# Artificial Intelligence : Linear Updates

chandru@iitpkd.ac.in

We want to understand the behaviour of iterative algorithms leading to temporal difference learning TD(0) algorithm. In what follows $\eta_t$ will be a sequence of identically distributed independent random variables which take $-1$ and $1$ with equal probability.

Step1 Averaging of only noise: Generate $t = 1000$ samples of $\eta_t, t = 1, \ldots, 1000$ and then compute the sample mean $\theta_t = \frac{\eta_1 + \ldots + \eta_t}{t}$. Plot $\eta_t$, and $\theta_t$, and visually check how the averaging happens.

Step2 Averaging of only noise using recursion: On the same samples of $\eta_t$, let $\theta_t = \theta_{t-1} + \alpha_t(\eta_t - \theta_{t-1})$, what happens with $\alpha_t = \frac{1}{t}$, and what happens with $\alpha_t = \frac{1}{t+1}$, or say $\alpha_t = \frac{1}{t+k}, k \geq 1$.

Step3 Recursion with constant step-size and only noise: On the same samples of $\eta_t$, let $\theta_t = \theta_{t-1} + \alpha_t(\eta_t - \theta_{t-1})$, what happens with $\alpha_t = \alpha$, for $\alpha = 2, 1, 0.1, 0.01$. Plot the results.

Step4 Averaging and Recursion with input plus noise: Let $\theta_t = \theta_{t-1} + \alpha_t(\theta_* + \eta_t - \theta_{t-1})$, now repeat Steps $1, 2, 3$. Let $\theta_*$ be any real number say $1$ or $-1.5$ etc.

In what follows $\eta_t \in \mathbb{R}^2$ will be a sequence of identically distributed independent random variables which take $-1$ and $1$ with equal probability, i.e., $\eta_t = (\eta_t(1), \eta_t(2))$ and each entry is independent of the other. In what follows $\theta_t \in \mathbb{R}^2$

Step1 Averaging of only noise: Generate $t = 1000$ samples of $\eta_t, t = 1, \ldots, 1000$ and then compute the sample mean $\theta_t = \frac{\eta_1 + \ldots + \eta_t}{t}$. Plot $\eta_t$, and $\theta_t$, and visually check how the averaging happens.

Step2 Averaging of only noise using recursion: On the same samples of $\eta_t$, let $\theta_t = \theta_{t-1} + \alpha_t(\eta_t - \theta_{t-1})$, what happens with $\alpha_t = \frac{1}{t}$, and what happens with $\alpha_t = \frac{1}{t+1}$, or say $\alpha_t = \frac{1}{t+k}, k \geq 1$.

Step3 Recursion with constant step-size and only noise: On the same samples of $\eta_t$, let $\theta_t = \theta_{t-1} + \alpha_t(\eta_t - \theta_{t-1})$, what happens with $\alpha_t = \alpha$, for $\alpha = 2, 1, 0.1, 0.01$. Plot the results.

Step4 Averaging and Recursion with input plus noise: Let $\theta_t = \theta_{t-1} + \alpha_t(\theta_* + \eta_t - \theta_{t-1})$, now repeat Steps $1, 2, 3$. Now $\theta_* \in \mathbb{R}^2$.

Note that from 1-dim to 2-dim nothing much should change for these set of experiments. Let us understand the 2-dim case bit more: in what follows $A \in \mathbb{R}^{2\times 2}$ and $b \in \mathbb{R}^2$. Our aim is to find out and check in experiment which values of $\alpha > 0$ gives convergence and for which values leads to divergent behaviour. We can plot $||\theta_t - \theta_*||_2^2$.

1. Let $\theta_t = \theta_{t-1} + \alpha(b - \theta_{t-1})$. This is same as $\eta_t = 0$ and $\theta_* = b$ in the Step $4$ of the previous set.
2. Let $\theta_t = \theta_{t-1} + \alpha(b - A\theta_{t-1})$. Note that if $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, then we get Step $1$ above. Now let $i)$ $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$, $ii)$ $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, $iii)$ $A = \begin{bmatrix} 1 & 0.1 \\ -0.1 & 1 \end{bmatrix}$, $iv)$ $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$, $v)$ $A = \begin{bmatrix} 1 & 10 \\ -10 & 1 \end{bmatrix}$, $vi)$ $A = \begin{bmatrix} 1 & 10 \\ 10 & 1 \end{bmatrix}$, $vii)$ $A = \begin{bmatrix} 1 & 0.1 \\ 0.1 & 1 \end{bmatrix}$. What happens in these cases for different $b$s, check with $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $b = \begin{bmatrix} 10 \\ 1 \end{bmatrix}$ (use your imagination).

3. Now consider Step 2, let us add noise, $\theta_t = \theta_{t-1} + \alpha(b - A\theta_{t-1} + \eta_t)$. Plot what happens.
4. Now try with diminishing step-sizes, i.e., $\alpha_t = \frac{c}{c_1+t}$, where $c = 0.1$ and vary $c_1 = 10, 100, 1000$ etc.

Let us now understand how $b - A\theta$ looks like. Pick the different $A$s and $b$s in the above set of experiments, and plot $(b - A\theta)$ as vectors in 2-dim.