Université Paris SaclayMSc AI & MSc Data Science.TC2-Optimization for Machine LearningFaïcel Chamroukhi

TD1 - November 2024.

Solution 1. General questions about optimization

- (a) **Objective function**: The function to be maximized or minimized in an optimization problem.
- (b) **Feasible set**: The set of all points that satisfy the constraints of the optimization problem.
- (c) **Optimal solution**: A point within the feasible set that minimizes or maximizes the objective function.

Solution 2. Convexity

(a) The function f(x) = |x| is convex.

A function $f : \mathbb{R} \to \mathbb{R}$ is convex if, for all $x_1, x_2 \in \mathbb{R}$ and for any $\lambda \in [0, 1]$:

$$f(\lambda x_1 + (1 - \lambda)x_2) \le \lambda f(x_1) + (1 - \lambda)f(x_2)$$

For f(x) = |x|, for any $x_1, x_2 \in \mathbb{R}$ and $\lambda \in [0, 1]$, we have:

$$f(\lambda x_1 + (1-\lambda)x_2) = |\lambda x_1 + (1-\lambda)x_2|$$

By the triangle inequality, we have:

$$|\lambda x_1 + (1 - \lambda)x_2| \le |\lambda x_1| + |(1 - \lambda)x_2| = \lambda |x_1| + (1 - \lambda)|x_2| = \lambda f(x_1) + (1 - \lambda)f(x_2) + (1 - \lambda)f$$

Threfore f(x) = |x| is convex.

- (b) For $f(x) = x^2$, the second derivative $f''(x) = 2 \ge 0$ for all $x \in \mathbb{R}$. Therefore, f(x) is convex.
- (c) $f(x,y) = x^2 + y^2$ is convex because the Hessian matrix $\nabla^2 f = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ is positive semidefinite : For $f : \mathbb{R}^2 \to \mathbb{R}$

$$\nabla^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

For the function $f(x, y) = x^2 + y^2$:

$$\frac{\partial^2 f}{\partial x^2} = 2, \quad \frac{\partial^2 f}{\partial y^2} = 2, \quad \frac{\partial^2 f}{\partial x \partial y} = 0$$

Therefore, the Hessian matrix is:

$$\nabla^2 f = \begin{bmatrix} 2 & 0\\ 0 & 2 \end{bmatrix} = 2I_2$$

To determine if f is convex, we need to verify if the Hessian matrix is positive semidefinite. A matrix is positive semidefinite if all its eigenvalues are non-negative.

In this case, the Hessian matrix is diagonal, the eigenvalues are then simply the diagonal elements, which are 2 and 2, so positive, which means that the Hessian matrix is positive semidefinite.

Since the Hessian matrix is positive semidefinite for all points $(x, y) \in \mathbb{R}^2$, the function $f(x, y) = x^2 + y^2$ is convex.

Solution 3. Differentiability

- (a) The derivative of $f(x) = x^3$ is $f'(x) = 3x^2$.
- (b) f(x) = |x| is not differentiable at x = 0 because the left-hand and right-hand derivatives at this point are not equal :

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

At x = 0, we have: $f'(0) = \lim_{h \to 0} \frac{|0+h| - |0|}{h} = \lim_{h \to 0} \frac{|h|}{h}$ When h > 0, |h| = h, so: $\lim_{h \to 0^+} \frac{|h|}{h} = \lim_{h \to 0^+} \frac{h}{h} = 1$ When h < 0, |h| = -h, so: $\lim_{h \to 0^-} \frac{|h|}{h} = \lim_{h \to 0^-} \frac{-h}{h} = -1$

$$\lim_{h \to 0^+} \frac{|h|}{h} \neq \lim_{h \to 0^-} \frac{|h|}{h}$$

Therefore, the derivative does not exist at x = 0.

(c) The gradient of $f(x, y) = x^2 + y^2 + xy$ is $\nabla f = (2x + y, 2y + x)^{\top}$, which is well-defined everywhere in \mathbb{R}^2 .

Solution 4. Optimality Conditions

(a) The minimum of $f(x) = x^2 + 4x + 4$ occurs at x = -2. This point satisfies the FOC f'(x) = 0 and satisfies SOC $f''(x) \ge 0$.

- To determine the point where f(x) achieves its minimum, we first need to find the critical points by taking the first derivative and setting it equal to zero (First-Order Optimality Condition):

$$f'(x) = \frac{d}{dx}(x^2 + 4x + 4) = 2x + 4$$

Then the critical point(s) correspond() to

$$f'(x) = 2x + 4 = 0$$

which gives x = -2, So the critical point is at x = -2.

- Verify the Second-Order Optimality Condition (we can also verify the first ones) : i.e. Determine the nature of the critical point: To determine whether the critical point x = -2 is a minimum, we can use the second derivative test. The second derivative of f(x) is:

$$f''(x) = \frac{d}{dx}(2x+4) = 2$$

Since f''(x) = 2 > 0 for all values of x, meaning that the critical point at x = -2 is a minimum.

The function $f(x) = x^2 + 4x + 4$ achieves its minimum at x = -2. The first-order optimality condition is satisfied because f(-2) = 0 and the second-order optimality condition is satisfied because the second derivative at $f''(-2) \ge 0$.