Optimization II – IP: valid Inequalities for IPs

ORIE 3310 – Oktay Gunluk

Last updated: February 12, 2024.

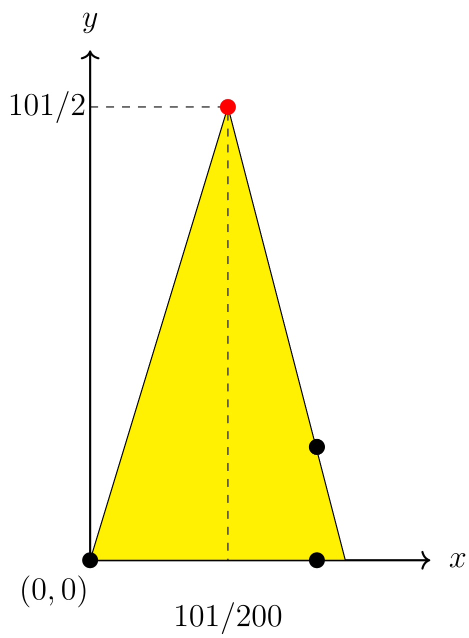
1 Valid inequalities and convex hulls

Example Consider the following simple IP:

.

IP1:

$$\begin{aligned} \max \quad & x + y \\ \text{s.t.} \quad & y \le 100x \\ & y + 100x \le 101 \\ & x, y \ge 0 \text{ and integer} \end{aligned}$$



If we solve the LP relaxation of this IP, the optimal solution is

$$x^* = \frac{101}{200}, \qquad y^* = \frac{101}{2} \qquad \text{and} \quad z^{LP} = 51.005 \longleftarrow \boldsymbol{U}$$

However, there are only 3 feasible integer points:

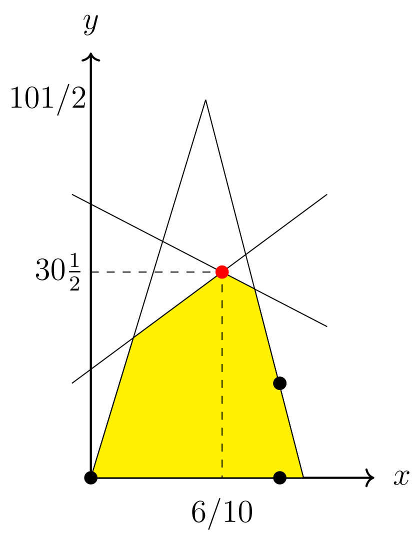
$$S = \Big\{(0,0), (1,0), (1,1)\Big\}$$

and $z^{IP} = 2$.

Therefore LP1 is very very optimistic and this creates problems in branch and bound because optimal LP solution is very far away from optimal IP solution.

(When we have more variables we cannot figure out what is happening by drawing a picture).

The question here is how to improve the formulation by adding more constraints to it while keeping all the feasible integer solutions in it.



Now consider adding two more constraints to the formulation to obtain IP2:

$$
\begin{aligned}
\max \quad & x + y \\
\text{s.t.} \quad & y \le 100x \\
& y + 100x \le 101 \\
& {\color{red} 2y - 80x \le 13} \\
& {\color{red} y + 60x \le 66.5} \\
& x, y \ge 0 \text{ and integer}
\end{aligned}
$$

Notice that the original integer points: $\{(0,1),(1,0),(1,1)\}$ are still feasible. Therefore, we have IP1 = IP2 (they have the same feasible (integer) solutions).

When we solve the LP relaxation of this new IP formulation, the optimal solution is

$$x^* = \frac{6}{10}, \qquad y^* = 30\frac{1}{2} \qquad \text{and} \quad z^{LP2} = 31.1 \longleftarrow \boldsymbol{U}$$

giving a much better upper bound (we are maximizing):

$$31 \approx z^{LP2} < z^{LP1} \approx 51$$

We conclude that LP2 is a better relaxation than LP1 for the IP (IP1=IP2).

Ideally, we would like to make the feasible set of the LP relaxation (yellow region) as small as possible while keeping all the integer points.

(If we exclude some integer points, we might be excluding a possible optimal solution to the IP. In this case we would not have a valid formulation for the original IP anymore.)

Valid inequalities: A linear inequality $a^T x \le b$ is a **valid inequality** for an IP if all integer solutions of the IP satisfy the inequality. (Note: the objective function does not play a role.)

The two inequalities we added to the formulation

$$2y - 80x \le 13 \quad \text{and} \quad y + 60x \le 66.5$$

are valid inequalities for the IP because they are satisfied by all feasible integer solutions:

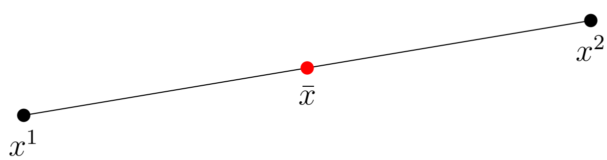
$$\Big\{(0,1),(1,0),(1,1)\Big\}$$

1.1 Convex hulls

Claim: If two points $x^1, x^2 \in \mathbb{R}^n$ satisfy a linear inequality, then all points on the line segment joining them, also satisfy that inequality:

$$\text{If} \qquad a^T x^1 \leq b \qquad \text{and} \qquad a^T x^2 \leq b \qquad \text{then} \qquad a^T \bar{x} \leq b$$

holds for all points $\bar{x}$ on the line segment $X = [x^1, x^2]$.



The set of points on the line segment joining $x^1, x^2 \in \mathbb{R}^n$ are

$$X \;=\; \Big\{x \in \mathbb{R}^n \;:\; x = \lambda x^1 + (1-\lambda)x^2, \text{ for some } \lambda \text{ such that } \; 1 \geq \lambda \geq 0\Big\}$$

or, equivalently

$$X \;=\; \Big\{x \in \mathbb{R}^n \;:\; x = x^2 + \lambda(x^1 - x^2), \text{ for some } \lambda \text{ such that } \; 1 \geq \lambda \geq 0\Big\}$$

The set X above is called the **convex hull** of x^1 and x^2:

$$\mathbf{conv}\Big(\underbrace{\{x^1, x^2\}}_{\text{a set of points}} \Big) \;=\; X$$

Notice that if an inequality is valid for both x^1 and x^2, for any $1 \geq \lambda \geq 0$ we have:

$$a^T x^1 \leq b \iff \lambda a^T x^1 \leq \lambda b \qquad \text{(because } \lambda \geq 0\text{)}$$

$$a^T x^2 \leq b \iff (1-\lambda) a^T x^2 \leq (1-\lambda) b \qquad \text{(because } 1-\lambda \geq 0\text{)}$$

$$\Longrightarrow \lambda a^T x^1 + (1-\lambda) a^T x^2 \leq \lambda b + (1-\lambda) b \iff a^T \underbrace{\big(\lambda x^1 + (1-\lambda)x^2\big)}_{\bar{x} \in X} \leq b$$

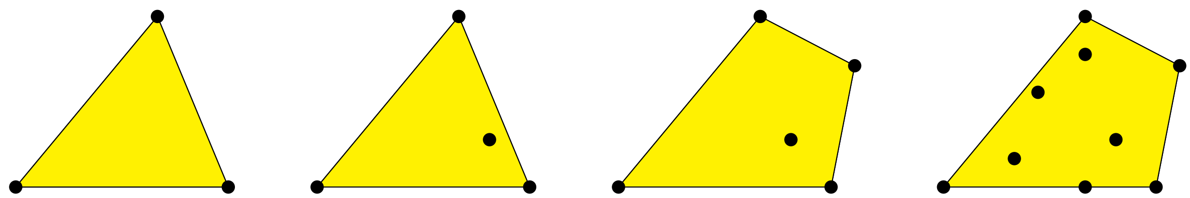
We just proved the above claim that if an an inequality is valid for both x^1 and x^2 then it is valid for **all** points in their convex hull.

More generally, if we have k points $x^1, \dots, x^k \in \mathbb{R}^n$ their convex hull is

$$X \;=\; \underbrace{\left\{ x \in \mathbb{R}^n \;:\; x = \sum_{j=1}^{k} \lambda_j x^j, \;\; \text{for some } \lambda \in \mathbb{R}^k \text{ such that } \;\; \sum_{j=1}^{k} \lambda_j = 1,\; \lambda \geq 0 \right\}}_{\textbf{convex hull of the points } \{x^1,\dots,x^k\}}$$

This is the set of points that can be obtained as a weighted combination of the points $x^1, \dots, x^k$ using non-negative weights that add up to 1.

In the pictures below, the yellow region gives the convex hull of the black points. It is the smallest polyhedron that contains all black points.



Claim: If an inequality $a^T x \leq b$ is satisfied by points $x^1, \dots, x^k \in \mathbb{R}^n$, then any point in their convex hull, also satisfies the inequality, i.e.

$$a^T \bar{x} \leq b \qquad \text{for all points } \bar{x} \in X$$

where

$$X \;=\; \underbrace{\left\{ x \in \mathbb{R}^n \;:\; x = \sum_{j=1}^{k} \lambda_j x^j, \;\; \text{for some } \lambda \text{ such that } \;\; \sum_{j=1}^{k} \lambda_j = 1,\; \lambda \geq 0 \right\}}_{\textbf{convex hull of the points } x^1,\dots,x^k}$$

Proof. If $\bar{x} \in X$ then $\bar{x} = \sum_{j=1}^{k} \lambda_j x^j$ for some $\lambda \in \mathbb{R}^k$ such that $\sum_{j=1}^{k} \lambda_j = 1$ and $\lambda \geq 0$.

Then:

$$\begin{array}{lclcl}
a^T x^1 \leq b & \iff & \lambda_1 a^T x^1 \leq \lambda_1 b & & (\text{note: } \lambda_1 \geq 0) \\
a^T x^2 \leq b & \iff & \lambda_2 a^T x^2 \leq \lambda_2 b & & (\text{note: } \lambda_2 \geq 0) \\
\vdots & & \vdots & & \\
a^T x^k \leq b & \iff & \lambda_k a^T x^k \leq \lambda_k b & & (\text{note: } \lambda_k \geq 0) \\
 & \implies & \displaystyle\sum_{j=1}^{k} \lambda_j a^T x^j \leq \sum_{j=1}^{k} \lambda_j b & \iff & a^T (\underbrace{\sum_{j=1}^{k} \lambda_j x^j}_{\bar{x}}) \leq b
\end{array}$$

[end of proof] $\square$

1.2 Valid inequalities for IPs

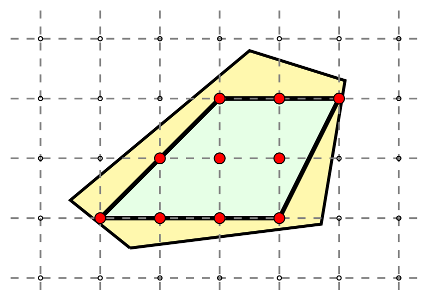
Consider an IP with n variables and m constraints:

$$\max \quad c^T x \;\text{ subject to } \quad \underbrace{\underbrace{Ax \le b, \quad x \ge 0}_{\text{feasible points to LP}}, \quad x \text{ integer}}_{\text{feasible points to IP}}$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$.

Valid inequalities: A linear inequality is a **valid** for an IP if all feasible integer solutions satisfy the inequality. Consequently, the inequality must be satisfied by all points in the convex hull of integer solutions.

- Yellow region is the LP relaxation
- Green region is the convex hull of feasible integer points.



1.3 Perfect formulations

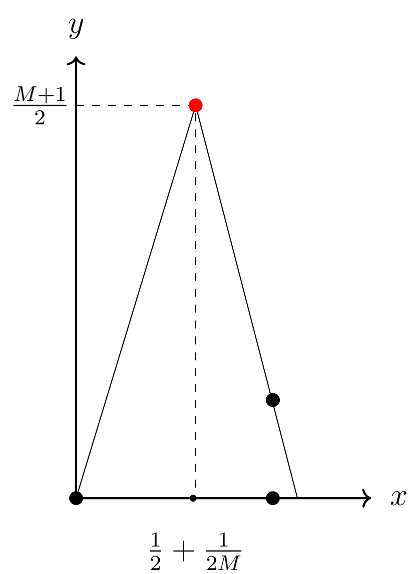
Example Consider the following simple IP where M is a big number:

$$\begin{aligned} \max \quad & x + y \\ \text{s.t.} \quad & y \le Mx \\ & y + Mx \le M + 1 \\ & x, y \ge 0 \text{ and integer} \end{aligned}$$

Feasible integer points are:

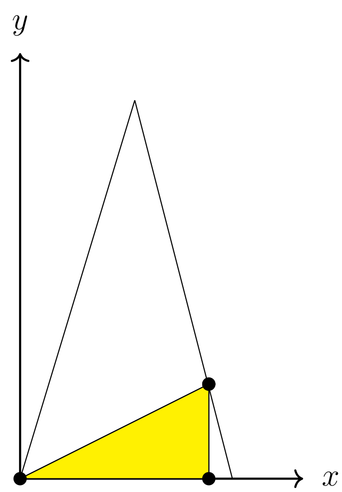
$$S = \Big\{(0,1), (1,0), (1,1)\Big\}$$

and $z^{IP} = 2$



We have $z^{LP} > M/2$ which is much bigger than $2 = z^{IP}$ when M is large (like 10^6).

Any valid inequality for this IP must be satisfied by all points in the convex hull of feasible integer points. The convex hull of $S = \{(0,1),(1,0),(1,1)\}$ is the yellow region below and it is defined by 3 inequalities:



Therefore, the best formulation of the IP is obtained by taking the convex hull of its feasible (integer) solutions:

$$
\begin{array}{rl}
\max & x + y \\
\text{s.t.} & \cancel{y \ \leq \ 100x} \quad \text{redundant now} \\
& \cancel{y + 100x \ \leq \ 101} \quad \text{redundant now} \\
& y - x \ \leq \ 0 \\
& x \ \leq \ 1 \\
& \cancel{x \geq 0}, \quad y \geq 0 \text{ and integer}
\end{array}
$$

- The convex hull of the integer points gives the best formulation in the sense that it has the smallest possible LP relaxation for a correct formulation. Anything smaller must exclude some feasible integer points.
- Such a formulation is called a **perfect formulation**. (it is usually very hard to obtain these perfect formulations...)
- All extreme points of perfect formulations are integral $\implies$ LP solves the IP!

Example: Remember the following sets:

$$K^1 = \left\{ \ x \in \{0,1\}^3 \ : 4x_1 + 5x_2 + 6x_3 \leq 7 \right\}$$

$$K^2 = \left\{ \ x \in \{0,1\}^3 : \begin{array}{r} x_1 + x_2 \leq 1 \\ x_1 + x_3 \leq 1 \\ x_2 + x_3 \leq 1 \end{array} \right\}$$

$$K^3 = \left\{ \ x \in \{0,1\}^3 \ : x_1 + x_2 + x_3 \leq 1 \right\}$$

where

$$K^1 = K^2 = K^3 = \Big\{ (0,0,0), (1,0,0), (0,1,0), (0,0,1) \Big\} \quad \longleftarrow \ \text{let's call this } K$$

The convex hull of these integer points is:

$$\mathbf{conv}(K) = \left\{x \in \mathbb{R}^3 : \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \lambda_0 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},\ \lambda_0+\lambda_1+\lambda_2+\lambda_3 = 1,\ \lambda \geq 0\right\}$$

in other words,

$$\mathbf{conv}(K) = \left\{x \in \mathbb{R}^3 \ :\ x_1 = \lambda_1,\ x_2 = \lambda_2,\ x_3 = \lambda_3,\quad \lambda_0 + \lambda_1 + \lambda_2 + \lambda_3 = 1,\quad \lambda \geq 0\right\}$$

or

$$\mathbf{conv}(K) = \left\{x \in \mathbb{R}^3 \ :\quad \lambda_0 + x_1 + x_2 + x_3 = 1,\quad x \geq 0,\ \lambda_0 \geq 0\right\}$$

or

$$\mathbf{conv}(K) = \left\{x \in \mathbb{R}^3 \ :\ x_1 + x_2 + x_3 \ \leq \ 1,\quad x \geq 0,\right\}$$

meaning K^3, without the upper bounds $x_i \leq 1$, is a perfect formulation!

(Note: Finding convex hulls of integer points is usually hard.)