Andre Milzarek · Zizhuo Wang · Haoxiang Yang · Spring Semester 2022

MAT 3007 – Optimization

Solutions 7

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!

$$\Omega_1 = \{x \in \mathbb{R}^n : \alpha \le (a^\top x)^3 \le \beta\}, \quad \alpha, \beta \in \mathbb{R},\ \alpha \le \beta,\ a \in \mathbb{R}^n,$$
$$\Omega_2 = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : x^\top x \le t^2\}.$$

b) Show that the hyperbolic set $\{x \in \mathbb{R}^2_+ : x_1 x_2 \ge 1\}$ is convex, where $\mathbb{R}^2_+ = \{x \in \mathbb{R}^2 : x \ge 0\}$.

Hint: Rewrite the condition "$x_1 x_2 \ge 1$" in a suitable way.

c) 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) Since the function $x \mapsto x^3$ is monotonically increasing, the set Ω_1 is equivalent to

$$\begin{aligned}\Omega_1 &= \{x \in \mathbb{R}^n : \mathrm{sign}(\alpha)|\alpha|^{\frac{1}{3}} \le a^\top x \le \mathrm{sign}(\beta)|\beta|^{\frac{1}{3}}\} \\ &= \{x \in \mathbb{R}^n : -a^\top x \le -\mathrm{sign}(\alpha)|\alpha|^{\frac{1}{3}}\} \cap \{x \in \mathbb{R}^n : a^\top x \le \mathrm{sign}(\beta)|\beta|^{\frac{1}{3}}\},\end{aligned}$$

where $\mathrm{sign}(x)$ denotes the sign-function ($\mathrm{sign}(x) = 1$ if $x > 0$, $\mathrm{sign}(x) = 0$ if $x = 0$, and $\mathrm{sign}(x) = -1$ if $x < 0$). Hence, Ω_1 is the intersection of two half-spaces and it can be immediately shown that Ω_1 is convex. (For instance, we can apply part c)).

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) We first notice $\Omega := \{x \in \mathbb{R}^2_+ : x_1 x_2 \ge 1\} = \{x \in \mathbb{R}^2_{++} : x_1 x_2 \ge 1\}$, where $\mathbb{R}^2_{++} = \{x \in \mathbb{R}^2 : x > 0\}$. Since the (natural) logarithm $\log(\cdot)$ is monotonically increasing, the condition $x_1 x_2 \ge 1$ is equivalent to

$$f(x_1, x_2) := -\log(x_1) - \log(x_2) \le 0.$$

The gradient and Hessian of f are given by:

$$\nabla f(x_1, x_2) = \begin{pmatrix} -x_1^{-1} \\ -x_2^{-1} \end{pmatrix}, \quad \nabla^2 f(x_1, x_2) = \begin{pmatrix} x_1^{-2} & 0 \\ 0 & x_2^{-2} \end{pmatrix}$$

Since $\nabla^2 f$ is positive definite on the open and convex set $\mathbb{R}^2_{++}$, it follows that f is convex on $\mathbb{R}^2_{++}$. This shows that the set Ω is convex.

c) 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 convex, then the composition $f \circ g : \mathbb{R}^n \to \mathbb{R}$, $(f \circ g)(x) = f(g(x))$ is convex.

b) Let $\Omega \subset \mathbb{R}^n$ be a convex set and suppose that $g : \Omega \to \mathbb{R}$ is convex and $f : I \to \mathbb{R}$ is convex 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 increasing, then $x \mapsto |f(x)|$ is a convex function on $\mathbb{R}$.

Solution :

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

b) *True.* We prove this result by using the basic definition of convexity. Let $x, y \in \Omega$ and $\lambda \in [0,1]$ be arbitrary. Using the convexity of g, we have $g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda)g(y)$. 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

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

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

c) *False.* Consider the following counterexample

$$f(x) = \begin{cases} \sqrt{x} & x \geq 0, \\ -\sqrt{|x|} & x \leq 0, \end{cases} \quad \text{and} \quad |f(x)| = \sqrt{|x|}.$$

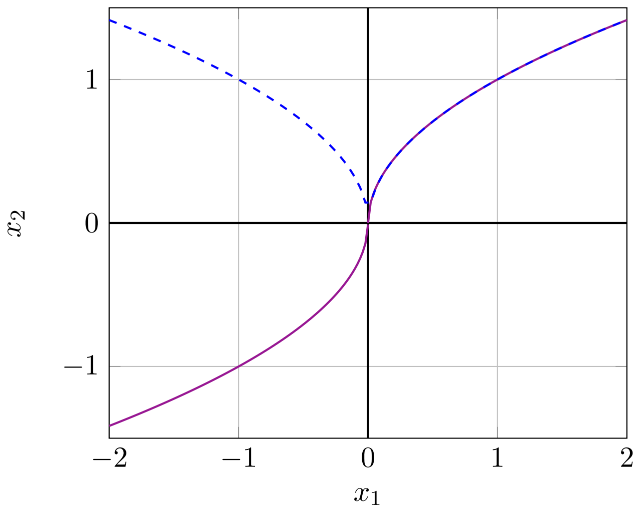
Using a sketch, we can easily verify that f is increasing. However, $|f(x)|$ is obviously not convex. (We have already seen that the square root is a concave function).

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 a norm on $\mathbb{R}^n$. Show that r is a convex function.

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



- $f : \mathbb{R}_{++} \to \mathbb{R}$, $f(x) := \sqrt{1 + x^{-2}}$, 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 $f(x) = \|x\|^3$ and define $g : \mathbb{R}^n \to \mathbb{R}$, $g(x) := \max_{y\in\mathbb{R}^n} y^\top x - f(y)$.

Show that g is well-defined, i.e, $g(x)$ exists for all x and satisfies $g(x) < \infty$. 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 and the positive homogeneity of the norm function r imply

$$r(\lambda x + (1 - \lambda)y) \le r(\lambda x) + r((1 - \lambda)y) = |\lambda|r(x) + |1 - \lambda|r(y) = \lambda r(x) + (1 - \lambda)r(y).$$

This shows that r is convex.

b) The mapping $f(x) = \sqrt{1 + x^{-2}}$ is twice continuously differentiable on the convex set $\mathbb{R}_{++}$ and it holds that

$$f'(x) = \frac{1}{2\sqrt{1 + x^{-2}}}\frac{-2}{x^3} = \frac{-1}{x^2\sqrt{1 + x^2}}, \quad f''(x) = \frac{2x\sqrt{1 + x^2} + \frac{x^3}{\sqrt{1+x^2}}}{x^4(1 + x^2)} = \frac{3x^2 + 2}{x^3(1 + x^2)\sqrt{1 + x^2}}.$$

Notice that we have $f''(x) > 0$ for all $x > 0$ and thus, f is convex on $\mathbb{R}_{++}$.

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

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

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 $f(x) = \|x\|^3$ is a convex function. Indeed, the norm $x \mapsto \|x\| \in \mathbb{R}_+$ is convex according to part a). Moreover, the mapping $x \mapsto x^3$ is convex and nondecreasing on $\mathbb{R}_+$. Hence, convexity of f follows from Problem 2b). Consequently, the function $y \mapsto y^\top x - f(y)$ is concave in y and the first-order optimality conditions are given by:

$$x - \nabla f(y) = 0 \quad\Longleftrightarrow\quad 3\|y\|^2 \cdot \frac{y}{\|y\|} = x \quad\Longleftrightarrow\quad 3\|y\| \cdot y = x.$$

Taking the norm, this yields $\|y\| = \sqrt{\|x\|/3}$ and $y = x/\sqrt{3\|x\|}$ (in the case $x = 0$, we obtain $y = 0$). The concavity of $y \mapsto y^\top x - f(y)$ ensures that this stationary point is a global maximum, i.e., we have

$$g(x) = \frac{\|x\|^2}{\sqrt{3\|x\|}} - \frac{\|x\|^3}{3\|x\|\sqrt{3\|x\|}} = \frac{2}{3\sqrt{3}}\|x\|^{\frac{3}{2}}.$$

This also shows that g is well-defined for all $x \in \mathbb{R}^n$.

Problem 4 (Geometric Programming): (approx. *25* points)

In this exercise, we discuss a class of nonconvex geometric programs that can be reformulated as convex optimization problems.

a) Let $a \in \mathbb{R}^n$ be given with $\sum_{i=1}^n a_i = 1$ and $a \geq 0$. Show that the matrix $A := \mathrm{diag}(a) - aa^\top$ is positive semidefinite. (Here, $\mathrm{diag}(a)$ is a $n \times n$ diagonal matrix with a on its diagonal).

Hint: The Cauchy-Schwarz inequality $x^\top y \leq \|x\|\|y\|$, $x, y \in \mathbb{R}^n$, can be helpful.

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

c) Convert the following optimization problem into a convex problem:

$$\begin{array}{ll} \min_{x\in\mathbb{R}^3} & \max\left\{\frac{x_1}{x_2}, \frac{\sqrt{x_3}}{x_2}\right\} \\ \text{subject to} & x_1^2 + \frac{2x_2}{x_3} \leq \sqrt{x_2}, \\ & \frac{x_1}{x_2} \geq x_3^2, \\ & x_1, x_2, x_3 \geq 0. \end{array} \tag{1}$$

Hint: Substitute the variables x_i in an appropriate way and apply the result of part b).

d) Use `CVX` (in `MATLAB` or `Python`) to solve problem (1).

Solution :

a) For all $h \in \mathbb{R}^n$, we have

$$h^\top Ah = h^\top \mathrm{diag}(a)h - (a^\top h)^2 = \sum\nolimits_{i=1}^n a_i h_i^2 - \left(\sum\nolimits_{i=1}^n \sqrt{a_i} \cdot \sqrt{a_i} h_i\right)^2$$
$$\geq \sum\nolimits_{i=1}^n a_i h_i^2 - \left(\left[\sum\nolimits_{i=1}^n a_i\right]^{\frac{1}{2}} \cdot \left[\sum\nolimits_{i=1}^n a_i h_i^2\right]^{\frac{1}{2}}\right)^2 = \sum\nolimits_{i=1}^n a_i h_i^2 - \sum\nolimits_{i=1}^n a_i h_i^2 = 0,$$

where we used the Cauchy-Schwarz inequality with $x_i := \sqrt{a_i}$ and $y_i := \sqrt{a_i} h_i$. This shows that A is positive semidefinite.

b) Defining $\exp(\mathbf{x}) := (\exp(x_1), ..., \exp(x_n))^\top \in \mathbb{R}^n$ and using the chain and product rule, it holds that

$$\nabla f(x) = \frac{1}{\sum_{i=1}^n \exp(x_i)} \begin{pmatrix} \exp(x_1) \\ \vdots \\ \exp(x_n) \end{pmatrix} = \frac{\exp(\mathbf{x})}{\mathbb{1}^\top \exp(\mathbf{x})}$$

and

$$\nabla^2 f(x) = \frac{1}{\mathbb{1}^\top \exp(\mathbf{x})} \mathrm{diag}(\exp(\mathbf{x})) - \frac{1}{(\mathbb{1}^\top \exp(\mathbf{x}))^2} \exp(\mathbf{x}) \exp(\mathbf{x})^\top.$$

Setting $a := \exp(\mathbf{x})/(\mathbb{1}^\top \exp(\mathbf{x}))$, we see that the Hessian of f has the format $\nabla^2 f(x) = \mathrm{diag}(a) - aa^\top$ for all x. Hence, by part a), we can infer that $\nabla^2 f(x)$ is positive semidefinite for all x and f is convex on $\mathbb{R}^n$.

c) We use the substitution $x_1 = \exp(y_1)$, $x_2 = \exp(y_2)$, and $x_3 = \exp(y_3)$ (this implicitly implies $x_1, x_2, x_3 > 0$). Then, the problem can be transformed to

$$\begin{array}{ll} \min_{y \in \mathbb{R}^3} & \max\left\{\exp(y_1 - y_2), \exp(\frac{1}{2}y_3 - y_2)\right\} \\ \text{subject to} & \exp(2y_1 - \frac{1}{2}y_2) + 2\exp(\frac{1}{2}y_2 - y_3) \leq 1, \\ & \exp(y_1 - y_2 - 2y_3) \geq 1, \end{array}$$

where we divided the first and second constraint by $\sqrt{x_2}$ and x_3^2, respectively. We now introduce a slack variable "$\exp(y_4)$" for the max-expression and introduce the constraints "$\exp(y_1 - y_2) \leq \exp(y_4)$" and "$\exp(\frac{1}{2}y_3 - y_2) \leq \exp(y_4)$". This yields

$$\begin{array}{ll} \min_{y \in \mathbb{R}^4} & \exp(y_4) \\ \text{subject to} & \exp(y_1 - y_2 - y_4) \leq 1, \\ & \exp(\frac{1}{2}y_3 - y_2 - y_4) \leq 1, \\ & \exp(2y_1 - \frac{1}{2}y_2) + 2\exp(\frac{1}{2}y_2 - y_3) \leq 1, \\ & \exp(y_1 - y_2 - 2y_3) \geq 1. \end{array}$$

Finally, taking the logarithm (this will not affect the minimization or the constraints since the logarithm is monotonically increasing), we obtain

$$\begin{array}{ll} \min_{y \in \mathbb{R}^4} & y_4 \\ \text{subject to} & y_1 - y_2 - y_4 \leq 0, \\ & \frac{1}{2}y_3 - y_2 - y_4 \leq 0, \\ & \log(\exp(2y_1 - \frac{1}{2}y_2) + \exp(\frac{1}{2}y_2 - y_3 + \log(2))) \leq 0, \\ & y_1 - y_2 - 2y_3 \geq 0. \end{array}$$

The third constraint can be written in the format "$f(Ay + b) \leq 0$", where $f(x_1, x_2) = \log(\exp(x_1) + \exp(x_2))$, and

$$A = \begin{pmatrix} 2 & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & -1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ \log(2) \end{pmatrix}.$$

By part b), f is convex and thus, the third constraint "$f(Ay + b) \leq 0$" defines a convex constraint. Since the objective function and all remaining constraints are linear, this reformulation of (1) is a convex problem.

d) Exemplary code can be found in Listing 1 & 2. The following solutions (for the original problem (1)) will be returned: $x_1 = 0.251966$, $x_2 = 0.216756$, $x_3 = 1.07816$, and `opt-val` $=$ 4.79039.

Listing 1: Problem 4d: `MATLAB` code

```
cvx_begin
cvx_precision high
variables y1 y2 y3 y4

minimize y4
subject to
        y1 - y2 - y4 <= 0;
        y3/2 - y2 - y4 <= 0;
        log(exp(2*y1-y2/2)+exp(y2/2-y3+log(2))) <= 0;
        y1 - y2 - 2*y3 >= 0;
cvx_end

x1 = exp(y1);
x2 = exp(y2);
x3 = exp(y3);
t  = exp(y4);

fprintf(1,'x1 = %g, x2 = %g, x3 = %g, opt-val = %g\n',x1,x2,x3,t);
```

Listing 2: Problem 4d: `Python` code

```
import cvxpy as cp
import numpy as np

y1 = cp.Variable()
y2 = cp.Variable()
y3 = cp.Variable()
y4 = cp.Variable()

con = [y1-y2-y4 <= 0]
con += [y3/2-y2-y4 <= 0]
con += [cp.exp(2*y1-y2/2)+2*cp.exp(y2/2-y3)<= 1]
con += [y1-y2-2*y3 >= 0]
prob=cp.Problem(cp.Minimize(y4),con)
prob.solve()

x1=np.exp(y1.value)
x2=np.exp(y2.value)
x3=np.exp(y3.value)
t=np.exp(y4.value)

print('x1 = %g, x2 = %g, x3 = %g, opt-val = %g'%(x1,x2,x3,t))
```