Convex Optimization - HW3

IGOR LUCINDO CARDOSO

October 20, 2025

Disclaimer

This work was assisted by ChatGPT in the following ways: improvement of text content, clarification of mathematical definitions, and rapid generation of graphs and tables.

Problem 1

1.1 — Determine the lifted cone $\mathcal{C}_U$:

$$\mathcal{C}_U = \{\lambda(x, 1) \in \mathbb{R}^3 \mid x \in U,\ \lambda > 0\}.$$

By substituting the definition of $U = \{x \in \mathbb{R}^2_{\geq 0} \mid \langle 1, x\rangle = 1\}$, we get

$$\mathcal{C}_U = \{\lambda(x, 1) \in \mathbb{R}^3 \mid \langle 1, x\rangle = 1,\ x \in \mathbb{R}^2_{\geq 0},\ \lambda > 0\}.$$

We can equivalently write

$$\mathcal{C}_U = \{(x, \lambda) \in \mathbb{R}^3 \mid \langle 1, x\rangle = \lambda,\ x \in \mathbb{R}^2_{\geq 0},\ \lambda > 0\}.$$

Thus,

$$\mathcal{C}_U = \{(x, \langle 1, x\rangle) \in \mathbb{R}^3 \mid \langle 1, x\rangle > 0,\ x \in \mathbb{R}^2_{\geq 0}\}.$$

1.2 — Compute the projection $\mathrm{Proj}_{\mathbb{R}^2}(\mathcal{C}_U)$:

$$\mathrm{Proj}_{\mathbb{R}^2}(\mathcal{C}_U) = \{x \in \mathbb{R}^2 \mid \exists\lambda > 0 \text{ such that } (x, \lambda) \in \mathcal{C}_U\}.$$

Substituting the definition of $\mathcal{C}_U$, we obtain

$$\mathrm{Proj}_{\mathbb{R}^2}(\mathcal{C}_U) = \{x \in \mathbb{R}^2 \mid x \in \mathbb{R}^2_{\geq 0},\ \langle 1, x\rangle > 0\}.$$

This simplifies to

$$\mathrm{Proj}_{\mathbb{R}^2}(\mathcal{C}_U) = \mathbb{R}^2_{\geq 0} \setminus \{0\}.$$

The projection $\mathrm{Proj}_{\mathbb{R}^2}(\mathcal{C}_U)$ corresponds to the nonnegative orthant excluding the origin, while U is the subset of points satisfying the equality $\langle 1, x\rangle = 1$. Therefore, U lies entirely within $\mathrm{Proj}_{\mathbb{R}^2}(\mathcal{C}_U)$:

$$U \subseteq \mathrm{Proj}_{\mathbb{R}^2}(\mathcal{C}_U).$$

1.3 — The visualization of the sets U and $\mathrm{Proj}_{\mathbb{R}^2}(\mathcal{C}_U)$ is shown in the figure below:

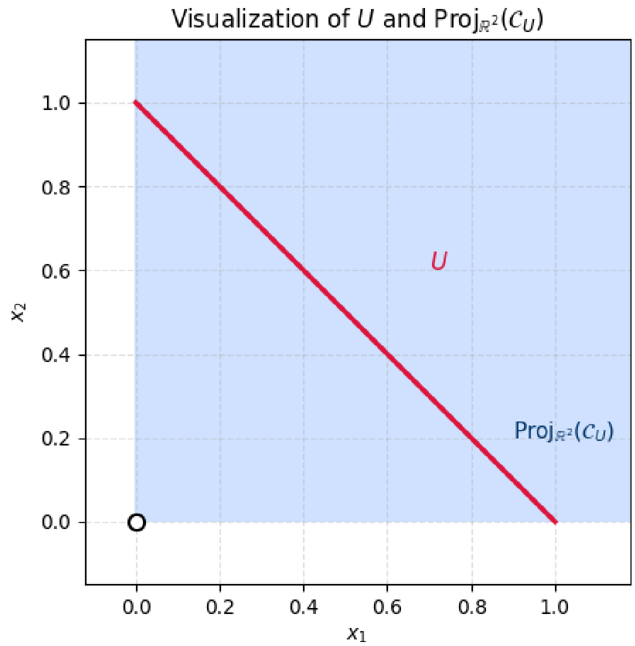

Figure 1: Visualization of the sets U and $\mathrm{Proj}_{\mathbb{R}^2}(\mathcal{C}_U)$ in $\mathbb{R}^2$.

Problem 2

We begin with the following claim, which characterizes the extreme rays of the homogenized cone associated with a polyhedron.

Claim 1. Given a polyhedron $P \subseteq \mathbb{R}^n$, the set of extreme rays of its homogenization $\mathcal{C}_P$ is

$$\operatorname{eray}(\mathcal{C}_P) = \{\, \lambda(x,1) \mid x \in \operatorname{ext}(P),\ \lambda > 0 \,\} \cup \{\, (d,0) \mid d \in \operatorname{eray}(P) \,\}.$$

Proof. We prove both inclusions.

($\subseteq$)
Let $(x,t) \in \operatorname{eray}(\mathcal{C}_P)$. We show that (x,t) must belong to one of the two sets described in the claim.
Case 1: $t > 0$.
Then $(x/t, 1) \in \mathcal{C}_P$, which implies $x/t \in P$. Suppose, for contradiction, that x/t is not an extreme point of P. Then there exist distinct $x_1, x_2 \in P$ such that

$$\frac{x_1 + x_2}{2} = \frac{x}{t}.$$

Consequently,

$$\frac{1}{2}(x_1,1) + \frac{1}{2}(x_2,1) = \left(\frac{x}{t}, 1\right).$$

Multiplying both sides by $t > 0$, we obtain

$$\frac{t}{2}(x_1,1) + \frac{t}{2}(x_2,1) = (x,t).$$

This expresses (x,t) as a nontrivial conic combination of two distinct elements of $\mathcal{C}_P$, contradicting the extremality of (x,t). Hence $x/t \in \operatorname{ext}(P)$, and (x,t) lies on the ray $\{\lambda(x/t,1) \mid \lambda > 0\}$.
Case 2: $t = 0$.
Then $(x,0) \in \mathcal{C}_P$ implies $x \in \operatorname{rec}(P)$. Suppose, for contradiction, that x is not an extreme ray of P. Then there exist distinct $d_1, d_2 \in \operatorname{rec}(P)$ such that

$$x = d_1 + d_2.$$

Consequently,

$$(x,0) = (d_1,0) + (d_2,0),$$

which again expresses $(x,0)$ as a nontrivial conic combination of other elements of $\mathcal{C}_P$, contradicting extremality. Therefore, $x \in \operatorname{eray}(P)$ and $(x,0)$ lies on the ray $\{\lambda(d,0) \mid \lambda > 0\}$ for some $d \in \operatorname{eray}(P)$.

In both cases, (x,t) belongs to one of the two families of rays described in the claim, and thus the inclusion ($\subseteq$) holds.

($\supseteq$)
We now prove that every ray of the two families described in the claim is indeed an extreme ray of $\mathcal{C}_P$.
Case 1: $x \in ext(P)$.
We show that the ray $\{\lambda(x,1) \mid \lambda > 0\}$ is extreme in $\mathcal{C}_P$. Suppose, for contradiction, that $(x,1)$ is not an extreme ray of $\mathcal{C}_P$. Then there exist distinct $(x_1,t_1), (x_2,t_2) \in \mathcal{C}_P$ such that

$$(x,1) = (x_1,t_1) + (x_2,t_2).$$

Since the last coordinate of $(x,1)$ is positive, we have $t_1, t_2 > 0$ and $t_1 + t_2 = 1$. Dividing both sides by $t_1 + t_2 = 1$, we obtain

$$x = t_1 \frac{x_1}{t_1} + t_2 \frac{x_2}{t_2}.$$

Because $(x_i, t_i) \in \mathcal{C}_P$ implies $\frac{x_i}{t_i} \in P$, this expresses x as a nontrivial convex combination of two distinct points in P, contradicting the fact that x is an extreme point. Therefore, $(x,1)$ must generate an extreme ray of $\mathcal{C}_P$.

Case 2: $d \in$ *eray*(P).
We show that the ray generated by $(d, 0)$ is extreme in $\mathcal{C}_P$. Suppose, for contradiction, that $(d, 0)$ is not extreme. Then there exist distinct $(x_1, t_1), (x_2, t_2) \in \mathcal{C}_P$ such that

$$(d, 0) = (x_1, t_1) + (x_2, t_2).$$

Since the last coordinate of $(d, 0)$ is zero, we must have $t_1 = t_2 = 0$. Hence, $(x_i, t_i) = (d_i, 0)$ for some $d_i \in \mathrm{rec}(P)$. This gives

$$d = d_1 + d_2,$$

which expresses d as a nontrivial conic combination of distinct recession directions of P, contradicting the fact that d is an extreme ray of P. Therefore, $(d, 0)$ must be an extreme ray of $\mathcal{C}_P$.

In both cases, the corresponding elements generate extreme rays of $\mathcal{C}_P$, and thus the inclusion $(\supseteq)$ holds. □

We can now prove that every polyhedron P admits a unique decomposition of the form

$$P = \mathrm{conv}(\mathrm{ext}(P)) + \mathrm{cone}(\mathrm{eray}(P)).$$

Proof. Let $\mathcal{C}_P$ denote the cone associated with P. By definition,

$$x \in P \iff (x, 1) \in \mathcal{C}_P.$$

Let $\mathrm{ext}(P) = \{x_i\}_{i \in [n]}$ and $\mathrm{eray}(P) = \{d_j\}_{j \in [m]}$. Since every element of a convex cone is a conic combination of its extreme rays, and by Claim 1, each element of $\mathcal{C}_P$ can be written as

$$(x, 1) = \sum_{i \in [n]} \alpha_i (x_i, 1) + \sum_{j \in [m]} \beta_j (d_j, 0), \qquad \alpha_i, \beta_j \geq 0.$$

Because the last coordinate of $(x, 1)$ equals 1, we must have

$$\sum_{i \in [n]} \alpha_i = 1.$$

Hence,

$$x = \sum_{i \in [n]} \alpha_i x_i + \sum_{j \in [m]} \beta_j d_j, \qquad \alpha_i, \beta_j \geq 0, \ \sum_{i \in [n]} \alpha_i = 1.$$

The first term represents a convex combination of the extreme points of P, that is, $\mathrm{conv}(\mathrm{ext}(P))$, while the second term represents a conic combination of the extreme rays of P, that is, $\mathrm{cone}(\mathrm{eray}(P))$. Therefore,

$$P = \mathrm{conv}(\mathrm{ext}(P)) + \mathrm{cone}(\mathrm{eray}(P)).$$

Uniqueness follows from the fact that the representation of a convex cone by its extreme rays is unique. □

Problem 3

3.3 — When running the simplex implementation on randomly generated instances (A, b, c) with $b, c \geq 0$, the resulting problems are typically *infeasible or unbounded*. This occurs because, under the constraint $Ax \geq b$ with $A, b \geq 0$, the feasible region tends to be either empty or unbounded in the positive orthant. In particular, if A has nonnegative entries, increasing x indefinitely preserves feasibility (since Ax only grows larger), leading to an unbounded objective value for maximization.

Instance	**Implementation (objval)**	**Gurobi (objval)**
1	$+\infty$	$+\infty$
2	$+\infty$	$+\infty$
3	$+\infty$	$+\infty$
4	$+\infty$	$+\infty$
5	$+\infty$	$+\infty$

However, when we generate instances with entries $a_{ij} \leq 0$ and $b \leq 0$, the problems become *feasible and bounded*. In this case, increasing x decreases the left-hand side of $Ax \geq b$, thus creating a closed and bounded feasible region.

Instance	**Implementation (objval)**	**Gurobi (objval)**
1	0.5988	0.5988
2	1.4479	1.4479
3	1.0953	1.0953
4	0.9362	0.9362
5	1.6408	1.6408

The objective values match across all feasible instances, confirming the correctness of the simplex implementation.

3.4 — Possible limitations of the current implementation include:

- **Degeneracy:** if multiple constraints are active at a basic feasible solution, the algorithm may experience cycling or stall without progress.
- **Pivot selection:** the implementation uses a naive pivot rule based on the most negative reduced cost, which can lead to cycling. More robust methods, such as *Bland's rule*, ensure termination even in degenerate cases.
- **Numerical stability:** repeated matrix inversions at each iteration may introduce rounding errors for ill-conditioned systems.

Problem 4

4.1 - We formulate the Generalized Assignment Problem as a binary integer program. The goal is to assign each job to exactly one agent while respecting the agents' capacity limits and minimizing the total assignment cost:

$$\begin{aligned}
\min_{x_{ij}} \quad & \sum_{i\in[n]}\sum_{j\in[m]} c_{ij}x_{ij} && && \text{(1a)}\\
\text{s.t.} \quad & \sum_{j\in[m]} x_{ij} = 1 && \forall\, i \in [n], && \text{(1b)}\\
& \sum_{i\in[n]} a_{ij}x_{ij} \le b_j && \forall\, j \in [m], && \text{(1c)}\\
& x_{ij} \in \{0,1\} && \forall\, i \in [n],\ j \in [m]. &&
\end{aligned}$$

Constraint (1b) ensures that each job is assigned to exactly one agent, while (1c) enforces each agent's resource limit b_j. The objective (1a) minimizes the total cost of assignments.

4.2 - We decompose the problem with respect to the capacity constraints (1c). For each agent $j \in [m]$, define its feasible assignment set:

$$A_j = \left\{ a_j = (a_{1j}, \ldots, a_{nj}) \;\middle|\; \sum_{i\in[n]} a_{ij}a_{ij} \le b_j,\ a_{ij} \in \{0,1\} \right\}.$$

Let $\bar{a}_j^p$, $p \in Q_j$, denote the extreme points of $\mathrm{conv}(A_j)$. Each extreme point (or *pattern*) $\bar{a}_j^p = (\bar{a}_{1j}^p, \ldots, \bar{a}_{nj}^p)$ satisfies

$$\sum_{i\in[n]} a_{ij}\bar{a}_{ij}^p \le b_j.$$

Then any feasible a_j can be expressed as a convex combination of these extreme points:

$$a_j = \sum_{p\in Q_j} \bar{a}_j^p \lambda_{jp}, \qquad \sum_{p\in Q_j} \lambda_{jp} = 1, \quad \lambda_{jp} \ge 0.$$

Substituting this representation into the original formulation yields the following equivalent problem:

$$\begin{aligned}
\min \quad & \sum_{j\in[m]}\sum_{p\in Q_j} \left(\sum_{i\in[n]} c_{ij}\bar{a}_{ij}^p \right) \lambda_{jp} && && \text{(2a)}\\
\text{s.t.} \quad & \sum_{j\in[m]}\sum_{p\in Q_j} \bar{a}_{ij}^p \lambda_{jp} = 1 && \forall\, i \in [n], && \text{(2b)}\\
& \sum_{p\in Q_j} \lambda_{jp} = 1 && \forall\, j \in [m], && \text{(2c)}\\
& \lambda_{jp} \ge 0 && \forall\, j \in [m],\ p \in Q_j. &&
\end{aligned}$$

Constraint (2b) ensures that each job is assigned exactly once across all agents and feasible patterns, while (2c) enforces that each agent selects one valid pattern $\bar{a}_j^p$ that satisfies its capacity limit.

4.3 - Now suppose all agents are identical, that is,

$$c_{ij} = c_i, \qquad a_{ij} = a_i, \qquad b_j = b, \quad \forall\, i \in [n],\ j \in [m].$$

In this case, each agent has the same feasible assignment set:

$$A = \left\{ a = (a_1, \ldots, a_n) \;\middle|\; \sum_{i\in[n]} a_i a_i \le b,\ a_i \in \{0,1\} \right\}.$$

Let $\bar{a}^p$, $p \in Q$, denote the extreme points of $\text{conv}(A)$, with components $\bar{a}_i^p \in \{0, 1\}$.Since all agents are identical, they all share the same pattern set Q.
Substituting this shared representation into the original problem gives:

$$\begin{aligned}
\min \quad & \sum_{j \in [m]} \sum_{p \in Q} \left(\sum_{i \in [n]} c_i \bar{a}_i^p \right) \lambda_{jp} && && \text{(3a)} \\
\text{s.t.} \quad & \sum_{j \in [m]} \sum_{p \in Q} \bar{a}_i^p \lambda_{jp} = 1 && \forall\, i \in [n] && \text{(3b)} \\
& \sum_{p \in Q} \lambda_{jp} = 1 && \forall\, j \in [m] && \text{(3c)} \\
& \lambda_{jp} \geq 0 && \forall\, j \in [m],\ p \in Q. &&
\end{aligned}$$

Here, each pattern $p \in Q$ represents a feasible subset of jobs whose total resource requirement does not exceed b, i.e.,

$$\sum_{i \in [n]} a_i \bar{a}_i^p \leq b.$$

Because all agents are identical, they share the same pool of patterns Q and differ only by their respective selection variables λ_{jp}.