ECE 586 BH: Interplay between Control and Machine Learning	Fall 2023
Homework 3	
Instructor: Bin Hu	Due date: Nov 2, 2023

1. This problem is about using control theory to study stochastic gradient descent (SGD) and SAGA. In the class, we talked about how to analyze SGD under the assumptions that f is m-strongly convex and f_i is L-smooth and convex. Here you are asked to analyze SGD under a different set of assumptions:

- f satisfies the "one-point convexity" condition: f has a unique global minimizer x^* and for all x, one has $(x-x^*)^\mathsf{T}\nabla f(x) \geq m\|x-x^*\|^2$ with $m>0$.
- f_i is L-smooth for all i.

Under the above assumptions, f and f_i are not convex in general. However, you can still obtain a convergence bound for SGD under these assumptions.

(a) (15 points) Suppose $v_k = \nabla f_{i_k}(w_k)$. Based on the above assumptions, the following two inequalities hold

$$\mathbb{E}\begin{bmatrix} v_k - x^* \\ w_k \end{bmatrix}^\mathsf{T} \begin{bmatrix} -2L^2 I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} v_k - x^* \\ w_k \end{bmatrix} \leq \frac{2}{n}\sum_{i=1}^n \|\nabla f_i(x^*)\|^2 = M$$

$$\mathbb{E}\begin{bmatrix} v_k - x^* \\ w_k \end{bmatrix}^\mathsf{T} \begin{bmatrix} 2mI & -I \\ -I & 0 \end{bmatrix} \begin{bmatrix} v_k - x^* \\ w_k \end{bmatrix} \leq 0$$

Use the above inequalities to prove that if there exists non-negative λ_1 and λ_2 such that

$$\begin{bmatrix} 1-\rho^2 & -\alpha \\ -\alpha & \alpha^2 \end{bmatrix} - \lambda_1 \begin{bmatrix} -2L^2 & 0 \\ 0 & 1 \end{bmatrix} - \lambda_2 \begin{bmatrix} 2m & -1 \\ -1 & 0 \end{bmatrix} \leq 0, \tag{1}$$

then SGD satisfies the bound

$$\mathbb{E}\|x_k - x^*\|^2 \leq \rho^{2k}\mathbb{E}\|x_0 - x^*\|^2 + \frac{\lambda_1 M}{1-\rho^2}$$

(b) (15 points) Use the above LMI condition to show that SGD satisfies the following bound:

$$\mathbb{E}\|x_k - x^*\|^2 \leq (1 - 2m\alpha + 2L^2\alpha^2)^k \mathbb{E}\|x_0 - x^*\|^2 + \frac{\alpha^2 M}{2m\alpha - 2L^2\alpha^2}$$

(c) (15 points) Finally, analyze SAGA under the assumption that f_i is L-smooth and m-strongly convex for all i. Under this assumption, f is also L-smooth and m-strongly convex. Suppose a uniform sampling is used, i.e. $p_i = \frac{1}{n}$. Set $n = 5$, $L = 10$, $m = 1$, and $\alpha = \frac{1}{3L}$. Show that SAGA converges at the rate $\rho^2 = 1 - \min\{\frac{1}{3n}, \frac{m}{3L}\}$ in this case.

2. (30 points) In this question we will revisit the problem of certifying control barrier functions (CBF) from expert demonstrations, but using an alternative approach to Lipschitz properties. Consider the control affine system

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t), \quad x(0) \in \mathbb{R}^n.$$

Suppose we have expert trajectory data $\mathcal{Z}_{dyn} := \{(x_i, u_i)\}_{i=1}^{N_1}$ such that the state data $\mathcal{X}_{safe} := \{x_i\}_{i=1}^{N_1}$ forms an ε-net over the set $\mathcal{D}$, a region of the safe set (i.e. for each $x \in \mathcal{D}$, there is some $x_i \in \mathcal{X}_{safe}$ such that $\|x - x_i\| \leq \varepsilon$). Similarly, we are given sampled data $\mathcal{X}_{unsafe} := \{x_i\}_{i=1}^{N_2}$ that forms an ε-net over an unsafe region $\mathcal{N}$ that we would like to avoid.

Now, suppose we are given a candidate continuously differentiable CBF $h : \mathbb{R}^n \to \mathbb{R}$ (perhaps parameterized by a neural-network), an extended class $\mathcal{K}$ function α and slack constant $\gamma > 0$ satisfying:

1) $h(x_i) \geq \gamma$ for all $x_i \in \mathcal{X}_{safe}$
2) $h(x_i) \leq -\gamma$ for all $x_i \in \mathcal{X}_{unsafe}$
3) $q(x_i, u_i) := \langle \nabla h(x_i), f(x_i) + g(x_i)u_i \rangle + \alpha(h(x_i)) \geq \gamma$ for all $(x_i, u_i) \in \mathcal{Z}_{dyn}$

In addition, assume that h and q are locally bounded near our data in the following sense:

$$\|x_i - x\| \leq \varepsilon \implies \|h(x_i) - h(x)\| \leq \delta(x_i, \varepsilon), \quad \forall x_i \in \mathcal{X}_{safe} \cup \mathcal{X}_{unsafe}$$

and

$$\|x_i - x\| \leq \varepsilon \implies \|q(x_i, u_i) - q(x, u_i)\| \leq \delta(x_i, \varepsilon), \quad \forall (x_i, u_i) \in \mathcal{Z}_{dyn}$$

Prove that if $\delta(x_i, \varepsilon) \leq \gamma$ for all $x_i \in \mathcal{X}_{safe} \cup \mathcal{X}_{unsafe}$, then:

a) $h(x) \geq 0$ for all $x \in \mathcal{D}$.
b) $h(x) \leq 0$ for all $x \in \mathcal{N}$.
c) $\sup_{u \in \mathcal{U}} \langle \nabla h(x), f(x) + g(x)u \rangle \geq -\alpha(h(x))$ for all $x \in \mathcal{D}$.

Hint: The proof is very similar to the one done in class using Lipschitz constants. In fact, if f is L-Lipschitz, we get a similar local bound:

$$\|x_i - x\| \leq \varepsilon \implies \|f(x_i) - f(x)\| \leq L\|x_i - x\| \leq L\varepsilon =: \delta(x_i, \varepsilon)$$

and so our local bound assumption is weaker than being Lipschitz.

3.(25 points) Consider the TD(0) with linear function approximation. Under policy π, suppose the state $\{s_k\}$ forms a Markov chain with the transition probability $p_{ij} = P(i_{k+1} = j | i_k = i)$. TD(0) iterates as

$$\theta_{k+1} - \theta_\pi = \theta_k - \theta_\pi + \varepsilon A_{i_k}(\theta_k - \theta_\pi) + \varepsilon(A_{i_k}\theta_\pi + b_{i_k}), \tag{2}$$

where $\sum_{i=1}^n p_i^\infty (A_i\theta_\pi + b_i) = 0$ with $p_i^\infty := \lim_{k\to\infty} P(i_k = i)$. Write out an analytical formula for the mean square TD error $\mathbb{E}\|\theta_k - \theta_\pi\|^2$ as a function of $\{A_i, b_i, p_{ij}\}$ and (θ_0, θ_π).