DDA6010 AS3

October 19, 2024

Problem 1

(a)

For X_1:

The set is defined as

$$X_1 = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : x^T x \leq t^2\}$$

which can be rewritten as

$$\|x\| \leq |t|$$

This set represents the region between two cones in $\mathbb{R}^{n+1}$, one for $t \geq 0$ and one for $t \leq 0$.

To test convexity, consider two points (x_1, t_1) and (x_2, t_2) in X_1 with $x_1 \neq x_2$ and $t_1 \neq t_2$. Take $\lambda \in [0,1]$ and compute the convex combination:

$$(x,t) = \lambda(x_1, t_1) + (1-\lambda)(x_2, t_2)$$

We need to check whether $(x,t) \in X_1$, i.e.,

$$\|x\|^2 \leq t^2$$

Counterexample:

Let $n = 1$ for simplicity. Choose:

$$x_1 = 1,\ t_1 = 1 \quad \text{and} \quad x_2 = -1,\ t_2 = -1$$

Their midpoint is:

$$x = 0, \quad t = 0$$

Now, $\|x\|^2 = 0$, and $t^2 = 0$, so $\|x\|^2 \leq t^2$ holds at the midpoint.

But now, consider:

$$x_1 = 2,\ t_1 = 2 \quad \text{and} \quad x_2 = 2,\ t_2 = -2$$

Their midpoint is:

$$x = 2, \quad t = 0$$

Now, $\|x\|^2 = 4$, but $t^2 = 0$, which is false.

Thus, the convex combination is not in X_1.

Conclusion for X_1: The set X_1 is **not convex**.

For X_2: The set is defined as

$$X_2 = \{x \in \mathbb{R}^n : (a^T x)^2 \leq \alpha\}, \quad \alpha \geq 0$$

This is a sublevel set of the convex function $f(x) = (a^T x)^2$, which is convex because it is the composition of convex functions.

Conclusion for X_2: The set X_2 is **convex**.

(b)

Any point x in the convex hull can be written as:

$$x = \lambda_1 e_1 + \lambda_2 e_2 + \lambda_3(-e_1) + \lambda_4(-e_2)$$

where $\lambda_i \geq 0$ and $\sum_{i=1}^{4} \lambda_i = 1$. Simplifying x, we get:

$$x = (\lambda_1 - \lambda_3, \lambda_2 - \lambda_4)$$

Let $\alpha = \lambda_1 - \lambda_3$ and $\beta = \lambda_2 - \lambda_4$, then:

$$|\alpha| \leq \lambda_1 + \lambda_3, \quad |\beta| \leq \lambda_2 + \lambda_4$$

Since $\sum_{i=1}^{4} \lambda_i = 1$, we have:

$$|\alpha| + |\beta| \leq 1$$

Thus, any point in the convex hull satisfies $|x_1| + |x_2| \leq 1$.

Conversely, any point $x \in \mathbb{R}^2$ with $|x_1| + |x_2| \leq 1$ can be expressed as a convex combination of $e_1, e_2, -e_1, -e_2$ by choosing appropriate λ_i.

Conclusion: The convex hull is exactly the set $\{x \in \mathbb{R}^2 : |x_1| + |x_2| \leq 1\}$.

(c)

The set S is the circle centered at $(1, 0)$ with radius 1:

$$S = \{(x_1, x_2) : (x_1 - 1)^2 + x_2^2 = 1\}$$

To find the conic hull of S, denoted cone(S), we parametrize points on S as:

$$x = (1 + \cos\theta, \sin\theta), \quad \theta \in [0, 2\pi)$$

Since S is a circle, thus we can express the conic hull by a conic combination which is $y = \lambda x$ with $\lambda \geq 0$. For all θ, $1 + \cos\theta \geq 0$, so the x_1-component of x is $x_1 = 1 + \cos\theta \geq 0$. The only point where $x_1 = 0$ is at $\theta = \pi$, which gives $x = (0, 0)$.
Thus, the conic hull is:

$$\text{cone}(S) = \{\lambda x : \lambda \geq 0,\, x \in S\} = \{(x_1, x_2) : x_1 > 0\} \cup \{(0, 0)\}$$

(d)

We will prove the statement by showing both implications:

1. If x is an extreme point of S, then $S \setminus \{x\}$ is convex.
2. If $S \setminus \{x\}$ is convex, then x is an extreme point of S.

Proof of (1)
Assume x is an extreme point of S. We need to show that $S \setminus \{x\}$ is convex.
Let $y, z \in S \setminus \{x\}$ and $\lambda \in [0, 1]$. Since S is convex and $y, z \in S$, the convex combination

$$w = \lambda y + (1 - \lambda) z$$

belongs to S.
We need to show that $w \in S \setminus \{x\}$, i.e., $w \neq x$.
Suppose, for contradiction, that $w = x$. Then

$$x = \lambda y + (1 - \lambda) z.$$

Since $y \neq x$ and $z \neq x$, and $\lambda \in (0, 1)$, this expresses x as a convex combination of two distinct points in S.
This contradicts the assumption that x is an extreme point of S. Therefore, $w \neq x$, and thus $w \in S \setminus \{x\}$.
Conclusion: $S \setminus \{x\}$ is convex.

Proof of (2)
Assume $S \setminus \{x\}$ is convex. We need to show that x is an extreme point of S.
Suppose, for contradiction, that x is **not** an extreme point of S. Then there exist $y, z \in S$, with $y \neq x$ and $z \neq x$, and $\lambda \in (0, 1)$ such that

$$x = \lambda y + (1 - \lambda) z.$$

Now, since $y, z \in S$ and $y \neq x$, $z \neq x$, we have $y, z \in S \setminus \{x\}$.
Since $S \setminus \{x\}$ is convex, the convex combination

$$w = \lambda y + (1 - \lambda) z$$

should belong to $S \setminus \{x\}$.

But $w = x$, which implies $x \in S \setminus \{x\}$, a contradiction.

Therefore, our assumption that x is not an extreme point must be false.

Problem 2

(a)

Proof of Convexity:Let $y_1, y_2 \in \mathbb{R}^n$ and $\lambda \in [0, 1]$.

Consider:

$$\sigma_C(\lambda y_1 + (1-\lambda)y_2) = \sup_{x \in C} (\lambda y_1 + (1-\lambda)y_2)^\top x.$$

By linearity of the dot product:

$$(\lambda y_1 + (1-\lambda)y_2)^\top x = \lambda y_1^\top x + (1-\lambda)y_2^\top x.$$

Thus:

$$\sigma_C(\lambda y_1 + (1-\lambda)y_2) = \sup_{x \in C} \left[\lambda y_1^\top x + (1-\lambda)y_2^\top x\right].$$

Since the supremum of a sum is less than or equal to the sum of suprema:

$$\sigma_C(\lambda y_1 + (1-\lambda)y_2) \leq \lambda \sup_{x \in C} y_1^\top x + (1-\lambda) \sup_{x \in C} y_2^\top x = \lambda \sigma_C(y_1) + (1-\lambda)\sigma_C(y_2).$$

Conclusion: σ_C is convex.

Proof of Closedness (Lower Semicontinuity):A function $f : \mathbb{R}^n \to (-\infty, \infty]$ is closed if its epigraph is a closed set.

The epigraph of σ_C is:

$$\text{epi}(\sigma_C) = \{(y, t) \in \mathbb{R}^n \times \mathbb{R} : t \geq \sup_{x \in C} y^\top x\}.$$

For each fixed $x \in C$, the function $y \mapsto y^\top x$ is continuous (linear in y). The supremum over $x \in C$ of continuous functions is *lower semicontinuous*.

Therefore, σ_C is lower semicontinuous, and its epigraph is closed.

Alternative Argument:

Since σ_C is convex, and convex functions are lower semicontinuous on the interior of their domain, and σ_C is finite everywhere (assuming C is bounded; if not, σ_C can take ∞), it follows that σ_C is closed.

Conclusion: σ_C is closed.

(b)

Let $x_1, x_2 \in \mathbb{R}^n$ and $\lambda \in [0, 1]$.

Consider $x = \lambda x_1 + (1-\lambda)x_2$. Compute $g(x)$:

$$g(x) = \sup_{y \in C} f(\lambda x_1 + (1-\lambda)x_2, y).$$

Since $f(\cdot, y)$ is convex, we have for each $y \in C$:

$$f(\lambda x_1 + (1-\lambda)x_2, y) \le \lambda f(x_1, y) + (1-\lambda) f(x_2, y).$$

Taking the supremum over $y \in C$ on both sides:

$$g(x) = \sup_{y \in C} f(\lambda x_1 + (1-\lambda)x_2, y) \le \sup_{y \in C} \left[\lambda f(x_1, y) + (1-\lambda) f(x_2, y)\right].$$

$$\le \lambda \sup_{y \in C} f(x_1, y) + (1-\lambda) \sup_{y \in C} f(x_2, y) = \lambda g(x_1) + (1-\lambda) g(x_2).$$

Conclusion: g is convex.

(c)

Let $x_1, x_2 \in \mathbb{R}^n$. Define:

- $y_1^* = y^*(x_1) = \arg\max_{y \in C} f(x_1, y)$.
- $y_2^* = y^*(x_2) = \arg\max_{y \in C} f(x_2, y)$.

Since $f(x, \cdot)$ is μ-strongly concave and C is convex, the maximizer $y^*(x)$ is unique for each x.

Step 1: Optimality ConditionsThe first-order optimality conditions for y_1^* and y_2^* are: 1. For y_1^*:

$$\langle \nabla_y f(x_1, y_1^*), y - y_1^* \rangle \le 0, \quad \forall y \in C.$$

2. For y_2^*:

$$\langle \nabla_y f(x_2, y_2^*), y - y_2^* \rangle \le 0, \quad \forall y \in C.$$

Step 2: Relate y_1^* and y_2^*Consider $y = y_2^*$ in the optimality condition for y_1^*:

$$\langle \nabla_y f(x_1, y_1^*), y_2^* - y_1^* \rangle \le 0.$$

Similarly, consider $y = y_1^*$ in the optimality condition for y_2^*:

$$\langle \nabla_y f(x_2, y_2^*), y_1^* - y_2^* \rangle \le 0.$$

Adding these two inequalities:

$$\langle \nabla_y f(x_1, y_1^*) - \nabla_y f(x_2, y_2^*), y_2^* - y_1^* \rangle \le 0.$$

Step 3: Decompose the Gradient DifferenceWe can decompose the difference $\nabla_y f(x_1, y_1^*) - \nabla_y f(x_2, y_2^*)$ as:

$$\nabla_y f(x_1, y_1^*) - \nabla_y f(x_2, y_2^*) = \left[\nabla_y f(x_1, y_1^*) - \nabla_y f(x_1, y_2^*)\right] + \left[\nabla_y f(x_1, y_2^*) - \nabla_y f(x_2, y_2^*)\right].$$

Step 4: Apply Strong Concavity and Smoothness**Strong Concavity in** y: Since $f(x, \cdot)$ is μ-strongly concave, its gradient with respect to y is μ-Lipschitz continuous:

$$\|\nabla_y f(x, y_1^*) - \nabla_y f(x, y_2^*)\| \leq \mu \|y_1^* - y_2^*\|.$$

Therefore, we have:

$$\langle \nabla_y f(x_1, y_1^*) - \nabla_y f(x_1, y_2^*), y_2^* - y_1^* \rangle \leq -\mu \|y_1^* - y_2^*\|^2.$$

Smoothness in x: Since f is L-smooth in x, the gradient $\nabla_y f(x, y)$ is Lipschitz in x:

$$\|\nabla_y f(x_1, y_2^*) - \nabla_y f(x_2, y_2^*)\| \leq L \|x_1 - x_2\|.$$

Using the Cauchy-Schwarz inequality:

$$\langle \nabla_y f(x_1, y_2^*) - \nabla_y f(x_2, y_2^*), y_2^* - y_1^* \rangle \leq L \|x_1 - x_2\| \|y_2^* - y_1^*\|.$$

Step 5: Combine the InequalitiesPutting everything together:

$$0 \geq \langle \nabla_y f(x_1, y_1^*) - \nabla_y f(x_2, y_2^*), y_2^* - y_1^* \rangle$$

$$= \langle \nabla_y f(x_1, y_1^*) - \nabla_y f(x_1, y_2^*), y_2^* - y_1^* \rangle + \langle \nabla_y f(x_1, y_2^*) - \nabla_y f(x_2, y_2^*), y_2^* - y_1^* \rangle$$

$$\leq -\mu \|y_1^* - y_2^*\|^2 + L \|x_1 - x_2\| \|y_2^* - y_1^*\|.$$

Rewriting the inequality:

$$-\mu \|y_1^* - y_2^*\|^2 + L \|x_1 - x_2\| \|y_1^* - y_2^*\| \geq 0.$$

Step 6: Solve for $\|y_1^* - y_2^*\|$Bring all terms to one side:

$$\mu \|y_1^* - y_2^*\|^2 \leq L \|x_1 - x_2\| \|y_1^* - y_2^*\|.$$

Assuming $y_1^* \neq y_2^*$ (if they are equal, the Lipschitz condition holds trivially), we can divide both sides by $\|y_1^* - y_2^*\|$:

$$\mu \|y_1^* - y_2^*\| \leq L \|x_1 - x_2\|.$$

Therefore:

$$\|y_1^* - y_2^*\| \leq \frac{L}{\mu} \|x_1 - x_2\| = \kappa \|x_1 - x_2\|.$$

Problem 3

Part 1: (a) $\Rightarrow$ (b)

Let $x, y \in \mathbb{R}^n$ and $\lambda \in [0, 1]$.

Since $1/f$ is concave:

$$\frac{1}{f(\lambda x+(1-\lambda)y)} \geq \lambda\frac{1}{f(x)} + (1-\lambda)\frac{1}{f(y)}.$$

Take reciprocals:

Because $f > 0$, the reciprocal function $s \mapsto 1/s$ is decreasing on $(0,\infty)$. Therefore, reversing the inequality:

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

Define:

$$H = \lambda\frac{1}{f(x)} + (1-\lambda)\frac{1}{f(y)}.$$

Then:

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

Now, take natural logarithms:

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

But $-\log(H)$ can be related to $\lambda\log(f(x))$ and $(1-\lambda)\log(f(y))$.

Consider the function $h(s) = -\log(s)$, which is convex on $(0,\infty)$ since $h''(s) = 1/s^2 > 0$.

Since H is a convex combination of $1/f(x)$ and $1/f(y)$:

$$H = \lambda a + (1-\lambda)b, \quad \text{where} \quad a = \frac{1}{f(x)}, \quad b = \frac{1}{f(y)}.$$

Applying Jensen's inequality to h:

$$-\log(H) \leq \lambda(-\log(a)) + (1-\lambda)(-\log(b)) = \lambda\log(f(x)) + (1-\lambda)\log(f(y)).$$

Therefore:

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

This shows that $\log(f)$ is convex.

Part 2: (b) $\Rightarrow$ (c)

Since $\log(f)$ is convex, we have:

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

Exponentiate both sides (since the exponential function preserves the inequality direction):

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

Now, observe that the right-hand side is less than or equal to:

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

This follows from the generalized Arithmetic-Geometric Mean inequality (given in the hint):

For $a, b \geq 0$ with $a + b = 1$:

$$af(x) + bf(y) \geq f(x)^a f(y)^b.$$

Substituting $a = \lambda$ and $b = 1 - \lambda$:

$$\lambda f(x) + (1-\lambda) f(y) \geq f(x)^\lambda f(y)^{1-\lambda}.$$

Thus, we have:

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

Therefore:

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

This inequality shows that f is convex.

Conclusion of Implications:

(a) $\Rightarrow$ (b): If $1/f$ is concave, then $\log(f)$ is convex.

(b) $\Rightarrow$ (c): If $\log(f)$ is convex, then f is convex.

Counterexample

Counterexample for (c) $\Rightarrow$ (b) Not TrueLet $f(x) = x^2$ on $x \geq 0$.

- f is convex on $[0, \infty)$.
- $\log(f(x)) = \log(x^2) = 2\log(x)$, which is concave on $(0, \infty)$ since the second derivative:

$$(\log(f(x)))'' = -\frac{2}{x^2} < 0.$$

Thus, f is convex, but $\log(f)$ is concave, not convex.

Counterexample for (b) $\Rightarrow$ (a) Not TrueLet $f(x) = e^{x^2}$.

- $\log(f(x)) = x^2$, which is convex.
- $f(x) = e^{x^2}$ is convex.
- $1/f(x) = e^{-x^2}$ is neither convex nor concave.

Compute the second derivative of $1/f(x)$:

$$\left(\frac{1}{f(x)}\right)'' = e^{-x^2}\left(4x^2 - 2\right).$$

- When $x = 0$, $\left(\frac{1}{f(x)}\right)'' = -2 < 0$.
- When $|x| > \sqrt{0.5}$, $\left(\frac{1}{f(x)}\right)'' > 0$.

Therefore, $1/f$ is neither convex nor concave on $\mathbb{R}$.

This shows that even though $\log(f)$ is convex, $1/f$ is not necessarily concave.

Problem 4

1. Necessity:

Assume x^* is a global minimizer of the problem.

That is, for all $x \in X$:

$$f(x^*) + \varphi(x^*) \leq f(x) + \varphi(x).$$

Goal: Show that:

$$\nabla f(x^*)^\top (x - x^*) + \varphi(x) - \varphi(x^*) \geq 0, \quad \forall x \in X.$$

Proof:

- Since f is convex and differentiable, it satisfies the first-order convexity inequality:

$$f(x) \geq f(x^*) + \nabla f(x^*)^\top (x - x^*), \quad \forall x \in X.$$

- Subtract $f(x^*) + \varphi(x^*)$ from both sides of the inequality defining x^* as a minimizer:

$$[f(x) - f(x^*)] + [\varphi(x) - \varphi(x^*)] \geq 0, \quad \forall x \in X.$$

- Using the first-order convexity inequality for f:

$$f(x) - f(x^*) \geq \nabla f(x^*)^\top (x - x^*), \quad \forall x \in X.$$

- Combine the inequalities:

$$\nabla f(x^*)^\top (x - x^*) + \varphi(x) - \varphi(x^*) \geq 0, \quad \forall x \in X.$$

Conclusion: The optimality condition holds.

2. Sufficiency:

Assume that:

$$\nabla f(x^*)^\top (x - x^*) + \varphi(x) - \varphi(x^*) \geq 0, \quad \forall x \in X.$$

Goal: Show that x^* is a global minimizer, i.e.,

$$f(x^*) + \varphi(x^*) \leq f(x) + \varphi(x), \quad \forall x \in X.$$

Proof:

- From the convexity and differentiability of f:

$$f(x) \geq f(x^*) + \nabla f(x^*)^\top (x - x^*), \quad \forall x \in X.$$

- Substitute this into the optimality condition:

$$[f(x) - f(x^*)] + [\varphi(x) - \varphi(x^*)] \geq 0, \quad \forall x \in X.$$

- **Rewriting:**

$$f(x) + \varphi(x) \geq f(x^*) + \varphi(x^*), \quad \forall x \in X.$$

Conclusion: x^* is a global minimizer of the problem.

Final Conclusion:

- **Necessity:** If x^* is a global minimizer, the optimality condition holds.
- **Sufficiency:** If the optimality condition holds, x^* is a global minimizer.

Therefore, x^* is a global solution to the optimization problem if and only if the optimality condition is satisfied:

$$x^* \in X, \quad \text{and} \quad \nabla f(x^*)^\top (x - x^*) + \varphi(x) - \varphi(x^*) \geq 0, \quad \forall x \in X.$$