The Chinese University of Hong Kong, Shenzhen
SDS · School of Data Science
Xiao Li · Junfeng Wu · Ming Yan · Fall Semester 2023

MAT 3007 – Optimization

Assignment 7

Due: 11:59pm, Nov. 24 (Friday), 2023

Instructions:

- Homework problems must be carefully and clearly answered to receive full credit. Complete sentences that establish a clear logical progression are highly recommended.
- Please submit your assignment on Blackboard.
- The homework must be written in English.
- Late submission will not be graded.
- Each student must not copy homework solutions from another student or from any other source.

Problem 1 (Convex Sets): (approx. *25* points)
In this exercise, we study convexity of various sets.

a) Verify whether the following sets are convex or not and explain your answer!

$$\Omega_1 = \{x \in \mathbb{R}^n : \alpha \le (a^\top x)^2 \le \beta\}, \quad \alpha, \beta \in \mathbb{R},\ 0 < \alpha \le \beta,\ a \in \mathbb{R}^n,$$
$$\Omega_2 = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : x^\top x \le t^2\}.$$

b) Decide whether the following statements are true or false. Explain your answer and either present a proof / verification or a counter-example.

- The intersection of two convex sets $\Omega_1, \Omega_2 \subset \mathbb{R}^n$ is always a convex set.
- Let $\Omega \subset \mathbb{R}^n$ be a convex set and suppose that the set $S := \{(x,t) \in \Omega \times \mathbb{R} : f(x) \le t\} \subset \mathbb{R}^n \times \mathbb{R}$ is convex. Then, $f : \Omega \to \mathbb{R}$ is a convex function.

Solution :

a) The set Ω_1 is not convex. To see this, let us set $n = 1$. Let $a = 1$, $\alpha = 1$ and $\beta = 4$, the set Ω_1 is

$$\Omega_1 = [1,2] \cup [-2,-1].$$

Let $x_1 = 1$ and $x_2 = -1$. Then, obviously $x_1, x_2 \in \Omega_1$, but the point $\frac{1}{2}x_1 + \frac{1}{2}x_2 = 0$ is not contained in Ω_1. Hence, Ω_1 cannot be convex.

The set Ω_2 is not convex. To see this, let us set $n = 1$ and $(x_1, t_1) = (1, 1)$, $(x_2, t_2) = (1, -1)$. Then, obviously $(x_1, t_1), (x_2, t_2) \in \Omega_2$, but the point $\frac{1}{2}(x_1, t_1) + \frac{1}{2}(x_2, t_2) = (1, 0)$ is not contained in Ω_2. Hence, Ω_2 cannot be convex.

b) The first statement is true: let $x, y \in \Omega_1 \cap \Omega_2$ and $\lambda \in [0,1]$ be arbitrary. Then, by convexity of Ω_1 and Ω_2, it follows $\lambda x + (1-\lambda)y \in \Omega_1$ and $\lambda x + (1-\lambda)y \in \Omega_2$. This shows $\lambda x + (1-\lambda)y \in \Omega_1 \cap \Omega_2$ and thus, the intersection $\Omega_1 \cap \Omega_2$ is a convex set.

We verify the second statement briefly. Let $x, y \in \Omega$ and $\lambda \in [0,1]$ be arbitrary. Then, we have $(x, f(x)), (y, f(y)) \in S$. By the convexity of S, we can infer $\lambda(x, f(x)) + (1-\lambda)(y, f(y)) = (\lambda x + (1-\lambda)y, \lambda f(x) + (1-\lambda)f(y)) \in S$. However, by definition of S, this means

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

Since $x, y \in \Omega$ and $\lambda \in [0,1]$ are arbitrary, this implies that f is convex on Ω.

Problem 2 (Convex Compositions): (approx. *20* points)

Either prove or find a counterexample for each of the following statements (you can assume that all functions are twice continuously differentiable if needed):

a) If $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are concave, then the composition $f \circ g : \mathbb{R}^n \to \mathbb{R}$, $(f \circ g)(x) = f(g(x))$ is concave.

b) Let $\Omega \subset \mathbb{R}^n$ be a convex set and suppose that $g : \Omega \to \mathbb{R}$ is concave and $f : I \to \mathbb{R}$ is concave and nondecreasing where $I \supseteq g(\Omega)$ is an interval containing $g(\Omega)$. Then, $f \circ g$ is convex.

c) If $f : \mathbb{R} \to \mathbb{R}$ is convex, then $x \mapsto |f(x)|$ is a convex function on $\mathbb{R}$.

Solution :

a) *False.* Let $n = 1$. Set $g(x) = -x^2$ and $f(x) = -x$. Both functions are obviously concave, but $f(g(x)) = x^2$ is a convex function.

b) *False.* We prove this result by using the basic definition of concavity. Let $x, y \in \Omega$ and $\lambda \in [0,1]$ be arbitrary. Using the concavity of g, we have $g(\lambda x + (1-\lambda)y) \geq \lambda g(x) + (1-\lambda)g(y)$. Since the interval I is convex and we have $g(\Omega) \subseteq I$, it follows $\lambda g(x) + (1-\lambda)g(y) \in I$. Moreover, since f is nondecreasing, we have

$$f(g(\lambda x + (1-\lambda)y)) \geq f(\lambda g(x) + (1-\lambda)g(y)) \geq \lambda f(g(x)) + (1-\lambda)f(g(y)),$$

where we used the concavity of f in the last step. This shows that $f \circ g$ is concave.

c) *False.* Consider the following counterexample

$$f(x) = x^2 - 1 \quad \text{and} \quad |f(x)| = \begin{cases} x^2 - 1 & |x| \geq 1, \\ 1 - x^2 & |x| \leq 1, \end{cases}.$$

Taking second order derivative, we know that f is convex, but $|f(x)|$ is not convex when $|x| \leq 1$.

Problem 3 (Convex Functions): (approx. *30* points)

In this exercise, convexity properties of different functions are investigated.

a) Let $r : \mathbb{R}^n \to \mathbb{R}$ be defined as $r(x) = \max_i |x_i|$. Show that r is a convex function.

b) Verify that the following functions are convex over the specified domain:

   – $f : \mathbb{R} \times \mathbb{R}_{++} \to \mathbb{R}$, $f(x) := x_1^2 / x_2$, where $\mathbb{R}_{++} := \{x \in \mathbb{R} : x > 0\}$.

- $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) := \frac{1}{2}\|Ax - b\|^2 + \mu\|Lx\|_\infty$, where $A \in \mathbb{R}^{m\times n}$, $L \in \mathbb{R}^{p\times n}$, $b \in \mathbb{R}^m$, and $\mu > 0$ are given and $\|y\|_\infty := \max_{i=1,...,p} |y_i|$, $y \in \mathbb{R}^p$.
- $f : \mathbb{R}^{n+1} \to \mathbb{R}$, $f(x, y) := \frac{\lambda}{2}\|x\|^2 + \sum_{i=1}^m \max\{0, 1 - b_i(a_i^\top x + y))\}$, where $a_i \in \mathbb{R}^n$ and $b_i \in \{-1, 1\}$ are given data points for all $i = 1, ..., m$ and $\lambda > 0$ is a parameter.

c) Let us set $f(x) = \|x\|_1 := \sum_{i=1}^n |x_i|$ and define $g : \mathbb{R}^n \to \mathbb{R}$, $g(x) := \max_{y\in\mathbb{R}^n} y^\top x - f(y)$.

Calculate $g(x)$ explicitly and verify that the function g is convex.

Solution :

a) Let $x, y \in \mathbb{R}^n$ and $\lambda \in [0, 1]$ be given. Then the triangle inequality implies

$$\begin{aligned} r(\lambda x + (1-\lambda)y) &= \max_i |\lambda x_i + (1-\lambda)y_i| \\ &\le \max_i |\lambda x_i| + \max_i |(1-\lambda)y_i| \\ &= \lambda r(x) + (1-\lambda)r(y). \end{aligned}$$

This shows that r is convex.

b) To show $f(x) = x_1^2/x_2$ is convex, we note that (for $x_2 > 0$), we have

$$\nabla^2 f(x) = \begin{pmatrix} 2/x_2 & -2x_1/x_2^2 \\ -2x_1/x_2^2 & 2x_1^2/x_2^3 \end{pmatrix} = \frac{2}{x_2^3}\begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}\begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}^T \succeq 0.$$

Hence, $f(x)$ is convex.

We now study $f(x) := \frac{1}{2}\|Ax - b\|^2 + \mu\|Lx\|_\infty$. Since $\frac{1}{2}\|\cdot\|^2$ is convex (the Hessian is the identity matrix), we know that $x \mapsto \frac{1}{2}\|Ax - b\|^2$ is convex as a composition of a linear and convex function. Moreover, part a) implies that the maximum-norm $\|\cdot\|_\infty$ is convex. Again $x \mapsto \|Lx\|_\infty$ is then a convex function. Together, this shows that f is convex.

Finally, let us define $g(x, y) = \frac{\lambda}{2}\|x\|^2$ and $g_i(x, y) = \max\{0, 1 - b_i(a_i^\top x + y))\}$. Then, f can be interpreted as the sum of the functions g and g_i, $i = 1, ..., m$ and convexity follows if each of the functions g, g_i, $i = 1, ..., m$ is convex. The mapping g_i is the maximum of the constant function $(x, y) \mapsto 0$ and of the affine-linear function $(x, y) \mapsto h_i(x, y) := 1 - b_i(a_i^\top x + y)$. Since both of these functions are convex (as linear mappings), the function g_i is convex. Finally, the Hessian of g is given by

$$\mathbb{R}^{(n+1)\times(n+1)} \ni \nabla^2 g(x, y) = \lambda \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \succeq 0.$$

This establishes convexity of f.

c) Defining $h_y(x) := y^\top x - f(y)$, we see that h_y is linear in x for every fixed $y \in \mathbb{R}^n$. Hence, $g(x) := \sup_{y\in\mathbb{R}^n} h_y(x)$ can be interpreted as the maximum/supremum of the infinite family $\{h_y(x)\}_{y\in\mathbb{R}^n}$. According to the lecture, since each h_y is convex, we can infer that g is a convex mapping (this observation is completely independent of f).

We now want to express $g(x)$ explicitly. First, let us notice that $f(x) = \|x\|_1$ is a convex function.

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= \|\lambda x + (1-\lambda)y\|_1 \\ &= \sum_{i=1}^n |\lambda x_i + (1-\lambda)y_i| \\ &\le \sum_{i=1}^n |\lambda x_i| + \sum_{i=1}^n |(1-\lambda)y_i| \\ &= \lambda f(x) + (1-\lambda)f(y) \end{aligned}$$

Consequently, the function $y \mapsto y^\top x - f(y)$ is concave in y and we can show that

$$g(x) = \begin{cases} 0 & \|x\|_\infty \le 1, \\ \infty & \text{otherwise}, \end{cases}.$$

To see this, let's consider $\|x\|_\infty > 1$ first, i.e., there exists j such that $x_j > 1$. Choose a special y such that it is all 0 except for the j-th positions and set $y_j = tx_j$ with $t > 0$. Then

$$x^T y - \|y\|_1 = \sum_j (x_j y_j - |y_j|) = \sum_j t(x_j^2 - |x_j|) \to \infty$$

as $t \to \infty$, which shows $g(x) = \infty$. Conversely, if $\|x\|_\infty \le 1$, then

$$x^T y - \|y\|_1 = \sum_{j=1}^n (x_j y_j - |y_j|) \le \sum_{j=1}^n (|x_j||y_j| - |y_j|) \le \sum_{j=1}^n (|y_j| - |y_j|) = 0.$$

Therefore, the maximum value is 0.

Problem 4: (approx. *25* points)

a) Let $A \in \mathbb{R}^{4\times 4}$ be a symmetric matrix with nonnegative components and $A1 = 1$, i.e., each row of the matrix A has sum 1. Prove that $I - A$ is positive semidefinite.

b) Let $a \in \mathbb{R}^n$. We define $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) := \log(1 + \exp(a^\top x))$. Show that f is a convex function.

c) Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, and $d \in \mathbb{R}$. Prove that the nonconvex optimization problem

$$\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & \dfrac{\|Ax - b\|}{c^\top x + d} \\ \text{subject to} & \|x\| \le 1, c^T x + d > 0 \end{array} \tag{1}$$

is equivalent to the convex optimization problem

$$\begin{array}{ll} \min\limits_{y\in\mathbb{R}^n, t} & \|Ay - bt\| \\ \text{subject to} & \|y\| \le t \\ & c^\top y + dt = 1 \end{array} \tag{2}$$

d) Use `CVX` (in `MATLAB` or `Python`) to solve problem (2) with

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \quad c = \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \quad d = 1$$

Solution :

a) Assume x is the eigenvector of A but not corresponding to eigenvalue 1, i.e. $Ax = \lambda x$. Denote i be the coordination corresponding to the largest absolute value of x. Then we can see that $\sum_j A_{ij} x_j = \lambda x_i$ and

$$|\lambda||x_i| \leq \sum_j A_{ij}|x_j| \leq \sum_j A_{ij}|x_i| = |x_i|.$$

Hence, $\lambda \leq 1$. In addition, 1 is also the eigenvalue, then $I - A$ is positive semidefinite.

b) We can see that

$$\nabla f(x) = \frac{a \exp(a^T x)}{1 + \exp(a^T x)}$$

and

$$\nabla^2 f(x) = \frac{aa^T \exp(a^T x)}{(1 + \exp(a^T x))^2} \succeq 0$$

Hence, $\nabla^2 f(x)$ is positive semidefinite for all x and f is convex on $\mathbb{R}^n$.

c) To show the equivalence, we can define

$$y = \frac{x}{c^T x + d}, \qquad t = \frac{1}{c^T x + d},$$

then we have $c^T y + dt = 1$. Note that

$$\|y\| = \left\| \frac{x}{c^T x + d} \right\| = \|tx\| = t\|x\|,$$

then we can see that $\|x\| \leq 1 \iff \|y\| \leq t$. In addition,

$$\|Ay - bt\| = \left\| A\frac{x}{c^T x + d} - b\frac{1}{c^T x + d} \right\| = \frac{\|Ax - b\|}{c^\top x + d}.$$

Hence, they have the same objective value. Therefore, they are equivalent.

d) Exemplary code can be found in Listing 1 & 2. The following solutions (for the original problem (1)) will be returned: $y_1 = 0$, $y_2 = 0.25$, $t = 0$, and `opt-val` $= 0$.

Listing 1: Problem 4d: MATLAB code

```
A = [1,2;3,4]; b = [2;4]; c = [4;3]; d = 1;
cvx_begin
cvx_precision high
variables y(2,1) t

minimize norm(A*y-b*t)
subject to
    norm(y) <= t;
    c'*y + d*t == 1;
cvx_end
```

Listing 2: Problem 4d: Python code

```
import cvxpy as cp
import numpy as np
A= np.array([[1,2],[3,4]])
b = np.array([[2],[4]])
c = np.array([[4],[3]])
d = 1

y = cp.Variable((2,1))
t = cp.Variable()

con = [cp.norm(y, 2) <= t]
con += [cp.sum(cp.multiply(c, y)) + d*t == 1]

obj = cp.Minimize(cp.norm(A @ y - b*t,2))
prob=cp.Problem(obj,con)
prob.solve()


print(y.value)
print(t.value)
```