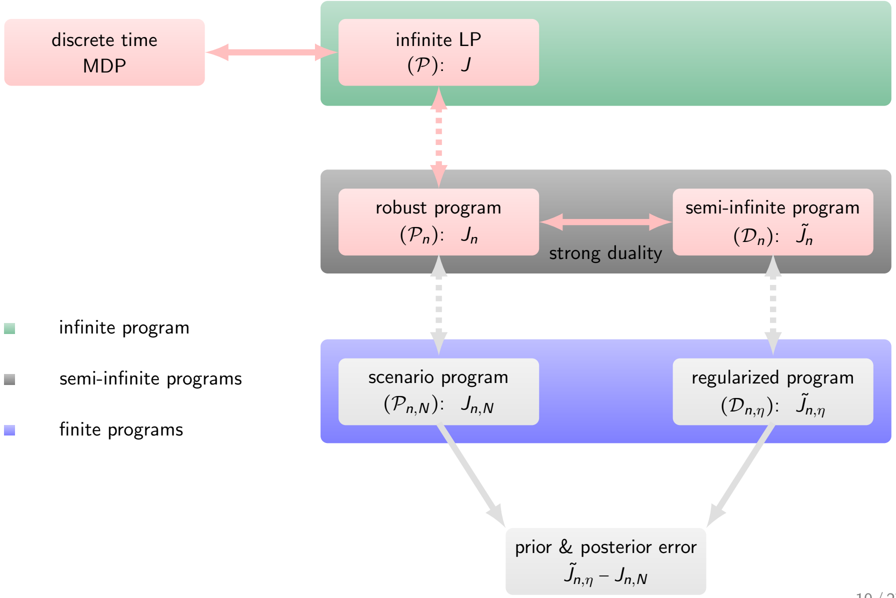
Corollary (Semi-infinite approximation)

Let $u^\star$ be the optimal value function and $\theta_{\mathcal{P}} > \|\psi\|_{\mathrm{L}}$. Then,

$$0 \leq J^{\mathrm{ac}} - J_n^{\mathrm{ac}} \leq (1 + \max\{L_Q, 1\}) \left\| u^\star - \Pi_{\mathbb{U}_n}(u^\star) \right\|_{\mathrm{L}}$$

- projection residual $\left\| u^\star - \Pi_{\mathbb{U}_n}(u^\star) \right\|_{\mathrm{L}}$
  - influenced by choice of basis functions
  - convergence rate $n^{-1/\dim(S)}$

Outline


discrete time
MDP
infinite LP
$(\mathcal{P})$: J
robust program
$(\mathcal{P}_n)$: J_n
strong duality
semi-infinite program
$(\mathcal{D}_n)$: $\tilde{J}_n$
scenario program
$(\mathcal{P}_{n,N})$: $J_{n,N}$
regularized program
$(\mathcal{D}_{n,\eta})$: $\tilde{J}_{n,\eta}$
prior & posterior error
$\tilde{J}_{n,\eta} - J_{n,N}$
infinite program
semi-infinite programs
finite programs

Finite approximation — randomized approach

- $\mathcal{A}x \succeq_{\mathbb{K}} b \;\Leftrightarrow\; \langle \mathcal{A}x - b, y \rangle \geq 0,\; \forall y \in \mathcal{K} := \mathcal{E}\{y \in \mathbb{K}^* : \|y\|_* = 1\}$
- $\{y_j\}_{j \leq N}$ i.i.d. samples on $(\mathcal{K}, \mathbb{P})$

Scenario program $(\mathcal{P}_{n,N})$

$$J_{n,N} := \begin{cases} \min\limits_{\alpha} & \alpha^\top \boldsymbol{c} \\ \text{s.t.} & \alpha^\top \mathcal{A}_n^* y_j \geq \langle b, y_j \rangle, \;\; \forall j \leq N \\ & \|\alpha\|_{\mathcal{R}} \leq \theta_{\mathcal{P}} \end{cases}$$

- $x^\star_{n,N} := \sum_{i=1}^{n} \alpha_i^\star x_i$
- $N(\varepsilon, \beta) := \min\left\{ N \in \mathbb{N} \;:\; \sum_{i=0}^{n-1} \binom{N}{i} \varepsilon^i (1-\varepsilon)^{N-i} \leq \beta \right\}$

Theorem (Finite approximation)

Let $\gamma\theta_{\mathcal{P}} > \|b\|$, for all ε, β in $(0,1)$ and $N \geq N(\varepsilon, \beta)$ we have

$$\mathbb{P}^N\left[0 \leq J_n - J_{n,N} \leq \left(\frac{2\theta_{\mathcal{P}} \|\boldsymbol{c}\|_{\mathcal{R}^*}}{\gamma\theta_{\mathcal{P}} - \|b\|} \right) h\left(x^\star_{n,N}, \varepsilon\right) \right] \geq 1 - \beta$$

Finite approximation — randomized approach (cont'd)

Theorem (Finite approximation)

Let $\gamma\theta_{\mathcal{P}} > \|b\|$, for all ε, β in $(0,1)$ and $N \geq N(\varepsilon, \beta, n)$ we have

$$\mathbb{P}^N\left[0 \leq J_n - J_{n,N} \leq \left(\frac{2\theta_{\mathcal{P}}\|\boldsymbol{c}\|_{\mathcal{R}^*}}{\gamma\theta_{\mathcal{P}} - \|b\|}\right) h\left(x^\star_{n,N}, \varepsilon\right)\right] \geq 1 - \beta$$

- extension of [Mohajerin Esfahani, S, Lygeros, TAC 2015]
- sampling complexity $N(\varepsilon, \beta, n) \sim \{1/\varepsilon,\ \log(1/\beta),\ n\}$ [Campi & Garatti, 2008] → extends to non-convex setting
- Tail bound function $h(x, \varepsilon)$
  - curse of dimensionality $h(x, \varepsilon) \sim (1/\varepsilon)^{\dim \mathcal{K}}$
  - candidate can be proposed, see also [Kanamori & Takeda, 2012]
  - asymptotic convergence $\sup_{x \in \mathbb{X}_n} h(x, \varepsilon) \xrightarrow[\varepsilon \to 0]{} 0$
- bounded dual optimizer in $(\mathcal{D}_n)$
  $$\|y^\star_n\|_* \leq \frac{2\theta_{\mathcal{P}}\|\boldsymbol{c}\|_{\mathcal{R}^*}}{\gamma\theta_{\mathcal{P}} - \|b\|}$$
- a priori bound available as well

Randomized approach — MDP setting

- $K := S \times A, \quad |K| := \sup\{\|k - k'\|_\infty \ : \ k, k' \in K\}$
- $\rho_n := \sup\{\|\alpha\|_{\ell_1} \ : \ \|\alpha\|_{\mathcal{R}} \leq 1\}$

Corollary (Finite approximation)

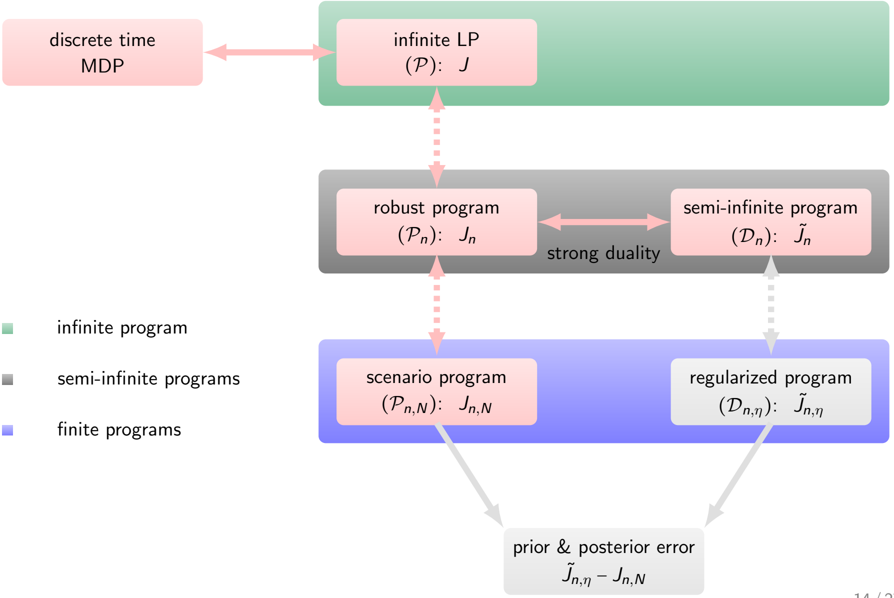
Let $\theta_{\mathcal{P}} > \|\psi\|_\infty$, for all ε, β in $(0,1)$ and $N \geq N(\varepsilon, \beta, n)$ we have

$$\mathbb{P}^N\left[0 \leq J^{\text{ac}}_{n,N} - J^{\text{ac}}_n \leq \left(\theta_{\mathcal{P}}\rho_n(L_Q+1) + \|\psi\|_{\mathrm{L}}\right)|K|\varepsilon^{1/\dim(K)}\right] \geq 1 - \beta$$

curse of dimensionality

- optimal choice of the regularization term $\theta_{\mathcal{P}}$
  - first approximation (projection residual): $\theta_{\mathcal{P}}$ large
  - second approximation (sample complexity): $\theta_{\mathcal{P}}$ small

Outline


discrete time
MDP
infinite LP
$(\mathcal{P})$: J
robust program
$(\mathcal{P}_n)$: J_n
strong duality
semi-infinite program
$(\mathcal{D}_n)$: $\tilde{J}_n$
infinite program
semi-infinite programs
finite programs
scenario program
$(\mathcal{P}_{n,N})$: $J_{n,N}$
regularized program
$(\mathcal{D}_{n,\eta})$: $\tilde{J}_{n,\eta}$
prior & posterior error
$\tilde{J}_{n,\eta} - J_{n,N}$

Structural convex optimization

- $\mathcal{A} := \left\{\alpha \in \mathbb{R}^n : \|\alpha\|_{\mathcal{R}} \le \theta_{\mathcal{P}}\right\}, \qquad \mathcal{Y} := \left\{y \in \mathbb{K}^* : \|y\|_* \le \theta_{\mathcal{D}}\right\},$

regularized program $(\mathcal{D}_{n,\eta})$

$$\tilde{J}_{n,\eta} := \sup_{y \in \mathcal{Y}} \left\{\left\langle b, y\right\rangle - \theta_{\mathcal{P}} \|\mathcal{A}_n^* y - \boldsymbol{c}\|_{\mathcal{R}^*} - \eta d(y)\right\},$$

- prox-function $d : \mathcal{Y} \to \mathbb{R}_+$, strongly convex, $\qquad \eta \in \mathbb{R}_+$
  - $y_\eta^\star(\alpha) := \arg\max_{y \in \mathcal{Y}} \left\{\left\langle b - \mathcal{A}_n \alpha, y\right\rangle - \eta d(y)\right\}$ analytically available
  - $\alpha \mapsto \mathcal{A}_n^* y_\eta^\star(\alpha)$ is Lipschitz continuous with a constant $\frac{L}{\eta}$
- convex duality

$$\tilde{J}_{n,\eta} = \inf_{\alpha \in \mathcal{A}} \alpha^\top \boldsymbol{c} + \left\langle b - \mathcal{A}_n \alpha, y_\eta^\star(\alpha)\right\rangle - \eta d\big(y_\eta^\star(\alpha)\big)$$

  - finite dimensional smooth convex optimization problem
  - use Nesterov's fast gradient schemes

Structural convex optimization (cont'd)

Fast Gradient Method

- For $k \geq 0$ do
  1. Define $r^{(k)} := \frac{\eta}{L}\Big(\boldsymbol{c} - \mathcal{A}_n^* y_\eta^\star\big(w^{(k)}\big)\Big)$;
  2. Compute $z^{(k)} := \mathbb{T}\Big(\sum_{j=0}^{k}\frac{j+1}{2}r^{(j)}, 0\Big), \quad \alpha^{(k)} := \mathbb{T}\Big(\frac{1}{\vartheta}r^{(k)}, w^{(k)}\Big)$;
  3. Set $w^{(k+1)} = \frac{2}{k+3}z^{(k)} + \frac{k+1}{k+3}\alpha^{(k)}$

Theorem (Finite approximation)

After k iterations, let $\hat{\alpha}_\eta := \alpha^{(k)}, \quad \hat{y}_\eta := \sum_{j=0}^{k}\frac{2(j+1)}{(k+1)(k+2)}y_\eta^\star\big(w^{(j)}\big)$

$$\Rightarrow J_{n,\eta}^{\mathrm{LB}} \leq J_n \leq J_{n,\eta}^{\mathrm{UB}} \quad \text{(a posteriori error)}$$

$$J_{n,\eta}^{\mathrm{LB}} := \big\langle b, \hat{y}_\eta\big\rangle - \theta_{\mathcal{P}}\|\mathcal{A}_n^*\hat{y}_\eta - \boldsymbol{c}\|_{\mathcal{R}^*}, \; J_{n,\eta}^{\mathrm{UB}} := \hat{\alpha}_\eta{}^\top\boldsymbol{c} + \sup_{y\in\mathcal{Y}}\big\langle b - \mathcal{A}_n\hat{\alpha}_\eta, y\big\rangle.$$

$$\text{For } k \geq F(\varepsilon, d, \mathcal{A}) \Rightarrow J_{n,\eta}^{\mathrm{UB}} - J_{n,\eta}^{\mathrm{LB}} \leq \varepsilon \quad \text{(a priori error)}$$

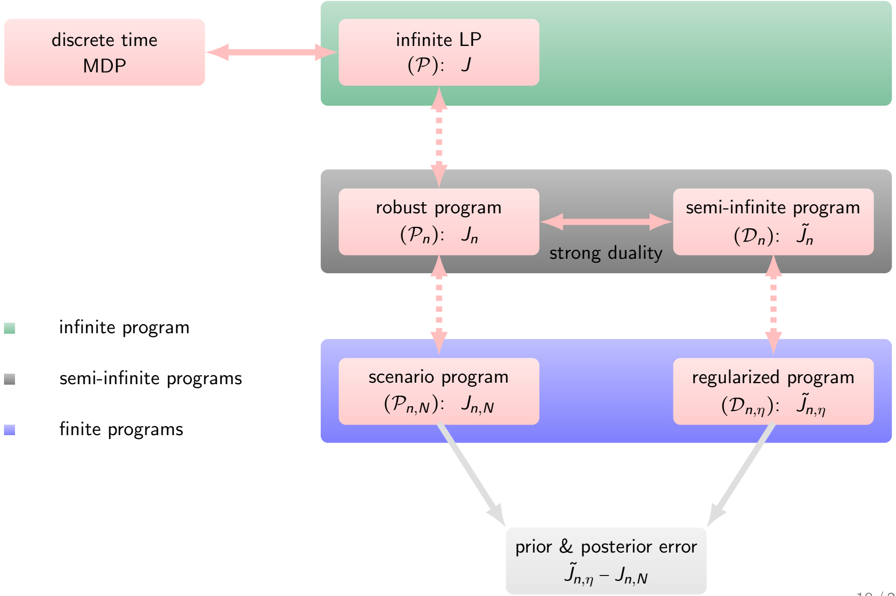
Structural convex optimization — MDP setting

- prox term $d \to$ relative entropy
  - $y^\star_\eta(\alpha) := \arg\max_{y\in\mathcal{Y}} \left\{ \left\langle b - \mathcal{A}_n\alpha, y \right\rangle - \eta d(y) \right\}$
    entropy maximization (analytically available)
  - $\alpha \mapsto \mathcal{A}_n^* y^\star_\eta(\alpha)$ Lipschitz cts. with $L = 4(\sup_{\|\alpha\|_{\mathcal{R}}\le 1} \|\alpha\|_{\ell_1})^2$

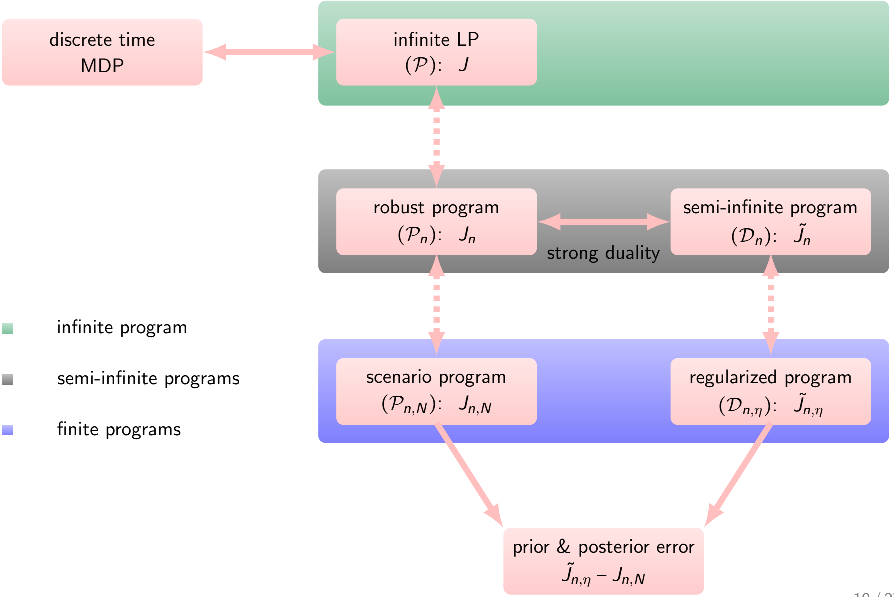
- complexity $\mathcal{O}\left(\varepsilon^{-1}\sqrt{\log(\varepsilon^{-1})}\right)$
  - computation of $\mathcal{A}_n^* y^\star_\eta$ (integration over $S \times A$)
  - choice of basis functions (e.g., polynomial optimization), QMC methods
  - radial basis functions [Kariotoglou, Margellos, Lygeros, 2016]

- fast gradient method with inexact gradient [Devolder & Nesterov'13]

Outline


discrete time MDP
infinite LP
$(\mathcal{P})$: J
robust program
$(\mathcal{P}_n)$: J_n
strong duality
semi-infinite program
$(\mathcal{D}_n)$: $\tilde{J}_n$
scenario program
$(\mathcal{P}_{n,N})$: $J_{n,N}$
regularized program
$(\mathcal{D}_{n,\eta})$: $\tilde{J}_{n,\eta}$
prior & posterior error
$\tilde{J}_{n,\eta} - J_{n,N}$
infinite program
semi-infinite programs
finite programs

Outline



discrete time MDP
infinite LP
$(\mathcal{P})$: J
robust program
$(\mathcal{P}_n)$: J_n
semi-infinite program
$(\mathcal{D}_n)$: $\tilde{J}_n$
strong duality
scenario program
$(\mathcal{P}_{n,N})$: $J_{n,N}$
regularized program
$(\mathcal{D}_{n,\eta})$: $\tilde{J}_{n,\eta}$
prior & posterior error
$\tilde{J}_{n,\eta} - J_{n,N}$
infinite program
semi-infinite programs
finite programs

Example 1: truncated LQG

- linear dynamics $s_{t+1} = \vartheta s_t + \rho a_t + \xi_t, \quad t \in \mathbb{N}$
- quadratic cost $\psi(s,a) = qs^2 + ra^2, \quad q \geq 0$ and $r > 0$
- $S = A = [-L, L]$
- $\{\xi_t\}_{t\in\mathbb{N}} \overset{\text{i.i.d.}}{\sim}$ truncated normal distribution

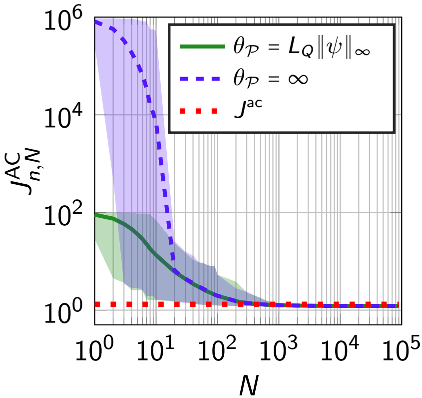


Figure: $n = 2$ basis functions

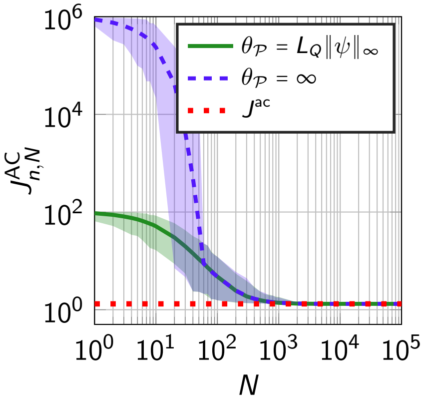


Figure: $n = 10$ basis functions

Example 1: truncated LQG (cont'd)

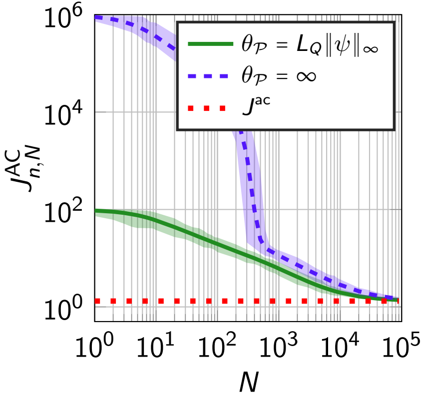


Figure: $n = 100$ basis functions

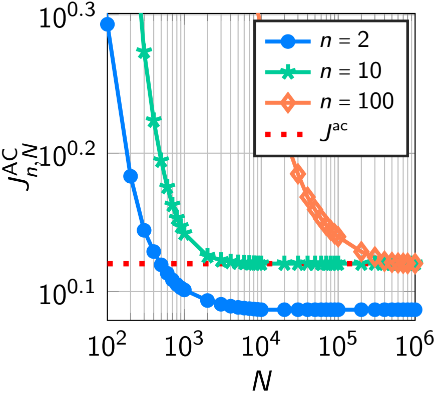


Figure: zoomed version of the average cost for different n

Example 1: truncated LQG (cont'd)

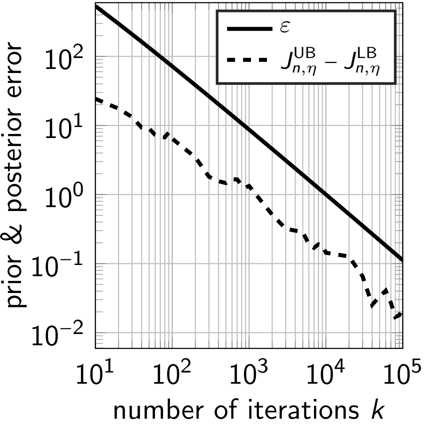


Figure: $n = 10$ basis functions

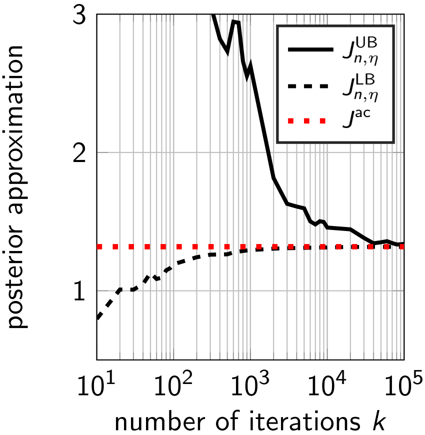


Figure: $n = 10$ basis functions

Example 2: fishery management

- Rickner model: $s_{t+1} = \vartheta_1 a_t \exp(-\vartheta_2 a_t + \xi_t), \quad t \in \mathbb{N}$
  - s_t population size at time t
  - a_t population to be left for spawning for the next season
  - $S = [\underline{\kappa}, \overline{\kappa}], \quad A(s) = [\underline{\kappa}, s]$
  - reward $\psi(a, s) = \varphi(s - a)$, $\varphi(z) := 3(z + 0.5)^{1/3} - (0.5)^{1/3}$
  - $\{\xi_t\}_{t\in\mathbb{N}} \overset{\text{i.i.d.}}{\sim}$ uniformly distributed on $[0, \lambda]$
  - consitency assumption: $\vartheta_1 a \exp(-\vartheta_2 a + \xi) \in [\underline{\kappa}, \overline{\kappa}]$ for all $(a, \xi) \in [\underline{\kappa}, \overline{\kappa}] \times [0, \lambda]$

Example 2: fishery management (cont'd)

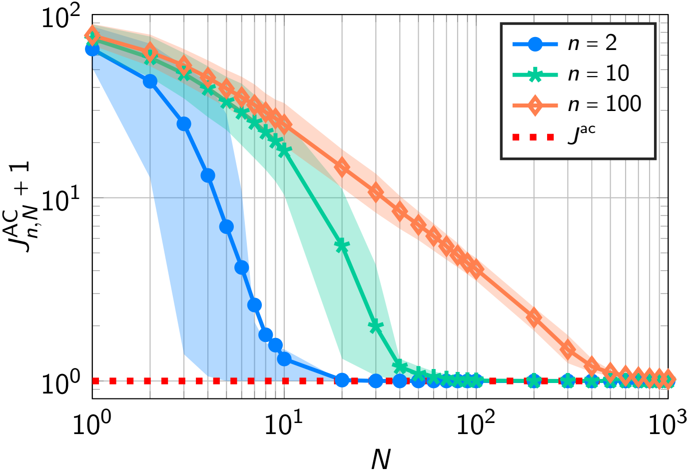


Figure: The objective performance $J^{ac}_{n,N}$ for 200 independent experiments

Example 2: fishery management (cont'd)

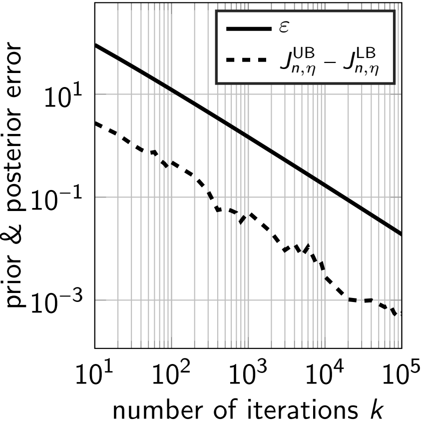


Figure: $n = 10$ basis functions

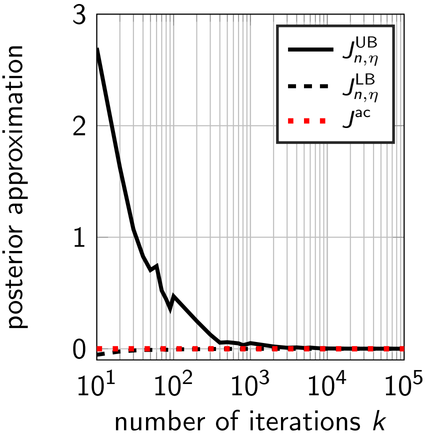


Figure: $n = 10$ basis functions

Conclusion & Future Work

arXiv:1701.06379

Conclusion

- approximation method for a class of infinite LPs
- LP formulation of MDPs on continuous spaces
  - constrained MDPs
  - using tools from mathematical programming

Outlook

- ε-approximating control policy
- how to sample (informative) state-action pairs
- unknown dynamics (reinforcement learning)

Acknowlegements

Peyman
Mohajerin Esfahani

Daniel Kuhn

John Lygeros

References:

- Mohajerin Esfahani, S., Kuhn and Lygeros, "From Infinite to Finite Programs: Explicit Error Bounds with Applications to Approximate Dynamic Programming", *arXiv:1701.06379*, 2017
- Mohajerin Esfahani, S., and Lygeros, "Performance bounds for the scenario approach and an extension to a class of non-convex programs", *IEEE TAC*, vol. 60, no. 1, 2015

Scenario-Based Approximation (general setting)

$$J^\star := \begin{cases} \min_x & c^\top x \\ \text{s.t.} & f(x,d) \le 0, \quad \forall d \in \mathcal{D} \\ & x \in \mathcal{X} \end{cases}$$

$$J^\star_N := \begin{cases} \min_x & c^\top x \\ \text{s.t.} & f(x,d_i) \le 0, \quad \forall i = 1, \cdots, N \\ & x \in \mathcal{X} \end{cases}$$

Assumptions

(i) f convex in x for each d

(ii) f bounded in d for each x

Number of samples for $\varepsilon, \beta \in [0,1]$ [Campi & Garatti'08]

$$N(\varepsilon,\beta) := \min\left\{ N \in \mathbb{N} \;\middle|\; \sum_{i=0}^{n-1} \binom{N}{i} \varepsilon^i (1-\varepsilon)^{N-i} \le \beta \right\}$$

Theorem (Mohajerin Esfahani, S., Lygeros, TAC 2015)

For all $N \ge N(\varepsilon,\beta)$, $\quad \mathbb{P}^N\left[J^\star - J^\star_N \in \left[0, I(\varepsilon)\right]\right] \ge 1-\beta$, where

$I(\varepsilon) := \min\left\{ \mathrm{L}h(\varepsilon), \; \max_{x\in\mathcal{X}} c^\top x - \min_{x\in\mathcal{X}} c^\top x \right\}$

Scenario Based Approximation (cont'd)

(i) Tail probability of the worst-case violation

$$p(x,\delta) := \mathbb{P}\Big[\sup_{v\in\mathcal{D}} f(x,v) - \delta < f(x,d)\Big].$$

(ii) uniform level-set bound $h(\cdot)$ of p if for all $\varepsilon \in [0,1]$

$$h(\varepsilon) \geq \sup\Big\{\delta \in \mathbb{R}_+ \;\Big|\; \inf_{x\in\mathcal{X}} p(x,\delta) \leq \varepsilon\Big\}.$$

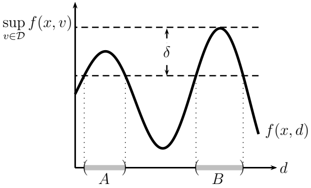


Figure: $p(x,\delta) = \mathbb{P}[A \cup B]$

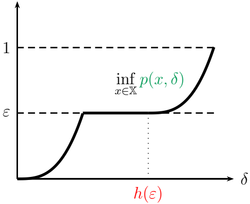


Figure: Uniform level set bound