SI151A
Convex Optimization and its Applications in Information Science, Fall 2021
Homework 2

Due on Oct 18, 2021, 23:59 UTC+8

1. Show that the following functions from $\mathbb{R}^n$ to $(-\infty, \infty]$ are convex:

(1) $f_1(x) = \ln(e^{x_1} + \cdots + e^{x_n})$ (10 points)

(2) $f_2(x) = \frac{1}{f(x)}$, where f is concave and $f(x)$ is a positive number for all x.(10 points)

(3) $f_3(x) = e^{\beta x^\top Ax}$, where A is a positive semidefinite symmetric $n \times n$ matrix and β is a positive scalar.(10 points)

Solution:

(1) We show that the Hessian of f_1 is positive semidefinite at all $x \in \mathbb{R}$. Let $\beta(x) = e^{x_1} + \cdots + e^{x_n}$. Then a straightforward calculation yields

$$z^\top \nabla^2 f_1(x) z = \frac{1}{\beta(x)^2} \sum_{i=1}^{n} \sum_{j=1}^{n} e^{x_i + x_j}(z_i - z_j)^2 \geq 0, \forall z \in \mathbb{R}^n.$$

(2) The function $f_2(x) = \frac{1}{f(x)}$ can be viewed as a composition $g(h(x))$ of the function $g(t) = -\frac{1}{t}$ for $t < 0$ and the function $h(x) = -f(x)$ for $x \in \mathbb{R}^n$. In this case, the g is convex and monotonically increasing in the set $\{t | t < 0\}$, while h is convex over $\mathbb{R}^n$. By the composition rule, it follows that the function $f_2(x) = \frac{1}{f(x)}$ is convex over $\mathbb{R}^n$.

(3) The function $f_3(x) = e^{\beta x^\top Ax}$ can be viewed as a composition $g(f(x))$ of the function $g(t) = e^{\beta t}$ for $t \in \mathbb{R}$ and the function $f(x) = x^\top Ax$ for $x \in \mathbb{R}^n$. In this case, g is convex and monotonically increasing over $\mathbb{R}$, while f is convex over $\mathbb{R}^n$ (since A is positive semidefintie). By the composition rule, it follows that the function $f_2(x) = e^{\beta x^\top Ax}$ is convex over $\mathbb{R}^n$.

2.

(1) Prove that the *entropy function*, defined as

$$f(x) = -\sum_{i=1}^{n} x_i \log(x_i)$$

with $\text{dom}(f) = \{x \in \mathbb{R}^n_{++} : \sum_{i=1}^n x_i = 1\}$, is strictly concave. (10 points)

(2) Let f be twice differentiable, with $\text{dom}(f)$ convex. Prove that f is convex if and only if

$$(\nabla f(x) - \nabla f(y))^\top (x - y) \geq 0,$$

for all x, y. (This property is called *monotonicity* of the gradient ∇f.)(10 points)

Solution:

(1) A straightforward calculation yields

$$\nabla^2 f = \begin{bmatrix} -\frac{1}{x_1} & 0 & \cdots & 0 \\ 0 & -\frac{1}{x_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\frac{1}{x_n} \end{bmatrix}$$

which is negative definite, thus f is strictly concave.

(2) "if":
Consider the function $g(t) = f(x + t(y - x))$ with $g'(t) = \nabla f(x + t(y - x))^\top (y - x)$.
If the monotonicity of the gradient ∇f holds, then

$$(\nabla f(x + t(y - x)) - \nabla f(x))^\top \cdot t(y - x) \geq 0.$$

If $t \geq 0$, we can further obtain

$$\begin{aligned} &(\nabla f(x + t(y - x)) - \nabla f(x))^\top \cdot (y - x) \geq 0. \\ \Longleftrightarrow &\nabla f(x + t(y - x))^\top (y - x) \geq \nabla f(x)^\top (y - x) \\ \Longleftrightarrow &g'(t) \geq g'(0). \end{aligned}$$

Therefore, by the fundamental theorem of calculus we have

$$f(y) = g(1) = g(0) + \int_0^1 g'(t)dt \geq g(0) + g'(0) = f(x) + \nabla f(x)^\top (y - x),$$

which is the first-order condition for convexity.

"only if":
If f is convex, by the first-order condition for convexity we have:

$$\begin{aligned} f(y) &\geq f(x) + \nabla f(x)^\top (y - x), \\ f(x) &\geq f(y) + \nabla f(y)^\top (x - y). \end{aligned}$$

Taking sum of the above two inequalities gives $(\nabla f(x) - \nabla f(y))^\top (x - y) \geq 0$.

3. The function $f(x, t) = -\log(t^2 - x^\top x)$, with $\textbf{dom}\ f = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : t > \|x\|_2\}$(i.e., the second-order cone), is convex. (The function f is called the logarithmic barrier function for the second-order cone.) This can be shown in many ways, for example by evaluating the Hessian and demonstrating that it is positive semidefinite. In this exercise you establish convexity of f using a relatively painless method, leveraging some composition rules and known convexity of a few other functions.

(1) Explain why $t - \left(\frac{1}{t}\right) u^\top u$ is a concave function on $\textbf{dom}\ f$. Hint: Use convexity of quadratic over linear function. (5 points)

(2) From this, show that $-\log\left(t - \left(\frac{1}{t}\right) u^\top u\right)$ is a convex function on $\textbf{dom}\ f$. (5 points)

(3) From this, show that f is convex. (5 points)

Solution:

(1) $\left(\frac{1}{t}\right) u^\top u$ is the quadratic over linear function, which is convex on textbfdom f. So $t - \left(\frac{1}{t}\right) u^\top u$ is concave, since it is a linear function minus a convex function.

(2) The function $g(u) = -\log u$ is convex and decreasing, so if u is a concave (positive) function, the composition rules tell us that $g \circ u$ is convex. Here this means $-\log\left(t - \left(\frac{1}{t}\right) u^\top u\right)$ is a convex function on $\textbf{dom}\ f$.

(3) We can write $f(x, t) = -\log(t - \left(\frac{1}{t}\right) x^\top x) - \log t$, which shows that f is a sum of two convex functions, hence convex.

4. Suppose that $g(x)$ is convex and $h(x)$ is concave. Suppose we restrict both functions into a closed, convex set C such that both $g(x)$ and $h(x)$ are always positive when $x \in C$. Prove that the function $f(x) = \frac{g(x)}{h(x)}$ is quasi-convex. (15 points)

Solution:

Choose some $\alpha \geq 0$ and consider the sub-level set $S_\alpha = \{x | f(x) \leq \alpha\}$. If $x \in S_\alpha$, then

$$\frac{g(x)}{h(x)} \leq \alpha \Longrightarrow g(x) \leq \alpha h(x) \Longrightarrow g(x) - \alpha h(x) \leq 0.$$

But $-h(x)$ is convex and $\alpha \geq 0$, and so $g(x) - \alpha h(x)$ is convex. Because the set S_α is a sub-level set of the convex function $g(x) - \alpha h(x)$, it is convex.

5. Each $X \in S^n_{++}$ has a unique *Cholesky factorization* $X = LL^\top$, where L is lower triangular, with $L_{ii} > 0$. Show that L_{ii} is a concave function of X (with domain S^n_{++}). (20 points)
Hint: L_{ii} can be expressed as $L_{ii} = (\omega - z^\top Y^{-1} z)^{\frac{1}{2}}$, where

$$\begin{bmatrix} Y & z \\ z^\top & \omega \end{bmatrix}$$

is the leading $i \times i$ submatrix of X.

Solution:

The function $f(z, Y) = z^\top Y^{-1} z$ with **dom** $f = \{(z, Y) | Y \succ 0\}$ is convex jointly in z and Y. To see this note that

$$(z, Y, t) \in \mathbf{epi} f \Longleftrightarrow Y \succ 0, \begin{bmatrix} Y & z \\ z^\top & t \end{bmatrix} \geq 0,$$

so **epi** f is a convex set. Therefore, $\omega - z^\top Y^{-1} z$ is a concave function of X. Since the squareroot is an increasing concave function, it follows from the composition rules that $L_{ii} = (\omega - z^\top Y^{-1} z)^{\frac{1}{2}}$ is a concave function of X.