RL exercise 07

Mengnan Wang

June 13, 2017

1 Normalized gradient

Gradient ascent with fixed step-size $alpha = 0.5$, shown in figure 1.

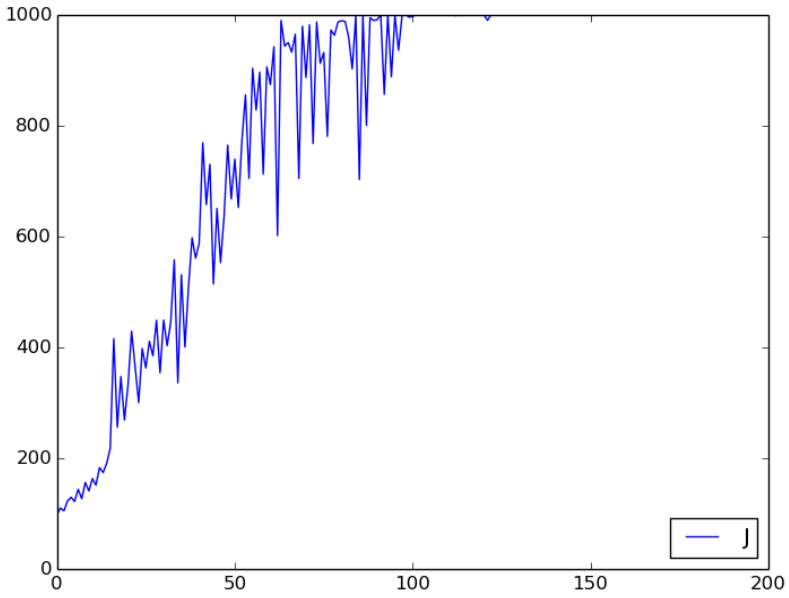

Figure 1: Fixed step size $alpha = 0.5$

2 Heuristic step-size

I first choose line search and wolfe condition to select step-size. While the result showed that wolfe condition was prone to 'jump' into local maximum but it was very stable inside local maximum, figure 2.

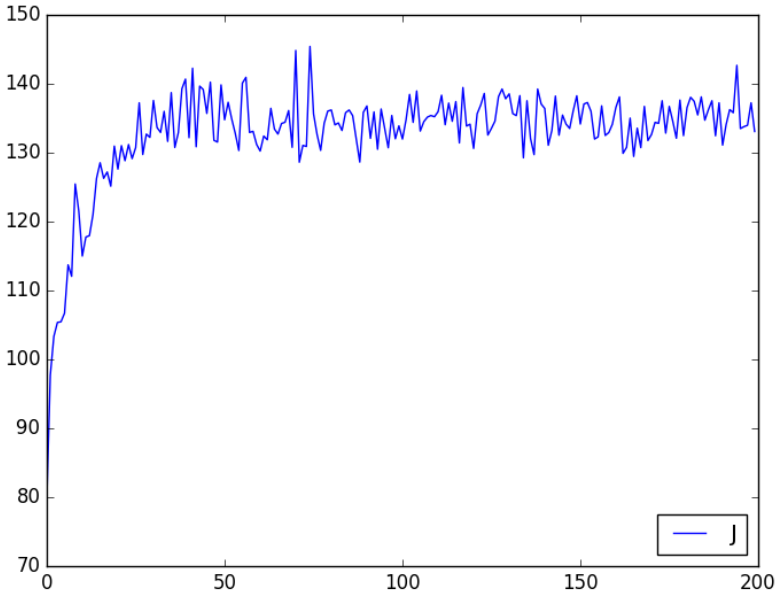

Figure 2: Wolfe condition and line search

Using decreasing step-size $\alpha_t = \frac{10}{t}$, the parameter converges quickly but it will fluctuate for a while, figure 3.

So i combine wolfe condition with $\alpha_t = \frac{10}{t}$. Since the maximal time step of each episode is 1000, the best expect reward should be 1000. With this prior knowledge, i first use $\alpha_t = \frac{10}{t}$ to choose step-size. After the expect reward is larger than 800. I switch to line search and wolfe condition for choosing step-size. The result shows that after expect reward is larger than 800, the parameter quickly converges to global optimum and stick to it, figure 4.

3 Adaptive step-size

I only run adaptive step-size for 20 iteration, since it converges very fast, figure 5.

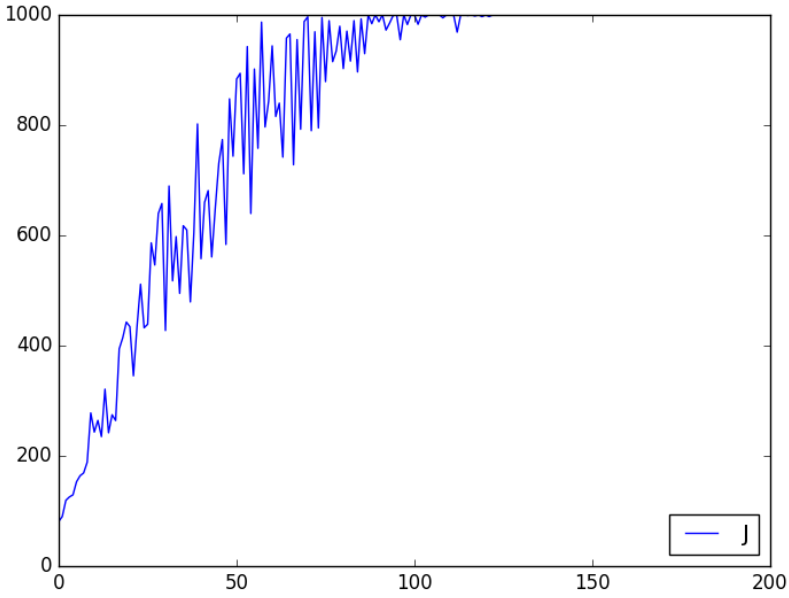

Figure 3: Decreasing step-size$\alpha_t = \frac{10}{t}$

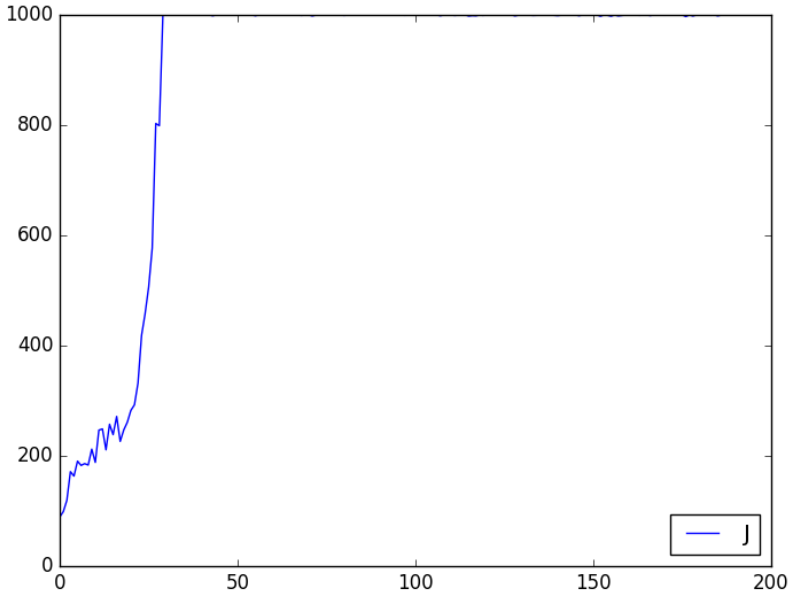

Figure 4: Combined step-size

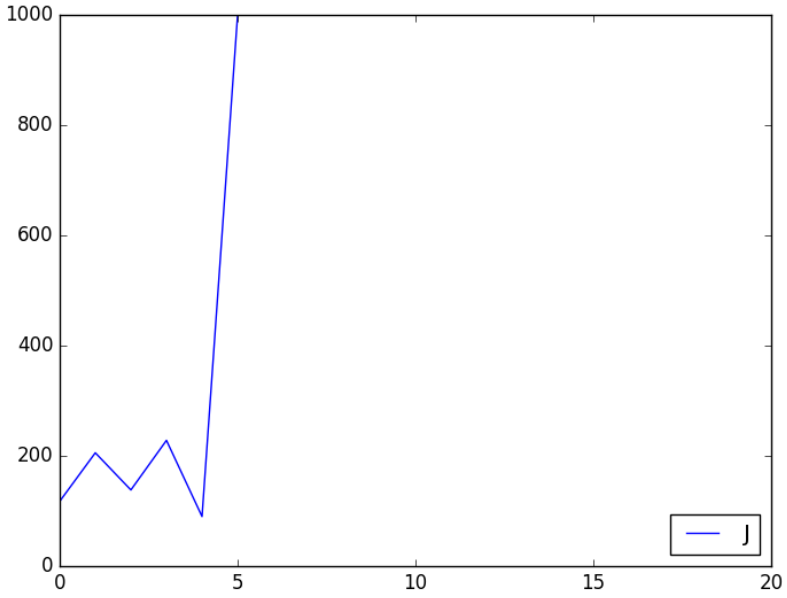

Figure 5: Adaptive step-size