Department of Mathematics, University of Utah

Introduction to Optimization

MATH 5770/6640, ME EN 6025 – Section 001 – Fall 2021

Homework 5 Solutions

Convex functions

Due November 16, 2021

Submit your homework assignment as a scanned copy **ON CANVAS**, to the `Homework 5` assignment.

Some of the exercises below are computational. The book problems are explained in Matlab. You need not use Matlab to complete the assignment; numerical simulation with any programming language is acceptable.

Text: *Introduction to Nonlinear Optimization*, Amir Beck,

Exercises:	#	7.1,
		7.3,
		7.7,
		7.10 (i), (ii), (iv), (v)
		7.25
Extra		P1,
		P2

7.1. For each of the following sets determine whether they are convex or not (explaining your choice).

(i) $C_1 = \left\{ \boldsymbol{x} \in \mathbb{R}^n \;:\; \|\boldsymbol{x}\|^2 = 1 \right\}$
(ii) $C_2 = \{ \boldsymbol{x} \in \mathbb{R}^n \;:\; \max_{i=1,\ldots,n} x_i \leq 1 \}$
(iii) $C_3 = \{ \boldsymbol{x} \in \mathbb{R}^n \;:\; \min_{i=1,\ldots,n} x_i \leq 1 \}$
(iv) $C_4 = \left\{ \boldsymbol{x} \in \mathbb{R}^n_{++} \;:\; \prod_{i=1}^n x_i \geq 1 \right\}$

Solution:

(i) C_1 is not convex. Let $\boldsymbol{x} \in \mathbb{R}^n$ be any vector with unit norm, so that $\boldsymbol{x} \in C_1$. Then $-\boldsymbol{x} \in C_1$ as well, but a convex combination of these two,

$$\frac{1}{2}\boldsymbol{x} + \frac{1}{2}(-\boldsymbol{x}) = \boldsymbol{0},$$

is clearly not in C_1 since it does not have unit norm.

(ii) C_2 is convex. One way to see this is to note that

$$f(\boldsymbol{x}) = \max_{i=1,\ldots,n} x_i,$$

is a convex function since it's the pointwise maximum of convex functions. (The function x_i is affine and hence convex for all i.) Since $C_2 = f^{-1}(-\infty, 1]) = \mathrm{Lev}(f, 1)$ is the level set of a convex function, it is also convex.

(iii) C_3 is not convex for any $n \geq 2$. (It is convex for $n = 1$.) To see this, let $\boldsymbol{x}$ and $\boldsymbol{y}$ be defined as,

$$\boldsymbol{x} = (1, 2, 2, 2, 2, \ldots), \qquad \boldsymbol{y} = (2, 1, 2, 2, 2, \ldots).$$

Note that $\boldsymbol{x}, \boldsymbol{y} \in C_3$. But a convex combination of them,

$$\frac{1}{2}\boldsymbol{x} + \frac{1}{2}\boldsymbol{y} = (1.5, 1.5, 2, 2, 2, \ldots),$$

which is not in C_3 since its minimum value is $1.5 > 1$.

(iv) C_4 is a convex set. Define $h(\boldsymbol{x}) = \sum_{j=1}^{n} - \log x_j$ on $\mathbb{R}^n_{++}$, and note tht h is convex on $\mathbb{R}^n_{++}$ since $-\log t$ is convex for $t > 0$ and h is a conic combination of n such convex functions. Note also that

$$\prod_{j=1}^{n} x_j \geq 1 \quad \Longleftrightarrow \quad \sum_{j=1}^{n}(-\log x_j) \leq 0,$$

so that $C_4 = h^{-1}((-\infty, 0]) = \text{Lev}(h, 0)$. Since C_4 is the level set of a convex function, it is a convex set.

7.3. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex as well as concave function. Show that f is an affine function; that is, there exist $\boldsymbol{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}$ such that $f(\boldsymbol{x}) = \boldsymbol{a}^T\boldsymbol{x} + b$ for any $\boldsymbol{x} \in \mathbb{R}^n$.

Solution: Define $b = f(\boldsymbol{x})$, and define,

$$g(\boldsymbol{x}) = f(\boldsymbol{x}) - b \quad \Longrightarrow \quad g(\boldsymbol{0}) = 0.$$

Since the constant function $-b$ is both convex and concave (for any value of b), then g is also both convex and concave, since it's a conic combination of two convex and concave functions. Since g is both convex and concave, then for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and any $\lambda \in (0, 1)$,

$$g(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) = \lambda g(\boldsymbol{x}) + (1-\lambda)g(\boldsymbol{y}).$$

Taking $\boldsymbol{y} = 0$ and noting that $g(\boldsymbol{0}) = 0$ implies,

$$\frac{1}{\lambda}g(\lambda\boldsymbol{x}) = g(\boldsymbol{x}), \qquad \lambda \in (0, 1). \tag{1}$$

Also, taking $\boldsymbol{x}$ and $\boldsymbol{y}$ as arbitrary and $\lambda = \frac{1}{2}$, we have,

$$\begin{aligned} g(\boldsymbol{x}) + g(\boldsymbol{y}) &= 2g\left(\frac{1}{2}(\boldsymbol{x} + \boldsymbol{y})\right) \\ &\overset{(1)}{=} g(\boldsymbol{x} + \boldsymbol{y}). \end{aligned}$$

As a consequence of $g(\boldsymbol{x}) + g(\boldsymbol{y}) = g(\boldsymbol{x} + \boldsymbol{y})$ and (1), we have,

$$g(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \alpha g(\boldsymbol{x}) + \beta g(\boldsymbol{y})$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, and all $\alpha, \beta \in \mathbb{R}$. Finally, with $e_j \in \mathbb{R}^n$ the cardinal unit vector in direction j, we have,

$$\begin{aligned} g(\boldsymbol{x}) = g\left(\sum_{j=1}^n x_j e_j\right) &= \sum_{j=1}^n x_j g(e_j) \\ &= \sum_{j=1}^n x_j a_j = \boldsymbol{a}^T \boldsymbol{x}, \end{aligned}$$

where we have defined the constants $a_j = g(e_j)$ for all j, and $\boldsymbol{a}$ has components a_j. Therefore, we have

$$f(\boldsymbol{x}) = g(\boldsymbol{x}) + f(0) = \boldsymbol{a}^T \boldsymbol{x} + b,$$

showing that f is affine.

7.7. Let $C \subseteq \mathbb{R}^n$ be a convex set. Let f be a convex function over C, and let g be a strictly convex function over C. Show that the sum function $f + g$ is strictly convex over C.

Solution: Let $\boldsymbol{x}, \boldsymbol{y} \in C$, and let $\lambda \in (0,1)$. By definition, we have,

$$\begin{aligned} f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &\le \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \\ g(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &< \lambda g(\boldsymbol{x}) + (1-\lambda) g(\boldsymbol{y}) \end{aligned}$$

Therefore,

$$\begin{aligned} (f+g)\left(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}\right) &= f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) + g(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \\ &\le \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) + g(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \\ &< \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) + \lambda g(\boldsymbol{x}) + (1-\lambda) g(\boldsymbol{y}) \\ &= \lambda (f+g)(\boldsymbol{x}) + (1-\lambda)(f+g)(\boldsymbol{y}), \end{aligned}$$

and hence $f + g$ is strictly convex.

7.10. Show that the following functions are convex over the specified domain C:
(i) $f(x_1, x_2, x_3) = -\sqrt{x_1 x_2} + 2x_1^2 + 2x_2^2 + 3x_3^2 - 2x_1 x_2 - 2x_2 x_3$ over $\mathbb{R}^3_{++}$
(ii) $f(\boldsymbol{x}) = \|\boldsymbol{x}\|^4$ over $\mathbb{R}^n$.
(iv) $f(\boldsymbol{x}) = \sqrt{\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} + 1}$ over $\mathbb{R}^n$ where $\boldsymbol{Q} \succeq \boldsymbol{0}$ is an $n \times n$ matrix.
(v) $f(x_1, x_2, x_3) = \max\{\sqrt{x_1^2 + x_2^2 + 20x_3^2 - x_1 x_2 - 4x_2 x_3 + 1}, (x_1^2 + x_2^2 + x_1 + x_2 + 2)^2\}$ over $\mathbb{R}^3$.

Solution:
(i) Define

$$f_1(x_1, x_2, x_3) = -\sqrt{x_1 x_2}, \qquad f_2(x_1, x_2, x_3) = 2x_1^2 + 2x_2^2 + 3x_3^2 - 2x_1 x_2 - 2x_2 x_3$$

so that $f = f_1 + f_2$. By direct computation, we have,

$$\nabla^2 f_1 = \frac{1}{4\sqrt{x_1 x_2}} \begin{pmatrix} \frac{x_2}{x_1} & -1 & 0 \\ -1 & \frac{x_1}{x_2} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \nabla^2 f_2 = \begin{pmatrix} 4 & -2 & 0 \\ -2 & 4 & -2 \\ 0 & -2 & 6 \end{pmatrix}$$

These formulas can be used to determine that $\nabla^2 f_1 \succeq \mathbf{0}$ for $\boldsymbol{x} \in \mathbb{R}^3_{++}$ and $\nabla^2 f_2 \succeq \mathbf{0}$. Thus, both f_1 and f_2 is convex. Since $f = f_1 + f_2$ is a conic combination of convex functions, then f is also convex.

(ii) Write f as $f(\boldsymbol{x}) = g(h(\boldsymbol{x}))$, where

$$g(t) = t^4, \qquad\qquad h(\boldsymbol{x}) = \|\boldsymbol{x}\|.$$

Since $\|\cdot\|$ is a norm, then h is a convex function with range $[0, \infty)$. The function g is convex, and for $t \in [0, \infty)$, the function $g(t)$ is also monotone increasing. Since f is a composition of a convex monotone increasing function (g) with a convex function (h), then f is convex.

(iv) Since $\boldsymbol{Q} \succeq \mathbf{0}$, then it has an orthogonal eigenvalue decomposition,

$$\boldsymbol{Q} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T,$$

where $\boldsymbol{V} \in \mathbb{R}^{n\times n}$ is an orthogonal matrix, and $\boldsymbol{\Lambda}$ is a diagonal matrix with non-negative values on the diagonal. Defining

$$\boldsymbol{A} \coloneqq \sqrt{\boldsymbol{\Lambda}}\boldsymbol{V}^T,$$

then we can write,

$$f(\boldsymbol{x}) = \sqrt{\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + 1} = \sqrt{\|\boldsymbol{A}\boldsymbol{x}\|^2 + 1} = h(g(\boldsymbol{x})),$$

where

$$g(\boldsymbol{x}) = \|\boldsymbol{A}\boldsymbol{x}\|, \qquad\qquad h(t) = \sqrt{t^2 + 1}.$$

A direct computation shows that

$$h''(t) = (t^2 + 1)^{3/2} > 0, \qquad\qquad h'(t) = \frac{t}{\sqrt{t^2 + 1}} > 0 \ (t \geq 0),$$

so that for $t \geq 0$, the function h is both monotone increasing and convex. The function g is the composition of an affine function with a norm,

$$g(\boldsymbol{x}) = \|\phi(\boldsymbol{x})\|, \qquad\qquad \phi(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x},$$

and since norms are convex, then g is convex. Finally, $f = h \circ g$ is the composition of a monotone increasing convex function with a convex function, and is hence convex.

(v) We write, $f(\boldsymbol{x}) = \max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\}$, where

$$f_1(\boldsymbol{x}) = \sqrt{x_1^2 + x_2^2 + 20x_3^2 - x_1x_2 - 4x_2x_3 + 1}, \quad f_2(\boldsymbol{x}) = (x_1^2 + x_2^2 + x_1 + x_2 + 2)^2,$$

so that f is the pointwise maximum of f_1 and f_2. Therefore, if we show that f_1 and f_2 are convex, this implies that f is convex. The function f_1 is convex since it can be written as,

$$f_1(\boldsymbol{x}) = h(q(\boldsymbol{x})), \qquad q(\boldsymbol{x}) = \sqrt{x_1^2 + x_2^2 + 20x_3^2 - x_1x_2 - 4x_2x_3} = \sqrt{\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}},$$

where h is the function defined in the solution to part (iv) of this question, and

$$\boldsymbol{A} = \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 1 & -2 \\ 0 & -2 & 20 \end{pmatrix} \succ \mathbf{0}.$$

Since $\boldsymbol{A} \succ \mathbf{0}$ and is symmetric, then by the same argument as in the solution to part (iv), there a matrix $\boldsymbol{B} \in \mathbb{R}^{3\times 3}$ such that $q(\boldsymbol{x}) = \sqrt{\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}} = \|\boldsymbol{B}x\|$, and hence q is convex. Since f_1 is the composition of h (a monotone increasing, convex function) with q (a convex function), then f_1 is convex.
The function f_2 is also convex since it can be written as,

$$f_2(\boldsymbol{x}) = g(p(\boldsymbol{x})), \qquad g(t) = t^2, \qquad p(\boldsymbol{x}) = x_1^2 + x_2^2 + x_1 + x_2 + 2.$$

We can immediately see that $g(t)$ is both monotone increasing and convex for $t \geq 0$. The function p satisfies,

$$\nabla^2 p = 2\boldsymbol{I} \succ \mathbf{0}, \qquad p(\boldsymbol{x}) = (x_1^2 + x_1 + 1) + (x_2^2 + x_2 + 1) > 0,$$

and hence is both convex and is strictly positive (hence we need only concern ourselves with $t > 0$ for the domain of g). Therefore, since f_2 is the composition of a monotone increasing and convex function with another convex function, then f_2 is convex.

7.25. Prove that if f and g are convex, twice differentiable, nondecreasing, and positive on $\mathbb{R}$, then the product fg is convex over $\mathbb{R}$. Show by an example that the positivity assumption is necessary to establish the convexity.

Solution: Since f and g are twice-differentiable, we can directly compute the second derivative of fg:

$$(fg)''(x) = f''(x)g(x) + g''(x)f(x) + 2f'(x)g'(x). \tag{2}$$

By assumption we have:

- f, g positive, convex $\Longrightarrow f''(x)g(x) \geq 0$, $g''(x)f(x) \geq 0$
- f, g nondecreasing $\Longrightarrow f'(x) \geq 0$, $g'(x) \geq 0$.

Using the above properties in (2) shows that $(fg)''(x) \geq 0$ for all x. I.e., the Hessian of fg is positive semi-definite everywhere, implying that fg is convex.

For an example demonstrating that the positivity assumption is necessary, consider

$$f(x) = e^x, \qquad g(x) = e^x - 4$$

Both f and g are (strictly) convex, twice-differentiable, and nondecreasing. But g is not positive on $\mathbb{R}$. Direct computation of $(fg)''$ results in,

$$(fg)''(x) = 4e^x(e^x - 1) < 0, \;\; x < 0.$$

Thus, over domain $x < 0$, the function (fg) is actually concave, not convex.

Additional problems:

P1. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *strongly convex* if there exists some $\alpha > 0$ such that $f(x) - \alpha\|x\|_2^2$ is convex. Show that strong convexity $\Rightarrow$ strict convexity $\Rightarrow$ convexity. In addition, give counterexamples to show that the reverse implications are not true.

Solution: We start by proving that a strongly convex function must be strictly convex: let f be strongly convex, i.e., there exists some $\alpha > 0$ such that $g(\boldsymbol{x}) = f(\boldsymbol{x}) - \alpha\|\boldsymbol{x}\|_2^2$ is convex. We first show that $h(\boldsymbol{x}) = \alpha\|\boldsymbol{x}\|_2^2$ is *strictly* convex: the Hessian is easily computed as,

$$\nabla^2 h(\boldsymbol{x}) = 2\boldsymbol{I} \succ \boldsymbol{0},$$

for all $\boldsymbol{x}$, establishing strict convexity of h. Turning back to the main goal, let $\lambda \in (0,1)$, and $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $\boldsymbol{x} \neq \boldsymbol{y}$ be arbitrary. We seek to show the definition of strict convexity:

$$f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) < \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}). \tag{3}$$

Since f is strongly convex, then g is convex, i.e.,

$$\begin{aligned} g(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &\leq \lambda g(\boldsymbol{x}) + (1-\lambda) g(\boldsymbol{y}) \\ f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) - \alpha\|\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\|_2^2 &\leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) - \lambda\alpha\|\boldsymbol{x}\|_2^2 - (1-\lambda)\alpha\|\boldsymbol{y}\|_2^2. \end{aligned}$$

Rearranging this last inequality yields,

$$\begin{aligned} f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) - \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) &\leq \alpha\|\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\|_2^2 - \lambda\alpha\|\boldsymbol{x}\|_2^2 - (1-\lambda)\alpha\|\boldsymbol{y}\|_2^2 \\ f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) - \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) &\leq h(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) - \lambda h(\boldsymbol{x}) - (1-\lambda) h(\boldsymbol{y}) < 0, \end{aligned}$$

where the last inequality is true since h is strictly convex. This last inequality is the desired relation (3).

We now seek to prove that a strictly convex function f is also convex. This is direct from the definition, since if f is strictly convex, then it satisfies (3) for any $\lambda \in (0,1)$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. This immediately implies that

$$f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}.$$

Finally, we furnish counterexamples for the two reverse implications: The function $f(x) = 0$ is convex by direction verification with the definition, but f is *not* strictly convex. (E.g., take $x = 0$, $y = 1$, $\lambda = \frac{1}{2}$.). The function $f(x) = x^4$ is strictly convex, but is not strongly convex since for any $\alpha > 0$, the function $g(x) = x^4 - \alpha x^2$ is not convex: It takes values

$$g(0) = 0, \qquad g\left(\pm\sqrt{\alpha/2}\right) = -\frac{\alpha^2}{4} < 0,$$

so that taking $x = -\sqrt{\alpha/2}$, $y = \sqrt{\alpha/2}$ and $\lambda = \frac{1}{2}$ violates the definition of convexity for g.

P2. (**6000-level students only**) Let $f : \mathbb{R} \to \mathbb{R}$ be strictly monotone (either increasing or decreasing) such that f has a well-defined functional inverse $f^{-1} : \mathbb{R} \to \mathbb{R}$. (I.e., the range of f is $\mathbb{R}$ and f is bijective.) Assume f is convex. If f is increasing, what can you say about f^{-1}? What about if f is decreasing?

Solution: Intuition about what happens is best learned by considering examples (e.g., $f(x) = \pm \log x$, $e^{\pm x}$, etc.) or by graphing functions (how is the graph of f^{-1} generated from the graph of f?). Here we'll just present the main results through proof.
Let $f : \mathbb{R} \to \mathbb{R}$ be strictly monotone (either increasing or decreasing), convex, and suppose $f^{-1} : \mathbb{R} \to \mathbb{R}$ is well-defined and exists. We seek to understand properties of f^{-1}. To that end, let $y_1, y_2 \in \mathbb{R}$ with $y_1 \neq y_2$, and define,

$$x_1 \coloneqq f^{-1}(y_1), \qquad\qquad x_2 \coloneqq f^{-1}(y_2),$$

and so by definition we also have $y_i = f(x_i)$, $i = 1, 2$. Note that also $x_1 \neq x_2$, since otherwise we would have $y_1 = y_2$, which is a contradiction. First note that f^{-1} is also strictly monotone (increasing if f is increasing, decreasing if f is decreasing):

Lemma 1. *Under the previously discussed assumptions on f, then f^{-1} is strictly monotone, and is increasing iff f is increasing. f^{-1} is decreasing iff f is decreasing.*

To see why this lemma is true, note that f strictly monotone increasing means,

$$z < w \Longrightarrow f(z) < f(w).$$

So suppose that $y_1 < y_2$. We could postulate some relationships between x_1 and x_2:

- $x_1 > x_2$: this cannot happen since f increasing would imply $y_1 > y_2$, which is a contradiction
- $x_1 = x_2$: this cannot happen since $y_1 \neq y_2$, as previously described

Since $x_1 \not\geq x_2$, we have $x_1 < x_2$, showing that f^{-1} is strictly monotone increasing. The same argument shows that if f is strictly monotone decreasing, then f^{-1} is strictly monotone decreasing, showing Lemma 1.
We now turn to exploring possible convexity of f^{-1}. For concreteness, we'll assume f is strictly monotone increasing (as opposed to decreasing). With y_1, y_2, x_1, x_2 as before, let $\lambda \in (0, 1)$ be arbitrary. Since f is convex, we have

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

We apply f^{-1} to both sides: since f^{-1} is strictly increasing, this results in

$$\lambda f^{-1}(y_1) + (1 - \lambda)f^{-1}(y_2) \leq f^{-1}\left(\lambda y_1 + (1 - \lambda)y_2\right).$$

Multiplying this result by -1, we obtain

$$(-f^{-1})\left(\lambda y_1 + (1 - \lambda)y_2\right) \leq \lambda(-f^{-1})(y_1) + (1 - \lambda)(-f^{-1})(y_2),$$

which establishes that $-f^{-1}$ is convex. I.e., when f is increasing, f^{-1} is *concave.* A similar argument shows that f^{-1} is convex if f is monotone decreasing. We can summarize this formally:

Proposition 1. *Suppose $f : \mathbb{R} \to \mathbb{R}$ is convex and strictly monotone, with $f^{-1} : \mathbb{R} \to \mathbb{R}$ well-defined. Then if f is increasing, f^{-1} is concave. If f is decreasing, then f^{-1} is convex.*