ECEN 743: Reinforcement Learning

Stochastic Approximation

Dileep Kalathil
Assistant Professor
Department of Electrical and Computer Engineering
Texas A&M University

References

- "Neuro-Dynamic Programming", D. Bertsekas and J. Tsitsiklis, Chapter 5

Stochastic Approximation

- Objective: Let h be an *unknown* function, $h : \mathbb{R}^d \to \mathbb{R}^d$.
  Solve the equation $h(\theta) = 0$ (to get the solution θ^*).

Stochastic Approximation

- Objective: Let h be an *unknown* function, $h : \mathbb{R}^d \to \mathbb{R}^d$.
  Solve the equation $h(\theta) = 0$ (to get the solution θ^*).

- Observations: Noisy measurements of $h(\cdot)$. We can sequentially select $(\theta_k)_{k \geq 1}$ and observe $z_k = [h(\theta_k) + w_k]$, where w_k is a noise term.
  The noise w_k is usually a martingale difference sequence, i.e., $\mathbb{E}[w_k | \mathcal{F}_{k-1}] = 0$, where $\mathcal{F}_{k-1} = \sigma(\theta_0, z_0, \ldots, \theta_{k-1}, z_{k-1})$

Stochastic Approximation

- Objective: Let h be an *unknown* function, $h : \mathbb{R}^d \to \mathbb{R}^d$.
  Solve the equation $h(\theta) = 0$ (to get the solution θ^*).

- Observations: Noisy measurements of $h(\cdot)$. We can sequentially select $(\theta_k)_{k\geq 1}$ and observe $z_k = [h(\theta_k) + w_k]$, where w_k is a noise term.
  The noise w_k is usually a martingale difference sequence, i.e., $\mathbb{E}[w_k|\mathcal{F}_{k-1}] = 0$, where $\mathcal{F}_{k-1} = \sigma(\theta_0, z_0, \ldots, \theta_{k-1}, z_{k-1})$

- Algorithm:
$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k[h(\theta_k) + w_k]$$

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k [h(\theta_k) + w_k]$$

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k [h(\theta_k) + w_k]$$

Example 1: Estimating mean: Let $(y_k)_{k \geq 1}$ be the i.i.d. samples of a random variable with mean θ^*. Let θ_k be the estimate of the mean after k step.

$$\theta_{k+1} = \frac{1}{k+1} \sum_{i=1}^{k+1} y_i = \theta_k + \frac{1}{k+1} (y_{k+1} - \theta_k)$$

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k [h(\theta_k) + w_k]$$

Example 1: Estimating mean: Let $(y_k)_{k \geq 1}$ be the i.i.d. samples of a random variable with mean θ^*. Let θ_k be the estimate of the mean after k step.

$$\theta_{k+1} = \frac{1}{k+1} \sum_{i=1}^{k+1} y_i = \theta_k + \frac{1}{k+1} (y_{k+1} - \theta_k)$$

This is stochastic approximation with

$$z_k =$$
$$h(\theta) =$$
$$w_k =$$
$$\alpha_k = 1$$

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k[h(\theta_k) + w_k]$$

Example 1: Estimating mean: Let $(y_k)_{k\geq 1}$ be the i.i.d. samples of a random variable with mean θ^*. Let θ_k be the estimate of the mean after k step.

$$\theta_{k+1} = \frac{1}{k+1}\sum_{i=1}^{k+1} y_i = \theta_k + \frac{1}{k+1}(y_{k+1} - \theta_k)$$

This is stochastic approximation with

$$\begin{aligned} z_k &= [h(\theta_k) + w_k] = (y_{k+1} - \theta_k) \\ h(\theta) &= \theta^* - \theta, \\ w_k &= y_{k+1} - \theta^*, \\ \alpha_k &= 1/(k+1). \end{aligned}$$

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k [h(\theta_k) + w_k]$$

Example 2: Stochastic Gradient Descent: Objective is to compute the minimum of a function $f(\cdot)$. Function is unknown, but noisy gradient is available. Stochastic gradient descent update equation is

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k [h(\theta_k) + w_k]$$

Example 2: Stochastic Gradient Descent: Objective is to compute the minimum of a function $f(\cdot)$. Function is unknown, but noisy gradient is available. Stochastic gradient descent update equation is

$$\theta_{k+1} = \theta_k - \alpha_k [\nabla f(\theta_k) + w_k]$$

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k [h(\theta_k) + w_k]$$

Example 3: Fixed Point Iteration: Objective is to compute the fixed point of a contraction mapping $F(\cdot)$. Noisy measurement of $F(\cdot)$ is available. Then, the stochastic fixed point iteration can be defined as

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k [h(\theta_k) + w_k]$$

Example 3: Fixed Point Iteration: Objective is to compute the fixed point of a contraction mapping $F(\cdot)$. Noisy measurement of $F(\cdot)$ is available. Then, the stochastic fixed point iteration can be defined as

$$\theta_{k+1} = \theta_k + \alpha_k [F(\theta_k) - \theta_k + w_k]$$

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k [h(\theta_k) + w_k]$$

Example 4: Q-learning as stochastic approximation: Q-learning update equation is

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha_t \left(r_t + \gamma \max_b Q_t(s_{t+1}, b) - Q_t(s_t, a_t)\right)$$

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k [h(\theta_k) + w_k]$$

Example 4: Q-learning as stochastic approximation: Q-learning update equation is

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha_t \left(r_t + \gamma \max_b Q_t(s_{t+1}, b) - Q_t(s_t, a_t)\right)$$

This can be written as $Q_{t+1} = Q_t + \alpha_t [h(Q_t) + w_t]$, where

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k [h(\theta_k) + w_k]$$

Example 4: Q-learning as stochastic approximation: Q-learning update equation is

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha_t \ (r_t + \gamma \ \max_b Q_t(s_{t+1}, b) - Q_t(s_t, a_t))$$

This can be written as $Q_{t+1} = Q_t + \alpha_t[h(Q_t) + w_t]$, where

$$h(Q) =$$

$$w_t =$$

$$z_t =$$

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k [h(\theta_k) + w_k]$$

Example 4: Q-learning as stochastic approximation: Q-learning update equation is

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha_t \ (r_t + \gamma \ \max_b Q_t(s_{t+1}, b) - Q_t(s_t, a_t))$$

This can be written as $Q_{k+1} = Q_k + \alpha_k [h(Q_k) + w_k]$, where

$$h(Q)(s,a) = F(Q)(s,a) - Q(s,a) = (r(s,a) + \gamma \ \mathbb{E}_{s' \sim P(\cdot|s,a)}[\max_b Q(s', b)]) - Q(s,a))$$

$$w_k = \gamma \ \max_b Q_t(s_{t+1}, b) - \gamma \ \mathbb{E}_{s' \sim P(\cdot|s,a)}[\max_b Q(s', b)])$$

$$z_t = h(Q_t) + w_t = (r_t + \gamma \ \max_b Q_t(s_{t+1}, b) - Q_t(s_t, a_t))$$

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k [h(\theta_k) + w_k]$$

Example 4: TD-learning: TD learning update equation is

$$V_{t+1}(s_t) = V_t(s_t) + \alpha_t(r_t + \gamma V_t(s_{t+1}) - V_t(s_t))$$

This can be written as $V_{t+1} = V_t + \alpha_t[h(V_t) + w_t]$, where

Stochastic Approximation: Example

- Stochastic approximation update equation

$$\theta_{k+1} = \theta_k + \alpha_k z_k = \theta_k + \alpha_k [h(\theta_k) + w_k]$$

Example 4: TD-learning: TD learning update equation is

$$V_{t+1}(s_t) = V_t(s_t) + \alpha_t (r_t + \gamma V_t(s_{t+1}) - V_t(s_t))$$

This can be written as $V_{t+1} = V_t + \alpha_t [h(V_t) + w_t]$, where

$$\begin{aligned} h(V) &= r_\pi + \gamma P_\pi V - V \\ w_t &= \gamma (V_t(s_{t+1}) - (P_\pi V_t)(s_{t+1})) \\ z_t &= r_t + \gamma V_t(s_{t+1}) - V_t(s_t) \end{aligned}$$

Convergence of Stochastic Approximation

- Proving the convergence of stochastic approximation algorithms requires sophisticated mathematical techniques

Convergence of Stochastic Approximation

- Proving the convergence of stochastic approximation algorithms requires sophisticated mathematical techniques
- One common proof approach is the ODE method

Convergence of Stochastic Approximation

- Proving the convergence of stochastic approximation algorithms requires sophisticated mathematical techniques
- One common proof approach is the ODE method
- The basic idea of ODE method is:

$$\theta_{k+1} = \theta_k + \alpha_k(h(\theta_k) + w_k) \quad \text{approximates the ode} \quad \dot{x}(t) = h(x(t))$$

Convergence of Stochastic Approximation

- Proving the convergence of stochastic approximation algorithms requires sophisticated mathematical techniques
- One common proof approach is the ODE method
- The basic idea of ODE method is:

$$\theta_{k+1} = \theta_k + \alpha_k(h(\theta_k) + w_k) \quad \text{approximates the ode} \quad \dot{x}(t) = h(x(t))$$

- In particular, under suitable assumptions, θ_k will converge to the globally asymptotically stable equilibrium of the ODE

Convergence of Stochastic Approximation

Theorem

Consider the stochastic approximation iteration $\theta_{k+1} = \theta_k + \alpha_k[h(\theta_k) + w_k]$. *Let* $\mathcal{F}_k = \sigma(\theta_0, w_0, \ldots, \theta_k, w_k)$. *We will make the following assumptions:*

Convergence of Stochastic Approximation

Theorem

Consider the stochastic approximation iteration $\theta_{k+1} = \theta_k + \alpha_k[h(\theta_k) + w_k]$. *Let* $\mathcal{F}_k = \sigma(\theta_0, w_0, \ldots, \theta_k, w_k)$. *We will make the following assumptions:*

1. $h : \mathbb{R}^d \to \mathbb{R}^d$ *is Lipschitz, i.e.,* $\|h(x) - h(y)\| \leq L\|x - y\|$

Convergence of Stochastic Approximation

Theorem

Consider the stochastic approximation iteration $\theta_{k+1} = \theta_k + \alpha_k[h(\theta_k) + w_k]$. *Let* $\mathcal{F}_k = \sigma(\theta_0, w_0, \ldots, \theta_k, w_k)$. *We will make the following assumptions:*

1. $h : \mathbb{R}^d \to \mathbb{R}^d$ *is Lipschitz, i.e.,* $\|h(x) - h(y)\| \leq L\|x - y\|$
2. $\mathbb{E}[w_{k+1}|\mathcal{F}_k] = 0, \forall k$

Convergence of Stochastic Approximation

Theorem

Consider the stochastic approximation iteration $\theta_{k+1} = \theta_k + \alpha_k[h(\theta_k) + w_k]$. *Let* $\mathcal{F}_k = \sigma(\theta_0, w_0, \ldots, \theta_k, w_k)$. *We will make the following assumptions:*

1. $h : \mathbb{R}^d \to \mathbb{R}^d$ *is Lipschitz, i.e.,* $\|h(x) - h(y)\| \leq L\|x - y\|$
2. $\mathbb{E}[w_{k+1}|\mathcal{F}_k] = 0, \forall k$
3. $\mathbb{E}[\|w_{k+1}\|^2|\mathcal{F}_k] \leq A + B\|\theta_k\|^2, \forall k$

Convergence of Stochastic Approximation

Theorem

Consider the stochastic approximation iteration $\theta_{k+1} = \theta_k + \alpha_k[h(\theta_k) + w_k]$. *Let* $\mathcal{F}_k = \sigma(\theta_0, w_0, \ldots, \theta_k, w_k)$. *We will make the following assumptions:*

1. $h : \mathbb{R}^d \to \mathbb{R}^d$ *is Lipschitz, i.e.,* $\|h(x) - h(y)\| \leq L\|x - y\|$
2. $\mathbb{E}[w_{k+1}|\mathcal{F}_k] = 0, \forall k$
3. $\mathbb{E}[\|w_{k+1}\|^2|\mathcal{F}_k] \leq A + B\|\theta_k\|^2, \forall k$
4. $\sum_k \alpha_k = \infty, \quad \sum_k \alpha_k^2 < \infty$

Convergence of Stochastic Approximation

Theorem

Consider the stochastic approximation iteration $\theta_{k+1} = \theta_k + \alpha_k[h(\theta_k) + w_k]$. *Let* $\mathcal{F}_k = \sigma(\theta_0, w_0, \ldots, \theta_k, w_k)$. *We will make the following assumptions:*

1. $h : \mathbb{R}^d \to \mathbb{R}^d$ *is Lipschitz, i.e.,* $\|h(x) - h(y)\| \leq L\|x - y\|$
2. $\mathbb{E}[w_{k+1}|\mathcal{F}_k] = 0, \forall k$
3. $\mathbb{E}[\|w_{k+1}\|^2|\mathcal{F}_k] \leq A + B\|\theta_k\|^2, \forall k$
4. $\sum_k \alpha_k = \infty, \quad \sum_k \alpha_k^2 < \infty$
5. *The ODE* $\dot{\theta} = h(\theta)$ *has a unique globally stable equilibrium* θ^*

Convergence of Stochastic Approximation

Theorem

Consider the stochastic approximation iteration $\theta_{k+1} = \theta_k + \alpha_k[h(\theta_k) + w_k]$. *Let* $\mathcal{F}_k = \sigma(\theta_0, w_0, \ldots, \theta_k, w_k)$. *We will make the following assumptions:*

1. $h : \mathbb{R}^d \to \mathbb{R}^d$ *is Lipschitz, i.e.,* $\|h(x) - h(y)\| \leq L\|x - y\|$
2. $\mathbb{E}[w_{k+1}|\mathcal{F}_k] = 0, \forall k$
3. $\mathbb{E}[\|w_{k+1}\|^2|\mathcal{F}_k] \leq A + B\|\theta_k\|^2, \forall k$
4. $\sum_k \alpha_k = \infty, \quad \sum_k \alpha_k^2 < \infty$
5. *The ODE* $\dot{\theta} = h(\theta)$ *has a unique globally stable equilibrium* θ^*
6. *The sequence* $(\theta_k)_{k\geq 1}$ *is bounded, i.e.,* $\sup_k \|\theta\| < \infty$

Convergence of Stochastic Approximation

Theorem

Consider the stochastic approximation iteration $\theta_{k+1} = \theta_k + \alpha_k[h(\theta_k) + w_k]$. *Let* $\mathcal{F}_k = \sigma(\theta_0, w_0, \ldots, \theta_k, w_k)$. *We will make the following assumptions:*

1. $h : \mathbb{R}^d \to \mathbb{R}^d$ *is Lipschitz, i.e.,* $\|h(x) - h(y)\| \leq L\|x - y\|$
2. $\mathbb{E}[w_{k+1}|\mathcal{F}_k] = 0, \forall k$
3. $\mathbb{E}[\|w_{k+1}\|^2|\mathcal{F}_k] \leq A + B\|\theta_k\|^2, \forall k$
4. $\sum_k \alpha_k = \infty, \quad \sum_k \alpha_k^2 < \infty$
5. *The ODE* $\dot{\theta} = h(\theta)$ *has a unique globally stable equilibrium* θ^*
6. *The sequence* $(\theta_k)_{k\geq 1}$ *is bounded, i.e.,* $\sup_k \|\theta\| < \infty$

Convergence of Stochastic Approximation

Theorem

Consider the stochastic approximation iteration $\theta_{k+1} = \theta_k + \alpha_k[h(\theta_k) + w_k]$. *Let* $\mathcal{F}_k = \sigma(\theta_0, w_0, \ldots, \theta_k, w_k)$. *We will make the following assumptions:*

1. $h : \mathbb{R}^d \to \mathbb{R}^d$ *is Lipschitz, i.e.,* $\|h(x) - h(y)\| \leq L\|x - y\|$
2. $\mathbb{E}[w_{k+1}|\mathcal{F}_k] = 0, \forall k$
3. $\mathbb{E}[\|w_{k+1}\|^2|\mathcal{F}_k] \leq A + B\|\theta_k\|^2, \forall k$
4. $\sum_k \alpha_k = \infty, \quad \sum_k \alpha_k^2 < \infty$
5. *The ODE* $\dot{\theta} = h(\theta)$ *has a unique globally stable equilibrium* θ^*
6. *The sequence* $(\theta_k)_{k \geq 1}$ *is bounded, i.e.,* $\sup_k \|\theta\| < \infty$

Then, $\theta_k \to \theta^*$ *almost surely, for any initial condition* θ_0