Review Report on

"A Generalized Reduced Linear Program for Markov Decision Processes"

by Chandrashekar Lakshminarayanan, Shalabh Bhatnagar, and Csaba Szepesvari

The authors considered an Approximate Dynamic Programming (ADP) method, called Generalized Reduced Linear Program (GRLP), to handle Markov Decision Processes (MDP) problems. Compared with Approximate Linear Program (ALP), the author claimed that the novel theoretical framework for GRLP could overcome the limitation of ALP and provide error bounds for the GRLP formulation. I carefully go through the proofs in this manuscript and still have two questions as listed in Major Comments. In a short summary, I think this is a nice work but the writing needs to be improved, and recommend a major revision to the following issues.

Major Comments:

1. In the proof of Lemma IV.6, in order to use Lemma IV.1, the authors let $A = W^T E\Phi$, $b = W^T HJ$, $c = e_i$, $b_0 = t\beta_\psi \psi$, and $x_0 = t\beta_\psi r_0$. Note that Lemma IV.1 requires $b_0 = Ax_0$ and the conclusion is $\min\{c^T Ax: Ax \geq b + b_0, x \in N'\} = \min\{c^T Ay: Ay \geq b, y \in N'\} + c^T b_0$. There are two inconsistences:

(1) $Ax_0 = W^T E\Phi t\beta_\psi r_0 = W^T E\left(t\beta_\psi \psi\right) = W^T E b_0$, which in general does not equal $b_0$ unless $W^T E$ is the identity matrix.

(2) $c^T Ax = e_i^T W^T E\Phi r$, which in general does not equal $e_i^T \Phi r$? Unless $W^T E$ is the identity matrix.

As to whether $W^T E = I$, based on Assumption III.1. (v), we know $W$ is a full rank $nd \times m$ matrix, and based on Definition II.1. (vi), we know that $E$ is a $nd \times n$ matrix. Thus, $W^T E$ is a $m \times n$ matrix, which is not the identity matrix if $m \neq n$. Even if $m = n$, $W^T E$ is not the identity matrix unless $W$ is very specifically chosen. The authors need to explain more why that is correct.

2. $r^*$ first appears in Lemma IV.2. But where is the definition of "$r^*$"?

Minor Comments:

1. In the statement of Lemma IV.4, "for any $J \in R^n$" appeared twice, however, one





Attachment uploaded Sat Nov 5 01:12:29 2016 PST

of them should be $\psi \in R^n$.

2. In the last sentence of the proof of Lemma IV.6, "$N' = N' + \alpha tr_0$ (from Assumption III.1-(vi))." The notation $\alpha$ should not appear here, because Assumption III.1-(vi) is $N' = N + tr_0$.

3. In the proof of Lemma IV.10, the first sentence is "We know from Lemma IV.9 that $\Phi\hat{r} = \hat{J} \geq \hat{V}$. It should be "from Lemma IV.8"

4. In the proof of Theorem IV.14, "Now $r'$ is feasible for the ALP in (7) by Lemma IV.13." Here it should be (6) instead of (7).

5. In the last sentence of the proof of Theorem IV.16, $\left|\left|\hat{V} - \hat{J}\right|\right|_{\infty,\psi} \leq \frac{1}{1-\beta_\psi}\left|\left|\hat{J} - \hat{\Gamma}\hat{J}\right|\right|$, here $\left|\left|\hat{J} - \hat{\Gamma}\hat{J}\right|\right|$ should be $\left|\left|\hat{J} - \hat{\Gamma}\hat{J}\right|\right|_{\infty,\psi}$.

6. I am not so comfortable with so many Lemmas and Theorems in the manuscript. It would be better if the authors can restructure the manuscript to emphasize the most important results as Theorems first and explain the necessity of Lemmas later.





Attachment uploaded Sat Nov 5 01:12:29 2016 PST