Linear Optimization Functions and Sets

Shixuan Zhang

ISEN 320-501, Fall 2023

1 Linear Optimization (LO) Representable Functions

Sometimes even when the objective function is nonlinear, the problem can be reformulated as a LO model. We call such objective functions as *LO representable functions*. To be precise, we consider the following optimization model

$$\begin{aligned} \min \quad & f(x) \\ \text{s. t.} \quad & Ax \le b, \\ & x \in \mathbb{R}^n, \end{aligned} \tag{1}$$

where $f(x)$ is a function on $\mathbb{R}^n$ defined by

$$f(x) = \max_{k=1,\dots,l} \{c_{k0} + c_{k1}x_1 + \cdots + c_{kn}x_n\}, \tag{2}$$

for some given number of pieces $l \in \mathbb{Z}_{\ge 1}$ and coefficients $c_{ki} \in \mathbb{R}$, $k = 1, \dots, l$, $i = 0, 1, \dots, n$. Clearly, f can be a nonlinear function. (Hint: take $n+1$ points such that the maximum at these points are not attained at the same index $k = 1, \dots, l$ and derive a contradiction.) Nevertheless, problem (1) can be rewritten as a LO model with an additional variable $y \in \mathbb{R}$

$$\begin{aligned} \min \quad & y \\ \text{s. t.} \quad & Ax \le b, \\ & y \ge c_{k0} + \sum_{i=1}^{n} c_{ki}x_i, \quad k = 1, \dots, l, \\ & x \in \mathbb{R}^n,\ y \in \mathbb{R}. \end{aligned} \tag{3}$$

To see this, note that

$$y \ge f(x) \iff y \ge c_{k0} + \sum_{i=1}^{n} c_{ki}x_i, \quad k = 1, \dots, l. \tag{4}$$

Thus the constraints involving $y \in \mathbb{R}$ in (3) is equivalent to $y \geq f(x)$. As y is not involved in any other constraint, an optimal solution (when it exists) must have $y = f(x)$. Therefore, (3) preserves the optimal value and any optimal solution in x that comes from (1).

Using such reformulation technique on the absolute value function $|x| = \max\{x, -x\}$, we can build a LO model in the following example.

Example 1. *A machine shop has a drill press and a milling machine which are used to produce two parts A and B. The required time (in minutes) per unit part on each machine is shown in the table below.*

	Drill press	Milling machine
A	3	4
B	5	3

The shop must produce at least 50 units in total (both A and B) and at least 30 units of part A and 20 units of B, and it can make at most 100 units of part A and 80 units of part B. Assume that the shop can make fractional amount of the parts. The goal is to minimize the absolute difference between the total running time of the drill press and that of the milling machine. Our two decision variables are

$$30 \leq x_1 \leq 100: \quad \textit{units of part A to be produced,}$$
$$20 \leq x_2 \leq 80: \quad \textit{units of part B to be produced.}$$

The linear constraint on the total units to be produced can be written as

$$x_1 + x_2 \geq 50.$$

The difference between the total running time of the drill press and that of the milling machine is

$$(3x_1 + 5x_2) - (4x_1 + 3x_2) = -x_1 + 2x_2.$$

To reformulate the absolute value function $|-x_1 + 2x_2| = \max\{-x_1 + 2x_2, x_1 - 2x_2\}$*, we need to introduce an additional decision variable*

$$y \in \mathbb{R}: \quad \textit{absolute difference between the total running times,}$$

and two additional linear constraints

$$y \geq -x_1 + 2x_2,$$
$$y \geq x_1 - 2x_2.$$

In summary, our LO model can be written as

$$
\begin{aligned}
\min \quad & y \\
\text{s.t.} \quad & y \geq -x_1 + 2x_2, \\
& y \geq x_1 - 2x_2, \\
& x_1 + x_2 \geq 50, \\
& 30 \leq x_1 \leq 100, \quad 20 \leq x_2 \leq 80, \quad y \in \mathbb{R}.
\end{aligned}
$$

We code the LO model in the script `model_machine.py` *and the output is displayed below.*

```
The minimum absolute difference is 0.00.
Units of part A to be produced = 40.00
Units of part B to be produced = 20.00
```

A natural question is how we can tell whether a nonlinear objective function $f(x)$ is LO representable or not. It turns out that the answer will depend on whether we are maximizing or minimizing our objective value. As we have seen above, for a minimization problem, any "finite-maximum" function (as defined in (2)) is LO representable. Using the same argument, a "finite-minimum" function can be reformulated in a LO maximization problem, as

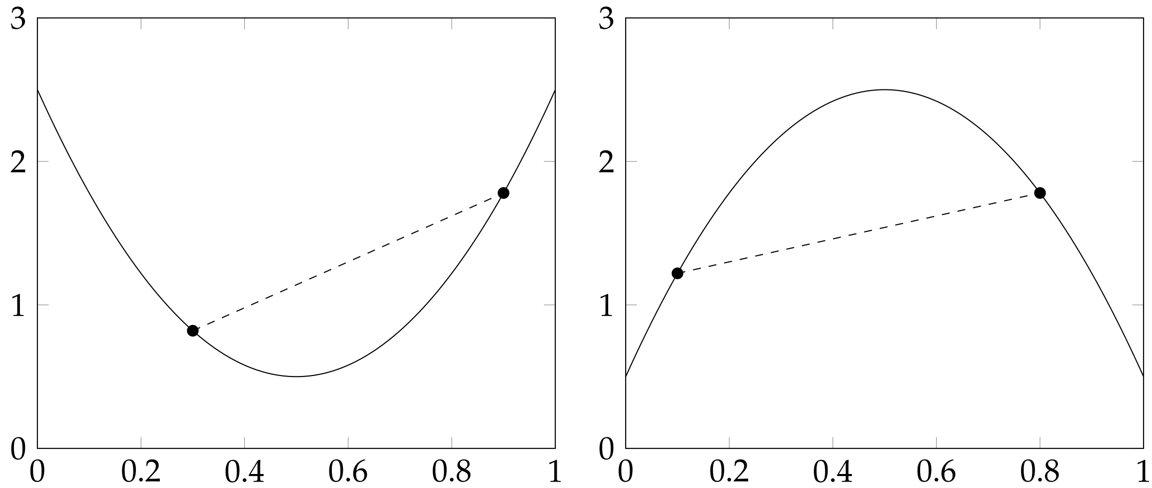
$$y \leq \min_{k=1,\ldots,l} \{c_{k0} + c_{k1}x_1 + \cdots + c_{kn}x_n\} \iff y \leq c_{k0} + \sum_{i=1}^{n} c_{ki}x_i, \quad k = 1, \ldots, l.$$

One important characterization of these maximum or minimum function is by *convexity* or *concavity* defined as follows.

Definition 1. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is* convex *if for any $x, y \in \mathbb{R}^n$ and any $0 \leq t \leq 1$, we have $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$. A function f is* concave *if $-f$ is convex.*

Geometrically, the definition says that if you take a line segment between any two points in the graph of your function and it stays above (resp. below) the graph, then it is convex (resp. concave). You can see simple examples in Figure 1. Intuitively speaking, a convex function bends "upward" (slope increasing in any direction), and a concave function bends "downward" (slope decreasing in any direction).

We claim that a maximum of linear functions is always convex. To see this, let L be an index set and $f(x) := \max_{k \in L}\{c_{k0} + c_{k1}x_1 + \cdots + c_{kn}x_n\}$. Then for any $x, y \in \mathbb{R}^n$



(a) a convex function

(b) a concave function

Figure 1: Illustration of convexity and concavity

and $0 \leq t \leq 1$, we have

$$
\begin{aligned}
& f(tx + (1-t)y) \\
&= \max_{k \in L}\{c_{k0} + c_{k1}(tx_1 + (1-t)y_1) + \cdots + c_{kn}(tx_n + (1-t)y_n)\} \\
&= \max_{k \in L}\{t(c_{k0} + c_{k1}x_1 + \cdots + c_{kn}x_n) + (1-t)(c_{k0} + c_{k1}y_1 + \cdots + c_{kn}y_n)\} \\
&\leq t \cdot \max_{k \in L}\{c_{k0} + c_{k1}x_1 + \cdots + c_{kn}x_n\} + (1-t) \cdot \max_{l \in L}\{c_{l0} + c_{l1}y_1 + \cdots + c_{ln}y_n\} \\
&= tf(x) + (1-t)f(y).
\end{aligned}
$$

Here, the first equality is derived directly by the definition of f; the second equality is derived by rearranging terms (and by splitting c_{k0} into $tc_{k0} + (1-t)c_{k0}$); the inequality here is due to the fact that we allow the maximum to be taken at different indices k and l; and the last equality is again by the definition of f. By reverting the direction of the inequality here, the argument shows that a minimum of linear functions is concave.

It is not enough by convexity or concavity alone to guarantee that the function is LO representable. For example, if the index set is infinite (such as $L = \mathbb{Z}$), then we might need infinitely many constraints in the reformulation (4) (as each index could correspond to one constraint). Thus we would also need the function to have finitely many "pieces" for it to be LO representable. For univariate functions, this can be defined as follows.

Definition 2. *A univariate function $f : \mathbb{R} \to \mathbb{R}$ is* piecewise linear *(with finitely many pieces), if there are points $-\infty = a_0 < a_1 < \cdots < a_l = +\infty$ such that on each interval $I_k := \{x \in \mathbb{R} : a_{k-1} < x < a_k\}$, $k = 1, \ldots, l$, $f(x)$ is an affine linear function, i.e., there exist*

$b_k, c_k \in \mathbb{R}$ *such that*

$$f(x) = b_k x + c_k, \quad \forall\, x \in I_k, \quad k = 1, \ldots, l.$$

Figure 2 illustrates some univariate piecewise linear functions on the interval $[0, 1]$. From the figures, we can see that a convex (resp. concave) piecewise linear function must have its dashed parts (i.e., the extension of each linear piece) below (resp. above) the function itself. In fact, the following statements are equivalent for a univariate piecewise linear function $f(x)$:

(i) $f(x)$ is convex;

(ii) $f(x) = \max_{k=1,\ldots,l}\{b_k x + c_k\}$;

(iii) the points and the coefficients in Definition 2 for $f(x)$ satisfy

$$f(a_k) = a_k b_k + c_k = a_k b_{k+1} + c_{k+1}, \text{ and } b_k \le b_{k+1} \quad \forall\, k = 1, \ldots, l-1.$$

The last statement essentially says that the function $f(x)$ is continuous and has non-decreasing slopes. A possible hint for any reader interested in the proof is that the definition of convexity for $f(x)$ implies that for any $h > 0$

$$\frac{f(x) - f(x-h)}{h} \le \frac{f(x+h) - f(x)}{h}, \quad \forall\, x \in \mathbb{R}.$$

This shows that the slope should be non-decreasing. Besides, taking limits of f from both sides towards a_k requires f to be continuous at a_k, for $k = 1, \ldots, l-1$. Similarly, the following statements are also equivalent for a piecewise linear function $f(x)$:

(i) $f(x)$ is concave;

(ii) $f(x) = \min_{k=1,\ldots,l}\{b_k x + c_k\}$;

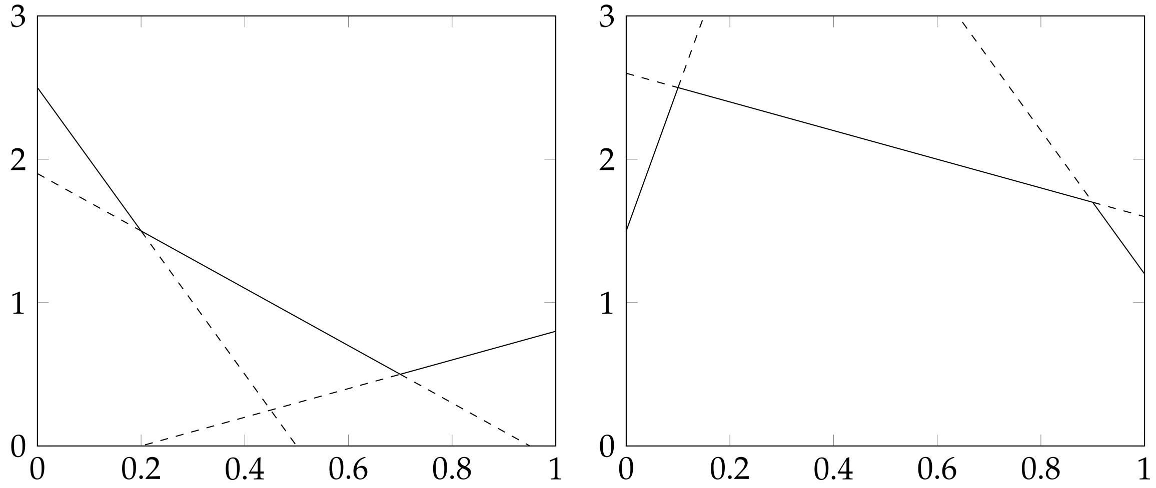
(iii) the points and the coefficients in Definition 2 for $f(x)$ satisfy

$$f(a_k) = a_k b_k + c_k = a_k b_{k+1} + c_{k+1}, \text{ and } b_k \ge b_{k+1} \quad \forall\, k = 1, \ldots, l-1.$$

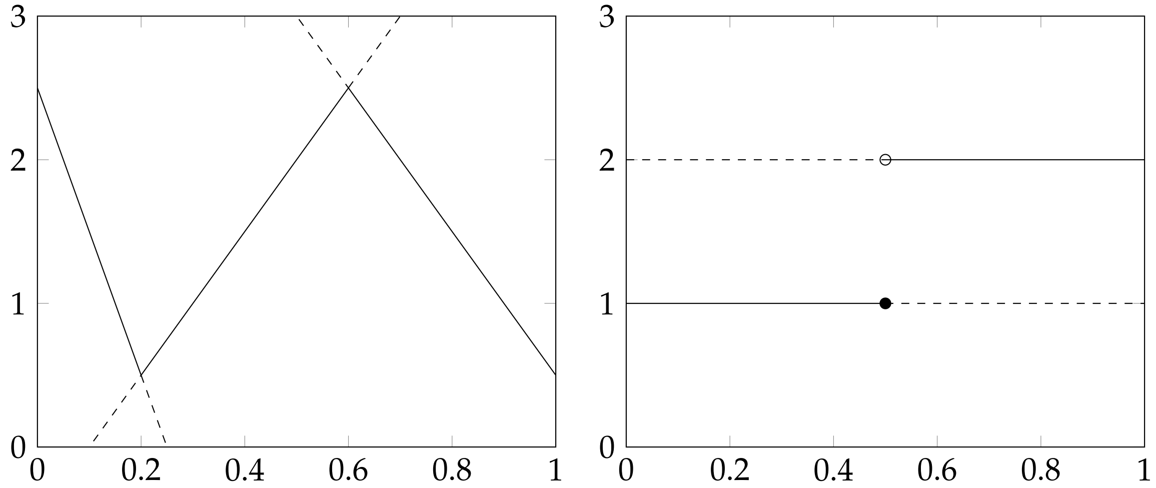
In practice, we may sometimes *approximate* a nonlinear objective function by piecewise linear functions. For example, given a nonlinear convex function $f(x)$ on an interval $[0, 1]$, we can take points $0 = a_0 < a_1 < \cdots < a_{l-1} < a_l = 1$, and then set

$$b_k = \frac{f(a_k) - f(a_{k-1})}{a_k - a_{k-1}}, \quad c_k = f(a_k) - a_k b_k, \quad k = 1, \ldots, l. \tag{5}$$

An illustration is plotted in Figure 3a. This procedure is often called *inner-approximation* (or *over-approximation*) of the nonlinear function f. Alternatively, if one can find differential information at points $0 \le a'_1 < a'_2 < \cdots < a'_l \le 1$, then an *outer-approximation* (or *under-approximation*) can be built as in Figure 3b.

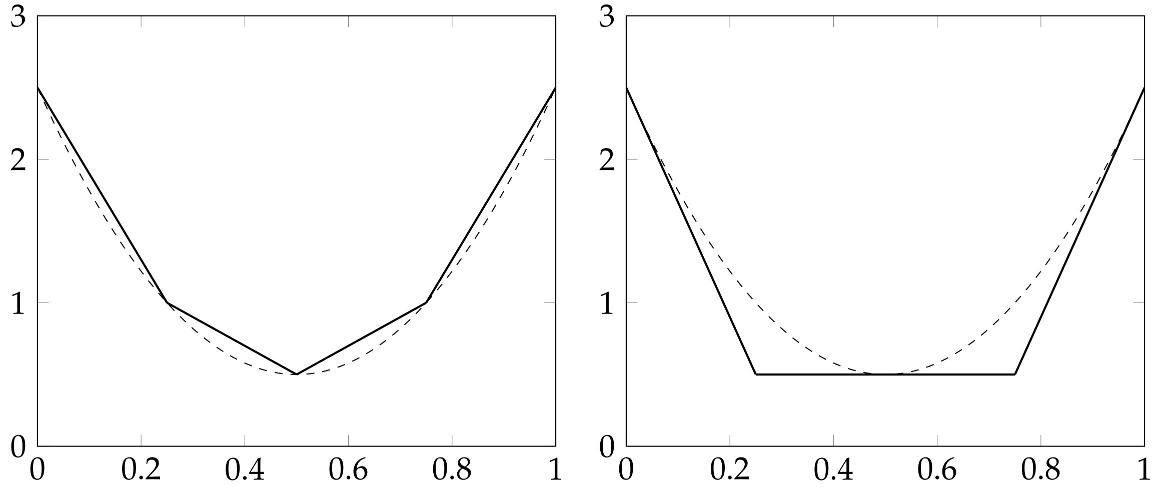


(a) a convex piecewise linear function (b) a concave piecewise linear function



(c) neither convex nor concave (d) neither convex nor concave

Figure 2: Illustration of piecewise linear functions



(a) a piecewise linear inner-approximation (b) a piecewise linear outer-approximation

Figure 3: Illustration of inner- and outer-approximations of a nonlinear function

Example 2. *An electric power grid operator wants to find a generation plan for two generators $i = 1$ and 2. The generation cost functions f_i for generators $i = 1, 2$ are described by two convex functions $f_1(x) = 2 + 0.5x + 0.01x^2$ and $f_2(x) = 3 + 0.4x + 0.02x^2$. The demand in the region is 10 MW for the next hour. Assume that there is no loss in the transmission. The goal is to minimize the total generation cost while meeting the demand. Let $x_i \geq$ denote the power generation from the generator $i = 1, 2$. The power demand constraint can then be written as*

$$x_1 + x_2 \geq 10.$$

Note that it suffices to consider generation within the range $[0, 10]$ for both generators. To handle the nonlinearity, we build inner-approximations for the generation cost functions f_1 and f_2 over $[0, 10]$, using (5) on the function values at the points $x_1, x_2 = 0, 5, 10$:

$$\begin{aligned} f_1(x_1) &\lessapprox \max\{0.55x_1 + 2,\ 0.65x_1 + 1.5\} \\ f_2(x_2) &\lessapprox \max\{0.5x_2 + 3,\ 0.7x_2 + 2\}. \end{aligned}$$

Therefore, by introducing additional variables $y_1, y_2 \in \mathbb{R}$ to be the approximate generation costs for generators $i = 1, 2$, we can write our approximate LO model as

$$\begin{aligned} \min \quad & y_1 + y_2 \\ \text{s.t.} \quad & x_1 + x_2 \geq 10, \\ & y_1 \geq 0.55x_1 + 2, \\ & y_1 \geq 0.65x_1 + 1.5, \\ & y_2 \geq 0.5x_2 + 3, \\ & y_2 \geq 0.7x_2 + 2, \\ & x_1, x_2 \geq 0,\ y_1, y_2 \in \mathbb{R}. \end{aligned}$$

We code the LO model in `model_generation.py` *and the output is displayed below.*

```
The minimum generation cost is 10.25.
Power generation at generator 1 = 5.00.
Approximate generation cost of generator 1 = 4.75.
Power generation at generator 2 = 5.00.
Approximate generation cost of generator 2 = 5.50.
```

We check that at the point $(x_1, x_2) = (5.0, 5.0)$, the actual generation cost $(f_1(x_1), f_2(x_2)) = (4.75, 5.5)$, which agrees with the our obtained approximate generation cost (y_1, y_2). This means that our approximation is tight *at the obtained solution and we have found an optimal solution exactly.*

2 LO Feasible Regions and Graphical Solutions

Other than the objective function, one may wonder what sets can be represented as the feasible region of a LO model. Such sets are known as *polyhedra*, which can be defined as follows.

Definition 3. (i) *A closed halfspace H in $\mathbb{R}^n$ is a subset*

$$H := \{x \in \mathbb{R}^n : a^\top x \le b\}$$

for some vector $a \in \mathbb{R}^n$ and real number $b \in \mathbb{R}$.

(ii) *A polyhedron (or a polyhedral set) in $\mathbb{R}^n$ is an intersection of finitely many closed halfspaces in $\mathbb{R}^n$.*

Recall that a general LO feasible region can be written as $X := \{x \in \mathbb{R}^n : Ax \le b\}$ for some matrix $A \in \mathbb{R}^{m\times n}$ and some vector $b \in \mathbb{R}^m$. It is then clear that X is a polyhedron because

$$X = \bigcap_{j=1}^{m} \left\{x \in \mathbb{R}^n : a_j^\top x \le b_j\right\},$$

where a_j is the j-th row vector of A. There could be multiple ways to represent a polyhedron as a LO feasible region. For example, $X := \{x \in \mathbb{R}^2 : 0 \le x_1, x_2 \le 1\}$ and $X' := \{x \in \mathbb{R}^2 : 0 \le x_1, x_2 \le 1,\ x_1 + x_2 \le 2\}$ are the feasible regions for two LO problems, but they represent the sample polyhedron, which is a square of side length 1.

Similar to LO representable functions, an important characterization of LO representable sets is convexity.

Definition 4. (i) *A set $X \subseteq \mathbb{R}^n$ is convex if for any two points $x, y \in X$ and any $0 \le t \le 1$, the point $tx + (1-t)y \in X$.*

(ii) *A closed convex set in $\mathbb{R}^n$ is an intersection of (possibly infinitely many) closed halfspaces in $\mathbb{R}^n$.*

Intuitively, a set is convex if we connect any two points in the set and the line segment would stay in the set. See Figure 4 for examples. To see that we are not abusing terminology, we show that a closed convex set is indeed convex as follows. Suppose J is a possibly infinite index set and

$$X = \bigcap_{j\in J} \left\{x \in \mathbb{R}^n : a_j^\top x \le b_j\right\}$$

is a closed convex set for vectors $a_j \in \mathbb{R}^n$ and real numbers $b_j \in \mathbb{R}$, $j \in J$. Take any points $x, y \in X$, which by definition satisfies

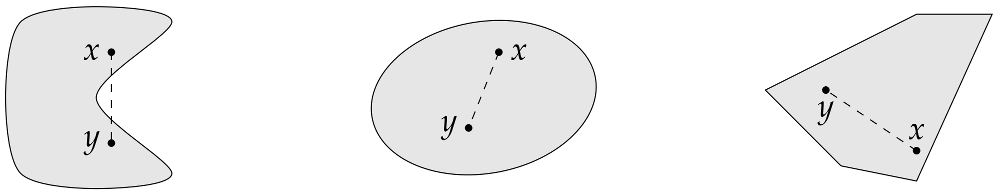
$$a_j^\top x \le b_j, \text{ and } a_j^\top y \le b_j, \quad \forall j \in J.$$

Thus using linearity, we see that

$$a_j^\top (tx + (1-t)y) = t \cdot (a_j^\top x) + (1-t) \cdot (a_j^\top y) \le tb_j + (1-t)b_j = b_j.$$

As this holds for any $j \in J$, we conclude that $tx + (1-t)y \in X$, which shows the convexity of X.

Using these definitions, it is clear that a polyhedron is a closed convex set. The converse is not necessarily true, which can be seen from a planar example in Figure 4b.



(a) Not convex (b) Convex, not polyhedral (c) Polyhedral

Figure 4: Non-example and examples of convex sets

Now we can finally answer the question about which functions are LO representable. For a minimization problem with decision variables $x \in \mathbb{R}^n$ and an auxiliary variable $y \in \mathbb{R}$, our reformulation technique (4) requires us to write the set $\{(x, y) \in \mathbb{R}^{n+1} : y \ge f(x)\}$ as a polyhedron (LO feasible region), so $f(x)$ must be a finite-maximum function, which is piecewise linear and convex. Similarly, for a maximization problem, our reformulation requires us to write the set $\{(x, y) \in \mathbb{R}^{n+1} : y \le f(x)\}$ as a polyhedron (LO feasible region), so $f(x)$ must be a finite-minimum function, which is piecewise linear and concave.

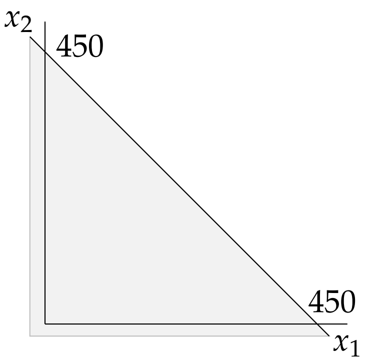
By interpreting polyhedra as intersection of halfspaces can help us visualize LO feasible regions. This is particularly useful for solving the LO problem when there are only two variables.

Example 3. *A company produces two types of baby carriers, non-reversible and reversible. Each non-reversible carrier sells for \$23, requires 2 linear yards of a solid color fabric, and costs \$8 to manufacture. Each reversible carrier sells for \$35, requires 2 linear yards of a printed fabric as well as 2 linear yards of a solid color fabric, and costs \$10 to manufacture. The company has 900 linear yards of solid color fabrics and 600 linear yards of printed fabrics available for its new carrier collection. It can spend up to \$4,000 on manufacturing the carriers. The demand is such that all reversible carriers made are projected to sell, whereas at most 350 non-reversible carriers can be sold. The goal of the company is to maximize its profit (e.g., the difference of revenues and expenses) resulting from manufacturing and selling the new carrier* *collection.*

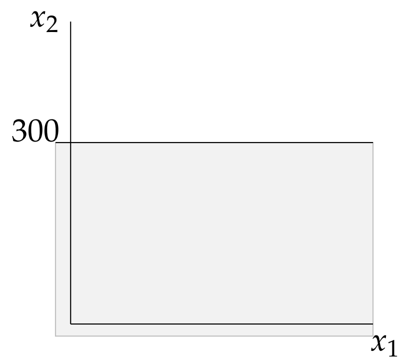
We define $x_1, x_2 \geq 0$ *to be the numbers of non-reversible and reversible carriers to manufacture. Then the LO model can be written as*

$$\begin{array}{llcrcll}
\max & 15x_1 & + & 25x_2 & & & \textit{(profit)} \\
\text{s. t.} & x_1 & + & x_2 & \leq & 450 & \textit{(solid color fabric constraint)} \\
 & & & x_2 & \leq & 300 & \textit{(printed fabric constraint)} \\
 & 4x_1 & + & 5x_2 & \leq & 2,000 & \textit{(budget constraint)} \\
 & x_1 & & & \leq & 350 & \textit{(demand constraint)} \\
 & & & x_1, x_2 & \geq & 0 & \textit{(nonnegativity constraints).}
\end{array}$$

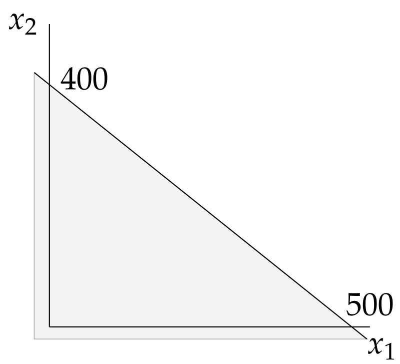
Each constraint can be plotted on the x_1-x_2 *plane as in Figure 5. Putting the constraints*



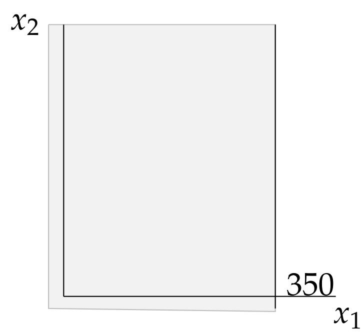
(a) solid fabric constraint $x_1 + x_2 \leq 450$



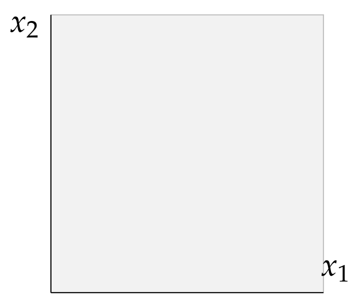
(b) printed fabric constraint $x_2 \leq 300$



(c) budget constraint $4x_1 + 5x_2 \leq 2,000$



(d) demand constraint $x_1 \leq 350$



(e) nonnegativity constraints

Figure 5: Constraints in the baby carrier problem

together, we can find optimal solutions by moving in the improving direction of the linear objective function $z = 15x_1 + 25x_2$, as shown in Figure 6. The optimal solution is $(x_1, x_2) = (125, 300)$.

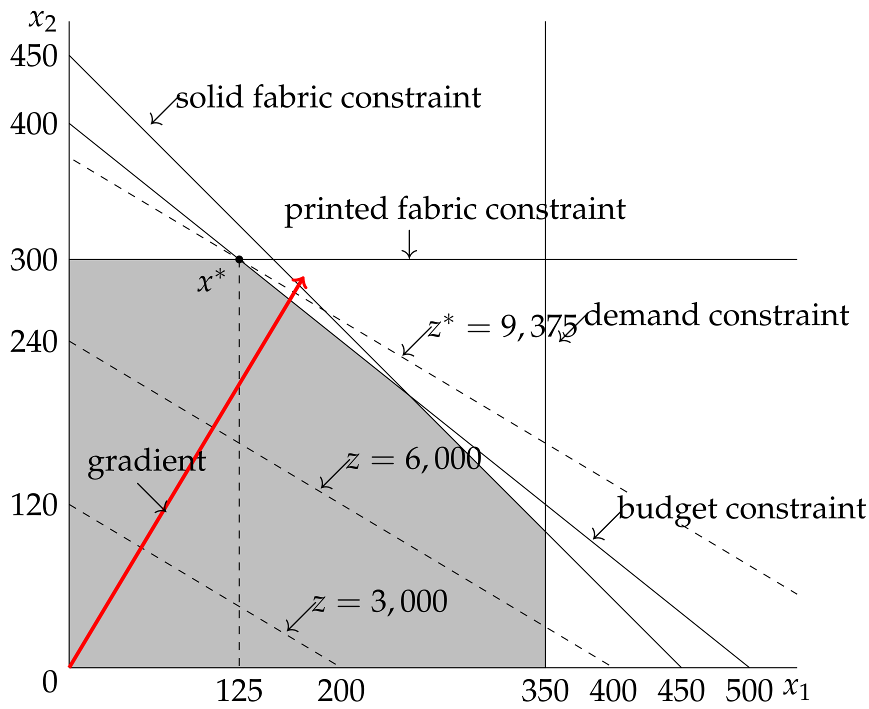


Figure 6: Feasible region and objective of the baby carriers problem

Example 4. *Now suppose in Example 3, the price of a non-reversible carrier is raised to \$28. The modified LO model becomes*

$$
\begin{array}{lrcrcll}
\max & 20x_1 & + & 25x_2 & & & \textit{(profit)} \\
\text{s. t.} & x_1 & + & x_2 & \leq & 450 & \textit{(solid color fabric constraint)} \\
& & & x_2 & \leq & 300 & \textit{(printed fabric constraint)} \\
& 4x_1 & + & 5x_2 & \leq & 2,000 & \textit{(budget constraint)} \\
& x_1 & & & \leq & 350 & \textit{(demand constraint)} \\
& & & x_1, x_2 & \geq & 0 & \textit{(nonnegativity constraints).}
\end{array}
$$

Note that the feasible region is the same while only the improving direction (gradient) is changed. The modified LO model can be plotted as in Figure 7. Now we can see that any points between $x^* = (125, 300)$ and $x' = (250, 200)$ are optimal and the optimal objective value is $z^* = 10,000$.

Example 5. *A retail store is planning an advertising campaign aiming to increase the number of customers visiting its physical location, as well as its online store. The store manager would like to advertise through a local magazine and through an online social network. The manager estimates that each 1,000 dollars invested in magazine ads will attract 100 new customers to*

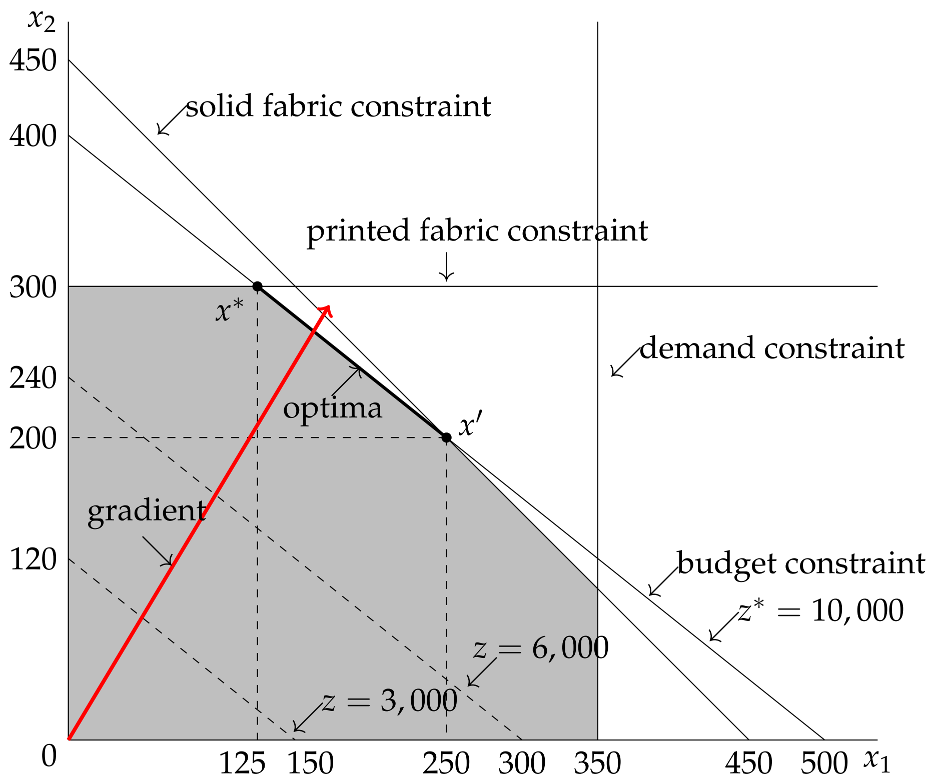


Figure 7: Feasible region and objective of the modified baby carriers problem

the store, as well as 500 new website visitors. In addition, each 1,000 dollars invested in online advertising will attract 50 new local store customers, as well as 1,000 new website visitors. The target for this campaign is to bring at least 500 new guests to the physical store and at least 5,000 new visitors to the online store. The decision variables are

$x_1 \geq 0$: *budget for magazine advertising (in thousands of dollars)*
$x_2 \geq 0$: *budget for online advertising (in thousands of dollars),*

and the LO model can be written as

$$
\begin{array}{llcrcll}
\min & x_1 & + & x_2 & & & \\
\text{s. t.} & 100x_1 & + & 50x_2 & \geq & 500 & \textit{(store visitors)} \\
 & 500x_1 & + & 1,000x_2 & \geq & 5,000 & \textit{(website visitors)} \\
 & & & x_1, x_2 & \geq & 0. & \textit{(nonnegativity)}
\end{array}
$$

We can plot the feasible region in Figure 8 and find that $(x_1, x_2) = (10/3, 10/3)$ *is the optimal solution with the optimal value* $z^* = 20/3$.

Instead of plotting the feasible regions manually, we can also use the `matplotlib` package in `Python` to (approximately) plot them. For example, we code the plotting procedure for Examples 3 and 5 in the scripts `plot_carriers.py` and `plot_advertising.py` and display the output in Figure 9.

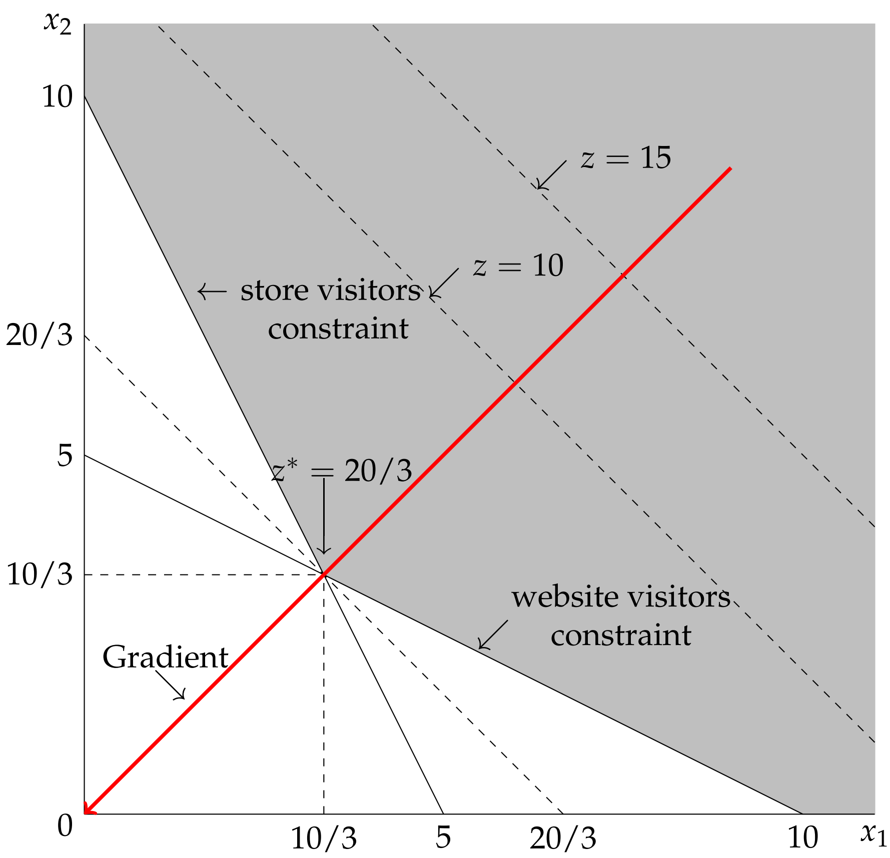


Figure 8: Feasible region and objective for advertising campaign problem

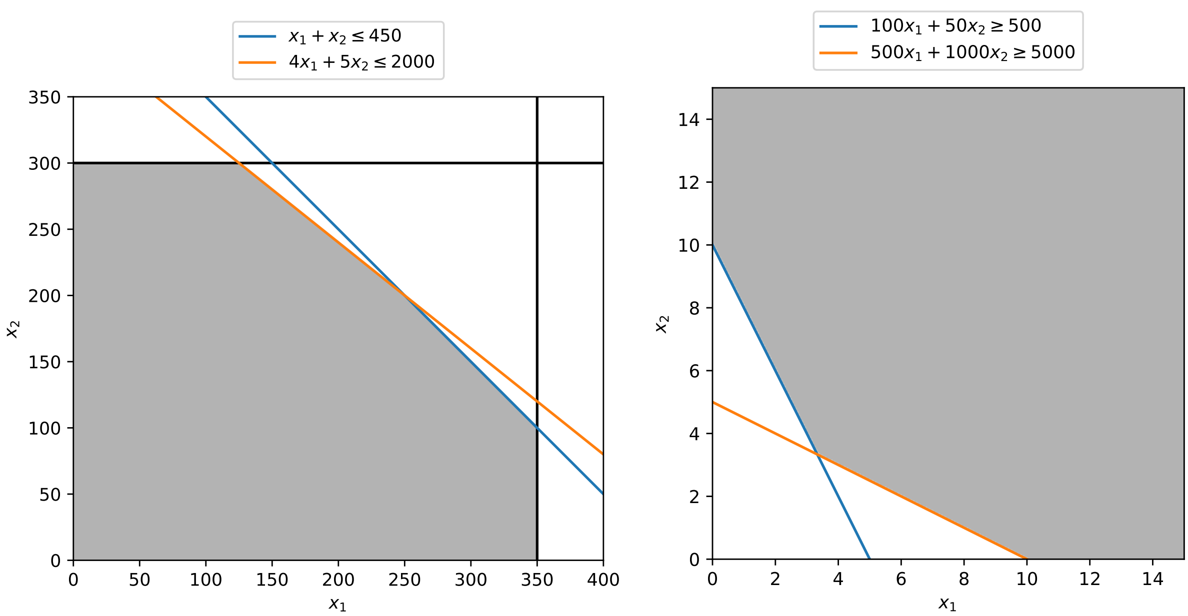


Figure 9: Feasible regions for Examples 3 and 5