MATH 173A: Homework #1

Due on Oct 15, 2024 at 23:59pm

Professor Cloninger

Ray Tsai

A16848188

Problem 1

Use the definition of convex functions to answer the following:

(a) Show that $f : \mathbb{R}^d \to \mathbb{R}$ given by $f(x_1, \ldots, x_d) = \|x\|_2^2 = \sum_{i=1}^d x_i^2$ is convex.

Proof. f is continuously differentiable, with $\nabla f(x) = 2x$. But then, for any $x, y \in \mathbb{R}^d$,

$$f(x) + \nabla f(x)^T (y - x) = x^T x + 2x^T (y - x) = 2x^T y - x^T x = \|y\|_2^2 - \|x - y\|_2^2 \leq f(y),$$

so f is convex. $\square$

(b) Show that $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = |x|$ is convex.

Proof. Let $x, y \in \mathbb{R}$. By the triangle inequality,

$$f(tx + (1-t)y) = |tx + (1-t)y| \leq t|x| + (1-t)|y| = tf(x) + (1-t)f(y),$$

for any $t \in [0, 1]$. The result now follows. $\square$

(c) For (b), show that f is not strictly convex.

Proof. Consider $x, y \geq 0$. Then,

$$f(tx + (1-t)y) = tx + (1-t)y = tf(x) + (1-t)f(y),$$

for all $t \in [0, 1]$. Hence, f is not strictly convex. $\square$

(d) Show that $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = \sqrt{|x|}$ is not convex.

Proof. Let $x = 1, y = 4$. Take $t = \frac{1}{2}$. Then,

$$f\left(t \cdot 1 + (1-t) \cdot 4\right) = f\left(\frac{5}{2}\right) = \sqrt{\frac{5}{2}},$$

but

$$tf(1) + (1-t)f(4) = \frac{1}{2} + 1 = \frac{3}{2} < \sqrt{\frac{5}{2}}.$$

Hence, f is not convex. $\square$

Problem 2

Use the definition of convex sets to answer the following:

(a) Show that if the sets S and T are convex, then $S \cap T$ is convex.

Proof. Let $x, y \in S \cap T$. Since S and T are convex, for any $t \in [0, 1]$, $tx+(1-t)y \in S$ and $tx+(1-t)y \in T$. But then $tx + (1 - t)y \in S \cap T$, so $S \cap T$ is convex. $\square$

(b) Show that the intersection of any number of convex sets is convex.

Proof. Let $S_1, S_2, \ldots, S_n$ be convex sets. We proceed by induction on $n \geq 2$. (a) yields the base case. For $n > 2$, $S_1 \cap S_2 \cap \ldots \cap S_{n-1}$ is convex by induction, and thus $S_1 \cap S_2 \cap \ldots \cap S_{n-1} \cap S_n$ is convex by (a). $\square$

(c) A hyperplane in $\mathbb{R}^d$ is a set of points of the form $\{x : a^T x = b\}$ where $a \in \mathbb{R}^d$ and $b \in \mathbb{R}$. Show that hyperplanes are convex.

Proof. Let Γ be a hyperplane $\{x : a^T x = b\}$ in $\mathbb{R}^d$. Let $x, y \in \Gamma$. Then, for any $t \in [0, 1]$,

$$a^T(tx + (1 - t)y) = t(a^T x) + (1 - t)(a^T y) = tb + (1 - t)b = b,$$

so $tx + (1 - t)y \in \Gamma$. Hence, Γ is convex. $\square$

Problem 3

Use the definition of convex functions and sets to answer the following. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and define the set

$$E_f = \{(x, w) \in \mathbb{R}^{n+1} \mid x \in \mathbb{R}^n,\, w \in \mathbb{R},\, f(x) \leq w\}.$$

(a) Show that for all $x \in \mathbb{R}^n$, $(x, f(x)) \in E_f$.

Proof. Put $w = f(x)$. Since $w = f(x) \geq f(x)$, $(x, f(x)) \in E_f$. □

(b) Show that if f is a convex function, then E_f is a convex set.

Proof. Let $(x_1, w_1), (x_2, w_2) \in E_f$. Since f is convex, for $t \in [0, 1]$.

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2) \leq tw_1 + (1-t)w_2,$$

But then $(tx_1 + (1-t)x_2, tw_1 + (1-t)w_2) \in E_f$, so E_f is convex. □

(c) Show conversely that if E_f is a convex set, then f is a convex function.

Proof. Let $x_1, x_2 \in \mathbb{R}^n$. Since E_f is convex,

$$t(x_1, f(x_1)) + (1-t)(x_2, f(x_2)) = (tx_1 + (1-t)x_2, tf(x_1) + (1-t)f(x_2)) \in E_f$$

for all $t \in [0, 1]$. But then $f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$, so f is convex. □

Problem 4

Find the gradient and Hessian of the following functions, and determine whether the functions are convex.

(a) $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x_1, x_2) = \frac{1}{2}x_1^4 + x_1 x_2 - e^{x_2}$.

Proof.

$$\nabla f(x) = (2x_1^3 + x_2, x_1 - e^{x_2}), \quad \nabla^2 f(x) = \begin{bmatrix} 6x_1^2 & 1 \\ 1 & -e^{x_2} \end{bmatrix}.$$

Since $\det(\nabla^2 f(0,0)) = -1 < 0$, $\nabla^2 f(x)$ is not positive semidefinite. It now follows that f is not convex, as f is twice continuously differentiable. □

(b) $f : \mathbb{R}^d \to \mathbb{R}$ given by $f(x) = \langle a, x \rangle^2 + \langle b, x \rangle$.

Proof. Since $f(x) = (a^T x)^2 + b^T x$, using the chain rule we have

$$\nabla f(x) = 2a^T xa + b, \quad \nabla^2 f(x) = 2aa^T.$$

Since $x^T \nabla^2 f(x) x^T = (a^T x)^T (a^T x) = \|a^T x\|^2 \geq 0$ for all $x \in \mathbb{R}^d$, $\nabla^2 f(x)$ is positive semidefinite, so f is convex. It now follows that f is not convex, as f is twice continuously differentiable. □

Problem 5

For each problem below, find the gradient and show your work.

(a) $f : \mathbb{R}^n \to \mathbb{R}$ for $f(x) = \|x\|_2^2$.

Proof. Since $f(x) = x^T x$,

$$\nabla f(x) = x^T(I + I^T) = 2x.$$

□

(b) $f : \mathbb{R}^n \to \mathbb{R}$ for $f(x) = \|Ax\|_2^2$ where $A \in \mathbb{R}^{m \times n}$.

Proof. Since $f(x) = x^T(A^T A)x$,

$$\nabla f(x) = 2x^T(A^T A) = 2(Ax)^T A = 2A^T Ax.$$

□

(c) $f : \mathbb{R}^n \to \mathbb{R}$ for $f(x) = \|Ax - b\|_2^2$ for $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

Proof. Since $f(x) = (Ax - b)^T(Ax - b) = \|Ax\|_2^2 - 2b^T Ax + \|b\|_2^2$, by (b)

$$\nabla f(x) = 2A^T Ax - 2A^T b = 2A^T(Ax - b).$$

□

(d) $f : \mathbb{R}^n \to \mathbb{R}$ for $f(x) = \|Ax - b\|_2^2 + \gamma\|x\|_2^2$ for $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ and $\gamma > 0$.

Proof. By (a) and (c),

$$\nabla f(x) = 2A^T(Ax - b) + 2\gamma x = 2(A^T A + \gamma I)x - 2A^T b.$$

□

Problem 6

This problem builds on the results from problem 5.

(a) For part 5(c), use the Hessian of $f(x)$ to show that f is convex. Under what conditions is f strictly convex?

Proof. By chain rule, $\nabla^2 f = 2A^T A$. But then

$$x^T(\nabla^2 f)x = 2(Ax)^T Ax = 2\|Ax\|_2^2 \geq 0$$

for all $x \in \mathbb{R}^n$, so $\nabla^2 f(x)$ is positive semidefinite, and thus f is convex. f strictly convex when $x^T \nabla^2 f(x) x = 2\|Ax\|_2^2 > 0$ for all $x \neq 0$. This is true when A has full rank. □

(b) For 5(d), show that $f(x)$ is always strictly convex.

Proof. It suffices to show that $\nabla^2 f = 2(A^T A + \gamma I)$ is positive definite. Since $\gamma > 0$

$$x^T(\nabla^2 f)x = 2(Ax)^T Ax + \gamma x^T x = 2\|Ax\|_2^2 + \gamma\|x\|_2^2 > 0,$$

for all $x \neq 0$. The result now follows. □