Lecture 14: Convergence Analysis

Course: Reinforcement Learning Theory
Instructor: Lei Ying
Department of EECS
University of Michigan, Ann Arbor

Convergence

Convergence Theorem

Consider an algorithm: $Y_{t+1} = Y_t + \beta_t S_t$. Let

- $\|\cdot\|$: Euclidean norm
- $\mathcal{F}_t = \{Y_0, \ldots, Y_t, S_0, \ldots, S_{t-1}, \beta_0, \ldots, \beta_t\}$

Assume

1. Step sizes β_t are non-negative, and

$$\sum_{t=0}^{\infty} \beta_t = \infty, \quad \sum_{t=0}^{\infty} \beta_t^2 < \infty$$

Convergence

Convergence Theorem

2. There exists a function $V : \mathbb{R}^n \to \mathbb{R}$ (Lyapunov function) such that:
   (a) $V(y) \geq 0 \, \forall y \in \mathbb{R}^n$
   (b) $V(y)$ is continuously differentiable and $\exists c > 0$ such that

$$\|\nabla V(y) - \nabla V(x)\| \leq c\|y - x\| \, \forall x, y \in \mathbb{R}^n$$

   (c) There exists a positive constant c' such that for all t,

$$c'\|\nabla V(Y_t)\|^2 \leq -\nabla V^T(Y_t)E[S_t|\mathcal{F}_t]$$

   (d) There exists k_1 and k_2 such that for all t,

$$E[\|S_t\|^2|\mathcal{F}_t] \leq k_1 + k_2\|\nabla V(Y_t)\|^2$$

Convergence

Convergence Theorem

Then we have the following with probability one:

- $V(Y_t)$ converges (as $t \to \infty$)
- $\lim_{t\to\infty} \nabla V(Y_t) = 0$
- Every limit point of Y_t is a stationary point of V.
- Stationary point y^* : $\nabla V(y^*) = 0$
- If V has a unique stationary point, then $Y_t \to y^*$ with probability one.

Example 1: Convergence of SGD

Example: minimize a function

$$V(x)$$

with the following stochastic gradient algorithm:

$$Y_{t+1} = Y_t - \beta_t(\nabla V(Y_t) + W_t)$$

Note that $S_t = -\nabla V(Y_t) - W_t$.

Assumptions:

- $\sum_{t=0}^{\infty} \beta_t = \infty, \quad \sum_{t=0}^{\infty} \beta_t^2 < \infty$ (condition 1 satisfied)
- $E[W_t | \mathcal{F}_t] = 0$
- $E[\|W_t\|^2 | \mathcal{F}_t] \leq A + B\|\nabla V(X_t)\|^2$ for some constants A and B.
- V is non-negative and has a Lipschitz continuous gradient (conditions 2(a) and 2(b) are satisfied).

Example 1: Convergence of SGD

Proof:

$$
\begin{aligned}
& \nabla V^T(Y_t) E[S_t | \mathcal{F}_t] \\
=& \nabla V^T(Y_t)(-\nabla V(Y_t) - E[W_t | \mathcal{F}_t]) \\
=& - \|\nabla V(Y_t)\|^2 && \text{(condition 2(c))}
\end{aligned}
$$

$$
\begin{aligned}
E[\|S_t\|^2 | \mathcal{F}_t] &= \|\nabla V(Y_t)\|^2 + E[\|W_t\|^2 | \mathcal{F}_t] + 2\nabla V^T(Y_t) \underbrace{E[W_t | \mathcal{F}_t]}_{=0} \\
&\leq \|\nabla V(Y_t)\|^2 + A + B\|\nabla V(Y_t)\|^2 \\
&= A + (B+1)\|\nabla V(Y_t)\|^2 && \text{(condition 2(d))}
\end{aligned}
$$

Thus, we conclude that

$$\lim_{t \to \infty} \nabla V(Y_t) = 0$$

Example 2: Incremental Gradient Methods (mini-batch)

Consider the following objective:

$$V(y) = \frac{1}{K}\sum_{k=1}^{K} V_k(y).$$

- Assume $\|\nabla V_k(y)\|^2 \leq C + \frac{D}{K}\|\sum_{k=1}^{K}\nabla V_k(y)\|^2 \quad \forall k, y.$

Example: Consider the training of NN with loss function $L(w)$

$$L(w) = \frac{1}{K}\sum_{k=1}^{K} L(x_k; w)$$

where x_k are data samples and w are the weight of the DNN.

Example 2: Incremental Gradient Methods (mini-batch))

Full gradient: $Y_{t+1} = Y_t - \frac{\beta_t}{K}\sum_{k=1}^{K} \nabla V_k(Y_t)$

Mini-batch (with size 1)

$$Y_{t+1} = Y_t - \beta_t \nabla V_{k(t)}(Y_t).$$

$k(t)$: a sequence of i.i.d. random variables, uniformly over $\{1, \ldots, K\}$

$$Y_{t+1} = Y_t - \frac{\beta_t}{K}\sum_{k=1}^{K} \nabla V_k(Y_t) - \beta_t(\nabla V_{k(t)}(Y_t) - \frac{1}{K}\sum_{k=1}^{K} \nabla V_k(Y_t))$$

Then,

$$Y_{t+1} = Y_t - \beta_t \nabla V(Y_t) - \beta_t W_t$$

$$W_t = \nabla V_{k(t)}(Y(t)) - \frac{1}{K}\sum_{k=1}^{K} \nabla V_k(Y_t)$$

$$E[\nabla V_{k(t)}(Y_t)|\mathcal{F}_t] = \frac{1}{K}\sum_{k=1}^{K} \nabla V_k(Y_t)$$

Example 2: Incremental Gradient Methods (mini-batch))

$$E[W_t|\mathcal{F}_t] = 0,$$

and

$$\begin{aligned}
E[\|W_t\|^2|\mathcal{F}_t] &= E[\|\nabla V_{k(t)}(Y_t)\|^2|\mathcal{F}_t] - \|E[\nabla V_{k(t)}(Y_t)|\mathcal{F}_t]\|^2 \\
&\leq E[\|\nabla V_{k(t)}(Y_t)\|^2|\mathcal{F}_t] \\
&\leq \max_k \|\nabla V_k(Y_t)\|^2 \\
&\leq C + \frac{D}{K}\|\sum_{k=1}^{K} \nabla V_k(Y_t)\|^2 \qquad \text{(Based on assumption)}
\end{aligned}$$

Thus, we conclude that

$$\lim_{t\to\infty} \nabla V(Y_t) = 0.$$

Reference

- Chapter 4.2 of Dimitri P. Bertsekas and John Tsitsiklis, *Neuro-Dynamic Programming*, Athena Scientific, 1996.

Acknowledgements: I would like to thank Alex Zhao for helping prepare the slides, and Honghao Wei and Zixian Yang for correcting typos/mistakes.