Automatica 97 (2018) 367–375

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Symmetry reduction for dynamic programming☆

John Maidens[a,*], Axel Barrau[b], Silvère Bonnabel[c], Murat Arcak[d]

[a] Department of Mechanical & Industrial Engineering, Ryerson University, 350 Victoria Street, Toronto, ON, M5B 2K3, Canada
[b] SAFRAN TECH, Groupe Safran, Rue des Jeunes Bois - Châteaufort, 78772 Magny Les Hameaux, France
[c] MINES ParisTech, PSL Research University, Centre for Robotics, 60 bd Saint-Michel, 75006 Paris, France
[d] Department of Electrical Engineering & Computer Sciences, University of California, Berkeley, 569 Cory Hall, Berkeley, CA, 94720, USA

ARTICLE INFO

Article history:
Received 7 April 2017
Received in revised form 9 January 2018
Accepted 25 July 2018

Keywords:
Optimal control
Dynamic programming
Invariant systems
Model reduction
Path planning
Medical applications

ABSTRACT

We present a method of exploiting symmetries of discrete-time optimal control problems to reduce the dimensionality of dynamic programming iterations. The results are derived for systems with continuous state variables, and can be applied to systems with continuous or discrete symmetry groups. We prove that symmetries of the state update equation and stage costs induce corresponding symmetries of the optimal cost function and the optimal policies. We then provide a general framework for computing the optimal cost function based on gridding a space of lower dimension than the original state space. This method does not require algebraic manipulation of the state update equations; it only requires knowledge of the symmetries that the state update equations possess. Since the method can be performed without any knowledge of the state update map beyond being able to evaluate it and verify its symmetries, this enables the method to be applied in a wide range of application problems. We illustrate these results on two six-dimensional optimal control problems that are computationally difficult to solve by dynamic programming without symmetry reduction.

1. Introduction

The dynamic programming algorithm for computing optimal control policies has, since its development, been known to suffer from the "curse of dimensionality" (Bellman, 1957). Its applicability in practice is typically limited to systems with four or five continuous state variables because the number of points required to grid a space of n continuous state variables increases exponentially with the state dimension n. This complexity has led to a collection of algorithms for approximate dynamic programming, which scale to systems with larger state dimension but lack the guarantees of global optimality of the solution associated with the original dynamic programming algorithm (Bellman & Dreyfus, 1959; Bertsekas, 2012; Powell, 2007, 2016).

In practice, many real-world systems exhibit symmetries that can be exploited to reduce the complexity of system models. Symmetry reduction has found applications in fields ranging from differential equations (Bluman & Kumei, 2013; Clarksonz & Mansfield, 1994) to model checking (Emerson & Sistla, 1996; Kwiatkowska, Norman, & Parker, 2006). In control engineering, symmetries have been exploited to improve control of mechanical systems (Bloch, Krishnaprasad, Marsden, & Murray, 1996; Bullo & Murray, 1999; Marsden, Montgomery, & Ratiu, 1990), develop more reliable state estimators (Barrau & Bonnabel, 2014), study the controllability of multiagent systems (Rahmani, Ji, Mesbahi, & Egerstedt, 2009) and to reduce the complexity of stability and performance certification for interconnected systems (Arcak, Meissen, & Packard, 2016; Rufino Ferreira, Meissen, Arcak, & Packard, 2017). Symmetry reduction has also been applied to the computation of optimal control policies for continuous-time systems in Grizzle and Marcus (1984) and Ohsawa (2013) and Markov decision processes (MDPs) in Zinkevich and Balch (2001) and Narayanamurthy and Ravindran (2007).

In this paper, we present a theory of symmetry reduction for the optimal control of discrete-time, stochastic nonlinear systems with continuous state variables. This reduction allows dynamic programming to be performed in a lower-dimensional state space. Since the computational complexity of a dynamic programming iteration increases exponentially with state dimension, this reduction significantly decreases computational burden. Further, our proposed method does not rely on an explicit transformation of the state update equations, making the method applicable in situations where such a transformation is difficult or impossible to find analytically.

☆ Research supported in part by the National Science Foundation, USA under grant ECCS-1405413. The material in this paper was partially presented at the 2017 American Control Conference, May. This paper was recommended for publication in revised form by Associate Editor Valery Ugrinovskii under the direction of Editor Ian R. Petersen.

* Corresponding author.
E-mail addresses: maidens@ryerson.ca (J. Maidens), axel.barrau@mines-paristech.fr (A. Barrau), silvere.bonnabel@mines-paristech.fr (S. Bonnabel), arcak@eecs.berkeley.edu (M. Arcak).

https://doi.org/10.1016/j.automatica.2018.08.024
0005-1098/

We present two theorems that summarize our method of symmetry reduction. Theorem 4 describes how symmetries of the system dynamics imply symmetries of the optimal cost and optimal policy functions. Theorem 6 then describes a method of computing the cost function based on reduced coordinate system that depends on fewer state variables.

This paper builds on the work we presented in the conference paper (Maidens, Barrau, Bonnabel, & Arcak, 2017). The most substantial improvement is the additional theoretical results presented in Sections 2.3 and 3.2. The conference version presented an *ad hoc* symmetry reduction for a magnetic resonance imaging (MRI) application, but did not provide a general methodology for computing the coordinate reduction. This paper addresses this shortcoming by presenting a general method based on the moving frame formalism, which leads to the general symmetry reduction formula presented in Theorem 6. Additionally, the MRI example has been reworked to match this new formalism, and the numerical implementation and graphs of the numerical solution have been improved. We have also included two new extensions of this formalism to the case of equivariant costs in Section 3.3 and to the synchronization problem of stochastic dynamic systems on matrix groups in Section 3.4, along with examples to illustrate the algorithm in these contexts.

This paper is organized as follows: in Section 2 we introduce notation and provide background information both on dynamic programming for optimal control, and on the mathematical theory of symmetries. In Section 3, we derive our main theoretical results, that is, we prove that control system symmetries induce symmetries of the optimal cost function and optimal control policy, and then leverage the result to present a general method of performing dynamic programming in reduced coordinates. In Section 4 we apply the algorithm to a cooperative control problem for two Dubins vehicles using a Lie group formulation. In Section 5 we apply symmetry reduction to compute the solution of an optimal control problem arising in dynamic MRI acquisition. Code to reproduce the computational results in this paper is available at https://github.com/maidens/Automatica-2017.

2. Dynamic programming and symmetries

In this section, we first recall the main features of dynamic programming for optimal control of stochastic discrete time systems. Then we introduce our problem and provide the reader with a primer on the classical theory of symmetries. We also introduce the notion of invariant control systems with invariant costs.

2.1. Dynamic programming for optimal control of stochastic systems

We begin by introducing dynamic programming for finite horizon optimal control following the notation of Bertsekas (2005). We consider a discrete-time dynamical system

$$x_{k+1} = f_k(x_k, u_k, w_k), \quad k = 0, 1, \dots, N-1 \tag{1}$$

where $x_k \in \mathcal{X} \subseteq \mathbb{R}^n$ is the system state, $u_k \in \mathcal{U} \subseteq \mathbb{R}^m$ is the control variable to be chosen at time k, $w_k \in \mathcal{W} \subseteq \mathbb{R}^\ell$ are independent continuous random variables each with density p_k, and $N \in \mathbb{Z}_+$ is a finite control horizon. Associated with this system is an additive cost function

$$g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k)$$

that we wish to minimize through our choice of u_k. We define a *control system* to be a tuple $\mathcal{S} = (\mathcal{X}, \mathcal{U}, \mathcal{W}, p, f, g, N)$ where $p = \prod_{k=0}^{N-1} p_k$ is the joint density of the random variables w_k.

We consider a class of control policies $\pi = \{\mu_0, \dots, \mu_{N-1}\}$ where $\mu_k : \mathcal{X} \to \mathcal{U}$ maps observed states to admissible control inputs. Given an initial state x_0 and a control policy π, we define the expected cost under this policy as

$$J_\pi(x_0) = \mathbb{E}\left[g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, \mu_k(x_k), w_k) \right].$$

An optimal policy π^* is defined as one that minimizes the expected cost:

$$J_{\pi^*}(x_0) = \min_{\pi \in \Pi} J_\pi(x_0)$$

where Π denotes the set of all admissible control policies. The optimal cost function, denoted $J^*(x_0)$, is defined to be the expected cost corresponding to an optimal policy.

As in Bertsekas (2005), we use min to denote the infimum value regardless of whether there is a policy $\pi \in \Pi$ that achieves a minimum. In the example problems presented in Sections 4 and 5, the existence of a minimum is guaranteed by compactness or finiteness arguments, respectively. In the general case, the optimal cost function J^* can be computed using the dynamic programming algorithm regardless of the existence of minimizers, but the existence of an optimal policy π^* requires that a minimum be achieved for each $x_k \in \mathcal{X}$.

We quote the following result due to Bellman from (Bertsekas, 2005):

Proposition 1 (*Dynamic Programming*)**.** *For every initial state x_0, the optimal cost $J^*(x_0)$ is equal to $J_0(x_0)$, given by the last step of the following algorithm, which proceeds backward in time from period $N-1$ to period 0:*

$$J_N(x_N) = g_N(x_N)$$

$$J_k(x_k) = \min_{u_k \in \mathcal{U}} \mathbb{E}\left[g_k(x_k, u_k, w_k) + J_{k+1}\Big(f_k(x_k, u_k, w_k)\Big) \right] \tag{2}$$
$$k = 0, 1, \dots, N-1,$$

where the expectation is taken with respect to the probability distribution of w_k defined by the density p. Furthermore, if there exists u_k^* minimizing the right hand side of (2) for each x_k and k, then the policy $\pi^* = \{\mu_0^*, \dots, \mu_{N-1}^*\}$ where $\mu_k^*(x_k) = u_k^*$ is optimal.

The intermediate functions $J_k(x_k)$ for $k > 0$ computed in this manner represent the optimal cost of the tail subproblem beginning at x_k. The optimal cost of the entire problem is given by the function $J^*(x_0) = J_0(x_0)$ obtained when this recursion terminates.

2.2. Invariant system with invariant costs

We first recall the definition of a transformation group for a control system, as in Martin, Rouchon, and Rudolph (2004), Jakubczyk (1998) and Respondek and Tall (2002). See Olver (1999) for the more general theory.

Definition 2 (*Transformation Group*)**.** A transformation group on $\mathcal{X} \times \mathcal{U} \times \mathcal{W}$ is set of tuples $h_\alpha = (\phi_\alpha, \chi_\alpha, \psi_\alpha)$ parametrized by elements α of a Lie group $\mathcal{G}$ having dimension r, such that the functions $\phi_\alpha : \mathcal{X} \to \mathcal{X}$, $\chi_\alpha : \mathcal{U} \to \mathcal{U}$ and $\psi_\alpha : \mathcal{W} \to \mathcal{W}$ are all C^1 diffeomorphisms and satisfy:

- $\phi_e(x) = x$, $\chi_e(u) = u$, $\psi_e(w) = w$ when e is the identity of the group $\mathcal{G}$ and
- $\phi_{a*b}(x) = \phi_a \circ \phi_b(x)$, $\chi_{a*b}(u) = \chi_a \circ \chi_b(u)$, $\psi_{a*b}(x) = \psi_a \circ \psi_b(x)$ for all $a, b \in \mathcal{G}$ where $*$ denotes the group operation and $\circ$ denotes function composition.