Covering Number: Analyses for Approximate Continuous-state POMDP Planning*

(Extended Abstract)

Zongzhang Zhang
Soochow University
zzzhang@suda.edu.cn

Quan Liu
Soochow University
quanliu@suda.edu.cn

ABSTRACT

To date, many theoretical results on discrete POMDPs have not yet been extended to continuous-state POMDPs, due to the infinite dimensionality of the belief space in a continuous-state case. In this paper, we define a distance in the ℓ_n-metric space with respect to a partitioning representation of the continuous-state space, and formalize the size of the search space reachable under inadmissible heuristics via the covering number concept. Together with existing proof techniques in discrete POMDPs, we use the covering number to analyze the computational complexity of approximate planning for some types of continuous-state POMDPs.

Keywords

POMDP; Continuous-state Space; Complexity Theory

1. INTRODUCTION

Recent developments in approximate solutions, e.g., point-based methods and online heuristic algorithms, have drastically improved the speed of discrete POMDP planning. On the theortical front, the covering number of a search space has been used to quantify the complexity of approximate discrete POMDP planning [2, 4]. However, most real world POMDP problems are naturally modeled by continuous-state spaces. In this case, the dimensionality of the belief space is infinite. Thus, simple generalizations of known approaches and complexity results for discrete-state models to continuous-state domains are not appropriate.

To find approximate solutions, existing approaches first approximately represent beliefs in the infinite-dimensional space by a low-dimensional discrete representation. Some of them share the idea to represent beliefs in the infinite-dimensional continuous belief space by a parametric form and solve the approximate POMDP on the parametric s-pace. The algorithm recently proposed in Brechtel et al. [1] is similar to ours, in that they also automatically learn a low-dimensional, discrete representation of the continuous-state space during the process of solving. The insight exploited in [1] is that states which are close in the continuous-state space will usually lead to similar outcomes and thus have similar α-vectors in most problems. Motivated by the work of [1], we propose a disjunct space representation, namely, space partitioning, with piecewise constant beliefs and α-functions. Then, we define a distance in the ℓ_n-metric space with respect to the space partitioning and use it to formalize the size of the search space reachable under inadmissible heuristics. Thus, we can use the covering number proof techniques in discrete POMDPs to analyze the computational complexity of approximate planning for some types of continuous-state POMDP from the covering number viewpoint.

*The work was supported in part by National Natural Science Foundation of China (61502323, 61272005, 61472262).

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

2. DEFINITIONS

Formally, a POMDP is a tuple (S, A, Z, T, R, Ω), where S, A, and Z, respectively, denote a state space, action space and observation space; T, R, and Ω are transition, reward, and observation functions, respectively. In continuous-state POMDPs, S is continuous, but A and Z are discrete.

The knowledge about the system's state is represented by a belief b in the belief space $\mathcal{B}$ with $b : S \to \mathbb{R}_{\geq 0}$ and $\int_{s\in S} b(s)ds = 1$. The initial belief b_0 is assumed to be known. After taking action a in belief b, the agent reaches the next belief $b^{a,z}$, computed by the Bayesian updating formula.

The goal of a POMDP solver is to find a policy $\pi : \mathcal{B} \to A$ maximizing the value $V : \mathcal{B} \to \mathbb{R}$ for b_0. The value associated with a policy π is given by $\mathbf{E}\Big[\sum_{t=0}^{\infty} \gamma^t R(b_t, \pi(b_t))\Big]$. The optimal value function $V^*(b)$ satisfies the Bellman equation: $V^*(b) = \max_{a\in A} Q^*(b, a)$, where $Q^*(b, a) = R(b, a) + \sum_{z\in Z} Pr(z|b, a)V^*(b^{a,z})$. So, $\pi^*(b) = \arg\max_{a\in A} Q^*(b, a)$. Additionally, $V^*(b)$ can be approximated arbitrary well by a peicewise-linear and convex function: $V^*(b) \approx \max_{\alpha\in\Gamma^*}(\alpha \cdot b) = \max_{\alpha\in\Gamma^*} \int_{s\in S} b(s)\alpha(s)ds$ with $S \subseteq \mathbb{R}^n$ and $\alpha : S \to \mathbb{R}^n$, where Γ^* is the finite set of linear α-functions.

We use V^L and V^U to represent the lower and upper bounds of V^*, respectively. Both V^L and V^U are assumed to be uniformly improvable. We also assume that we are provided with heuristics, f and y, that can provide initial values V^L_f to the lower bound of V^* and V^U_y to the upper bound of V^*. The function $y(b)$ may not be the true upper bound. That is, $y(b)$ is allowable to be less than $V^*(b)$ for some b in the belief space $\mathcal{B}$. Similarly, Q^L, Q^U, Q^L_f and Q^U_y are used for the corresponding bounds of Q^*.

A search space B is a subspace of $\mathcal{B}$ and can be represented as a belief tree $\mathcal{T}_B$ rooted at the initial belief b_0. The reachable belief space under inadmissible heuristics, $\mathcal{R}^{y,U}_{f,L}(b_0)$, is defined as the set of beliefs that are reachable from b_0 by going through action branches a that satisfy

$Q_y^U(b,a) \geq V^*(b)$ and observation branches z that satisfy $V_y^U(b) - V_f^L(b) - \frac{\epsilon}{\gamma^{d_b}} > 0$.

To split the continuous-state space into finite and abstraction states, we adopt the following definition of a state-space partitioning. Such a kind of partitioning works well in bounding the error between the value function of the induced discrete POMDP and the true value function of the original continuous-state POMDP.

Definition 1. *A μ-partitioning of a continuous-state space S, $\mathcal{K}_\mu = \{K_\mu^1, K_\mu^2, \cdots K_\mu^{|\mathcal{K}_\mu|}\}$, satisfies that*

1. $S \subseteq \bigcup_{i=1}^{|\mathcal{K}_\mu|} K_\mu^i$;
2. $K_\mu^i \bigcap K_\mu^j = \emptyset$, *where* $i,j \in \{1,2,\cdots,|\mathcal{K}_\mu|\}$ *and* $i \neq j$;
3. *for any* $s_i, s_j \in K_\mu^m$ *and any* $\alpha \in \Gamma^*$, $|\alpha(s_i) - \alpha(s_j)| \leq \mu$, *where* $i,j \in \{1,2\}$ *and* $m = \{1,2,\cdots,|\mathcal{K}_\mu|\}$.

The partitioning $\mathcal{K}_\mu$ induces an ℓ_n-distance metric on the belief space $\mathcal{B}$, as below:

Definition 2. *Suppose that $\mathcal{K}_\mu$ be a μ-partitioning of the continuous-state space S. The distance between any two beliefs b and b' in the infinite-dimensional ℓ_n-metric space $\mathcal{B}$ with respect to $\mathcal{K}_\mu$ is*

$$d_{n,\mathcal{K}_\mu}(b,b') = \Big(\sum_{K\in\mathcal{K}_\mu} \Big| \int_{s\in K} b(s)ds - \int_{s\in K} b'(s)ds \Big|^n\Big)^{\frac{1}{n}}.$$

The new metric measures the difference in probability mass for subsets of states rather than individual states. The mathematical definition of the *covering number* of a set of points in the ℓ_n-metric space with respect to the space partitioning $\mathcal{K}_\mu$ is recapitulated as follows:

Definition 3. *Given an infinite-dimensional ℓ_n-metric space $\mathcal{X}$, a δ-cover of a set $B \subseteq \mathcal{X}$ is a set of points $C \subseteq \mathcal{X}$ such that for every point $b \in B$, there is a point $c \in C$ with $d_{n,\mathcal{K}_\mu}(b,c) \leq \delta$. The δ-covering number of B, denoted by $\mathcal{C}_B(\delta)$, is the size of the smallest δ-cover of B.*

Intuitively, the covering number of a space B is equal to the minimum number of balls of radius δ needed to cover B. We denote $\mathcal{C}_{f,L}^{y,U}(\delta)$ as the δ-covering number of $\mathcal{R}_{f,L}^{y,U}(b_0)$.

3. COMPLEXITY RESULTS

The following Lipschitz condition is satisfied by the optimal value function on the ℓ_n-metric space with respect to the space partitioning $\mathcal{K}_\mu$.

Lemma 1. *Let $\mathcal{K}_\mu$ be a μ-partitioning of the continuous-state space S, and let $\mathcal{B}$ be the corresponding belief space over S in the ℓ_n-metric space. For any b, $b' \in \mathcal{B}$, if $d_{n,\mathcal{K}_\mu}(b,b') \leq \delta$, then $|V^*(b) - V^*(b')| \leq \frac{R_{\max}}{1-\gamma}\rho_n\delta + 2\mu$. Here*

$$\rho_n = \begin{cases} 1 & n = 1 \\ \sqrt{|\mathcal{K}_\mu|} & 1 < n \leq 2 \\ |\mathcal{K}_\mu| & 2 < n \leq \infty. \end{cases} \quad (1)$$

The effect of the μ-partitioning over the state space in Lemma 1 is to obtain a finite-dimensional belief representation and constrain the error caused by dimensional reduction, that is, the difference of optimal values between two beliefs in the same μ-region splitted by $\mathcal{K}_\mu$, is no more than 2μ. Lemma 1 is fundamental here because it provides a way of connecting the covering number to POMDP planning complexity: for any two beliefs in an infinite-dimensional space of continuous probability distributions, if their distance in the ℓ_n-metric space with respect to $\mathcal{K}_\mu$ is small, then their optimal values are also similar.

Using Lemma 1, we can establish a theoretical connection between the covering number and the planning complexity of continuous-state POMDPs, as below:

Theorem 1. *Assume that $\mathcal{K}_\mu$ be a μ-partitioning of the continuous-state space S, where $\mu \leq \frac{(1-\gamma)\epsilon}{6\gamma}$. Suppose that the computational complexity of $\int_{s\in K} b(s)ds$ for any $K \in \mathcal{K}_\mu$ is $O(1)$. For any $b_0 \in \mathcal{B}$, let $\mathcal{C}_{f,L}^{y,U}(\delta)$ be the δ-covering number of $\mathcal{R}_{f,L}^{y,U}(b_0)$ in the ℓ_n-metric space with respect to $\mathcal{K}_\mu$. Define $\zeta = \max\{0, \max_{b\in\mathcal{R}(b_0)}[V^*(b_0) - V_y^U(b)]\}$. Given constant $\epsilon > 0$, an approximation $V(b_0)$ of $V^*(b_0)$, with error $|V^*(b_0) - V(b_0)| \leq 2\epsilon + \gamma\zeta$, can be found in time $O\Big(h \cdot \mathcal{C}_{f,L}^{y,U}(\delta/2)^2\Big)$, where $h = \log_\gamma \frac{(1-\gamma)\epsilon}{3R_{\max}}$, $\delta = \frac{(1-\gamma)^2\epsilon}{3\gamma\rho_n R_{\max}}$, V_f^L is used as an initial lower bound, and the inadmissible heuristic function V_y^U is used for the initial upper bound. Here, ρ_n is defined as Equation 1.*

The theorem provides insights in the three aspects: (1) under some conditions an approximately optimal solution over the ℓ_n-metric space with respect to $\mathcal{K}_\mu$ can be computed in time at most quadratic polynomial in the covering number of $\mathcal{R}_{f,L}^{y,U}(b_0)$; (2) prior knowledge can be used as heuristics to reduce the search space size to make POMDP problems easier to solve; and (3) a near optimal policy is still computable when using an inadmissible heuristic as the upper bound of V^*. These insights might be helpful in designing more efficient methods for continuous-state POMDPs.

Finally, we discuss the two conditions in Theorem 1. The first one is that the computational complexity of finding a μ-partitioning of the continuous-state space is $O\big(h \cdot \mathcal{C}_{f,L}^{y,U}(\delta/2)^2\big)$. This is satisfiable under the assumption: there exists a positive constant θ such that, for any two states s and s' satisfying $|s-s'| \leq \theta$, we have $|\alpha(s) - \alpha(s')| \leq \mu$. We think such an assumption can be met in most continuous-state problems. When $|s-s'| < +\infty$, $|\mathcal{K}_\mu|$ equals to $\lceil 1/\epsilon * \max_{s,s'} |s-s'| \rceil$ in this case. The second one that requiring $\int_{s\in K} b(s)ds$ to be computed in $O(1)$ is satisfiable when the beliefs can be represented by Gaussian mixtures [3]. In this case, the integral $\int_{s\in K} b(s)ds$, for a continuous region K, has a closed form.

REFERENCES

[1] S. Brechtel, T. Gindele, and R. Dillmann. Solving continuous POMDPs: Value iteration with incremental learning of an efficient space representation. In *Proc. ICML-13*, pages 370–378, 2013.

[2] D. Hsu, W. Lee, and N. Rong. What makes some POMDP problems easy to approximate. In *Proc. NIPS-07*, pages 689–696, 2007.

[3] J. Porta, N. Vlassis, M. Spaan, and P. Poupart. Point-based value iteration for continuous POMDPs. *Journal of Machine Learning Research*, 7:2329–2367, 2006.

[4] Z. Zhang, D. Hsu, and W. Lee. Covering number for efficient heuristic-based POMDP planning. In *Proc. ICML-14*, pages 28–36, 2014.