MAT3007 - Assignment 7

Yohandi

November 24, 2023

Problem 1: Convex Sets

a)
- To prove that $\Omega_1 = \{x \in R^n : \alpha \leq (a^T x)^2 \leq \beta\}, \alpha, \beta \in R, 0 < \alpha \leq \beta, a \in R^n$ is not a convex set, we can simply find any two different points $x_1, x_2 \in \Omega_1$ and a $\lambda \in [0, 1]$ such that $\lambda x_1 + (1-\lambda)x_2 \notin \Omega_1$. Consider the following counterexample: Suppose we have $x_1, x_2 \in \Omega_1$, where $a^T x_1 = \sqrt{\alpha}$ and $a^T x_2 = -\sqrt{\alpha}$. Then, for a $\lambda = \frac{1}{2}$, $(a^T(\lambda x_1 + (1-\lambda)x_2))^2 = (\lambda\sqrt{\alpha} - (1-\lambda)\sqrt{\alpha})^2 = ((2\lambda - 1)\sqrt{\alpha})^2 = 0 < \alpha$. This implies that $\frac{1}{2}x_1 + \frac{1}{2}x_2 \notin \Omega_1$; hence, Ω_1 is not a convex set.
- To prove the convexity of $\Omega_2 = \{(x, t) \in R^n \times R : x^T x \leq t^2\}$, we let a function $g_t(x) = x^T x - t^2 = \|x\|^2 - t^2$ for a fixed t. Note that $x^T x$ is a convex function as it is a quadratic form with a positive semi-definite matrix (the identity matrix). As the term $-t^2$ is a constant in this particular case, we conclude that g_t is a convex function; by theorem, this further implies that $\Omega_2 = \{(x, t) \in R^n \times R : g_t(x) \leq 0\}$ is formulated as a convex set.

b)
- The statement is *true*. Let $x_1, x_2 \in \Omega_1 \cap \Omega_2$ and $\lambda \in [0, 1]$, then:
  - $x_1, x_2 \in \Omega_1 \Rightarrow \lambda x_1 + (1-\lambda)x_2 \in \Omega_1$
  - $x_1, x_2 \in \Omega_2 \Rightarrow \lambda x_1 + (1-\lambda)x_2 \in \Omega_2$

  , implying that $\lambda x_1 + (1-\lambda)x_2 \in \Omega_1 \cap \Omega_2$. This verifies that $\Omega_1 \cap \Omega_2$ is also a convex set.
- The statement is *false*. Consider a case where $n = 1$ and $f(x) = x^3$, for which f is not a convex function; however, $(x, 1) \in S, \forall x \in [0, 1]$, for which Ω and S are convex sets. This disproves the claim of f being a convex function. A few claims were taken from the Lecture 14, page 18.

Problem 2: Convex Compositions

a) The statement is *false*. One of the counterexamples is as follows:

Consider $f(x) = -x$ and $g(x) = \sqrt{x}$, for which both functions are concave. However, $f(g(x)) = -\sqrt{x}$ is convex. The functions' convexity and concavity claims are taken from the Lecture 14, page 9 and 11.

b) The statement is *false* (not necessarily *true*). The proof is as follows:

Let $x_1, x_2 \in \Omega$ and $\lambda \in [0,1]$, then $g(\lambda x_1 + (1-\lambda)x_2) \geq \lambda g(x_1) + (1-\lambda)g(x_2)$ due to the concavity of g. Subsequently, $(f \circ g)(\lambda x_1 + (1-\lambda)x_2) = f(g(\lambda x_1 + (1-\lambda)x_2)) \geq f(\lambda g(x_1) + (1-\lambda)g(x_2))$ due to the non-decreasing property of f. As $g(\Omega) \subseteq I$, we conclude that $g(x_1), g(x_2) \in I \Rightarrow f(\lambda g(x_1) + (1-\lambda)g(x_2)) \geq \lambda f(g(x_1)) + (1-\lambda)f(g(x_2)) = \lambda(f \circ g)(x_1) + (1-\lambda)(f \circ g)(x_2)$ using the concavity of f.

As a result, it is shown that $(f \circ g)(\lambda x_1 + (1-\lambda)x_2) \geq \lambda(f \circ g)(x_1) + (1-\lambda)(f \circ g)(x_2)$, which implies $f \circ g$ is concave (not necessarily convex unless $f \circ g$ is an affine function).

c) The statement is *false*. One of the counterexamples is as follows:

Consider $f(x) = -\sqrt{x}$, for which the function is convex. As $\sqrt{x} \geq 0 \Rightarrow -\sqrt{x} \leq 0$. This means $|f(x)| = -f(x)$, which implies $|f(x)|$ is concave.

Problem 3: Convex Functions

a) The triangle inequality, $|a+b| \leq |a| + |b|$, will be used in the proof. The proof is as follows:

Let $p, q \in R^n$ and $\lambda \in [0,1]$, then:

$$\begin{aligned} r(\lambda p + (1-\lambda)q) &= \max_i |\lambda p_i + (1-\lambda)q_i| \\ &\leq \max_i |\lambda p_i| + |(1-\lambda)q_i| \\ &\leq \max_i |\lambda p_i| + \max_i |(1-\lambda)q_i| \\ &= \lambda \max_i |p_i| + (1-\lambda) \max_i |q_i| \\ &= \lambda r(p) + (1-\lambda)r(q) \end{aligned}$$

Hence, r is shown to be a convex function.

b) For the followings, $\|\cdot\|^2$ will often appear and claimed to be convex. Proof will be provided here:

For $x, y \in R^n$ and $\theta \in [0,1]$:

$$\begin{aligned} \|\theta x + (1-\theta)y\|^2 &\leq (\theta \|x\| + (1-\theta)\|y\|)^2 \\ &= \theta^2 \|x\|^2 + 2\theta(1-\theta)\|x\|\|y\| + (1-\theta)^2 \|y\|^2 \\ &\leq \theta \|x\|^2 + (1-\theta)\|y\|^2 \end{aligned}$$

By the triangle inequality and the arithmetic-geometric mean inequality, $\|\cdot\|^2$ is convex.

- For a function $f(x) = \frac{x_1^2}{x_2}$, the Hessian matrix is derived as follows:

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

Then, the eigenvalues λ_1, λ_2 satisfy: $\lambda_1 + \lambda_2 = \frac{2x_1^2+2x_2^2}{x_2^3}$ and $\lambda_1\lambda_2 = 0$. One of them is 0 and the other one is $\frac{2x_1^2+2x_2^2}{x_2^3}$, which is positive as $x_2 > 0$. This shows the positive semi-definite of the Hessian matrix for f, which implies that f is convex.

- Let's claim that $\frac{1}{2}\|Ax-b\|^2$ is convex. As $Ax-b$ is an affine function and $\frac{1}{2}\|z\|^2$ is a convex quadratic function, the composition is convex under affine transformations and non-negative weighted sums, the claim is true. Then, by part a), $\|Lx\|_\infty = \max_i |Lx_i|$ is shown to be convex. By convex function properties taken from Lecture 14, page 15, f is convex.
- As $\|x\|^2$ is a convex quadratic function, so is $\frac{\lambda}{2}\|x\|^2$ provided that $\lambda > 0$, for which the λ constraint is guaranteed in this problem. Then, for a function, say $g_i(x,y) = \max\{0, 1 - b_i(a_i^T x + y))\}$, g_i is mapped to either 0 or $1 - b_i(a_i^T x + y)$, which is an affine function. In their linear mappings, both functions are convex, which implies that g_i is convex as well. Now, as $\sum_{i=1}^m g_i(x,y)$ is a summation over m convex functions, along with the addition of a convex function $\frac{\lambda}{2}\|x\|^2$, we deduce that f is convex by convex function properties taken from Lecture 14, page 15.

c) Let $h_y(x) = y^T x - f(y)$, then for a fixed y, h_y is linear. This implies that $g(x) = \sup_{y \in R^n} h_y(x)$, i.e., $g(x)$ is a supremum of $h_y(x), \forall y \in R^n$. As h_y is convex and forms a set g for all $y \in R^n$, g is convex by Lecture 14, page 17.

For each x, $g(x)$ can be computed by finding the y that maximizes $y^T x - \|y\|_1$. Since $\|y\|_1$ is the sum of the absolute values of the components of y, the maximum of $y^T x - \|y\|_1$ occurs when each component y_i has the same sign as the corresponding x_i and is at most as large in absolute value.

Therefore, for each i, the optimal y_i is $y_i = \text{sign}(x_i)$ if $|x_i| \leq 1$ and $y_i = x_i$ otherwise. This results in:

$$g(x) = \sum_{i:|x_i|\leq 1} x_i^2 + \sum_{i:|x_i|>1} |x_i|$$

Problem 4

(a) As A is a symmetric matrix, so is $I - A$. By the definition of positive semidefinite matrix, proving $I - A$ being positive semidefinite is equivalent to proving $x^T(I-A)x \geq 0, \forall x \neq 0$.

As A has all its components non-negative and the sum for each of its rows equal to 1, $x^T Ax$ is essentially the sum of the components of x. Let $x' = Ax$, then $|x'_i| \leq |x_i|, \forall i = \{1, 2, 3, 4\}$. As the elements of $x^T x$ and $x^T x'$ are greater than or equal to 0, we conclude that $x^T Ax$ is less than or equal to $x^T x$ by the previous statement. This implies that $x^T(I - A)x = x^T x - x^T Ax \geq 0$ holds true, which further implies that $I - A$ is positive semidefinite.

(b) We are interested in the Hessian matrix of f.

$$
\begin{aligned}
f(x) &= \log(1 + \exp(a^T x)) \\
\Rightarrow \nabla f(x) &= \frac{\exp(a^T x)}{1 + \exp(a^T x)} a \\
\Rightarrow \nabla^2 f(x) &= \frac{\exp(a^T x)}{(1 + \exp(a^T x))^2} a a^T
\end{aligned}
$$

As $\exp(a^T x)$ is larger than 0 and aa^T is an outer matrix that is a positive semidefinite matrix; this implies that $\nabla^2 f(x)$ is also a positive semidefinite matrix, which further implies that f is convex.

(c) For a t, let $y = xt$. Then the problem is equivalent to:

$$
\begin{aligned}
\min_{x \in R^n, t} \quad & \frac{\|Axt - bt\|}{c^T xt + dt} \\
\text{subject to} \quad & \|xt\| \leq t, c^T xt + dt > 0
\end{aligned}
$$

, which is also equivalent to

$$
\begin{aligned}
\min_{y \in R^n, t} \quad & \frac{\|Ay - bt\|}{c^T y + dt} \\
\text{subject to} \quad & \|y\| \leq t, c^T y + dt > 0
\end{aligned}
$$

The set of points (y, t) such that $c^T y + dt > 0$ forms a convex set because the condition describes a half-space and a norm ball, both of which are convex sets. The condition $c^T y + dt = 1$, which defines a hyperplane, is also a convex set. We can "set" up the value of c and d accordingly and find that:

$$
\begin{aligned}
\min_{y \in R^n, t} \quad & \|Ay - bt\| \\
\text{subject to} \quad & \|y\| \leq t, c^T y + dt = 1
\end{aligned}
$$

With these transformations, the original nonconvex problem is shown to be equivalent to a convex problem. The convex problem is easier to solve because it can be cast into a standard convex optimization form, for which the algorithms exist.

(d)
```python
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)
t = cp.Variable()

problem = cp.Problem(cp.Minimize(cp.norm(A @ y - b * t)), [cp.norm(y)
    <= t, c.T @ y + d * t == 1])
problem.solve()

print("(y, t) := (", y.value, ", " , t.value,  ")", sep = '')
print("The minimum value is ", problem.value)

```

```
(y, t) := ([3.96574855e-10 2.50000000e-01], 0.24999999982376303)
The minimum value is  1.9160503543852565e-10

```