ENMG 616 Advanced Optimization & Techniques

Homework Assignment 1

Due Date: October 08, 2020

Problem 1

We first compute the gradient and set it to zero to find the set of stationary points.

$$\nabla f(x,y) = \begin{bmatrix} 4x\left(x^2-4\right) \\ 2y \end{bmatrix} = 0.$$

Hence, the set of stationary points are: $(0,0),\ (-2,0),\ (2,0)$. We next compute the Hessian to determine whether these points are local minima.

$$\nabla^2 f(0,0) = \begin{bmatrix} -16 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{(Indefinite)}, \qquad \nabla^2 f(2,0) = \nabla^2 f(-2,0) = \begin{bmatrix} 32 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{(PSD)}.$$

It directly follows that $(-2,0)$ and $(2,0)$ are two local minima (more specifically they are global here), and $(0,0)$ is a stationary solution.

Problem 2

Let $f(x,y)$ denote the objective. Then,

$$\nabla f(x,y) = \begin{bmatrix} 2x+\beta y+1 \\ 2y+\beta x+2 \end{bmatrix}$$

and

$$\nabla^2 f(x,y) = \begin{bmatrix} 2 & \beta \\ \beta & 2 \end{bmatrix}.$$

Case 1: $\beta = 2$ or -2. In this case, $\nabla f(x,y) = 0$ has no solution. So $f(x,y)$ does not have stationary points, and the objective function is unbounded.

Case 2: $\beta \neq \pm 2$. By solving $\nabla f(x,y) = 0$, we obtain

$$x = \frac{2-2\beta}{\beta^2-4} \quad \text{and} \quad y = \frac{4-\beta}{\beta^2-4}$$

which is the unique stationary point. When $\beta \in (-2,2)$, the Hessian matrix is positive semi-definite, and thus the stationary point is a global minimum. When $\beta > 2$ or $\beta < -2$, the objective function is unbounded. The stationary point in that case is not a global minimum.

Problem 3

Consider the following optimization problem

$$\min_{x\geq 0} f(x) =\triangleq \frac{1}{x} + x.$$

Since $\nabla^2 f(x) = 2/x^3 + 1 \geq 0$ for all $x > 0$, then the function is convex on the feasibility set. To find the global minimum, it suffices to solve

$$\nabla f(x) = \frac{-1}{x^2} + 1 = 0 \implies x = 1.$$

Hence,

$$2 = f(1) \leq f(x) = \frac{1}{x} + x \text{ for all } x \geq 0.$$

Problem 4

Consider the optimization problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \triangleq e^{x_1} + e^{x_2} + \ldots + e^{x_n} \\ s.t. \quad & \mathbf{x} \in \mathcal{X} \triangleq \left\{ x \in \mathbb{R}^n \,\middle|\, \sum_{i=1}^n x_i = s \right\} \end{aligned}$$

a. We first show that the objective function is convex by showing its hessian is positive semi-definite.

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} = \begin{cases} e^{x_i} & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

It follows that

$$\mathbf{v}^T \nabla^2 f(\mathbf{x}) \mathbf{v} = \sum_{i=1}^n v_i^2 e^{x_i} \geq 0 \quad \forall v \in \mathbb{R}^n \Rightarrow \nabla^2 f(\mathbf{x}) \succeq 0.$$

We next show that $\mathcal{X}$ is convex. For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\alpha \in [0, 1]$, we have

$$\begin{aligned} \sum_{i=1}^n [\alpha x_i + (1-\alpha) y_i] &= \sum_{i=1}^n \alpha x_i + \sum_{i=1}^n (1-\alpha) y_i \\ &= \alpha \sum_{i=1}^n x_i + (1-\alpha) \sum_{i=1}^n y_i \\ &= \alpha s + (1-\alpha) s \\ &= s \end{aligned}$$

This shows that the optimization problem is convex.

b. Problem becomes:

$$\begin{aligned}\min_{x} \quad & \sum_{i=1}^{n-1} e^{x_i} + e^{s-\sum_{i=1}^{n-1} x_i} \\ s.t. \quad & \mathbf{x} \in \mathbb{R}^{n-1}\end{aligned}$$

c. Let $g(\mathbf{x}) = \sum_{i=1}^{n-1} e^{x_i} + e^{\gamma-\sum_{i=1}^{n-1} x_i}$. We can see that g is an affine transformation of f which implies that function g is convex. Therefore, any global optimum point should satisfy the first order optimality condition.

$$\nabla f(\mathbf{x}) = 0 \quad \Rightarrow e^{x_i} = e^{s-\sum_{i=1}^{n-1} x_i} \quad \text{i=2, \ldots, n-1}$$

$$\Rightarrow x_i = s - \textstyle\sum_{i=1}^{n-1} x_i \quad \text{i=2, \ldots, n-1.}$$

Summing over all x_i, we get

$$\sum_{i=1}^{n-1} x_i = (n-1)s - (n-1)\sum_{i=1}^{n-1} x_i,$$

which implies

$$\sum_{i=1}^{n-1} x_i = \frac{n-1}{n} s.$$

Substituting back, we obtain

$$x_i = \frac{s}{n},$$

which is a global minimum.

d.

$$f(\mathbf{x}) = \sum_{i=1}^{n} y_i \geq f(\mathbf{x}^*) = n\, e^{s/n} \implies \frac{1}{n}\sum_{i=1}^{n} y_i \geq \left(\prod_{i=1}^{n} y_i\right)^{1/n}.$$

Problem 5

a. We first compute the gradient

$$\nabla f(x,y) = \begin{bmatrix} \dfrac{\sqrt{y}}{2\sqrt{x}} \\ \\ \dfrac{\sqrt{x}}{2\sqrt{y}} \end{bmatrix} \quad x, y > 0.$$

Then,

$$\nabla^2 f(x,y) = \begin{bmatrix} \dfrac{-\sqrt{y}}{4x\sqrt{x}} & \dfrac{1}{4\sqrt{xy}} \\ \\ \dfrac{1}{4\sqrt{xy}} & \dfrac{-x}{4y\sqrt{y}} \end{bmatrix}.$$

It follows that

$$\begin{aligned}
\mathbf{v}^T\,\nabla^2 f\mathbf{v} &= \frac{-\sqrt{y}}{4x\sqrt{x}}v_1^2 + \frac{v_1v_2}{2\sqrt{xy}} - \frac{-\sqrt{x}}{4y\sqrt{y}}v_2^2 \\
&= \frac{-1}{4\sqrt{xy}}\left[\frac{y}{x}v_1^2 - 2v_1v_2 + \frac{x}{y}v_2^2\right] \\
&= \frac{-1}{4\sqrt{xy}}\left[\sqrt{\frac{y}{x}}v_1 - \sqrt{\frac{x}{y}}v_2\right]^2 \leq 0.
\end{aligned}$$

Thus $\nabla^2 f$ is negative semi-definite which directly implies that f is concave.

b. Consider the following two points $p_1 = (1,1,1)$ and $p_2 = (0,0,0)$. Then for $\alpha = 1/2$, we have

$$f(\alpha p_1 + (1-\alpha)p_2) = f(1/2,1/2,1/2) = \sqrt{1/8} < 1/2 = \alpha\, f(p_1) + (1-\alpha)f(p_2).$$

Hence, f is not concave.
Now consider the following two points $p_1 = (3,1,2)$ and $p_2 = (1,3,4)$. Then for $\alpha = 1/2$, we have

$$f(\alpha p_1 + (1-\alpha)p_2) = f(2,2,3) = \sqrt{12} > \frac{\sqrt{6}+\sqrt{12}}{2} = \alpha\, f(p_1) + (1-\alpha)f(p_2).$$

Hence, f is not convex. Thus f is neither convex nor concave.

c. Since an affine transformation of a convex functions is convex, showing convexity of $f(\cdot)$ is equivalent to showing $g(\mathbf{y}) = \log(e^{y_1} + ... + e^{y_n})$ is convex. We show the convexity of $g(\cdot)$, by showing $m(t) = g(\mathbf{y} + t\mathbf{h})$ is convex in t for any fixed $\mathbf{y}, \mathbf{h}$. Notice that

$$m'(t) = \frac{\sum_{i=1}^n h_i e^{y_i + th_i}}{\sum_{i=1}^n e^{y_i+th_i}}$$

and

$$\begin{aligned}
m''(t) &= \frac{\left(\sum_i h_i^2 e^{y_i+th_i}\right)\left(\sum_i e^{y_i+th_i}\right) - \left(\sum_i h_i e^{y_i+th_i}\right)^2}{\left(\sum_i h_i e^{y_i+th_i}\right)^2} \\
&= \frac{\sum_{i=1}^n a_i^2 \sum_{i=1}^n b_i^2 - \left(\sum_{i=1}^n a_i b_i\right)^2}{\left(\sum_{i=1}^n a_i b_i\right)^2} \\
&= \frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2 - \langle a,b\rangle^2}{\langle a,b\rangle^2} \\
&\geq 0.
\end{aligned}$$

Here $a_i = h_i\sqrt{e^{y_i+th_i}}$, $b_i = \sqrt{e^{y_i+th_i}}$, and the last inequality holds by Cauchy-Schwartz. Hence, f is convex.

Problem 6

a. Notice that

$$\nabla^2 f(x) = 6x^2 - 2,$$

which is negative for $x \in \left[\frac{-1}{3}, \frac{1}{3}\right]$. Hence f is not convex.

b. By setting the gradient to zero, we get

$$\nabla f(x) = 2x(x^2 - 1) = 0 \implies x = 0, \text{ or } x = \pm 1.$$

c. Since $\nabla^2 f(0) = -2 < 0$ and $\nabla^2 f(1) = \nabla^2 f(-1) = 4 > 0$, then $x = 0$ is a stationary point (not local), and $x = \pm 1$ are local minima. The optimal objective value is 0 which is achieved at $x = \pm 1$. Hence, all local are global.

Problem 7

a. Consider $\mathbf{x}_1 = \mathbf{x}_0$, $\mathbf{x}_2 = -\mathbf{x}_0$, and $\alpha = 0.5$. Then

$$f\big(\alpha(\mathbf{x}_1) + (1-\alpha)\mathbf{x}_2\big) = f(\mathbf{0}) = \frac{1}{2}\|A\|_F^2 > 0.$$

Moreover,

$$\alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) = \frac{\alpha}{2}\|A - A\|_F^2 + \frac{1-\alpha}{2}\|A - A\|_F^2 = 0.$$

Since $f\big(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2\big) > \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2)$, the function is not convex.

b. By setting the gradient of the objective function to zero, we get

$$\nabla f(\mathbf{x}) = -2(\mathbf{A} - \mathbf{x}\mathbf{x}^T)\mathbf{x} = \mathbf{0}.$$

The set of stationary points:

$$\{\mathbf{x} \,|\, \mathbf{x} = \mathbf{0} \text{ or } \mathbf{A} = \mathbf{x}\mathbf{x}^T\}.$$

c. The optimal objective value is 0 which is achieved at any stationary point that satisfies $\mathbf{x} = \pm\mathbf{x}_0$. We still need to show that $\mathbf{x} = \mathbf{0}$ is not a local minimum.
Computing the hessian of the objective function, we get

$$\nabla^2 f(\mathbf{x}) = -2\mathbf{A} + 4\mathbf{x}\mathbf{x}^T + 2\|\mathbf{x}\|_2^2 \mathbf{I}_n.$$

Consider a point $\mathbf{x} = \mathbf{0}$, and the direction $\mathbf{v} = \mathbf{x}_0$ we obtain

$$\mathbf{v}^T \nabla^2 f(\mathbf{0})\mathbf{v} = -2\mathbf{x}_0\mathbf{A}\mathbf{x}_0 = -2\|\mathbf{x}_0\|_2^2 < 0.$$

Thus $(\mathbf{0}, \mathbf{0})$ is not a local minimum.

We have showed that $\mathbf{0}$ is not a local minimum. Moreover, we should that the stationary points $\pm\mathbf{x}_0$ are local and global. Thus every local minimum of the problem is global.