

Relaxing the Approximate Linear Program by Restricting the Dual Space

October 16, 2011

1. Motivation

[dFvR03] studied the approximate linear programming approach (ALP) to approximate dynamic programming. It has been subsequently noted that the feasible set of the ALP is restrictive. [DVVM09] and [WB10] list two explicit ways of generalizing ALP. In this paper we show that both of these methods can be obtained from the dual problem of the ALP by restring the feasible set of the dual. This motivates other ways of restricting the dual and finding more ways of restricting the dual to obtain relaxed primal problems.

2. Dual Shrinkage

Let \mathcal{S} be the finite set of states for a Markov decision process. Let the per stage cost be $g_{x,a}$ and let it be an infinite horizon problem with discount α . Also let $\Phi : \mathcal{S} \rightarrow \mathbb{R}^m$ be the feature map. In ALP we solve the program,

$$(1) \quad \begin{aligned} & \text{maximize} && \sum_{x \in \mathcal{S}} \nu_x r^\top \Phi(x) \\ & \text{subject to} && r^\top \Phi(x) \leq g_{a,x} + \alpha \mathbf{E}_{x,a}[r^\top \Phi(Y)] \quad \forall x \in \mathcal{S}, a \in \mathcal{A}(x), \\ & && r \in \mathbb{R}^m. \end{aligned}$$

Here $\mathbf{E}_{x,a}[f(Y)] \triangleq \sum_{y \in \mathcal{S}} p(x,y,a)f(y)$.

Let $\lambda_{x,a}$ be the dual variable associated with each constraint. Then, the dual problem is,

$$(2) \quad \begin{aligned} & \text{minimize} && \lambda^\top g \\ & \text{subject to} && \sum_{x,a} \lambda_{x,a} (\Phi(x) - \alpha \mathbf{E}_{x,a}[\Phi(Y)]) = \sum_x c_x \Phi(x) \\ & && \lambda \geq 0, \lambda \in \mathbb{R}^{|\mathcal{S}||\mathcal{A}|}. \end{aligned}$$

Let us think of shrinking the dual feasible set by adding constraints of the type $\lambda \in \mathcal{C}$. We show that two choices of these sets give us two ADP methods in the literature.

2.1. Smoothed Approximate Linear Program

[DVVM09] discuss the following relaxation to the ALP,

$$(3) \quad \begin{aligned} & \text{maximize} && \sum_{x \in \mathcal{S}} \left(\nu_x r^\top \Phi(x) + \frac{2}{1-\alpha} \pi_x s_x \right) \\ & \text{subject to} && r^\top \Phi(x) \leq g_{a,x} + \alpha \mathbf{E}_{x,a}[r^\top \Phi(Y)] + s_x \quad \forall x \in \mathcal{S}, a \in \mathcal{A}(x), \\ & && s \geq 0, s \in \mathbb{R}^{|\mathcal{S}||\mathcal{A}|}, r \in \mathbb{R}^m. \end{aligned}$$

We note that this is equivalent to adding the constraint,

$$\sum_{a \in \mathcal{A}} \lambda_{x,a} \leq \frac{2}{1-\alpha} \pi_x, \quad \forall x \in \mathcal{S},$$

to the dual (2).

2.2. Iterated Bellman constraints

[WB10] come up with a problem,

$$(4) \quad \begin{aligned} & \text{maximize} && \sum_{x \in \mathcal{S}} \nu_x r^{(0)\top} \Phi(x) \\ & \text{subject to} && r^{(i)\top} \Phi(x) \leq g_{a,x} + \alpha \mathbb{E}_{x,a}[r^{((i+1)\%m)\top} \Phi(Y)] \quad \forall x \in \mathcal{S}, a \in \mathcal{A}(x), 0 \leq i \leq M, \\ & && r^{(i)} \in \mathbb{R}^m, \quad \forall 0 \leq i \leq M. \end{aligned}$$

Let the dual variables of this program be denoted by $\lambda^{(i)}$. The dual of this program is,

$$(5) \quad \begin{aligned} & \text{minimize} && \sum_{i=0}^{M-1} \lambda^{(i)\top} g \\ & \text{subject to} && \sum_{x,a} \left(\lambda_{x,a}^0 \Phi(x) - \lambda_{x,a}^{M-1} \alpha \mathbb{E}_{x,a}[\Phi(Y)] \right) = \sum_x c_x \Phi(x) \\ & && \sum_{x,a} \lambda_{x,a}^i \Phi(x) = \lambda_{x,a}^{i-1} \alpha \mathbb{E}_{x,a}[\Phi(Y)], \quad \forall 1 < i < M, \\ & && \lambda^{(i)} \geq 0, \lambda^{(i)} \in \mathbb{R}^{|\mathcal{S}||\mathcal{A}|}, \quad \forall 0 \leq i < M. \end{aligned}$$

This is a restricted version of (2) since any feasible point $\{\lambda^{(i)}\}$ of (5) we have that $\sum_i \lambda^{(i)}$ is feasible for (2) with the same objective.

3. Directions

1. Try other choices of restrictions \mathcal{C} . Check $\|\lambda\| \leq C$ for different norms.
2. If pairs of (x, a) are known to be suboptimal set $\lambda_{x,a} = 0$.

References

- [dFvR03] D. P. de Farias and B. van Roy. The linear programming approach to approximate dynamic programming. *OPERATIONS RESEARCH*, 51(6):850–865, 2003.
- [DVVM09] V. F. Desai V. V., Farias and C. C. Moallemi. Approximate dynamic programming via a smoothed linear program. To appear in *Operations Research*, 2009.
- [WB10] Y. Wang and S. Boyd. Approximate dynamic programming via a iterated bellman inequalities. Manuscript, 2010.