Quiz 1: Multi-Agent Machine Learning

Course Code: CS6007 Instructor: Dr. Avishek Ghosh

Date: August 11, 2025

Q1. [Convexity] [5 marks] Let $S_1, S_2 \subset \mathbb{R}^{m+n}$ be convex. Define the *partial sum* set

$$S = \left\{ (x,\, y_1 + y_2) \;\middle|\; x \in \mathbb{R}^m,\ y_1, y_2 \in \mathbb{R}^n,\ (x, y_1) \in S_1,\ (x, y_2) \in S_2 \right\}.$$

Is S is convex? Provide reasoning.

Solution: Show that S is convex: for any $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}) \in S$ and any $\theta \in [0,1]$,

$$\big(\theta x^{(1)} + (1-\theta)x^{(2)},\ \theta y^{(1)} + (1-\theta)y^{(2)}\big) \in S.$$

Representation of the two points. Because $(x^{(i)}, y^{(i)}) \in S$, there exist $y_1^{(i)}, y_2^{(i)} \in \mathbb{R}^n$ such that

$$(x^{(i)}, y_1^{(i)}) \in S_1, \qquad (x^{(i)}, y_2^{(i)}) \in S_2, \qquad y^{(i)} = y_1^{(i)} + y_2^{(i)}, \quad i = 1, 2.$$

Convex combination. Fix $\theta \in [0,1]$ and set

$$x_\theta = \theta x^{(1)} + (1-\theta)x^{(2)}, \quad y_{1,\theta} = \theta y_1^{(1)} + (1-\theta)y_1^{(2)}, \quad y_{2,\theta} = \theta y_2^{(1)} + (1-\theta)y_2^{(2)}.$$

Convexity of S_1 and S_2 gives

$$(x_\theta, y_{1,\theta}) \in S_1, \qquad (x_\theta, y_{2,\theta}) \in S_2.$$

Summing the y-components.

$$y_{1,\theta} + y_{2,\theta} = \theta(y_1^{(1)} + y_2^{(1)}) + (1-\theta)(y_1^{(2)} + y_2^{(2)}) = \theta y^{(1)} + (1-\theta)y^{(2)}.$$

Conclusion. Thus

$$\big(x_\theta,\ y_{1,\theta} + y_{2,\theta}\big) = \big(\theta x^{(1)} + (1-\theta)x^{(2)},\ \theta y^{(1)} + (1-\theta)y^{(2)}\big) \in S,$$

so S is convex.

Q2.[Smoothness and Strong Convexity] [2.5+2.5 marks] .

(a) Suppose that $f : \mathbb{R}^d \to \mathbb{R}$ is α-strongly convex, and $g : \mathbb{R}^d \to \mathbb{R}$ is β-smooth. Prove that the function $h(x) = f(x) - g(x)$ is convex if $\alpha \geq \beta$?

(b) Is the converse true i.e is there exist a function f which is $\alpha-$strongly convex and g which is $\beta-$ smooth such that $h(x) = f(x) - g(x)$ is convex but $\alpha < \beta$? Either prove or construct a counter-example.

Solution:

(a): You can prove this directly using second order characterization of convex functions. Indeed,

$$\nabla^2 h(x) = \nabla^2 f(x) - \nabla^2 g(x) \succeq (\alpha - \beta)I.$$

Hence, h is convex, i.e., $\nabla^h(x) \succeq 0$ if $\alpha \geq \beta$.

We can prove this without the second order condition as well. The proof is given below. **Given:** $f : \mathbb{R}^d \to \mathbb{R}$ is α-strongly convex, i.e.,

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2}\|y - x\|^2, \quad \forall x, y \in \mathbb{R}^d.$$

and $g : \mathbb{R}^d \to \mathbb{R}$ is β-smooth, i.e.,

$$g(y) \leq g(x) + \langle \nabla g(x), y - x \rangle + \frac{\beta}{2}\|y - x\|^2, \quad \forall x, y \in \mathbb{R}^d.$$

To Show: The function $h(x) = f(x) - g(x)$ is convex if $\alpha \geq \beta$.
Proof:
For $h(x)$ to be convex, we need to show that for all $x, y \in \mathbb{R}^d$,

$$h(y) \geq h(x) + \langle \nabla h(x), y - x \rangle.$$

The function $h(x) = f(x) - g(x)$ has the gradient:

$$\nabla h(x) = \nabla f(x) - \nabla g(x).$$

Using the strong convexity of f, we have:

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2}\|y - x\|^2.$$

Using the smoothness of g, we have:

$$g(y) \leq g(x) + \langle \nabla g(x), y - x \rangle + \frac{\beta}{2}\|y - x\|^2.$$

To form an inequality for $h(y) = f(y) - g(y)$, we combine the two inequalities:

$$\begin{aligned} &h(y) = f(y) - g(y) \\ &\geq \left[f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2}\|y - x\|^2\right] - \left[g(x) + \langle \nabla g(x), y - x \rangle + \frac{\beta}{2}\|y - x\|^2\right]. \end{aligned}$$

Simplifying this, we obtain:

$$h(y) \geq h(x) + \langle \nabla f(x) - \nabla g(x), y - x \rangle + \frac{\alpha - \beta}{2}\|y - x\|^2.$$

So when $\alpha \geq \beta$ then the term $\frac{\alpha-\beta}{2}\|y - x\|^2 \geq 0$ this implies

$$h(y) \geq h(x) + \langle \nabla f(x) - \nabla g(x), y - x \rangle$$

Since the inequality holds for all $x, y \in \mathbb{R}^d$ when $\alpha \geq \beta$, the function $h(x) = f(x) - g(x)$ is convex.

$\square$

(b): Converse: If $h(x) = f(x) - g(x)$ is convex, does it imply that $\alpha \geq \beta$?
Answer: The converse is *not* necessarily true.
Counterexample:
Counterexample:
Consider the functions:

$$f(x) = x^2 + \frac{3}{2}\sin x, \quad g(x) = \frac{3}{2}\sin x.$$

1. **Strong Convexity of $f(x)$:** The function $f(x) = x^2 + \frac{3}{2}\sin x$ is strongly convex.compute the second derivative:

$$f''(x) = 2 + \frac{3}{2}\cos x.$$

Since $\cos x$ oscillates between -1 and 1, the minimum value of $f''(x)$ is:

$$f''(x) \geq 2 - \frac{3}{2} = \frac{1}{2}.$$

Therefore, $f(x)$ is $\alpha = \frac{1}{2}$-strongly convex.

2. **Smoothness of $g(x)$:** The function $g(x) = \frac{3}{2}\sin x$ is smooth. To find its smoothness parameter β, we compute the second derivative:

$$g''(x) = -\frac{3}{2}\sin x.$$

Since $\sin x$ oscillates between -1 and 1, the maximum absolute value of $g''(x)$ is:

$$|g''(x)| \leq \frac{3}{2}.$$

This means that $g(x)$ is $\beta = \frac{3}{2}$-smooth.

3. **Forming the Function $h(x) = f(x) - g(x)$:**

$$h(x) = \left(x^2 + \frac{3}{2}\sin x\right) - \frac{3}{2}\sin x = x^2.$$

4. **Checking Convexity of $h(x)$:** The function $h(x) = x^2$ is a convex function since its second derivative is:

$$h''(x) = 2,$$

which is positive for all x. Therefore, $h(x)$ is convex.

5. **Conclusion:** In this example, $f(x) = x^2 + \frac{3}{2}\sin x$ is $\alpha = \frac{1}{2}$-strongly convex, and $g(x) = \frac{3}{2}\sin x$ is $\beta = \frac{3}{2}$-smooth. Here, $\alpha < \beta$, yet the function $h(x) = f(x) - g(x) = x^2$ is convex. This shows that the convexity of $h(x)$ does not imply $\alpha \geq \beta$. Hence, the converse is false.

Note: There are many example you can construct , for example Consider:

$$f(x) = x^T Q_1 x, \quad g(x) = x^T Q_2 x,$$

where the matrices Q_1 and Q_2 are defined as:

$$Q_1 = \begin{bmatrix} 7 & 0 \\ 0 & 3 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 6 & 0 \\ 0 & 2 \end{bmatrix}.$$

□

Q3. Gradient Descent [1+3+1 marks] Let $f : \mathbb{R}^d \to \mathbb{R}$ with $f(x) = \frac{1}{2}\|x - z\|_2^2$. We want to use Gradient Descent with step size $\alpha^k = \frac{1}{2}$.

(a) Write the Gradient Descent Update.

(b) Show that Gradient Descent converges to z, i.e., $\lim_{T\to\infty} x^T = z$.

(c) Comment on the rate of convergence.

Solution: Let $f : \mathbb{R}^d \to \mathbb{R}$ be given by

$$f(x) \;=\; \tfrac{1}{2}\,\|x - z\|_2^2, \qquad z \in \mathbb{R}^d \text{ fixed.}$$

(a) *Gradient–descent update.* The gradient is $\nabla f(x) = x - z$. With step size $\alpha_k = \frac{1}{2}$,

$$x^{k+1} \;=\; x^k - \alpha_k \nabla f(x^k) \;=\; x^k - \tfrac{1}{2}\,(x^k - z) \;=\; \tfrac{1}{2}\,x^k + \tfrac{1}{2}\,z.$$

(b) *Convergence to the minimiser z:* Unrolling the recursion,

$$x^{k+1} = \left(\tfrac{1}{2}\right)^{k+1} x^0 + \frac{1}{2}\,z \sum_{j=0}^{k} \left(\tfrac{1}{2}\right)^j = \left(\tfrac{1}{2}\right)^{k+1} x^0 + z\left(1 - \left(\tfrac{1}{2}\right)^{k+1}\right).$$

Hence $\lim\limits_{k\to\infty} x^k = z$.

(c) *Rate of convergence.* The error norm decays geometrically:

$$\|x^k - z\| = \left(\tfrac{1}{2}\right)^k \|x^0 - z\|.$$

Thus the algorithm converges *exponentially* (often called *linear* or *geometric* convergence) with contraction factor $\frac{1}{2}$.

End of Quiz