
Correction:

Convergent Temporal-Difference Learning with Arbitrary Smooth Function Approximation

Hamid R. Maei
University of Alberta
Edmonton, AB, Canada

Csaba Szepesvári*
University of Alberta
Edmonton, AB, Canada

Shalabh Bhatnagar
Indian Institute of Science
Bangalore, India

Doina Precup
McGill University
Montreal, QC, Canada

David Silver
University of Alberta,
Edmonton, AB, Canada

Richard S. Sutton
University of Alberta,
Edmonton, AB, Canada

1. There is a sign error in equations (11) and (12). The correct equations are:

$$h(\theta, u) = \mathbb{E}[(\delta - \phi^\top u) \nabla^2 V_\theta(s) u], \quad (11)$$

$$-\frac{1}{2} \nabla J(\theta) = \mathbb{E}[(\phi - \gamma \phi') \phi^\top w] - h(\theta, w) = \mathbb{E}[\delta \phi] - \gamma \mathbb{E}[\phi' \phi^\top w] - h(\theta, w). \quad (12)$$

2. The last equation in the proof of theorem 1 (after the term “Finally, observe that:”) has a sign error. The correct equation is:

$$\begin{aligned} -\mathbb{E}[(\gamma \phi' - \phi) \phi^\top w] &= \mathbb{E}[(\phi - \gamma \phi') \phi]^\top (\mathbb{E}[\phi \phi^\top]^{-1} \mathbb{E}[\delta \phi]) \\ &= \mathbb{E}[\delta \phi] - \mathbb{E}[\gamma \phi' \phi^\top] (\mathbb{E}[\phi \phi^\top]^{-1} \mathbb{E}[\delta \phi]) = \mathbb{E}[\delta \phi] - \mathbb{E}[\gamma \phi' \phi^\top w]. \end{aligned}$$

*On leave from MTA SZTAKI, Hungary.