Assignment 1 - CS703 Optimisation and Computing

HOANG Thi Linh

March 10, 2025

1 Convex sets

1.1

Show that a convex combination of a finite number of elements of a convex set also belongs to that set.

Proof. Let C be a convex set and $x_1, x_2, ..., x_n \in C$. Suppose we have $\theta_1, \theta_2, .., \theta_n$ with $\theta_i \geq 0$ and $\sum_{i=1}^{n} \theta_i = 1$. We must show that:

$$x = \sum_{i=1}^{n} \theta_i x_i \in C$$

Now we prove that by induction.
Suppose that it's true for $n = r$, then will prove that it also holds for $n = r + 1$.
Since C is convex the result is true, trivially, for $n = 1$ and by definition for $n = 2$. Consider the convex combination $\theta_1 x_1 + \theta_2 x_2 + ...\theta_{r+1} x_{r+1}$ and define $\Theta := \sum_{i=1}^{r} \theta_i$. Since $1 - \Theta = \theta_{r+1}$, we have

$$\left(\sum_{i=1}^{r} \frac{\theta_i}{\Theta} x_i \right) + \theta_{r+1} x_{r+1} = \Theta \left(\sum_{i=1}^{r} \frac{\theta_i}{\Theta} x_i \right) + \big(1 - \Theta\big) x_{r+1}.$$

Note that $\sum_{i=1}^{r} \frac{\theta_i}{\Theta} = 1$ and so, by the induction hypothesis $\left(\sum_{i=1}^{r} \frac{\theta_i}{\Theta} x_i \right) \in C$. Since $x_{r+1} \in C$, it follows that the right-hand side is a convex combination of two points of C and hence lies in C. $\square$

1.2

$$S_1 = \{x \in \mathbb{R}^2 \mid \|x\|_\infty \leq 1\} \quad \text{and} \quad S_2 = \{x \in \mathbb{R}^2 \mid \|x\|_2 \leq 1\}.$$

(a). $S_1 \cup S_2$ is convex. **True**.
For any $x \in S_2$ we have

$$\|x\|_2 \leq 1 \quad \Rightarrow \quad |x_1| \leq \|x\|_2 \leq 1 \quad \text{and} \quad |x_2| \leq \|x\|_2 \leq 1,$$

which implies $x \in S_1$. Thus, $S_2 \subset S_1$ and consequently,

$$S_1 \cup S_2 = S_1.$$

Since S_1 (the square $[-1, 1] \times [-1, 1]$) is convex, the union is convex.
(b). $S_1 \cap S_2$ is convex. **True**
Both S_1 and S_2 are convex sets. Moreover, since $S_2 \subset S_1$, it follows that $S_1 \cap S_2 = S_2$, which is convex.
(c) $S_1 \setminus S_2$ is convex. **False**.
The set $S_1 \setminus S_2$ represents the region in the square S_1 that lies outside the circle S_2. One can easily find two points in $S_1 \setminus S_2$ whose connecting line segment passes through the interior of S_2, thereby leaving $S_1 \setminus S_2$. Hence, this set is not convex.

1.3

Prove that the hyperbolic set $S = \{x_1, x_2\} \in \mathbb{R}^2_+ \mid x_1x_2 \geq 1\}$ is convex.

Proof. If $x_1x_2 \geq 1$, then neither x_1 nor x_2 can be zero; that is, every point in S actually satisfies $x_1 > 0$ and $x_2 > 0$. Hence, we can equivalently write

$$S = \{(x_1, x_2) \in \mathbb{R}^2_{++} \mid x_1x_2 \geq 1\}.$$

Taking logarithms (which is valid since $x_1, x_2 > 0$), we have

$$\ln(x_1x_2) = \ln x_1 + \ln x_2 \geq 0.$$

Since the function $\ln x$ is concave on $\mathbb{R}_{++}$, its superlevel set

$$\{(x_1, x_2) \in \mathbb{R}^2_{++} \mid \ln x_1 + \ln x_2 \geq 0\}$$

is convex. Therefore, the set S is convex. □

2 Convex functions

2.1

With $0 < \alpha \leq 1$, we consider the function:

$$f(x) = \frac{x^\alpha - 1}{\alpha}, x \geq 0$$

We have the first derivative of $f(x)$ with respect to x:

$$\nabla f(x) = x^{\alpha-1},$$

and second derivative:

$$\nabla^2 f(x) = (\alpha - 1)x^{\alpha-2}.$$

When $0 < \alpha \leq 1$, we have $\alpha - 1 \ < 0$, so that $\nabla^2 f(x) < 0$ for all $x > 0$. Then, $f(x)$ is strictly concave. When $\alpha = 1$, $f(x) = x - 1$ is a linear function and it is both convex and concave on its domain.

2.2

Let

$$f_1(x) = \prod_{i=1}^{n} x_i^{\theta_i} \quad \text{and} \quad f_2(x) = \sum_{i=1}^{n} \theta_i x_i,$$

with the weights satisfying $\theta_i \geq 0$ and $\sum_{i=1}^n \theta_i = 1$.
(a) Showing that $f_1(x) \leq f_2(x)$:
Taking the natural logarithm of f_1 we obtain

$$\ln f_1(x) = \sum_{i=1}^{n} \theta_i \ln x_i.$$

Since the logarithm function is concave on $\mathbb{R}_{++}$, Jensen's inequality implies

$$\sum_{i=1}^{n} \theta_i \ln x_i \leq \ln\Big(\sum_{i=1}^{n} \theta_i x_i\Big) = \ln f_2(x).$$

Exponentiating both sides gives

$$f_1(x) \leq f_2(x),$$

which is the weighted version of the arithmetic mean–geometric mean inequality.
(b) Showing that f_1 is concave:
Since

$$\ln f_1(x) = \sum_{i=1}^{n} \theta_i \ln x_i$$

is a concave function (as a weighted sum of concave functions), it follows that $f_1(x)$ is *log-concave*. In this setting, one can show that the weighted geometric mean f_1 is indeed concave on $\mathbb{R}^n_{++}$.

2.3

Let x be a real-valued random variable taking values in $\{a_1, a_2, \ldots, a_n\}$ with $a_1 < a_2 < \cdots < a_n$ and with probabilities

$$\mathbb{P}(x = a_i) = p_i, \quad \text{for } i = 1, \ldots, n.$$

The functions below are defined on the probability simplex

$$\{p \in \mathbb{R}^n_+ \mid \sum_{i=1}^{n} p_i = 1\}.$$

(i) $f(p) = \mathbb{E}[x] = \sum_{i=1}^{n} p_i a_i$:
This is an affine function of p and is therefore both convex and concave.
(ii) $f(p) = \mathbb{P}(x \geq \alpha) = \sum_{a_i \geq \alpha} p_i$:
This is also a linear (affine) function of p and is both convex and concave.
(iii) $f(p) = \sum_{i=1}^{n} p_i \ln p_i$:
Since the function $\phi(p) = p \ln p$ is convex for $p > 0$ (as its second derivative $\phi''(p) = 1/p > 0$), the sum $f(p)$ is convex.
(iv) $f(p) = \mathrm{Var}(x) = \mathbb{E}[x^2] - \big(\mathbb{E}[x]\big)^2$:
Note that $\mathbb{E}[x^2] = \sum_{i=1}^{n} p_i a_i^2$ is linear in p, and $\big(\mathbb{E}[x]\big)^2$ is convex (being the square of an affine function). The negative of a convex function is concave, so the difference is concave. Hence, $\mathrm{Var}(x)$ is concave as a function of p.

3 Duality

3.1

Consider the primal problem

$$\begin{aligned} \min_{x_1, x_2} \quad & x_1 - 2x_2 \\ \text{s.t.} \quad & 2x_1 - x_2 = -5, \\ & 2x_1 - x_2 = 3. \end{aligned}$$

(a) Derivation of the Dual:
Suppose we have Lagrange multipliers ν_1 and ν_2 for the two equality constraints. The Lagrangian is

$$\begin{aligned} L(x_1, x_2, \nu_1, \nu_2) &= x_1 - 2x_2 + \nu_1\big(2x_1 - x_2 + 5\big) + \nu_2\big(2x_1 - x_2 - 3\big) \\ &= x_1 - 2x_2 + 2\nu_1 x_1 - \nu_1 x_2 + 5\nu_1 + 2\nu_2 x_1 - \nu_2 x_2 - 3\nu_2 \\ &= \Big(1 + 2\nu_1 + 2\nu_2\Big)x_1 - \Big(2 + \nu_1 + \nu_2\Big)x_2 + 5\nu_1 - 3\nu_2. \end{aligned}$$

For the Lagrangian to be bounded below (i.e., for the infimum over x_1 and x_2 to be finite), the coefficients of x_1 and x_2 must vanish:

$$1 + 2\nu_1 + 2\nu_2 = 0 \quad \text{and} \quad 2 + \nu_1 + \nu_2 = 0.$$

These conditions imply

$$\nu_1 + \nu_2 = -\frac{1}{2} \quad \text{and} \quad \nu_1 + \nu_2 = -2,$$

which is a contradiction.

(b) Primal and Dual Optimal Values:

Since the constraints $2x_1 - x_2 = -5$ and $2x_1 - x_2 = 3$ are inconsistent, the primal problem is infeasible (its optimal value is $+\infty$). In contrast, no λ can render the Lagrangian bounded below, and thus the dual optimal value is $-\infty$.

(c) Duality Gap:

There is an infinite duality gap since the primal is infeasible while the dual is unbounded.

3.2

Consider the problem

$$\begin{array}{ll} \min\limits_{x \in \mathbb{R}} & x^2 + 2x + 4 \\ \text{s.t.} & x^2 - 5x \leq -4. \end{array}$$

(a) Convexity:

The objective function $x^2 + 2x + 4$ is convex (being a quadratic function with a positive coefficient on x^2). The constraint can be rewritten as

$$x^2 - 5x + 4 \leq 0.$$

Since $x^2 - 5x + 4$ is also a quadratic function with a positive quadratic coefficient, its sublevel set is convex. Hence, the problem is convex.

(b) Derivation of the Dual:

Introduce a Lagrange multiplier $\lambda \geq 0$ for the inequality constraint. The Lagrangian is

$$L(x, \lambda) = x^2 + 2x + 4 + \lambda(x^2 - 5x + 4) = (1 + \lambda)x^2 + (2 - 5\lambda)x + (4 + 4\lambda).$$

Since $1 + \lambda > 0$ for $\lambda \geq 0$, the minimizer in x is obtained by differentiating with respect to x and setting the derivative to zero:

$$\frac{\partial L}{\partial x} = 2(1 + \lambda)x + (2 - 5\lambda) = 0 \quad \Longrightarrow \quad x^* = -\frac{2 - 5\lambda}{2(1 + \lambda)}.$$

Substituting x^* back into the Lagrangian gives the dual function:

$$g(\lambda) = 4 + 4\lambda - \frac{(2 - 5\lambda)^2}{4(1 + \lambda)}.$$

Thus, the dual problem is:

$$\begin{array}{ll} \max\limits_{\lambda} & 4 + 4\lambda - \frac{(2-5\lambda)^2}{4(1+\lambda)} \\ \text{s.t.} & \lambda \geq 0. \end{array}$$

(c) KKT Conditions:

The Karush-Kuhn-Tucker (KKT) conditions for this problem are:

1. **Primal feasibility:** $x^2 - 5x + 4 \leq 0$.
2. **Dual feasibility:** $\lambda \geq 0$.
3. **Complementary slackness:** $\lambda(x^2 - 5x + 4) = 0$.
4. **Stationarity:** $\dfrac{\partial L}{\partial x} = 2(1 + \lambda)x + (2 - 5\lambda) = 0$.

(d) Solving the KKT Conditions:

From the complementary slackness condition we can have two case $\lambda = 0$ and $x^2 - 5x + 4 = 0$.

Case 1: Inactive contraint $\lambda = 0$ then following the stationary codition we have

$$2x + 2 = 0 \Longrightarrow x = -1$$

This solution does not hold the primial fesibility condition.
Case 2: the constraint is active $x^2 - 5x + 4 = 0$ so that $\lambda > 0$. Factoring the constraint gives

$$(x-1)(x-4) = 0 \quad \Longrightarrow \quad x = 1 \quad \text{or} \quad x = 4.$$

Now, using the stationarity condition:

- For $x = 1$:
$$2(1+\lambda)(1) + (2-5\lambda) = 2 + 2\lambda + 2 - 5\lambda = 4 - 3\lambda = 0 \quad \Longrightarrow \quad \lambda = \frac{4}{3}.$$
This solution satisfies $\lambda > 0$ and the constraint.

- For $x = 4$:
$$2(1+\lambda)(4) + (2-5\lambda) = 8 + 8\lambda + 2 - 5\lambda = 10 + 3\lambda = 0 \quad \Longrightarrow \quad \lambda = -\frac{10}{3},$$
which is not feasible since $\lambda \geq 0$.

Thus, the valid solution is $x^* = 1$ and $\lambda^* = \frac{4}{3}$.
(e) Optimal Objective Value:
Substitute $x^* = 1$ into the objective function:

$$f(1) = 1^2 + 2 \cdot 1 + 4 = 7.$$

(f) Slater's Condition and Strong Duality:
Slater's condition requires the existence of a strictly feasible point. For example, $x = 2$ yields

$$2^2 - 5 \cdot 2 + 4 = 4 - 10 + 4 = -2 < 0,$$

so a strictly feasible point exists. Therefore, Slater's condition is satisfied, and strong duality holds.

4 Modeling and Solving Optimization Problems

4.1 Reformulating Norm Approximations as Linear Programs

ℓ_∞-Norm Approximation: We want to solve

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_\infty.$$

Introduce an auxiliary variable $t \geq 0$ and rewrite the problem as

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{s.t.} \quad & (Ax-b)_i \leq t, \quad i = 1, \ldots, m, \\ & -(Ax-b)_i \leq t, \quad i = 1, \ldots, m. \\ & t \geq 0 \end{aligned}$$

This formulation is equivalent to the original problem since minimizing t forces the maximum absolute deviation $\|Ax - b\|_\infty$ to be as small as possible while the constraints ensure that every component satisfies $|(Ax-b)_i| \leq t$.

ℓ_1-Norm Approximation: For the problem

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_1 = \min_x \sum_{i=1}^{m} |(Ax - b)_i|,$$

introduce auxiliary variables $t_i \geq 0$ for $i = 1, \ldots, m$ such that

$$\begin{array}{rl} \min\limits_{x, t_1, \ldots, t_m} & \sum_{i=1}^{m} t_i \\ \text{s.t.} & (Ax - b)_i \leq t_i, \quad i = 1, \ldots, m, \\ & -(Ax - b)_i \leq t_i, \quad i = 1, \ldots, m. \\ & t \geq 0 \end{array}$$

Here, the auxiliary variables t_i capture the absolute values $|(Ax - b)_i|$, so the LP is equivalent to the original ℓ_1 minimization problem.

4.2 Transportation Network Problem

Consider a directed network $G = (V, E)$ where each edge $(i, j) \in E$ has an associated transportation cost c_{ij} (per unit of goods) and a capacity d_{ij} (upper limit on the goods transported through that edge). The objective is to transport r units of goods from a source node s to a target node t at minimum cost.

(a) LP Formulation: Let x_{ij} denote the flow on edge (i, j). The linear program is formulated as:

$$\begin{array}{rl} \min\limits_{x} & \sum\limits_{(i,j) \in E} c_{ij}\, x_{ij} \\ \text{s.t.} & \sum\limits_{j:\,(v,j) \in E} x_{vj} - \sum\limits_{i:\,(i,v) \in E} x_{iv} = b_v, \quad \forall\, v \in V, \\ & 0 \leq x_{ij} \leq d_{ij}, \quad \forall\, (i, j) \in E. \end{array}$$

Here, the node supply/demand b_v is defined as

$$b_v = \begin{cases} r, & \text{if } v = s, \\ -r, & \text{if } v = t, \\ 0, & \text{otherwise.} \end{cases}$$

The first constraints enforce the required net flow at the source and target nodes, while the second ensures flow conservation at all other nodes. The capacity constraints guarantee that the flow on each edge does not exceed its limit.

(b) Solve the primal problem and show the optimal

First, the given network is labeled as Figure 1. Then, add the decision variables, objective, and constraints of flows, and solve the problem by CVXOPT.

(c) Extract the value of dual variables.

The codes for parts (b) and (c) are provided in the CS703_Assignment_1.ipynb that is attached to this assignment.

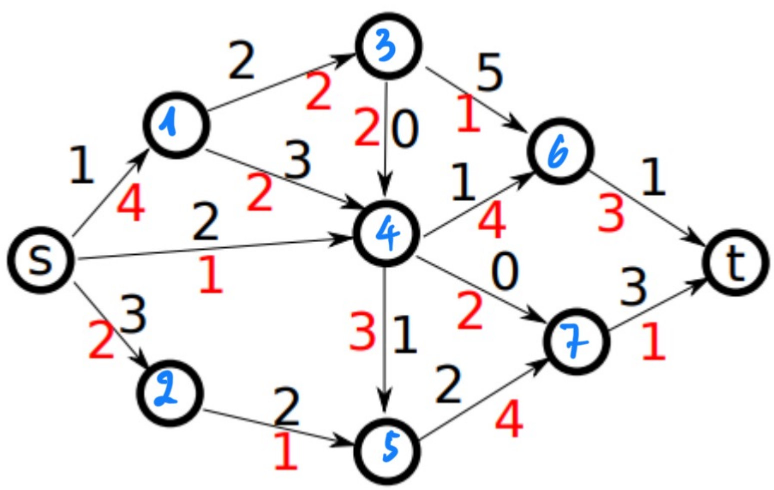

Figure 1: The transportation network with labels