# Section 04: Convexity and Gradient Descent

## 1. Convexity

We've seen multiple equivalent definitions of what it means for a function to be convex. For today we'll be primarily using this one:

**Definition 1** (convex functions)**.** *A function* $f : \mathbb{R}^d \to \mathbb{R}$ *is* ***convex*** *on a set* $A$ *if for all* $x, y \in A$ *and* $\lambda \in [0, 1]$*:*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

(a) Which of the following functions are convex? (Hint: draw a picture)
(i) $|x|$ (ii) cos$(x)$ (iii) $x^2$

(b) Suppose you know that $f$ and $g$ are convex functions on a set $A$. Show that $h(x) := \max\{f(x), g(x)\}$ is also convex of $A$.

(c) Does the same result hold for $h(x) = \min\{f(x), g(x)\}$? If so, give a proof. If not, provide convex functions $f, g$ such that $h$ is not convex.

(d) Convex functions are useful because local minima are always global minima. Informally explain why this has to be true (a picture might help)

## 2. Other Definitions of Convexity

Recall from the homework, that we also sometimes talk about a **set** being convex:

**Definition 2** (convex set)**.** *A set* $A$ *is* ***convex*** *if for all* $x, y \in A$ *and all* $\lambda \in [0, 1]$*, the point* $\lambda x + (1 - \lambda)y$ *is also in* $A$*.*

(a) On the homework, you use prove statements related to the following fact:

a function $f$ defined on a convex set $A \subseteq \mathbb{R}^n$ is convex on $A$ if and only if the set $S = \{(x, z) \in \mathbb{R}^{n+1} : z \geq f(x), x \in A\}$ is convex.

In this part, we'll prove the statement in general.

(b) Suppose $A$ and $B$ are convex sets. Is $A \cap B$ convex? Is $A \cup B$ convex? Either prove or give a counter-example.

(c)

**Definition 3** (concave functions)**.** *We say a function* $f$ *is* ***concave*** *on a set* $A$ *if for all* $x, y \in A$ *and all* $\lambda \in [0, 1]$

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y)$$

Show that if $f(x)$ is convex on $A$ then $g(x) := -f(x)$ is concave on $A$.

(d) Can a function be both convex and concave on the same set? If so, give an example. If not, describe why not.

(e) There is another definition of convex functions if we know the function is twice differentiable:

**Definition 4** (convexity, second order condition). *A twice-differentiable function is convex if* $f''(x) \geq 0$ *for all* $x$

Use the definition above to show $-\log(x)$ is convex.

(f) Use the fact that $-\log(x)$ is convex to show the "arithmetic mean-geometric mean (AMGM) inequality": If $a, b \geq 0$ then $\sqrt{ab} \leq \frac{a+b}{2}$.

(g) Show that if $f$ is convex, then $g(x) = f(ax + b)$ (where $a, b$ are real numbers) is also convex.

## 3. Gradient Descent

We will now examine gradient descent algorithm and study the effect of learning rate $\alpha$ on the convergence of the algorithm. Recall from lecture that Gradient Descent takes on the form of $x_{t+1} = x_t - \alpha \nabla f$

(a) Suppose we want to minimize the function $f(x) = x^2 - 4x + 6$. Run gradient descent by hand to compute using $\alpha = 0.5, 0.1, 10$. For each value of alpha, what was the observation? Don't worry about completing the computation, the goal here is for you to notice a trend following each $\alpha$ value picked.

(b) Consider the two variable function $f(x, y) = x^2y^2 + x^2 - 10x$. Starting from the point $(2, 3)$ run gradient descent with a step size of $0.1$.