Approximation of Constrained Average Cost Markov Control Processes

Tobias Sutter, Peyman Mohajerin Esfahani, and John Lygeros

Automatic Control Laboratory

Conference on Decision and Control, 2014

ETH zürich

Constrained Markov Decision Processes (MDPs)

- Markov control model $(X, A, \{A(x) | x \in X\}, Q, c, d, \ell)$
- long-run average cost

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

- constrained MDP

$$\rho_{\min} = \begin{cases} \inf\limits_{\nu, \pi} & J_c(\pi, \nu) \\ \text{s.t.} & J_d(\pi, \nu) \leq \ell \end{cases}$$

  - multi-objective problems
  - MDP under limited information constraints
- simplifying assumptions (for this talk only)
  - $X, A \subset \mathbb{R}^n$ compact
  - $\{A(x) | x \in X\} = A$

Linear Programming Formulation

$$\text{Primal LP} \quad \mathrm{P}: \quad J^\star := \begin{cases} \min_{\mu} & \langle \mu, c \rangle \\ \text{s.t.} & \langle \mathrm{L}\mu, u \rangle = 0 \quad \forall u \in \mathcal{C}(X) \\ & \langle \mu, d \rangle \le \ell \\ & \mu \in \mathcal{P}(X \times A) \end{cases}$$

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

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

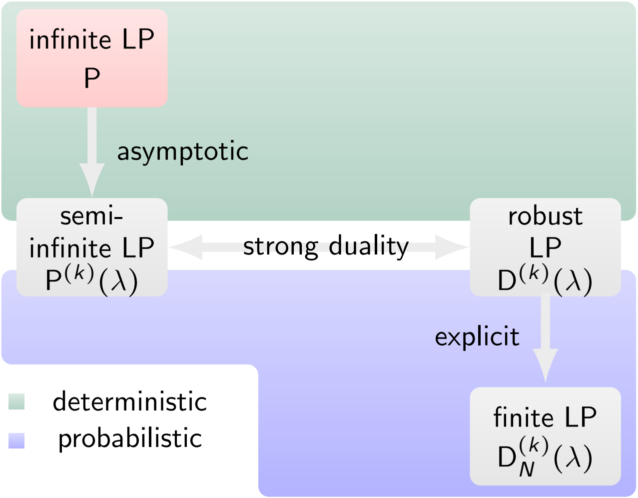
Theorem (LP equivalence)

Under certain technical assumptions, (P) is solvable and $\rho_{\min} = J^\star$

How to approximate (P)?

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

Outline


infinite LP
P
asymptotic
semi-
infinite LP
$P^{(k)}(\lambda)$
strong duality
robust
LP
$D^{(k)}(\lambda)$
explicit
finite LP
$D_N^{(k)}(\lambda)$
deterministic
probabilistic

Step1 — From infinite to semi-infinite LP

- basis functions $\{u_1, \cdots, u_k\}$

Semi-inf. LP $\quad \mathrm{P}^{(k)}(\lambda): \quad J_k^\star(\lambda) := \begin{cases} \min\limits_{\mu,\eta} & \langle \mu, c\rangle + \lambda\eta \\ \text{s.t.} & \left|\langle \mathrm{L}\mu, u_i\rangle\right| \leq \eta \quad \forall i \leq k \\ & \langle \mu, d\rangle \leq \ell + \eta \\ & \mu \in \mathcal{P}(X \times A), \ \eta \in \mathbb{R}_{\geq 0} \end{cases}$

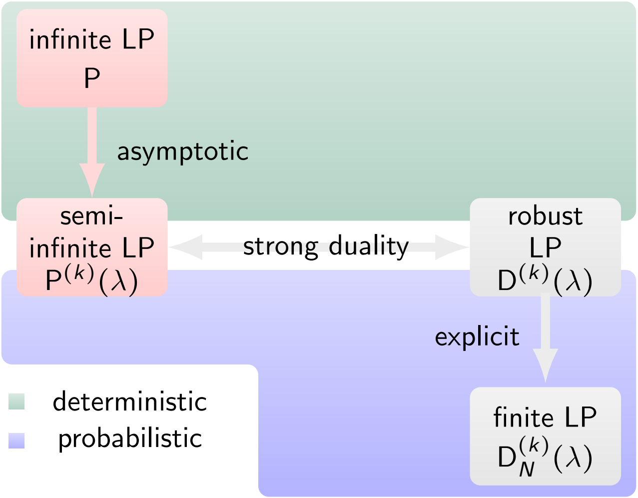
Theorem (convergence)

(i) For every $k \in \mathbb{N}$ and $\lambda \geq 0$, $\mathrm{P}^{(k)}(\lambda)$ is solvable

(ii) $\lim_{k,\lambda\to\infty} J_k^\star(\lambda) = J^\star$

- inspired by [Hernández-Lerma & Lasserre'98]
- penalization term $\lambda\eta$ provides a priori bound on dual variables
- explicit error bound & choice of basis functions ($\to$ Outlook)

Outline


infinite LP
P
asymptotic
semi-infinite LP
$P^{(k)}(\lambda)$
strong duality
robust LP
$D^{(k)}(\lambda)$
explicit
finite LP
$D_N^{(k)}(\lambda)$
deterministic
probabilistic

Duality — Robust LP

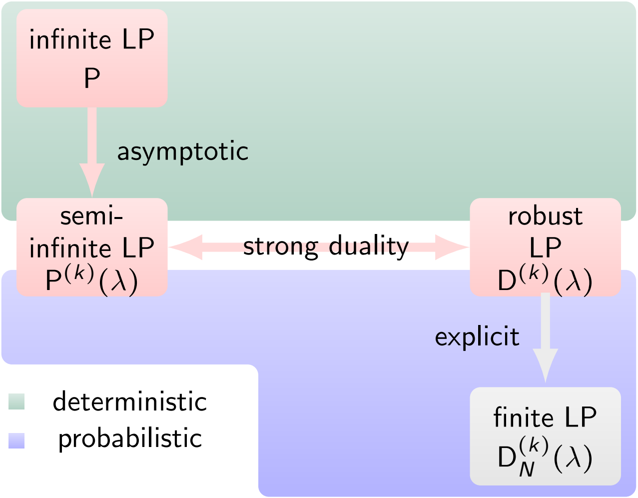
Robust LP $\quad \mathrm{D}^{(k)}(\lambda)$:

$$J_k^\star(\lambda) = \begin{cases} \max\limits_{\rho,\gamma,\alpha,\beta} & \rho - \gamma\ell \\ \text{s.t.} & \rho + \sum_{i=1}^k(\alpha_i - \beta_i)\mathrm{L}^* u_i(x,a) \\ & \qquad \le \gamma d(x,a) + c(x,a) \quad \forall (x,a) \in X \times A \\ & \gamma + \sum_{i=1}^k(\alpha_i + \beta_i) \le \lambda \\ & \rho \in \mathbb{R},\ \gamma \in \mathbb{R}_{\ge 0},\ \alpha,\beta \in \mathbb{R}^k_{\ge 0}, \end{cases}$$

Lemma (strong duality)

For every $k \in \mathbb{N}$ and $\lambda \ge 0$, there is no duality gap between $\mathrm{P}^{(k)}(\lambda)$ and $\mathrm{D}^{(k)}(\lambda)$

Outline



infinite LP
P
asymptotic
semi-infinite LP
$P^{(k)}(\lambda)$
strong duality
robust LP
$D^{(k)}(\lambda)$
explicit
finite LP
$D_N^{(k)}(\lambda)$
deterministic
probabilistic

Step2 — From semi-infinite to finite LP

Method 1 — Scenario Based Approximation

- randomly sample finite number of constraints
- solve finite LP
- probabilistic, explicit error bound
- discussed in detail in the next slides

Method 2 — Adaptive Approximation Scheme

- construct "important" constraints by solving an auxiliary optimization problem
- cutting plane method
- determinisitic, explicit error bound
- see proceedings for details

Scenario-Based Approximation (general setting)

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

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

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 et al., TAC 2015, to appear)

For all $N \ge N(\varepsilon,\beta)$, $\quad \mathbb{P}^N\Big[J^\star - J_N^\star \in \big[0, I(\varepsilon)\big]\Big] \ge 1-\beta$, 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 $\varepsilon \in [0,1]$

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

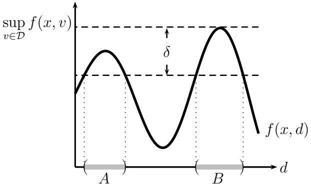


Figure : $p(x,\delta) = \mathbb{P}[A \cup B]$

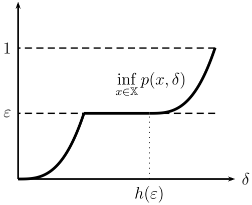


Figure : Uniform level set bound

Step2 — From semi-infinite to finite LP (cont'd)

Scenario LP $\quad \mathrm{D}_N^{(k)}(\lambda):$

$$J^\star_{k,N}(\lambda) = \begin{cases} \max\limits_{\rho,\gamma,\alpha,\beta} & \rho - \gamma\ell \\ \text{s.t.} & \rho + \sum_{i=1}^k (\alpha_i - \beta_i)\mathrm{L}^* u_i(x_j, a_j) \\ & \qquad \le \gamma d(x_j, a_j) + c(x_j, a_j) \quad \forall j = 1, \cdots, N \\ & \gamma + \sum_{i=1}^k (\alpha_i + \beta_i) \le \lambda \\ & \rho \in \mathbb{R},\ \gamma \in \mathbb{R}_{\ge 0},\ \alpha, \beta \in \mathbb{R}^k_{\ge 0}, \end{cases}$$

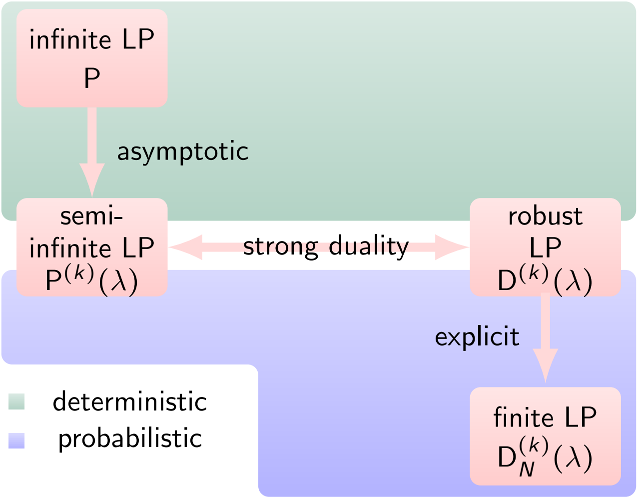
Corollary

Given $\varepsilon, \beta \in [0,1]$ and $\lambda \ge 0$. For all $N \ge N(\varepsilon, \beta)$

$$\mathbb{P}^N\Big[J^\star_{k,N}(\lambda) - J^\star_k(\lambda) \in \big[0, h(\varepsilon)\big]\Big] \ge 1 - \beta$$

- Corollary to [Mohajerin Esfahani et al., TAC 2015, to appear]
- Min-max problem $\Rightarrow$ Lipschitz constant $L = 1$
- $h(\varepsilon)$ can be proposed such that $h(\varepsilon) \to 0$ as $\varepsilon \to 0$ [Kanamori & Takeda, 2012]

Outline


infinite LP
P
asymptotic
semi-infinite LP
$P^{(k)}(\lambda)$
strong duality
robust LP
$D^{(k)}(\lambda)$
explicit
finite LP
$D_N^{(k)}(\lambda)$
deterministic
probabilistic

LQG-Example

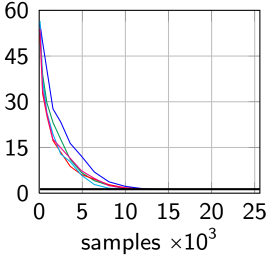


Figure : $k = 7$

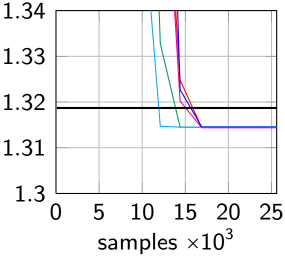


Figure : $k = 7$ (zoomed)

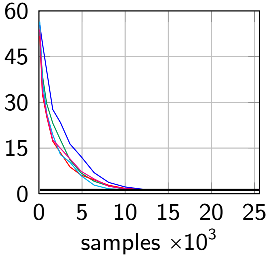


Figure : $k = 11$

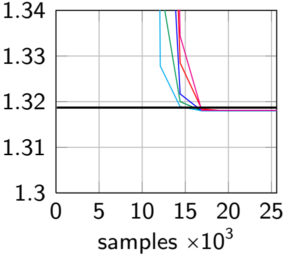


Figure : $k = 11$ (zoomed)

Summary/Outlook

- Linear programming formulation of MDPs on continuous spaces
  - constrained MDPs
  - using tools from mathematical programming

- Two stage approximation method
  - fist step with asymptotic error bounds
  - second step with explicit approximation error (scenario based approximation, cutting plane method)

arXiv: soon ☺

- Outlook
  - explicit error bound for first approximation step
  - optimal choice of basis functions
  - derivation of ε-approximating policies

Appendix 1 — Adaptive Approximation Scheme

- constraint of maximum violation

$$\delta(\rho,\gamma,\alpha,\beta) := \sup_{(x,a)\in X\times A} \left\{ \rho + \sum_{j=1}^{k} (\alpha_j - \beta_j) \mathrm{L}^{\star} u_j(x,a) - \gamma d(x,a) - c(x,a) \right\}$$

Algorithm 1: Cutting Plane Method

Step 1: Set $m > 0$, $(X \times A)_m := \{(x_1, a_1), \cdots, (x_m, a_m)\}$, $\varepsilon \in \mathbb{R}_{\geq 0}$ arbitrary small

Step 2: Solve $\mathrm{D}^{(k)}_{(X\times A)_m}(\lambda)$, get $(\rho^m, \gamma^m, \alpha^m, \beta^m)$

Step 3: Calculate $\delta(\rho^m, \gamma^m, \alpha^m, \beta^m)$ and denote its maximizer by (x^{m+1}, a^{m+1})

Step 4: If $\delta(\rho^m, \gamma^m, \alpha^m, \beta^m) < \varepsilon$, stop and output $\rho^m - \gamma^m \ell$ as the solution

Step 5: Set $(X \times A)_{m+1} := (X \times A)_m \cup \{(x^{m+1}, a^{m+1})\}$, update $m := m + 1$, then go to Step 2

LQG-Example for Adaptive Approximation

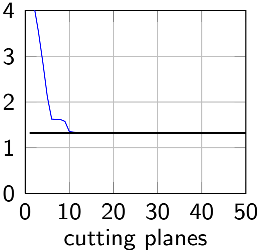


Figure : $k = 7$

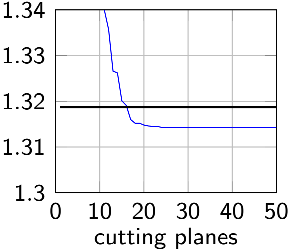


Figure : $k = 7$ (zoomed)

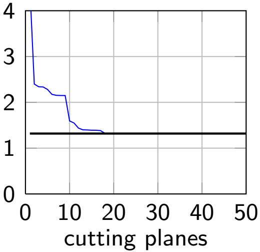


Figure : $k = 11$

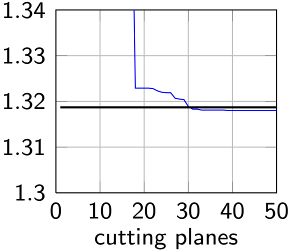


Figure : $k = 11$ (zoomed)