APPM 5360, Spring 2023 - Written Homework 3 and 4

Eappen Nelluvelil; Collaborators: Bisman, Kalyan, Tyler, Jack, Logan

February 17, 2023

1. Find a 2D function $f(x, y)$ such that $x \mapsto f(x, y)$ is convex for every y, and $y \mapsto f(x, y)$ is convex for every x, but f is not a convex function (that is, it is not jointly convex in (x, y)).

   Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = xy$. For any fixed $x_0 \in \mathbb{R}$, $f(x_0, y) = x_0 y$ is convex as it is linear, and similarly, for any fixed $y_0 \in \mathbb{R}$, $f(x, y_0) = y_0 x$ is also convex.

   However, f is not jointly convex in x and y. Consider the points $(-1, 1)$ and $(1, -1)$. We have that

   $$f\left(\frac{1}{2}(-1, 1) + \frac{1}{2}(1, -1)\right) = f(0, 0) = 0,$$

   but we have that

   $$\begin{aligned}\frac{1}{2} f(-1, 1) + \frac{1}{2} f(1, -1) &= \frac{1}{2}(-1)(1) + \frac{1}{2}(1)(-1) \\ &= -1,\end{aligned}$$

   which shows that f is not convex.

2. Problem 3.14: We say the function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is *convex-concave* if $f(x, z)$ is a concave function of z, for each fixed x, and a convex function of x, for each fixed z. We also require its domain to have the product form $\mathrm{dom}(f) = A \times B$, where $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}^m$ are convex.

   (a) Give a second-order condition for a twice-differentiable function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ to be convex-concave, in terms of its Hessian $\nabla^2 f(x, z)$.

   A second-order condition for a twice-differentiable function to be convex-concave is that for every fixed z, $\nabla^2 f(x, z)$ is symmetric, positive semi-definite, and for every fixed x, $\nabla^2 f(x, z)$ is symmetric, negative-definite.

   (b) Suppose that $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is convex-concave and differentiable, with $\nabla f(\widetilde{x}, \widetilde{z}) = 0$. Show that the saddle-point property holds: for all x, z, we have

   $$f(\widetilde{x}, z) \leq f(\widetilde{x}, \widetilde{z}) \leq f(x, \widetilde{z}).$$

   Show that this implies that f satisfies the *strong max-min property*:

   $$\sup_z \inf_x f(x, z) = \inf_x \sup_z f(x, z).$$

   (and their common value is $f(\widetilde{x}, \widetilde{z})$).

   We have that $\nabla f(\widetilde{x}, \widetilde{z}) = 0$. Since f is concave for every fixed x, which, in this case, we take our fixed x to be $\widetilde{x}$, we have that $f(\widetilde{x}, z) \leq f(\widetilde{x}, \widetilde{z})$. This can be seen by applying the first-order concavity condition to f at $(\widetilde{x}, \widetilde{z})$, i.e.,

   $$\begin{aligned}f(\widetilde{x}, z) &\leq f(\widetilde{x}, \widetilde{z}) + \nabla f(\widetilde{x}, \widetilde{z})^T ((\widetilde{x}, z) - (\widetilde{x}, \widetilde{z})) \\ &\leq f(\widetilde{x}, \widetilde{z}).\end{aligned}$$

Furthermore, since f is convex for every fixed z, which, in this case, we take our fixed z to be $\widetilde{z}$, we have that $f\left(\widetilde{x}, \widetilde{z}\right) \leq f\left(x, \widetilde{z}\right)$. This can be seen by applying the first-order convexity condition to f at $\left(\widetilde{x}, \widetilde{z}\right)$, i.e.,

$$\begin{aligned} f\left(x, \widetilde{z}\right) &\geq f\left(\widetilde{x}, \widetilde{z}\right) + \nabla f\left(\widetilde{x}, \widetilde{z}\right)\left(\left(x, \widetilde{z}\right) - \left(\widetilde{x}, \widetilde{z}\right)\right) \\ &= f\left(\widetilde{x}, \widetilde{z}\right). \end{aligned}$$

We have that $f\left(\widetilde{x}, z\right) \leq f\left(\widetilde{x}, \widetilde{z}\right)$ for any z. Then taking the supremum over all z, we have that $\sup_z f\left(\widetilde{x}, z\right) = f\left(\widetilde{x}, \widetilde{z}\right)$.

Similarly, we have that $f\left(\widetilde{x}, \widetilde{z}\right) \leq f\left(x, \widetilde{z}\right)$ for any x. Then taking the infimum over all x, we have that $f\left(\widetilde{x}, \widetilde{z}\right) \leq \inf_x \left(x, \widetilde{z}\right)$.

Thus, we have that $\sup_z f\left(\widetilde{x}, z\right) = \inf_x f\left(x, \widetilde{z}\right)$. From this, it follows that

$$\inf_x \sup_z f\left(x, z\right) \leq \sup_z \inf_x f\left(x, z\right).$$

To show the inequality the other way, we use the minimax inequality to get that

$$\sup_z \inf_x f\left(x, z\right) \leq \inf_x \sup_z f\left(x, z\right).$$

Thus, we have that

$$\sup_z \inf_x f\left(x, z\right) = \inf_x \sup_z f\left(x, z\right).$$

(c) Now suppose that $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is differentiable, but not necessarily convex-concave, and the saddle-point property holds at $\widetilde{x}, \widetilde{z}$:

$$f\left(\widetilde{x}, z\right) \leq f\left(\widetilde{x}, \widetilde{z}\right) \leq f\left(x, \widetilde{z}\right).$$

for all x, z. Show that $\nabla f\left(\widetilde{x}, \widetilde{z}\right) = 0$.

We have that $f\left(\widetilde{x}, z\right) \leq f\left(\widetilde{x}, \widetilde{z}\right)$ for any z, which implies that $\frac{\partial f}{\partial z}_{\left(\widetilde{x}, z\right)} = 0$ for all z. Furthermore, we also have that $f\left(\widetilde{x}, \widetilde{z}\right) \leq f\left(x, \widetilde{z}\right)$ for any x, which implies that $\frac{\partial f}{\partial x}_{\left(x, \widetilde{z}\right)} = 0$ for all x.

Then, evaluating the two partials at $\left(\widetilde{x}, \widetilde{z}\right)$, we have that $\nabla f\left(\widetilde{x}, \widetilde{z}\right) = \langle 0, 0 \rangle$, as desired.

3. Problem 3.16 (d): let $f\left(x_1, x_2\right) = \frac{x_1}{x_2}$ on $\mathbb{R}^2_{++}$. Determine if this function is convex, concave, quasiconvex and/or quasiconcave. Justify your answers.

   The Hessian of $f\left(x, y\right)$ is given by $\nabla^2 f\left(x, y\right) = \begin{bmatrix} 0 & -\frac{1}{y^2} \\ -\frac{1}{y^2} & \frac{2x}{y^3} \end{bmatrix}$. Taking the determinant of this Hessian, we have that the determinant is $-\frac{1}{y^4}$, which means that the matrix is indefinite, i.e., the two eigenvalues differ in sign. Thus, f is neither convex nor concave.

   If we look at the level sets of f, i.e., given an $\alpha \in \mathbb{R}^{++}$, we have that $f\left(x, y\right) = \alpha \implies \frac{x}{y} = \alpha \implies x = \alpha y$, which is a convex set. Furthermore, the domain of f is also convex, which implies that f is quasilinear, i.e., f is both quasiconvex and quasiconcave.

4. Problem 3.18 (a): prove that $f\left(X\right) = \text{trace}\left(X^{-1}\right)$ is convex on $\text{dom}\left(f\right) = S^n_{++}$.

   We will verify concavity by considering an arbitrary line, given by $Z + tV$, where $Z \in S^n_{++}$, $V \in S^n$.

We define $g(t) = f(Z + tV)$, where t is restricted so that $(Z + tV) \in S^n_{++}$. We have that

$$\begin{aligned}
g(t) &= \operatorname{tr}\left((Z+tV)^{-1}\right) \\
&= \operatorname{tr}\left(\left(Z^{\frac{1}{2}}Z^{\frac{1}{2}} + tZ^{-\frac{1}{2}}Z^{\frac{1}{2}}VZ^{-\frac{1}{2}}Z^{\frac{1}{2}}\right)^{-1}\right) \\
&= \operatorname{tr}\left(Z^{-1}\left(I + tZ^{-\frac{1}{2}}VZ^{-\frac{1}{2}}\right)^{-1}\right) \\
&= \operatorname{tr}\left(Z^{-1}\left(I + PDP^{-1}\right)^{-1}\right) \quad \text{where we made use of the fact that } Z^{-\frac{1}{2}}VZ^{-\frac{1}{2}} \text{ is diagonalizable} \\
&= \operatorname{tr}\left(Z^{-1}\left(PP^{-1} + tPDP^{-1}\right)^{-1}\right) \\
&= \operatorname{tr}\left(Z^{-1}\left(P(I+tD)P^{-1}\right)^{-1}\right) \\
&= \operatorname{tr}\left(Z^{-1}(I+tD)^{-1}\right) \\
&= \sum_{i=1}^{n} \frac{\left(Z^{-1}\right)_{ii}}{1+t\lambda_i}.
\end{aligned}$$

We take the first and second derivatives of g with respect to t:

$$\begin{aligned}
g'(t) &= \sum_{i=1}^{n} -\frac{\left(Z^{-1}\right)_{ii}\lambda_i}{(1+t\lambda_i)^2} \\
g''(t) &= \sum_{i=1}^{n} \frac{2\left(Z^{-1}\right)_{ii}\lambda_i^2}{(1+t\lambda_i)^3}.
\end{aligned}$$

We have that $g''(t) \geq 0$, which implies that f is convex.

5. Problem 3.36 (a): Derive the conjugate of $f(x) = \max_{i=1,\ldots,n} x_i$ on $\mathbb{R}^n$.

   The conjugate of this function is

   $$f^*(y) = \sup_{x\in\mathbb{R}^n} \left(y^T x - f(x)\right).$$

   We consider two cases for $y \in \mathbb{R}^n$:

   (a) $y \geq 0$ Here, we consider three subcases:

      i. $\sum_{i=1}^n y_i < 1$
      In this case, we take $x \in \mathbb{R}^n$ to be $x = -t \cdot 1$, where $t > 0$. Then, $(y) = y^T x - \max_i x_i = -t\sum_{i=1}^n y_i + t$, which goes to ∞ as $t \to \infty$ since $\sum_{i=1}^n y_i < 1$.

      ii. $\sum_{i=1}^n y_i > 1$
      In this case, we take $x \in \mathbb{R}^n$ to be $x = t \cdot 1$, where $t > 0$. Then, $(y) = y^T x - \max_i x_i = t\sum_{i=1}^n y_i - t$, which goes to ∞ as $t \to \infty$ since $\sum_{i=1}^n y_i > 1$.

      iii. $\sum_{i=1}^n y_i = 1$
      In this case, we have that

      $$\begin{aligned}
      \left(y^T x - \max_i x_i\right) &= \sum_{i=1}^n y_i x_i - \max_i x_i \\
      &\leq x_k - x_k \quad (\text{where } x_k = \max_i x_i) \\
      &= 0.
      \end{aligned}$$

      Since we are taking the supremum over all $x \in \mathbb{R}^n$, we have that $f^*(y) = 0$.

(b) $y_i < 0$ for some $i \in [n]$

If $y_i < 0$, we take $x \in \mathbb{R}^n$ such that $x_i = -t$ and $x_k = 0$ for all $k \neq i$, where $t > 0$. Then, we have that $y^T x - \max_i x_i = y_i x_i - 0$, and since we are taking the supremum over all $x \in \mathbb{R}^n$, we simply let $t \to \infty$ to get that $f^*(y) = \infty$.

Thus, we have that

$$f^*(y) = \begin{cases} 0 & \text{if } y \geq 0, \sum_{i=1}^n y_i = 1, \\ \infty & \text{if } y \geq 0, \sum_{i=1}^n y_i < 1, \\ \infty & \text{if } y \geq 0, \sum_{i=1}^n y_i > 1, \\ \infty & \text{if } y_i < 0 \text{ for some } i \in [n]. \end{cases}$$