December 2, 2008

Errata
"Asynchronous Stochastic Approximation and Q-Learning "
J. N. Tsitsiklis,
Machine Learning, Vol. 16, No. 4, 1994, pp. 185-202.

The proof of Lemma 9 is incorrect as written. A corrected version, essentially the same as the one given in D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-dynamic Programming,* Athena Scientific, 1996, Proposition 5.6, is as follows.

The definition of F_{iu}^{π} in p. 200 should be

$$F_{iu}^{\pi}(Q) = E[c_{iu}] + \sum_{j\neq 1} p_{ij}(u) Q_{j,\pi(j)}, \qquad i \neq 1, u \in U(i).$$

Then, consider a Markov chain with states (i, u) and with the following dynamics: from any state (i, u), we move to state $(j, \pi(j))$, with probability $p_{ij}(u)$; in particular, subsequent to the first transition, we are always at a state of the form $(i, \pi(i))$ and the first component of the state evolves according to π. Let us identify all states of the form $(1, u)$, with a single (absorbing) state. Because π was assumed proper for the original problem, it follows that the system with states (i, u) also evolves according to a proper policy. The transition probability matrix for this chain, after deleting the row and column associated with the absorbing state, has a maximal eigenvalue strictly less than one. By the Perron-Frobenius theorem, there exists a positive vector w with components $w_{i,u}$ and some $\gamma \in [0, 1)$ such that

$$\sum_{j\neq 1} p_{ij}(u) w_{j,\pi(j)} \leq \gamma w_{i,u}, \qquad \forall\ i \neq 1.$$

Therefore, for any vectors Q and Q', we have

$$\begin{aligned} \frac{|F_{iu}^{\pi}(Q) - F_{iu}^{\pi}(Q')|}{w_{i,u}} &\leq \frac{1}{w_{i,u}} \sum_{j\neq 1} p_{ij}(u) w_{j,\pi(j)} \frac{|Q_{j,\pi(j)} - Q'_{j,\pi(j)}|}{w_{j,\pi(j)}} \\ &\leq \gamma \max_{j\neq 1, u\in U(j)} \frac{|Q_{ju} - Q'_{ju}|}{w_{j,v}}. \end{aligned}$$

The rest of the argument remains as given in the paper.