HOMEWORK # 3
Junyi Ji
February 22, 2023

I finished this assignment in 6 hours, haha. So proud of myself. Trying to survive the projects and classes.

1. Determine whether or not the following functions are convex in their given domains:

   (a) $f(x) = x_1x_2$ in the domain $x \in \mathbb{R}^2$

   (b) $f(x) = e^{x_1+x_2}$ in the domain $x \in \mathbb{R}^2$

   (c) $f(x) = x_1^4 + x_2^4 - x_1^2x_2^2$ in the domain $x \in \mathbb{R}^2$

   (d) $f(x) = x_1^3 + x_2^3$ in the domain $x \in \mathbb{R}^2_{++}$

   (e) $f(x) = \tan x$ for the domain $x \in (0, 1)$ (boundary points not included)

   (f) $f(x) = \frac{1}{x_1x_2}$ in the domain $x \in \mathbb{R}^2_{++}$.

   (g) $f(x) = \frac{x_1}{x_2}$ in the domain $x \in \mathbb{R}^2_{++}$.

   (h) $f(x) = x_1^{\alpha}x_2^{1-\alpha}$, where $0 \le \alpha \le 1$ and in the domain $x \in \mathbb{R}^2_{++}$.

   Answer:

   (a) Convex. Use the 2-order condition. The Hessian matrix is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \succ 0$.

   (b) Convex. Use the 2-order condition. The Hessian matrix is $\begin{bmatrix} e^{x_1+x_2} & e^{x_1+x_2} \\ e^{x_1+x_2} & e^{x_1+x_2} \end{bmatrix} \succeq 0$.

   (c) Convex. It is equivalent to the function $f(y) = y_1^2 + y_2^2 - y_1y_2$ in the domain $y \in \mathbb{R}_+$. The Hessian matrix for $f(y)$ is $\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \succ 0$.

   (d) Convex. The Hessian matrix is $\begin{bmatrix} 6x_1x_2 & 0 \\ 0 & 6x_1x_2 \end{bmatrix} \succ 0, x \in \mathbb{R}^2_{++}$.

   (e) Convex. The Hessian matrix is $2\frac{sinx}{cos^3x} > 0, x \in (0, 1)$.

   (f) Convex. The Hessian matrix is $\begin{bmatrix} \frac{2}{x_1^3x_2} & \frac{1}{x_1^2x_2^2} \\ \frac{1}{x_1^2x_2^2} & \frac{2}{x_1x_2^3} \end{bmatrix} \succ 0, x \in \mathbb{R}_{++}$, since $det(\begin{bmatrix} \frac{2}{x_1^3x_2} & \frac{1}{x_1^2x_2^2} \\ \frac{1}{x_1^2x_2^2} & \frac{2}{x_1x_2^3} \end{bmatrix}) = \frac{3}{x_1^4x_2^4} > 0, x \in \mathbb{R}^2_{++}$.

   (g) Nonconvex. The Hessian matrix is $\begin{bmatrix} 0 & -\frac{1}{x_2^2} \\ -\frac{1}{x_2^2} & \frac{2x_1}{x_2^2} \end{bmatrix}$, which is not positive semidefinite for $x \in \mathbb{R}^2_{++}$.

   (h) Concave. The Hessian matrix is $\begin{bmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & \alpha(\alpha-1)x_1^{\alpha}x_2^{-\alpha-1} \end{bmatrix} \preceq 0, x \in \mathbb{R}^2_{++}, \alpha \in [0, 1]$.

2. Via the definition of a convex function (Jensen's inequality)

$$f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta) f(x_2)$$

for $\theta \in [0,1]$, prove that the quadratic function

$$f(x) = \frac{1}{2} x^\top Q x - x^\top b + c$$

for all $x \in \mathbb{R}^n$ is convex if the symmetric matrix Q is positive semidefinite.

Proof:

To prove the quadratic function is convex, just need to prove that if the matrix $\boldsymbol{Q} \succeq 0$, $F(x_1, x_2) = f(\theta x_1 + (1-\theta)x_2) - \theta f(x_1) - (1-\theta) f(x_2) \leq 0$. Here, $x_1, x_2 \in \mathbb{R}^n$, $\theta \in [0,1]$.

$$\begin{aligned} 2F(x_1, x_2) &= (\theta x_1 + (1-\theta)x_2)^\top \boldsymbol{Q} (\theta x_1 + (1-\theta)x_2) - \theta x_1^\top \boldsymbol{Q} x_1 - (1-\theta) x_2^\top \boldsymbol{Q} x_2 && (1) \\ &= (\theta^2 - \theta) x_1^\top \boldsymbol{Q} x_1 + (\theta^2 - \theta) x_2^\top \boldsymbol{Q} x_2 - 2(\theta^2 - \theta) x_1^\top \boldsymbol{Q} x_2 && (2) \\ &= (\theta^2 - \theta)(x_1 - x_2)^\top \boldsymbol{Q} (x_1 - x_2) && (3) \end{aligned}$$

Note:

- It is obvious that the term of $x^\top b + c$ will be eliminated when conducting the minus;
- $x_1^\top \boldsymbol{Q} x_2 = x_2^\top \boldsymbol{Q} x_1$;
- $(\theta^2 - \theta) \leq 0, \theta \in [0,1]$;
- $x^\top \boldsymbol{Q} x \geq 0$, if $\boldsymbol{Q} \succeq 0$;

Hence, $F(x_1, x_2) \leq 0$. The quadratic function is convex in the defined domain.

---

3. Is the product of two convex functions convex? If yes, prove it; if not, give a counterexample.

Answer: No. A simple example is that $f(x) = x^2$, $g(x) = \log x$, $f(x)$ and $g(x)$ are all convex functions. However, $f(x)g(x) = x^2 \log x$ is nonconvex.

---

4. A given function $f : \mathbb{R}^n \to R$ is assumed to be continuously differentiable. Also, assume that $f(x)$ **is concave on a convex set** $\mathcal{X}$. Given the aforementioned properties of $f(x)$, prove that for all $x_1, x_2 \in \mathcal{X}$, $f(x)$ satisfies this property:

$$f(x_2) \leq f(x_1) + Df(x_1)(x_2 - x_1).$$

Hint: *Back to basics—what is the basic definition of a derivative?*

Proof:

According to the definition of the concave function, for any $x, y \in \mathcal{X}$, and any $\theta \in [0, 1]$:

$$f((1-\theta)x + \theta y) \geq (1-\theta)f(x) + \theta f(y) \tag{4}$$

We could have the following inequalities.

$$f(x + \theta(y - x)) \geq f(x) + \theta(f(y) - f(x)) \tag{5}$$

$$\frac{f(x + \theta(y - x)) - f(x)}{\theta} \geq (f(y) - f(x)) \tag{6}$$

$$\frac{f(x + \theta(y - x)) - f(x)}{\theta(y - x)} \geq \frac{(f(y) - f(x))}{y - x} \tag{7}$$

$$\lim_{\theta \to 0} \frac{f(x + \theta(y - x)) - f(x)}{\theta(y - x)} = Df(x) \geq \lim_{\theta \to 0} \frac{(f(y) - f(x))}{y - x} = \frac{(f(y) - f(x))}{y - x} \tag{8}$$

Then, the inequality is derived: $Df(x) \geq \frac{(f(y)-f(x))}{y-x}$. To replace x with x_1 and y with x_2:

$$f(x_2) \leq f(x_1) + Df(x_1)(x_2 - x_1) \tag{9}$$

---

5. Show that the set Ω given by $\Omega = \{y \in \mathbb{R}^2; \|y\|^2 \leq 4\}$ is convex, where $\|y\|^2 = y^\top y$.

Proof:

For any $y_1, y_2 \in \Omega$:

$$y_1^\top y_1 \leq 4 \tag{10}$$
$$y_2^\top y_2 \leq 4 \tag{11}$$

$$
\begin{aligned}
||\theta y_1 + (1-\theta)y_2|| &= (\theta y_1 + (1-\theta)y_2)^\top (\theta y_1 + (1-\theta)y_2) && (12)\\
&= \theta^2 y_1^\top y_1 + (1-\theta)^2 y_2^\top y_2 + \theta(1-\theta)(y_1^\top y_2 + y_2^\top y_1) && (13)\\
&\leq \theta^2 y_1^\top y_1 + (1-\theta)^2 y_2^\top y_2 + \theta(1-\theta)(y_1^\top y_1 + y_2^\top y_2) && (14)\\
&\leq 4(\theta^2 + (1-\theta)^2 + 2\theta(1-\theta)) = 4 && (15)
\end{aligned}
$$

Hence, the set Ω given by $\Omega = \{y \in \mathbb{R}^2; \|y\|^2 \leq 4\}$ is convex.

---

6. Given a multivariable function $f(x)$, many optimization solvers use the following algorithm to solve $\min_x f(x)$:

   (a) Choose an initial guess, $x^{(0)}$

   (b) Choose an initial real, symmetric positive definite matrix $H^{(0)}$

   (c) Compute $d^{(k)} = -H^{(k)}\nabla_x f(x^{(k)})$

   (d) Find $\beta^{(k)} = \arg\min_\beta f(x^{(k)} + \beta^{(k)}d^{(k)}),\ \beta \geq 0$

   (e) Compute $x^{(k+1)} = x^{(k)} + \beta^{(k)}d^{(k)}$

   For this problem, we assume that the given function is a typical quadratic function ($x \in \mathbb{R}^n$), as follows:

   $$f(x) = \frac{1}{2}x^\top Qx - x^\top b + c, \quad Q = Q^\top \succ 0.$$

   Answer the following questions:

   (a) Find $f(x^{(k)} + \beta^{(k)}d^{(k)})$ for the given quadratic function.

   (b) Obtain $\nabla_x f(x^{(k)})$ for $f(x)$.

   (c) Using the chain rule, and given that $\beta^{(k)} = \arg\min_\beta f(x^{(k)} + \beta^{(k)}d^{(k)})$, find a closed form solution for $\beta^{(k)}$ in terms of the given matrices ($H^{(k)}, \nabla f(x^{(k)}), d^{(k)}, Q$).

   (d) Since it is required that $\beta^{(k)} \geq 0$, provide a sufficient condition related to $H^{(k)}$ that guarantees the aforementioned condition on $\beta^{(k)}$.

   Answer:

   (a) $f(x^{(k)} + \beta^{(k)}d^{(k)})$:

   $$f(x^{(k)} + \beta^{(k)}d^{(k)}) = \frac{1}{2}(x^{(k)} + \beta^{(k)}d^{(k)})^\top Q(x^{(k)} + \beta^{(k)}d^{(k)}) - (x^{(k)} + \beta^{(k)}d^{(k)})^\top b + c \tag{16}$$

   (b) $\nabla_x f(x^{(k)})$:

   $$\nabla_x f(x^{(k)}) = Qx^{(k)} - b \tag{17}$$

   (c) See the proof below:
   To minimize $f(x^{(k)} + \beta^{(k)}d^{(k)})$, the derivative of $f(x^{(k)} + \beta^{(k)}d^{(k)})$ equals to 0:

   $$\nabla_{\beta^{(k)}} f(x^{(k)} + \beta^{(k)}d^{(k)}) = d^{(k)}(Qx^{(k)} + Q\beta^{(k)}d^{(k)} - b) = 0 \tag{18}$$

   Hence,

   $$\beta^{(k)}QH^{(k)}\nabla_x f(x^{(k)}) = Qx^{(k)} - b = \nabla_x f(x^{(k)}) \tag{19}$$

   Hence,

   $$\beta^{(k)} = {H^{(k)}}^{-1}Q^{-1} \tag{20}$$

   (d) If $H^{(k)} \succeq 0$, it is clear that $\beta^{(k)} \geq 0$ since $Q \succ 0$.

7. For the following function, find the set of values for β such that the function is convex.

$$f(x,y,z) = x^2 + y^2 + 5z^2 - 2xz + 2\beta xy + 4yz$$

Answer: The Hessian matrix is:

$$H = \begin{bmatrix} 1 & \beta & -1 \\ \beta & 1 & 2 \\ -1 & 2 & 5 \end{bmatrix} \succeq 0 \tag{21}$$

For the leading principle minors, $1, 1, 5 \geq 0$.
$det(\begin{bmatrix} 1 & \beta \\ \beta & 1 \end{bmatrix}) = 1 - \beta^2 \geq 0, det(\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}) = 1 \geq 0, det(\begin{bmatrix} 1 & 1 \\ 1 & 5 \end{bmatrix}) = 4 \geq 0.$
$det(\begin{bmatrix} 1 & \beta & -1 \\ \beta & 1 & 2 \\ -1 & 2 & 5 \end{bmatrix}) = -5\beta^2 - 4\beta \geq 0.$
Hence, $\beta \in [-0.8, 0]$.

---

8. Prove that the following set given by

$$\Psi = \{x : x^\top Px \leq 1\}$$

is in fact a convex one given that the matrix P is a symmetric positive definite matrix.
Proof:
For any $x_1, x_2 \in \Psi$:

$$x_1^\top Px_1 \leq 1 \tag{22}$$
$$x_2^\top Px_2 \leq 1 \tag{23}$$

Lemma 1. Since P is a symmetric positive definite matrix, we have:

$$(x_2 - x_1)^\top P(x_2 - x_1) \geq 0 \tag{24}$$
$$x_2^T Px_2 + x_1^T Px_1 - (x_2^T Px_1 + x_1^T Px_2) \geq 0 \tag{25}$$
$$x_2^T Px_1 + x_1^T Px_2x \leq x_2^T Px_2 + x_1^T Px_1 \tag{26}$$

For any $x_1, x_2 \in \Psi$ and $\theta \in [0, 1]$, using **Lemma 1** in Equation 27.

$$C(\theta) = (\theta x_1 + (1-\theta)x_2)^\top Q(\theta x_1 + (1-\theta)x_2) \tag{27}$$
$$= \theta^2 x_1^\top Qx_1 + (1-\theta)^2 x_2^\top Qx_2 + \theta(1-\theta)(x_1^\top Qx_2 + x_2^\top Qx_1) \tag{28}$$
$$\leq \theta^2 x_1^\top Qx_1 + (1-\theta)^2 x_2^\top Qx_2 + \theta(1-\theta)(x_1^\top Qx_1 + x_2^\top Qx_2) \tag{29}$$
$$\leq 1 \tag{30}$$

Hence, the set Ψ is in fact a convex one given that the matrix P is a symmetric positive definite matrix.

---

9. Now given the previous problem, prove that

$$g(x) = x_1^2 + 2x_1x_2 + 3x_2^2 - 5x_1 + 6x_2 + 10$$

is a convex function by writing $g(x) = x^\top Px + b^\top x + c$ and then using the argument that the sum of convex function is a convex function.

Proof:

$$g(x) = x^\top \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} x + \begin{bmatrix} -5 & 6 \end{bmatrix}^\top x + 10 \tag{31}$$

It is convex since the matrix $P \succ 0$.

$$g(x) = m(x) + n(x) + l(x) + k(x) \tag{32}$$

$$\begin{cases} m(x) = x_1^2 \\ n(x) = 3x_2^2 \\ l(x) = 2x_1x_2 \\ n(x) = -5x_1 + 6x_2 + 10 \end{cases} \tag{33}$$

It is clear that $m(x), n(x), l(x), k(x)$ are all convex functions. Then $g(x)$ is a convex function.

---

10. A family of concave utility functions. For $0 < \alpha \leq 1$ let

$$u_\alpha(x) = \frac{x^\alpha - 1}{\alpha},$$

with dom $u_\alpha = \mathbf{R}_+$. We also define $u_0(x) = \log x$ (with dom $u_0 = \mathbf{R}_{++}$).

(a) Show that for $x > 0$, $u_0(x) = \lim_{\alpha \to 0} u_\alpha(x)$.

(b) Show that u_α are concave, monotone increasing, and all satisfy $u_\alpha(1) = 0$.

These functions are often used in economics to model the benefit or utility of some quantity of goods or money. Concavity of u_α means that the marginal utility (i.e., the increase in utility obtained for a fixed increase in the goods) decreases as the amount of goods increases. In other words, concavity models the effect of satiation.

Answer:

(a) Using L'Hopital's Law,

$$\lim_{\alpha \to 0} \frac{x^\alpha - 1}{\alpha} = \lim_{\alpha \to 0} \frac{(x^\alpha - 1)'}{\alpha'} = \lim_{\alpha \to 0} \frac{\log x x^\alpha}{1} = \log x = u_0(x) \tag{34}$$

(b) See the proof below.

$$f'(x) = x^{\alpha - 1} \geq 0 \tag{35}$$

$$f''(x) = (\alpha - 1)x^{\alpha - 2} \leq 0 \tag{36}$$

$$u_\alpha(1) = 0 \tag{37}$$

Hence, u_α are concave, monotone increasing, and all satisfy $u_\alpha(1) = 0$.

---

11. Adapt the proof of concavity of the log-determinant function in §3.1.5 to show the following.

   (a) $f(X) = \mathbf{tr}\left(X^{-1}\right)$ is convex on $\mathbf{dom}\, f = \mathbf{S}^n_{++}$.

   (b) $f(X) = (\det X)^{1/n}$ is concave on $\mathbf{dom}\, f = \mathbf{S}^n_{++}$.

   Proof:

   (a) See the proof below.

$$
\begin{aligned}
g(t) &= \mathbf{tr}\left((Z+tV)^{-1}\right) && (38)\\
&= \mathbf{tr}\left(Z^{-1}\left(I+tZ^{-1/2}VZ^{-1/2}\right)^{-1}\right) && (39)\\
&= \mathbf{tr}\left(Z^{-1}Q(I+t\Lambda)^{-1}Q^T\right) && (40)\\
&= \mathbf{tr}\left(Q^TZ^{-1}Q(I+t\Lambda)^{-1}\right) && (41)\\
&= \sum_{i=1}^{n}\left(Q^TZ^{-1}Q\right)_{ii}(1+t\lambda_i)^{-1} && (42)
\end{aligned}
$$

   Since $1+t\lambda_1$ is convex, $f(X)$ is convex.

   (b) See the proof below.

$$
\begin{aligned}
g(t) &= (\det(Z+tV))^{1/n} && (43)\\
&= \left(\det Z^{1/2}\det\left(I+tZ^{-1/2}VZ^{-1/2}\right)\det Z^{1/2}\right)^{1/n} && (44)\\
&= (\det Z)^{1/n}\left(\prod_{i=1}^{n}(1+t\lambda_i)\right)^{1/n} && (45)
\end{aligned}
$$

   Since $(\prod_{i=1}^{n} x_i)^{1/n}$ is concave, $f(X)$ is concave.
