Infinite-Dimensional Sums-of-Squares for Optimal Control

E. Berthier, J. Carpentier, A. Rudi and F. Bach – Inria, ENS, PSL Research University

Optimal Control as a Linear Program

• We consider a **finite-horizon optimal control problem**, with compact state and action spaces $\mathcal{X} \subset \mathbb{R}^d$ and $\mathcal{U} \subset \mathbb{R}^p$:

$$V^*(t_0, x_0) = \inf_{u(\cdot)} \int_{t_0}^{T} L(t, x(t), u(t))\mathrm{d}t + M(x(T))$$
$$\forall t \in [t_0, T],\ \dot{x}(t) = f(t, x(t), u(t)), \quad x(0) = x_0. \qquad \text{(OCP)}$$

Let μ_0 be a probability measure on $\mathcal{X}$. We are interested in the value of the stochastic initial point problem $\mathbb{E}_{x_0 \sim \mu_0}[V^*(0, x_0)]$.

• Under some convexity assumptions, this problem is equivalent to finding a **maximal subsolution** of the Hamilton-Jacobi-Bellman equation:

$$\sup_{V \in C^1([0,T]\times\mathcal{X})} \int V(0, x_0)\mathrm{d}\mu_0(x_0)$$
$$\forall (t, x, u),\ \frac{\partial V}{\partial t}(t, x) + L(t, x, u) + \nabla V(t, x)^\top f(t, x, u) \geq 0$$
$$\forall x,\ V(T, x) \leq M(x). \qquad \text{(P)}$$

The **optimal value function** V^* is a solution of HJB, with equality in the 2nd constraint, and in the 1st one at the optimal controller $u^*(t, x)$.

• V is searched in a **linearly parameterized** set of functions:

$$\mathcal{F} = \{(t, x) \mapsto V_\theta(t, x) = \theta^\top \psi(t, x) + M(x) \mid \theta \in \mathbb{R}^m\},$$

with $\psi(T, .) = 0$. If $V^* \in \mathcal{F}$ then (P) reduces to a **linear program**. **However**, this LP has a **dense set of constraints** of the form:

$$\forall (t, x, u), \quad H(t, x, u) \geq 0.$$

Baseline: Subsampling Inequalities

A **simple relaxation** of problem (P) is to subsample a finite number of inequalities, and obtain an LP (penalized to avoid unbounded solutions):

$$\sup_{\theta \in \mathbb{R}^m} \frac{1}{n} \sum_{i=1}^{n} V_\theta(0, x^{(i)}) - \lambda_\theta \|\theta\|_2^2$$
$$\forall i \in I, \quad H_\theta(t^{(i)}, x^{(i)}, u^{(i)}) \geq 0. \qquad \text{(LP)}$$

An informative example

It is not straightforward to quantify the effect of this relaxation as a function of the number of subsampled inequalities. Yet a particular instance of the optimal control problem is the **global optimization** problem:

$$\sup c \ \text{ s.t. } \forall y \in \mathbb{R}^p,\ g(y) - c \geq 0.$$

Subsampling inequalities *v.s.* equalities

Subsampled inequalities lead to the approximation $\min g \simeq \min\{g(x_i)\}$. It requires $O(\varepsilon^{-p})$ samples to approximate $\min g$ with precision ε. If $g \in C^s(\mathbb{R}^p)$ is **smooth**, this rate can be improved to $O(\varepsilon^{-p/s})$ with a sum-of-squares representation and subsampled equalities.

Sum-of-Squares Representations

In the case of **polynomial** f, L and M, the state-of-the-art method is to strengthen the constraint $H \geq 0$ by $H(t, x, u) = \sum_j p_j(t, x, u)^2$, where the p_j are polynomials. This is a polynomial sum-of-squares (SoS), and leads to **Lasserre's hierarchy** of SDPs, converging to the value of (P).

Going beyond polynomials

We extend this to non-polynomial **smooth** problems, by representing non-negative functions as SoS in a reproducing kernel Hilbert space (RKHS) $\mathcal{H}$, *e.g.* a Sobolev space, with positive-definite kernel k, and possibly **infinite-dimensional** embedding $\Phi(y) = k(y, \cdot)$.

A SoS Hamiltonian

We consider a kernel k on $[0, T] \times \mathcal{X} \times \mathcal{U}$. We define the **strengthening**:

$$\sup_{\substack{V \in C^1([0,T]\times\mathcal{X}),\\ \mathcal{A} \in \mathbb{S}_+(\mathcal{H})}} \int V(0, x_0)\mathrm{d}\mu_0(x_0)$$
$$\forall (t, x, u),\ H(t, x, u) = \langle \Phi(t, x, u), \mathcal{A}\Phi(t, x, u)\rangle. \qquad \text{(KSOS)}$$

This problem gives a **lower-bound** on the value of (P). It coincides with (P) if there exists $\mathcal{A} \in \mathbb{S}_+(\mathcal{H})$ such that, at the optimal V^*:

$$\forall (t, x, u),\ H^*(t, x, u) = \langle \Phi(t, x, u), \mathcal{A}\Phi(t, x, u)\rangle.$$

Exact SoS representations for certain smooth problems

Theorem 1 (informal) *Let $s \in \mathbb{N}$, $s \geq 1$. Assume that:*

• *f is **control-affine**: $\forall (t, x, u) \in [0, T] \times \mathcal{X} \times \mathcal{U}$,*

$$f(t, x, u) = g(t, x) + B(t, x)u;$$

• *L is strongly convex: $\nabla_u^2 L(t, x, u) \succcurlyeq \rho I$ for some $\rho > 0$;*

• *L and B and V^* are **sufficiently smooth**;*

Then H^* is a SoS of p smooth functions $(w_j)_{1 \leq j \leq p} \in C^s(\Omega)$:

$$\forall (t, x, u) \in \Omega, \quad H^*(t, x, u) = \sum_{j=1}^{p} w_j(t, x, u)^2.$$

Subsampling Equalities

We can then subsample the SoS equality constraints. Using generic properties of an RKHS, we obtain a **semi-definite program** of the form:

$$\sup_{B \succcurlyeq 0, \theta \in \mathbb{R}^m} c^\top \theta - \lambda_\theta \|\theta\|_2^2 - \lambda \mathrm{Tr}(B) + C$$
$$\text{such that } \forall i \in \{1, \ldots, n\},\ b_i + a_i^\top \theta = (\Phi_i)^\top B \Phi_i. \qquad \text{(SDP)}$$

• The $\Phi_i \in \mathbb{R}^n$ are computed from the Cholesky decomposition of the **kernel matrix**, with entries $[K]_{ij} = k\left((t^{(i)}, x^{(i)}, u^{(i)}), (t^{(j)}, x^{(j)}, u^{(j)})\right)$.

• The **regularization** parameter $\lambda > 0$ allows for subsampling to recover the non-subsampled program when $n \to \infty$, and $\lambda \to 0$ at the proper rate. In the limit $\lambda \to 0$, we **recover the LP** formulation.

Practical Algorithm

• We add a **slack variable** $\delta \in \mathbb{R}^n$ to improve stability and account for imperfect modeling of V^* in $\mathcal{F}$, γ is typically chosen large.

• We also add a **log barrier** function on B, with small ε, that allows to solve the dual of the following problem:

$$\sup_{B \succcurlyeq 0, \theta, \delta} c^\top \theta - \lambda_\theta \|\theta\|_2^2 - \lambda \mathrm{Tr}(B) - \gamma \|\delta\|^2 + \varepsilon \log\det B + C$$
$$\text{such that } \quad \forall i \in \{1, \ldots, n\},\ b_i + a_i^\top \theta = (\Phi_i)^\top B \Phi_i + \delta_i.$$

• The dual is solved with **damped Newton**'s method, a parameter-free algorithm with cost $O(n^3)$ in time and $O(n^2)$ in space for each iteration.

Numerical Example

We solve a **linear quadratic regulator** in dimensions $d = 2$, $p = 1$, for which the optimal value function and controller are known in closed-form. In particular, we know that:

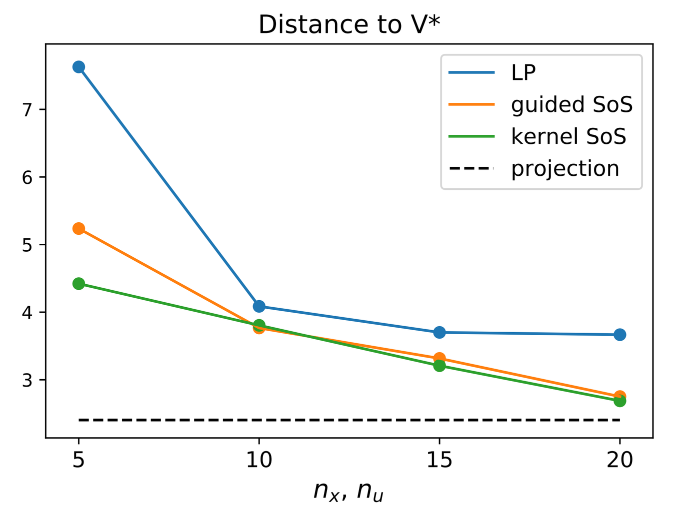
$$H^*(t, x, u) = (u + K(t)x)^\top R(u + K(t)x).$$

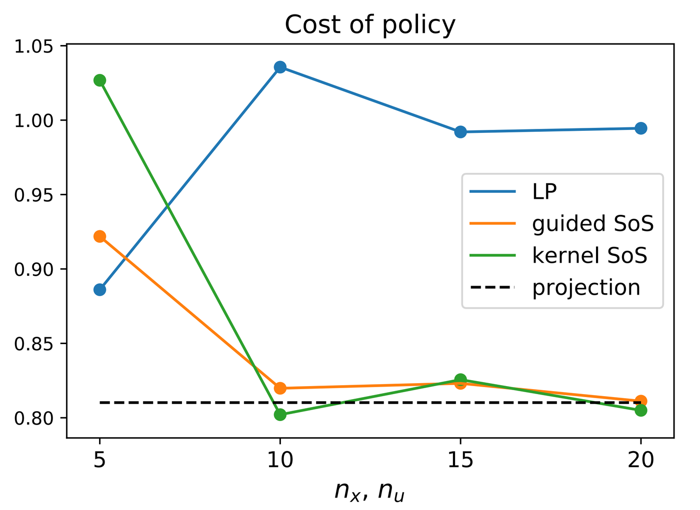
We compare **three methods**, with $n = n_t n_x n_u$ sampled points:

- the LP with subsampled inequalities;
- the SDP with fixed embedding Ψ instead of Φ ("guided SoS");
- the SDP formulation ("kernel SoS") with kernel:

$$k((t, x, u), (t', x', u')) = \langle u, u'\rangle/100 + \langle x, x'\rangle \times \exp(-|t - t'|).$$

The results are given in terms of **accuracy** of the approximation of the optimal value function and **cost** of the greedy policy obtained from V.





References

[1] J.-B. Lasserre, D. Henrion, C. Prieur, and E. Trélat. Nonlinear optimal control via occupation measures and LMI-relaxations. *SIAM J. on Control and Optim.*, 47(4):1643–1666, 2008.

[2] U. Marteau-Ferey, F. Bach, and A. Rudi. Non-parametric models for non-negative functions. In *Advances in Neural Information Processing Systems*, 2020.

[3] A. Rudi, U. Marteau-Ferey, and F. Bach. Finding global minima via kernel approximations. Technical Report 2012.11978, arXiv, 2020.

[4] R. Vinter. Convex duality and nonlinear optimal control. *SIAM J. on Control and Optim.*, 31(2):518–538, 1993.

Download the paper