Lecture 11 - Exact methods for POMDPS

OBJECTIVE: In this lecture we show that the value function for finite horizon POMDPS is piece-wise linear convex (PWLC). This important property can be exploited in the design of POMDP algorithms for finite state spaces.

◇ FINITE HORIZON PIECE-WISE LINEAR CONVEXITY

In the early days of POMDPS, Smallwood and Sondik showed that the value function for a finite horizon POMPD can be expressed as a piece-wise linear combination of beliefs.

$$V_t(b) = \max_{\alpha \in \Gamma_t} \sum_{s \in \mathcal{S}} \alpha(s) b(s)$$

where the $\alpha \in \Gamma_t = \{\alpha_0, \alpha_1, \ldots, \alpha_m\}$ define $|\mathcal{S}|$-dimensional hyper-planes. This is illustrated below for a two state problem, where we only need to consider one belief $b = b(s_1)$.

⋆

◇ INDUCTION PROOF OF PWLC

We will prove the PWLC of the value function by induction. First let us consider the base case. When there are no steps to go, the value function simplifies to:

$$V_0(b) = \max_a \sum_s r(s,a)b(s) = \max_{\alpha^a \in \Gamma_0} \sum_{s \in \mathcal{S}} \alpha^a(s)b(s)$$

⋆ Proof:

For the induction step, assume that the result is true:

$$V_{t-1}(b) = \max_{\alpha \in \Gamma_{t-1}} \sum_s \alpha(s)b(s)$$

and recall our Bellman recursion:

$$V_t(b) = \max_a \left\{ \sum_s b(s)r(s,a) + \gamma \sum_y p(y|a,b)V_{t-1}(b_y^a(\mathbf{s}')) \right\}$$

By substituting the PWLC expression for V_{t-1} into Bellman's recursion, we shall prove that $V_t(b)$ is also PWLC.

⋆ Proof:

⋆ Proof:

◇ EXACT ALGORITHMS

In the previous proof we defined the quantities $\alpha^a(s) \in \Gamma_t^{ar}$ and $\alpha^{ay}(s) \in \Gamma_t^{ay}$ as follows

$$\begin{aligned}\alpha^a(s) &= r(a,s) \\ \alpha^{ay}(s) &= \gamma \sum_{s'} p(s'|a,s)p(y|s',a)\alpha(s')\end{aligned}$$

The sets $\{\Gamma_t^{ar}, \Gamma_t^{ay}\}$ of α-vectors is the first thing we compute. Next, we know from the proof that value function using these vectors is:

$$V_t(b) = \max_a \left\{ \sum_s \alpha^a(s)b(s) + \sum_y \max_{\alpha \in \Gamma_{t-1}} \sum_s \alpha^{ay}(s)b(s) \right\}$$

This tells us that the thing to do next is to sum the α-vectors

over y. this is done with the cross-sum operator:

$$\Gamma_t^a = \Gamma_t^{ar} \bigoplus \Gamma_t^{a,y1} \bigoplus \Gamma_t^{a,y2} \bigoplus \cdots$$

Here the symbol $\bigoplus$ denotes the cross-sum operator. When applied to two sets $A = \{a_1, a_2, \ldots, a_m\}$ and $B = \{b_1, b_2, \ldots, b_n\}$, it produces $C = \{a_1+b_1, a_1+b_2, \ldots, a_1+b_n, a_2+b_1, \ldots, a_m+b_n\}$.

The next step is to take the union of sets:

$$\Gamma_t = \bigcup_a \Gamma_t^a$$

Finally, can use the PWLC definition of the value function in order to compute it. We will look at a simple example next. Note however that some algorithms (such as incremental pruning) avoid computing all the α-vectors by focusing only on the active ones.

◇ EXAMPLE FROM PINEAU'S THESIS

⋆ Proof:

⋆ Proof: