Optimistic Planning Algorithms For Constrained Optimal Control Problems

Olivier Bokanowksi[1], Nidhal Gammoudi[2] , Hasnaa Zidani[3]

[1] Université Paris-Diderot (Paris 7), France
olivier.bokanowski@math.univ-paris-diderot.fr
[2] Institut Polytechnique de France
nidhal.gammoudi@ensta-paris.fr
[3] INSA Rouen, France
hasnaa.zidani@insa-rouen.fr

Mots-clés : *Reinforcement Learning, Optimistic Planning, Optimal Control, State Constraints.*

1 Introduction

In this work, we propose to solve state-constrained finite-horizon nonlinear optimal control problems by means of optimistic planning methods. A main strength of this approach is the relation between the convergence rates to the optimal solution and the computational resources allowed, which is established using some ideas of bandit theory and reinforcement learning [4].

2 Problem formulation and motivations

For a given finite time horizon $T > 0$ and a nonlinear dynamics f, consider the following state-constrained optimal control problem : :

$$\begin{aligned} v(t,x) := \inf_{a(\cdot)\in\mathcal{A}} & \int_t^T \ell(y^a_{t,x}(s), a(s))ds + \Phi(y^a_{t,x}(T)) \\ \text{s.t.} \quad & y^a_{t,x}(t) = x \\ & \dot{y}^a_{t,x}(s) = f(y^a_{t,x}(s), a(s)), \ \text{ a.e. } \ s \in [t,T], \\ & y^a_{t,x}(s) \in \mathcal{K}, \quad \forall s \in [t,T] \quad \text{(state constraints)} \end{aligned} \tag{1}$$

where $t \in [0,T]$, $x \in \mathbb{R}^d$, ℓ and Φ are respectively the distributed and the final cost functions, the input variable $a(\cdot) \in \mathcal{A}$, the set of measurable functions from $[0,T]$ taking values in A, a compact set of $\mathbb{R}^m$, $m \geq 1$, and $\mathcal{K}$ is a closed subset of $\mathbb{R}^d$.

In general, when $\mathcal{K} \neq \mathbb{R}^d$, the value function v may be discontinuous and cannot be characterized through Hamilton-Jacobi equations unless some controllability assumptions are satisfied. For this reason, we follow the *level set approach*, introduced in [2], which consists in characterizing v by means of the following auxiliary control problem free of state constraints :

$$w(t,x,z) \quad = \quad \inf_{a(\cdot)\mathcal{A}} \left\{ \Big(\int_t^T \ell(y^a_{t,x}(s), a(s))ds \ + \ \Phi(y^a_{t,x}(T)) \ - \ z \Big) \bigvee \Big(\max_{s\in[t,T]} g(y^a_{t,x}(s)) \Big) \right\}, \tag{2}$$

where $(t,x,z) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}$ and g is a regular function verifying : $x \in \mathcal{K} \iff g(x) \leq 0$.

It is known that, under some convexity assumption on f, v can be determined by the following relation (see [2]) :

$$v(t,x) = \inf\{z \in \mathbb{R} \ \mid \ w(t,x,z) \leq 0\}. \tag{3}$$

3 Contributions

Our contribution in this work [1] consists in developing optimistic planning approaches to deal with finite horizon problems in presence of state constraints. In particular, those algorithms are exploited to approximate the auxiliary problem (2) and hence to get an approximation of the optimal solution for the original state-constrained problem (1) by means of the relation (3). While classical methods for calculating the value function are generally based on a discretization in the state space, optimistic planning algorithms have the advantage of using adaptive discretization in the control space. These approaches are therefore very suitable for control problems where the dimension of the control variable m is low and allow to deal with problems where the dimension of the state space d can be very high. Our algorithms also have the advantage of providing, for given computing resources, the best control strategy whose performance is as close as possible to optimality while its corresponding trajectory comply with the state constraints. Furthermore, we characterize the convergence rates to the optimal solution as a function of the computational resources used. Moreover, we exploit some ideas from *Model Predictive Control* [5] to propose an amelioration of the implementation of those algorithms. Finally, we illustrate the relevance of our algorithms on several nonlinear optimal control problems even in high dimensions of the state (until $d = 10^3$) which is impossible to handle with classical numerical methods.

Références

[1] O. Bokanowski, N. Gammoudi and H. Zidani, *Optimistic planning algorithms for state-constrained optimal control problems*, Preprint submitted to Computers and Mathematics with Applications, 2020.

[2] A. Altarovici, O. Bokanowski and H. Zidani, *A general Hamilton-Jacobi framework for non-linear state-constrained control problems*, ESAIM : Control, Optimisation and Calculus of Variations, 2013, pp. 337–357.

[3] L. Buşoniu, E. Páll and R. Munos, *Continuous-action planning for discounted infinite-horizon nonlinear optimal control with Lipschitz values*, Automatica, 2018, pp. 100–108.

[4] R. Munos, *From bandits to Monte-Carlo Tree Search : The optimistic principle applied to optimization and planning*, 2014.

[5] L. Grüne and P. Jürgen, *Nonlinear model predictive control theory and algorithms*, 2011.