Paper 1822: Concentration Bounds for Two Timescale Stochastic Approximation with Applications to Reinforcement Learning

Consider the stochastic approximation given by

$$\begin{aligned} \theta_{n+1} &= \theta_n + \alpha_n(0 - q\theta_n), \\ w_{n+1} &= w_n + \beta_n(0 - w_n), \end{aligned} \tag{1}$$

with step sizes $\alpha_n = \frac{1}{n}$, say $0 < q < 0.1$ and $\theta_n \in \mathbf{R}$, $w_n \in \mathbf{R}$, $\beta_n = \frac{1}{\sqrt{n}}$. We can put (1) into the framework in the paper by setting $h_1 = 0 - 0.1\theta$, $h_2 = 0 - 1\,w$, $M^{(i)}_{n+1} = 0$, $i = 1, 2$. In this case $\theta^* = 0$ and $\lambda(\theta) = 0$. The related ordinary differential equation is then $\dot{\theta}(t) = -q\theta(t)$ whose solution is $\theta(t) = \theta(0)e^{-qt}$ with probability 1. Assume $\theta(0) = \theta_0 = 1$ for the rest of the discussion. For the aforementioned step size $t(n) = \sum_{0 \le k \le n} \frac{1}{k} \approx \log n$ and hence

$$\theta(t) = e^{-qt} \Rightarrow \theta_n \approx e^{-q \log n} = \frac{1}{n^q} \tag{2}$$

Thus the eigenvalue shows up in the rate in the exponent in (2). A tighter analysis can be done on the original sequence say for $n > n_0$

$$\begin{aligned} \theta_n &= \Pi_{k=0}^{n-n_0}(1 - \frac{q}{k+n_0})\theta_{n_0} \\ |\theta_n| &= \Pi_{k=0}^{n-n_0}(1 - \frac{q}{k+n_0})|\theta_{n_0}| \end{aligned}$$

Now using the fact that for n_0 sufficiently large $e^{-2\frac{q}{k+n_0}} \le (1 - \frac{q}{k+n_0}) \le e^{-\frac{q}{k+n_0}}$ we have

$$|\theta_{n+n_0}| \ge |\theta_{n_0}| e^{-2q\sum_{k=0}^{n-n_0} \frac{1}{k+n_0}} \tag{3}$$

$$|\theta_{n+n_0}| \ge |\theta_{n_0}| e^{-2q \log(n-n_0)} \tag{4}$$

$$|\theta_{n+n_0}| \ge |\theta_{n_0}| (\frac{n_0}{n})^{2q} \tag{5}$$

and the eigenvalue shows up in the form as $\frac{1}{n^{2q}}$.

We now apply Corollary 2 to the stochastic approximation in (1) by setting $\epsilon_2 = \infty$. Now we have

$$Pr\{|\theta_n| > \epsilon_1 | G'_{n_0}\} \le C[e^{-n_0 c'_1 \epsilon_1^2} + e^{-\sqrt{n_0} c'_2 \epsilon_1^2 / \log(n_0)}]$$

Considering the fact that out of the two exponents on the right hand side of the above equation, the second term is dominant for large n_0 we have

$$Pr\{|\theta_n| > \epsilon_1 | G'_{n_0}\} \le C[e^{-n_0 c'_1 \epsilon_1^2} + e^{-\sqrt{n_0} c'_2 \epsilon_1^2 / \log(n_0)}] \approx C e^{-\sqrt{n_0} c'_2 \epsilon_1^2 / \log(n_0)}$$

Now let $\delta = C e^{-\sqrt{n_0} c'_2 \epsilon_1^2 / \log(n_0)}$,

$$Pr\left\{ |\theta_n| > \sqrt{\frac{\log(n_0)\log(\frac{C}{\delta})}{c'_2\sqrt{n_0}}} \middle| G'_{n_0} \right\} \le \delta. \tag{6}$$

Since $n \geq n_1 \geq n_0$, we have equation (6) starts holding for all $n \geq n_1 \geq n_0$ with high probability (i.e., $1-\delta$). So for sufficiently large $n \geq n_0$, θ_n reaches 0 at rate roughly $\frac{1}{n_0^{\frac{1}{4}}}$. However, from (2) and (3) we expect a rate $\frac{1}{n^q}$ where q is the eigenvalue, which can be arbitrarily small.

Issue: c_1' and c_2' (in Corollary 2 of the paper) might be dependent on q_1 and q_2 which are actually the minimum eigenvalues as mentioned in lines $489-490$ above equation (28) of the submission. However, c_1' and c_2' are not showing up in the exponent of n_0 in (6).