Math 173A Homework 1 Solutions

Jon Pham

October 19, 2019

1. Let $f(x) = x^4 - 2x^3 + x^2$ be defined as in the problem.

   (a) For smooth, easily differentiable functions, the best strategy is to use the second derivative test from Calculus. Compute $f'(x) = 4x^3 - 6x^2 + 2x = 2x(2x^2 - 3x + 1) = 2x(2x - 1)(x - 1)$. Your critical points are then, $x_1 = 0, x_2 = \frac{1}{2}, x_3 = 1$.
   Compute $f''(x) = 12x^2 - 12x + 2$. We have that

   $$\begin{aligned} f''(x_1) &= 2 > 0 \\ f''(x_2) &= -1 < 0 \\ f''(x_3) &= 2 > 0 \end{aligned}$$

   The conclusion is that x_1, and x_3 are local minimizers, and x_2 is a local maximizer. Thus, the minima of f is $f(x_1) = f(x_3) = 0$, while the maxima of f is $f(x_2) = \frac{1}{16}$.

   (b) Write $f(x) = x^2(x^2 - 2x + 1) = x^2(x - 1)^2 = [x(x - 1)]^2$. We see that $f(x)$ is a square, so it must satisfy $f(x) \geq 0$ for all $x \in \mathbb{R}$. Because it actually achieves the lower bound at the points x_1 and x_3, the minima of 0 is a global one.
   The maxima achieved at x_2 is clearly local because $f(100) > f(x_2)$. You should of course, compute this to make sure, but for a solutions guide over the weekends, I'm leaving it at proof by inspection.

2. Note that even though we have better methods to prove convexity at this point, the problem explicitly asks that we use the definition, which means proving the convex inequality.

   (a) Define $f(x) = (x - 1)^2$, and let $x, y \in \mathbb{R}, \lambda \in [0, 1]$ form an arbitrary convex combination. Then we have that

   $$\begin{aligned} f(\lambda x + (1 - \lambda)y) &= (\lambda x + (1 - \lambda)y - 1)^2 \\ &= (\lambda(x - 1) + (1 - \lambda)(y - 1))^2, \text{ using } 1 = \lambda + (1 - \lambda) \\ &= (\lambda \hat{x} + (1 - \lambda)\hat{y})^2, \text{ a change of variables} \\ &\leq \lambda \hat{x}^2 + (1 - \lambda)\hat{y}^2, \text{ because } x^2 \text{ is a convex function} \\ &= \lambda f(x) + (1 - \lambda) f(y), \text{ reversing our change of variables} \end{aligned}$$

   This inequality thus holds for arbitrary convex combination, and we are done.

   (b) Define $f(x_1, x_2) = x_1^2 + x_2^2$. Let $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$ and $\lambda \in [0, 1]$ form an arbitrary convex combination. Then we have

   $$\begin{aligned} f(\lambda x + (1 - \lambda)y) &= (\lambda x_1 + (1 - \lambda)y_1)^2 + (\lambda x_2 + (1 - \lambda)y_2)^2 \\ &\leq \lambda x_1^2 + (1 - \lambda)y_1^2 + \lambda x_2^2 + (1 - \lambda)y_2^2 \\ &= \lambda(x_1^2 + x_2^2) + (1 - \lambda)(y_1^2 + y_2^2) \\ &= \lambda f(x) + (1 - \lambda) f(y) \end{aligned}$$

   where the first inequality was achieved using two applications of the fact that x^2 is a convex function.

(c) Define $f(x) = \sqrt{x}$ on $\mathbb{R}^+$. Since we are showing the function is NOT convex by definition, it's enough to find a counterexample to the convex inequality. We will choose the points $1, 4 \in \mathbb{R}^+$ and $\lambda = \frac{1}{2}$. Compute

$$\begin{aligned} f(\frac{1}{2} + 2) &= \sqrt{\frac{5}{2}} = \sqrt{\frac{10}{4}} \\ &> \sqrt{\frac{9}{4}} \\ &= \frac{3}{2} = \frac{1}{2} + 1 \end{aligned}$$

Thus the convex inequality does not hold.

3. It's enough to just prove the second part of the problem, since the first part trivially follows. Let C_i denote convex sets indexed by $i \in I$, and define $S := \bigcap_{i \in I} C_i$. We will show that S is a convex set.

   To do this, let $x, y \in S$ and $\lambda \in [0, 1]$. Then we have $x, y \in C_i$ for all $i \in I$; by definition of C_i, which are all convex, we must have $\lambda x + (1 - \lambda)y \in C_i$ for all $i \in I$ as well. This satisfies the definition of intersection, so $\lambda x + (1 - \lambda)y \in S$, as desired.

4. For this problem, it's enough to just prove the second part since the first part is just the special case of $\alpha = 0, \beta = 1, n = 2$, and $a = (1, 1)$. Let T be defined as in the problem, and let $x, y \in T$ and $\lambda \in [0, 1]$ form an arbitrary convex combination. To show membership in T, we need to verify the defining inequalities. To do this, compute

$$\begin{aligned} \langle a, \lambda x + (1 - \lambda)y \rangle &= \langle a, \lambda x \rangle + \langle a, (1 - \lambda)y \rangle, \text{ using dot product properties.} \\ &= \lambda \langle a, x \rangle + (1 - \lambda)\langle a, y \rangle \end{aligned}$$

   By our assumption that $x, y \in T$, we have that

$$\begin{aligned} \lambda \langle a, x \rangle + (1 - \lambda)\langle a, y \rangle &\leq \lambda\beta + (1 - \lambda)\beta = \beta \\ \lambda \langle a, x \rangle + (1 - \lambda)\langle a, y \rangle &\geq \lambda\alpha + (1 - \lambda)\alpha = \alpha \end{aligned}$$

   Thus our convex combination satisfies the defining inequalities of T and is in the set, proving convexity.

5. We know that for a twice differentiable function, the Hessian being positive semidefinite is a necessary and sufficient condition for convexity. All of these functions are easily seen to be twice differentiable, so we will simply compute the Hessian.

   (a) Define $f(x_1, x_2) = (x_1 - x_2)^2$. Compute

$$\begin{aligned} \nabla f &= (2(x_1 - x_2), -2(x_1 - x_2))^T \\ \nabla^2 f &= \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \succcurlyeq 0 \end{aligned}$$

   This is positive semidefinite because its eigenvalues are all non-negative (0 and 4). Thus f is a convex function.

   (b) Define $f(x, y) = x^3 - 2xy - y^6$. Compute

$$\begin{aligned} \nabla f &= (3x^2 - 2y, -2x - 6y^5)^T \\ \nabla^2 f &= \begin{bmatrix} 6x & -2 \\ -2 & -30y^4 \end{bmatrix} \end{aligned}$$

   At the point $(x, y) = (1, 1)$, the Hessian is NOT positive semidefinite, because it has a negative determinant. Thus f is not a convex function.

(c) Define $f(x_1, x_2, x_3) = cos(x_1) + x_2 x_3^2$. Compute

$$\nabla f = (-sin(x_1), x_3^2, 2x_2 x_3)^T$$
$$\nabla^2 f = \begin{bmatrix} -cos(x_1) & 0 & 0 \\ 0 & 0 & 2x_3 \\ 0 & 2x_3 & 2x_2 \end{bmatrix}$$

At the point $(x_1, x_2, x_3) = (0, -1, 0)$, the matrix is NOT positive semidefinite, because it has negative eigenvalues. Thus f is not a convex function.

6. Recall from lecture that the gradient is the direction of greatest increase (Lecture Notes 1, page 21). Evaluating each of them at the origin yields $\vec{0}$ for all three functions.