

1 Linear programming – 3D example

1.1 Setup

Three variables $(x_1, x_2, x_3)^\top = \mathbf{x}$ with (a) non-negativity constraints, i.e., $x_i \geq 0$ for $i = \{1, 2, 3\}$ and (b) 5 bounds as follows:

- (b.1) $\mathbf{a}_1^\top \mathbf{x} \leq b_1$ with $\mathbf{a}_1^\top = (0, -4, 1)$ and $b_1 = 8$
- (b.2) $\mathbf{a}_2^\top \mathbf{x} \leq b_2$ with $\mathbf{a}_2^\top = (1, 0, -1)$ and $b_2 = 2$
- (b.3) $\mathbf{a}_3^\top \mathbf{x} \leq b_3$ with $\mathbf{a}_3^\top = (1, 2, 0)$ and $b_3 = 4$
- (b.4) $\mathbf{a}_4^\top \mathbf{x} \leq b_4$ with $\mathbf{a}_4^\top = (1, 4, 1)$ and $b_4 = 8$
- (b.5) $\mathbf{a}_5^\top \mathbf{x} \leq b_5$ with $\mathbf{a}_5^\top = (0, 0, 1)$ and $b_5 = 6$

Note that constraint (b.5) amounts to limiting x_3 “from above” so that, together with non-negativity constraint (a.3), we have $0 \leq x_3 \leq 6 = b_5$.

1.2 Equation system with slack variables

We work with 5 slack variables $s_j \geq 0$, $j = \{1, 2, 3, 4, 5\}$, $\mathbf{s} = (s_1, s_2, s_3, s_4, s_5)^\top$, for the 5 bounds (b.1)–(b.5) in addition to the 3 variables x_i , $i = \{1, 2, 3\}$, to setup 5 equations:

$$\mathbf{A} \mathbf{x} = \mathbf{b} - \mathbf{s} \quad \text{with} \quad \mathbf{A} = \begin{pmatrix} \mathbf{a}_1^\top \\ \mathbf{a}_2^\top \\ \vdots \\ \mathbf{a}_5^\top \end{pmatrix} \in \mathbb{R}^{5 \times 3} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_5 \end{pmatrix} \in \mathbb{R}^{5 \times 1}$$

1.3 Objective function

In the following, we aim at maximizing $f(\mathbf{x}) = \mathbf{c}^\top \mathbf{x}$ with $\mathbf{c}^\top = (1, 1, 1)$.

1.4 Starting point

We start the process of finding the optimal solution $\hat{\mathbf{x}}$ with basic feasible solution $\mathbf{p}_0 = \mathbf{0} = (0, 0, 0)^\top$. Clearly, \mathbf{p}_0 fulfills all non-negativity constraints (a.1)–(a.3) as well as all bounds (b.1)–(b.5), since all $b_j \geq 0$ and all $\mathbf{a}_j^\top \mathbf{p}_0 = \mathbf{a}_j^\top \mathbf{0} = 0 \leq b_j$.

Since $\mathbf{A} \mathbf{p}_0 = \mathbf{0}$, we easily find the values of all slack variables by $\mathbf{s}_0 = \mathbf{b} - \mathbf{A} \mathbf{p}_0 = \mathbf{b}$. We also find the value of our objective function f at \mathbf{p}_0 : $f(\mathbf{p}_0) = \mathbf{c}^\top \mathbf{p}_0 = 0$. To keep track, we bring all of this, i.e., the current approximation of the solution, the according slack variables, and the according objective function, together in one track vector $\mathbf{t}_0^\top = (0 \ 0 \ 0 \mid 8 \ 2 \ 4 \ 8 \ 6 \mid 0)$.

As long as we stick to basic feasible solutions \mathbf{p}_k , always three of the first 8 components of \mathbf{t}_k will be zero. Furthermore, we continue only as long as we do not decrease the last component of the track vector, i.e., the objective function.

1.5 First iteration

In each iteration, we can (possibly) “drive up” one of the three zero-components (among the first eight \mathbf{t} -components), bringing another one “down” to 0. In the first iteration, we can choose one from three zero-components, i.e., from the first three, representing x_1 , x_2 , and x_3 – in subsequent iterations, we avoid “going back” by again increasing the component that we just reduced to 0 in the last step, of course. Also, we check, whether the next step will contribute to further increasing the objective function.

The first iteration is, as indicated, as little special (and also a bit easier, altogether). To choose, which x_i to “drive up”, we could check their influence on the objective function (and choose the one x_i that contributes the most). In this example, all x_i contribute equally (since $\mathbf{c}^\top = (1, 1, 1)$), so we simply choose x_1 for the first iteration, keeping $x_2 = x_3 = 0$. If the set of all feasible solutions is finite, i.e., nowhere “open to infinity”, then we can assume that at least one slack variable will decrease (linearly) as we increase x_1 . Our goal with this first iteration is to increase x_1 as much as possible, necessarily bringing “down” this slack variable to 0, accordingly. We thus search (in this first iteration) for a new, basic feasible solution $\mathbf{p}_1 = (p_{1,1}, 0, 0)^\top$, with maximally large $p_{1,1} > 0$, while not violating $\mathbf{s}_1 \geq \mathbf{0}$.

To check (and to find $p_{1,1}$), we write $\mathbf{s}_1 = \mathbf{b} - \mathbf{A} \mathbf{p}_1 = \mathbf{b} - \mathbf{A} (p_{1,1}, 0, 0)^\top$:

$$\mathbf{s}_1 = \begin{pmatrix} s_{1,1} \\ s_{1,2} \\ s_{1,3} \\ s_{1,4} \\ s_{1,5} \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \\ 4 \\ 8 \\ 6 \end{pmatrix} - \begin{pmatrix} 0 & -4 & 1 \\ 1 & 0 & -1 \\ 1 & 2 & 0 \\ 1 & 4 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_{1,1} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 8 \\ 2 - p_{1,1} \\ 4 - p_{1,1} \\ 8 - p_{1,1} \\ 6 \end{pmatrix}$$

We see that three slack variables will decrease as $p_{1,1}$ increases and we also see that it is the second slack variable, $s_{1,2}$, that limits how much we can increase $p_{1,1}$ (without violating $\mathbf{s}_1 \geq \mathbf{0}$). In fact, it is also clear that the maximal value for $p_{1,1}$ is 2 (according to the $2 - p_{1,1} \geq 0$ requirement).

We thus have found the next-best approximation as $\mathbf{p}_1 = (2, 0, 0)^\top$. Evaluating the above used equation for \mathbf{s}_1 (with $p_{1,1} = 2$) and $f(\mathbf{p}_1) = \mathbf{c}^\top \mathbf{p}_1$, we can also fill in the track vector for this first iteration: $\mathbf{t}_1^\top = (2 \ 0 \ 0 \mid 8 \ 0 \ 2 \ 6 \ 6 \mid 2)$. According to the design of this first iteration, we (again) find that three (out of the first eight) components are 0 (now x_2 , x_3 , and s_2). We also see that we have successfully increased the objective function.

1.6 Second iteration – the first “typical” one

We continue with the principle that we (try to) “drive up” one of the three zero-components in \mathbf{t}_1 – but only doing so, if we can further increase the objective function, also (so we have to check that first). We won’t attempt to “drive up” s_2 , though, as this would bring us back to the initial approximation $\mathbf{p}_0 = \mathbf{0}$. In the following, we first “forget” about this, including s_2 also among the candidates for \mathbf{t} -components that may be “driven up”, showing that this would anyway lead to a decreasing objective function (not an option). Accordingly, we first consider the \mathbf{t} -components corresponding to x_2 , x_3 , and s_2 as candidates, doing the following steps one-by-one (in each case of considering one of the three candidates, we keep the other two as they are, i.e., at zero):

1. We use three equations from above, corresponding to the three candidates, to express \mathbf{x} in terms of the one “to-be-driven-up” candidate c , i.e., $\mathbf{x}(c)$, assuming the other two to remain 0.
2. We then consider the expected influence on the objective function by computing df/dc – if that derivative is negative (f would decrease with increasing c), we discard this candidate c .

Eventually, we choose one non-discarded candidate c – possibly the one with the largest df/dc . For this candidate c , we then check, how the other \mathbf{t} -components, which weren’t considered as possible candidates up-front, change as we “drive up” c , taking a closer look at those that would decrease (we need to ensure that we do not violate the non-negativity constraints for all x_i and s_j). The one decreasing component that “hits” 0 “first” determines the maximal value to which we can “drive up” c .

1.6.1 Should we “drive up” x_2 ?

We use the following three equations, involving \mathbf{x} (including x_2 and x_3) and s_2 :

$$\begin{aligned} -x_2 &= 0 - x_2 && // \ x_2 \text{ is its own slack variable, along with } x_2 \geq 0 \rightarrow -x_2 \leq 0 \\ -x_3 &= 0 && // \ x_3 \text{ is kept zero} \dots \\ \mathbf{a}_2^\top \mathbf{x} &= b_2 && // \text{constraint (b.2), but with } s_2 = 0 \end{aligned}$$

These three equations can be written as

$$\mathbf{L} \mathbf{x} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix} \mathbf{x} = \mathbf{R} - (x_2, 0, 0)^\top = \begin{pmatrix} 0 \\ 0 \\ b_2 \end{pmatrix} - \begin{pmatrix} x_2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -x_2 \\ 0 \\ 2 \end{pmatrix}$$

We thus get

$$\mathbf{x}(x_2) = \mathbf{L}^{-1} \left(\mathbf{R} - (x_2, 0, 0)^\top \right) = \mathbf{L}^{-1} \mathbf{R} - \mathbf{L}^{-1} (x_2, 0, 0)^\top \quad \text{with } \mathbf{L}^{-1} = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}$$

This leads to

$$f(\mathbf{x}(x_2)) = \mathbf{c}^\top \mathbf{x}(x_2) = \mathbf{c}^\top \mathbf{L}^{-1} \left(\mathbf{R} - (x_2, 0, 0)^\top \right)$$

and to

$$df/dx_2 = -\mathbf{c}^\top \mathbf{L}^{-1} (1, 0, 0)^\top = -(1, 1, 1) (0, -1, 0)^\top = 1 > 0$$

This shows that we can keep x_2 as a candidate for a \mathbf{t} -component to be “driven up”.

1.6.2 Should we maybe “drive up” x_3 ?

We use the following three equations:

$$\begin{aligned} -x_2 &= 0 && // x_2 \text{ is kept zero in this case...} \\ -x_3 &= 0 - x_3 && // x_3 \text{ is its own slack variable, along with } x_3 \geq 0 \rightarrow -x_3 \leq 0 \\ \mathbf{a}_2^\top \mathbf{x} &= b_2 && // \text{constraint (b.2), but with } s_2 = 0 \end{aligned}$$

We rewrite these equations, again, as

$$\mathbf{L} \mathbf{x} = \mathbf{R} - (0, x_3, 0)^\top \quad // \mathbf{L} \text{ and } \mathbf{R} \text{ as above (section 1.6.1)}$$

We get

$$\mathbf{x}(x_3) = \mathbf{L}^{-1} \left(\mathbf{R} - (0, x_3, 0)^\top \right)$$

and

$$f(\mathbf{x}(x_3)) = \mathbf{c}^\top \mathbf{L}^{-1} \left(\mathbf{R} - (0, x_3, 0)^\top \right)$$

as well as

$$df/dx_3 = -\mathbf{c}^\top \mathbf{L}^{-1} (0, 1, 0)^\top = -(1, 1, 1) (-1, 0, -1)^\top = 2 > 0$$

Accordingly, we can also keep x_3 as a candidate for a \mathbf{t} -component to be “driven up”. In fact, it may give us more benefit as $df/dx_3 = 2 > 1 = df/dx_2$.

1.6.3 Should we still “drive up” s_2 (against all odds)?

Knowing that this won’t work, we still look at this “candidate” for illustration purposes:

$$\begin{aligned} -x_2 &= 0 && // x_2 \text{ is kept zero in this case...} \\ -x_3 &= 0 && // x_3 \text{ is also kept zero in this case...} \\ \mathbf{a}_2^\top \mathbf{x} &= b_2 - s_2 && // \text{constraint (b.2)...} \end{aligned}$$

This leads to

$$\begin{aligned} \mathbf{L} \mathbf{x} &= \mathbf{R} - (0, 0, s_2)^\top \\ \mathbf{x}(s_2) &= \mathbf{L}^{-1} \left(\mathbf{R} - (0, 0, s_2)^\top \right) \\ f(\mathbf{x}(s_2)) &= \mathbf{c}^\top \mathbf{L}^{-1} \left(\mathbf{R} - (0, 0, s_2)^\top \right) \\ df/ds_2 &= -\mathbf{c}^\top \mathbf{L}^{-1} (0, 0, 1)^\top = -(1, 1, 1) (1, 0, 0)^\top = -1 < 0 \end{aligned}$$

Any attempt to “drive up” s_2 here would result in decreasing the objective function (we don’t).

1.6.4 So let’s “drive up” x_3 !

When “driving up” x_3 , we expect at least one \mathbf{t} -component (that was not a candidate for being “driven up”, i.e., x_1, s_1, s_3, s_4 , or s_5) to decrease (x_2 and s_2 are kept equal to 0, while x_3 is “driven up”). We can express them depending on x_3 to check:

$$\begin{aligned} \mathbf{x}(x_3) &= \mathbf{L}^{-1} \mathbf{R} - \mathbf{L}^{-1} (0, x_3, 0)^\top \\ \mathbf{s}(x_3) &= \mathbf{b} - \mathbf{A} \mathbf{x}(x_3) \end{aligned}$$

This leads to

$$\mathbf{x}(x_3) = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} -x_3 \\ 0 \\ -x_3 \end{pmatrix} = \begin{pmatrix} 2 + x_3 \\ 0 \\ x_3 \end{pmatrix}$$

and

$$\mathbf{s}(x_3) = \begin{pmatrix} 8 \\ 2 \\ 4 \\ 8 \\ 6 \end{pmatrix} - \begin{pmatrix} 0 & -4 & 1 \\ 1 & 0 & -1 \\ 1 & 2 & 0 \\ 1 & 4 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 + x_3 \\ 0 \\ x_3 \end{pmatrix} = \begin{pmatrix} 8 - x_3 \\ 0 \\ 2 - x_3 \\ 6 - 2x_3 \\ 6 - x_3 \end{pmatrix} \quad \begin{array}{l} // \text{decreasing to 0 for } x_3 = 8 \\ // \text{constant 0 (we required } s_2 = 0) \\ // \text{decreasing to 0 for } x_3 = 2 \\ // \text{decreasing to 0 for } x_3 = 3 \\ // \text{decreasing to 0 for } x_3 = 6 \end{array}$$

To not violate the non-negativity constraints, 2 is the maximal value up to which we can “drive up” x_3 – at $x_3 = 2$, slack variable s_3 vanishes (required for “arriving” at a new basic feasible solution).

With all this, we now “arrived” at the next-best approximation $\mathbf{p}_2 = \mathbf{x}(x_3) = (4, 0, 2)^\top$ with a further increased objective function, $f(\mathbf{p}_2) = 6$. Our updated track vector is $\mathbf{t}_2^\top = (4 \ 0 \ 2 \mid 6 \ 0 \ 0 \ 2 \ 4 \mid 6)$.

1.7 Third iteration

Three \mathbf{t} -components are zero (x_2 , s_2 , and s_3) and we consider two of them as candidates for driving them up: x_2 and s_2 – driving up s_3 (just vanished) would lead to a decreased objective function (not good). We start with identifying the relevant equations (to find \mathbf{L} and \mathbf{R}):

$$\begin{aligned} -x_2 &= 0 - x_2 \\ \mathbf{a}_2^\top \mathbf{x} &= b_2 - s_2 \\ \mathbf{a}_3^\top \mathbf{x} &= b_3 - s_3 \end{aligned} \Rightarrow \mathbf{L} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 1 & 2 & 0 \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix}; \quad \mathbf{L}^{-1} = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 0 & 0 \\ 2 & -1 & 1 \end{pmatrix}$$

Checking candidate x_2 , we compute $df/dx_2 = -\mathbf{c}^\top \mathbf{L}^{-1} (1, 0, 0)^\top = -(1, 1, 1) (2, -1, 2)^\top = -3 < 0$.
 Checking candidate s_2 , we compute $df/ds_2 = -\mathbf{c}^\top \mathbf{L}^{-1} (0, 1, 0)^\top = -(1, 1, 1) (0, 0, -1)^\top = 1 > 0$.
 Checking candidate s_3 , we compute $df/ds_3 = -\mathbf{c}^\top \mathbf{L}^{-1} (0, 0, 1)^\top = -(1, 1, 1) (1, 0, 1)^\top = -2 < 0$.
 Accordingly, there's only one way to further increase f : drive up s_2 !

$$\begin{aligned} \mathbf{x}(s_2) &= \mathbf{L}^{-1} \mathbf{R} - \mathbf{L}^{-1} (0, s_2, 0)^\top = \begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ -s_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 2 + s_2 \end{pmatrix} \\ \mathbf{s}(s_2) &= \mathbf{b} - \mathbf{A} \mathbf{x}(s_2) = \begin{pmatrix} 8 \\ 2 \\ 4 \\ 8 \\ 6 \end{pmatrix} - \begin{pmatrix} 2 + s_2 \\ 2 - s_2 \\ 4 \\ 6 + s_2 \\ 2 + s_2 \end{pmatrix} = \begin{pmatrix} 6 - s_2 \\ s_2 \\ 0 \\ 2 - s_2 \\ 4 - s_2 \end{pmatrix} \quad \begin{array}{l} // \text{ decreasing to 0 for } s_2 = 6 \\ // \text{ just (increasing) } s_2, \text{ of course} \\ // \text{ kept zero } \dots \\ // \text{ decreasing to 0 for } s_2 = 2 \\ // \text{ decreasing to 0 for } s_2 = 4 \end{array} \end{aligned}$$

We drive s_2 up to the value of 2 (s_4 vanishes). This leads to $\mathbf{t}_3^\top = (4 \ 0 \ 4 \mid 4 \ 2 \ 0 \ 0 \ 2 \mid 8)$.

1.8 Fourth iteration

Three \mathbf{t} -components are zero (x_2 , s_3 , and s_4) and we consider two of them as candidates for driving them up: x_2 and s_3 – driving up s_4 (just vanished) would lead to a decreased objective function (not good). We start with identifying the relevant equations (to find \mathbf{L} and \mathbf{R}):

$$\begin{aligned} -x_2 &= 0 - x_2 \\ \mathbf{a}_3^\top \mathbf{x} &= b_3 - s_3 \\ \mathbf{a}_4^\top \mathbf{x} &= b_4 - s_4 \end{aligned} \Rightarrow \mathbf{L} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 2 & 0 \\ 1 & 4 & 1 \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} 0 \\ 4 \\ 8 \end{pmatrix}; \quad \mathbf{L}^{-1} = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 0 & 0 \\ 2 & -1 & 1 \end{pmatrix}$$

Checking candidate x_2 , we compute $df/dx_2 = -\mathbf{c}^\top \mathbf{L}^{-1} (1, 0, 0)^\top = -(1, 1, 1) (2, -1, 2)^\top = -3 < 0$.
 Checking candidate s_3 , we compute $df/ds_3 = -\mathbf{c}^\top \mathbf{L}^{-1} (0, 1, 0)^\top = -(1, 1, 1) (1, 0, -1)^\top = 0$.

While there's no way to further increase f , we could drive up s_3 without decreasing f , at least: All feasible solutions along this path, until “hitting” the next basic feasible solution, would be optimal (with the same value of the objective function). Even though this is not necessary, since we have one optimal solution already, i.e., $\mathbf{p}_3 = (4, 0, 4)^\top$, we still do this next step here to find all optimal solutions:

$$\begin{aligned} \mathbf{x}(s_3) &= \mathbf{L}^{-1} \mathbf{R} - \mathbf{L}^{-1} (0, s_3, 0)^\top = \begin{pmatrix} 4 \\ 0 \\ 4 \end{pmatrix} - \begin{pmatrix} s_3 \\ 0 \\ -s_3 \end{pmatrix} = \begin{pmatrix} 4 - s_3 \\ 0 \\ 4 + s_3 \end{pmatrix} \\ \mathbf{s}(s_3) &= \mathbf{b} - \mathbf{A} \mathbf{x}(s_3) = \begin{pmatrix} 8 \\ 2 \\ 4 \\ 8 \\ 6 \end{pmatrix} - \begin{pmatrix} 4 + s_3 \\ -2s_3 \\ 4 - s_3 \\ 8 \\ 4 + s_3 \end{pmatrix} = \begin{pmatrix} 4 - s_3 \\ 2 + 2s_3 \\ s_3 \\ 0 \\ 2 - s_3 \end{pmatrix} \quad \begin{array}{l} // \text{ decreasing to 0 for } s_3 = 4 \\ // \text{ increasing} \\ // \text{ just (increasing) } s_3, \text{ of course} \\ // \text{ kept zero } \dots \\ // \text{ decreasing to 0 for } s_3 = 2 \end{array} \end{aligned}$$

We can drive s_3 to the value of 2 (s_5 vanishes). This leads to $\mathbf{t}_4^\top = (2 \ 0 \ 6 \mid 2 \ 6 \ 2 \ 0 \ 0 \mid 8)$.

Any solution $\hat{\mathbf{x}}(t) = (4-t, 0, 4+t)^\top$ with $t \in [0, 2]$ is optimal and the (optimal) value of the objective function for these solutions is $f(\hat{\mathbf{x}}(t)) = \mathbf{c}^\top (4-t, 0, 4+t)^\top = 8$.

To double-check, we can have a quick look (no reasons for any hopes here, though!): the vanished \mathbf{t} -components are x_2 , s_4 , and s_5 (s_5 just vanished in the last step). We use the following equations:

$$\begin{aligned} -x_2 &= 0 - x_2 \\ \mathbf{a}_4^\top \mathbf{x} &= b_4 - s_4 \\ \mathbf{a}_5^\top \mathbf{x} &= b_5 - s_5 \end{aligned} \quad \Rightarrow \quad \mathbf{L} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 4 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} 0 \\ 8 \\ 6 \end{pmatrix}; \quad \mathbf{L}^{-1} = \begin{pmatrix} 4 & 1 & -1 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Checking candidate x_2 , we compute $df/dx_2 = -\mathbf{c}^\top \mathbf{L}^{-1} (1, 0, 0)^\top = -(1, 1, 1)(4, -1, 0)^\top = -3 < 0$.
 Checking candidate s_4 , we compute $df/ds_4 = -\mathbf{c}^\top \mathbf{L}^{-1} (0, 1, 0)^\top = -(1, 1, 1)(1, 0, 0)^\top = -1 < 0$.

It's now clear that we can only decrease the objective function by doing one more step.

To sum up, we can repeat all track vectors:

$$\begin{aligned} \mathbf{t}_0^\top &= (0 \ 0 \ 0 \mid 8 \ 2 \ 4 \ 8 \ 6 \mid 0) \\ \mathbf{t}_1^\top &= (2 \ 0 \ 0 \mid 8 \ 0 \ 2 \ 6 \ 6 \mid 2) \\ \mathbf{t}_2^\top &= (4 \ 0 \ 2 \mid 6 \ 0 \ 0 \ 2 \ 4 \mid 6) \\ \mathbf{t}_3^\top &= (4 \ 0 \ 4 \mid 4 \ 2 \ 0 \ 0 \ 2 \mid 8) \\ \mathbf{t}_4^\top &= (2 \ 0 \ 6 \mid 2 \ 6 \ 2 \ 0 \ 0 \mid 8) \end{aligned}$$

We see (again), how we increased the objective function from $0 = f(\mathbf{p}_0)$ to $8 = f(\mathbf{p}_3) = f(\mathbf{p}_4) = f(\hat{\mathbf{x}}(t))$. It's also interesting to see that driving up x_2 never promised enough. Furthermore, one can look at the distribution of slack (from iteration to iteration).

Finally (for now and here), we can also double-check with MatLab:

```
linprog(-[1 1 1], A, b, [], [], [0 0 0], [])
after A = [0 -4 1; 1 0 -1; 1 2 0; 1 4 1; 0 0 1]
and b = [8; 2; 4; 8; 6]
indeed returns (4, 0, 4)^\top = \hat{\mathbf{x}}(0) as result.
```

1.9 A more general approach

The above described approach works as long as exactly three, non-agreeing constraints fully determine a basic feasible solution. The above exercised example, in fact, also comes with a basic feasible solution where four constraints agree without slack: $(0, 2, 0)^\top$

- clearly, all non-negativity constraints on the three x_i are fulfilled
- further, $\mathbf{A} (0, 2, 0)^\top = (-8, 0, 4, 8, 0)^\top \leq \mathbf{b}$
- since we find (at least) three constraints that are fulfilled “without slack”, we accept $(0, 2, 0)^\top$ as basic feasible solution – in fact, we have $x_1 = x_3 = s_3 = s_4 = 0$, i.e., four such constraints

To also deal well with such situations, we need to generalize the approach a little bit (check yourself, why that's necessary¹). To do so, we reframe the problem, also, such that we can handle the non-negativity constraints on the x_i together with the bounds (b.1)–(b.5): Formulating the problem on the basis of \mathbf{t} works (“track vector” from above) – we “just” need to remember that $t_i = x_i$ for $i \in \{1, 2, 3\}$ and $t_{j+3} = s_j$ for $j \in \{1, 2, \dots, 5\}$. We keep the objective function in the last component of \mathbf{t} , i.e., $t_9 = f(\mathbf{x})$.

To make this work, we need to adapt \mathbf{A} and \mathbf{b} such that $\mathbf{A} \mathbf{x} = \mathbf{b} - \mathbf{t}$, bringing the variables x_i and the slack variables s_j together into one setup with $\mathbf{A} \in \mathbb{R}^{9 \times 3}$ and $\mathbf{t}, \mathbf{b} \in \mathbb{R}^{9 \times 1}$:

- the right-hand side \mathbf{b} (or $\mathbf{b} - \mathbf{t}$ as set up above) is easy to start with: the non-negativity constraints on the x_i amount to axis-parallel planes including the origin $\mathbf{0} = (0, 0, 0)^\top$ with formulae $\mathbf{e}_i^\top \mathbf{x} = 0$ with the three \mathbf{e}_i being aligned with the three axes, leading to $b_i = 0$ for $i \in \{1, 2, 3\}$; the next five components of \mathbf{b} are taken directly from bounds (b.1)–(b.5), i.e., $b_{j+3} = {}^{1.1}b_j$ (the “1.1” superscript “on the left” denotes b -values as set up above in Section 1.1); the last component of \mathbf{b} , i.e., b_9 , can be safely set to 0; altogether, this gives $\mathbf{b}^\top = (0, 0, 0, 8, 2, 4, 8, 6, 0)$ in this example
- the first three rows of \mathbf{A} and \mathbf{b} correspond to the non-negativity constraints on the three x_i , i.e., $x_i \geq 0$; to get $b_i - x_i$ on the right-hand side (as set up above), we get the first three rows of \mathbf{A} as $a_{i,k} = -\delta_{i,k}$ for $i, k \in \{1, 2, 3\}$; the next five rows of \mathbf{A} are then given by $a_{j+3,i} = {}^{1.1}a_{j,i}$ for $j \in \{1, 2, \dots, 5\}$ and $i \in \{1, 2, 3\}$; the last row of \mathbf{A} needs to be set to $a_{9,i} = -c_i$ for $i \in \{1, 2, 3\}$

¹ hint: check the dimensions of \mathbf{L}

Writing the whole system out for this example, we get

$$\mathbf{A} \mathbf{t} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ \hline 0 & -4 & 1 \\ 1 & 0 & -1 \\ 1 & 2 & 0 \\ 1 & 4 & 1 \\ 0 & 0 & 1 \\ \hline -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \hline 8 \\ 2 \\ 4 \\ 8 \\ 6 \\ \hline 0 \end{pmatrix} - \begin{pmatrix} t_1 = x_1 \\ t_2 = x_2 \\ t_3 = x_3 \\ \hline t_4 = s_1 \\ t_5 = s_2 \\ t_6 = s_3 \\ t_7 = s_4 \\ t_8 = s_5 \\ \hline t_9 = f(\mathbf{x}) \end{pmatrix} = \mathbf{b} - \mathbf{t}$$

Note that the left-hand side represents all needed linear relationships between the x_i (for the non-negativity constraints on x_i , the five bounds, and the objective function), while the right-hand side represents the actual bounds minus all corresponding “slack” (to account for the inequalities).

We can also rewrite to express all the “slack” \mathbf{t} – including the x_i and $f(\mathbf{x})$, of course – as the difference between the upper bounds and $\mathbf{A} \mathbf{x}$: $\mathbf{t} = \mathbf{b} - \mathbf{A} \mathbf{x}$ (should not become negative ever during the iteration).

For the following, it also makes sense to visualize (in our minds) that the above system is describing a bounded region in 3D (over \mathbf{x}) where the outer boundary is piece-wise planar. We assume (a) that the problem is given such that only finite feasible solutions exist – the region is actually bounded on all sides, i.e., nowhere “open to infinity” – and (b) that at least one (basic) feasible solution exists, i.e., that the problem is not over-constrained ($\mathbf{p}_0 = \mathbf{0} = (0, 0, 0)^\top$ is one such basic feasible solution in our case).

It also helps to visualize that every row k in $\mathbf{A} \mathbf{x} = \mathbf{b}$ represents a (potentially bounding) plane around the bounded region according to $\mathbf{a}_{k,\dots} \mathbf{x} = b_k$, where $\mathbf{a}_{k,\dots}^\top$ is the outwards-facing normal of that plane². Basic feasible solutions are vectors \mathbf{x}_{bfs} that satisfy (at least) three actually bounding planes, i.e., they are the corners/vertices of the piece-wise planar bounding geometry, satisfying the (at least) three according plane equations $\mathbf{a}_{k,\dots} \mathbf{x} = b_k$, meaning that the according slack variables t_k are all 0.

For the iteration, we stick to the principle that we move from one basic feasible solution \mathbf{p}_i to the next-best one, \mathbf{p}_{i+1} , via an edge of the bounding geometry. These edges are linear sets of feasible solutions $\mathbf{p}_i + t \mathbf{v}$ with non-negative t that satisfy (at least) two of the plane equations that both \mathbf{p}_i and \mathbf{p}_{i+1} satisfy. As these edges are part of the bounding geometry, none of the feasible solutions along them violate any of the $t_k \geq 0$, of course (at least two of them are equal to 0). With all this, we are ready to formulate our generalized approach (per iteration from the currently best approximation \mathbf{p}_i):

1. find all $k \in \{1, 2, \dots, 8\}$ such that $\mathbf{a}_{k,\dots}^\top \mathbf{p}_i = b_k$, i.e., with $t_k = 0$ (find all planes containing \mathbf{p}_i), resulting in $n \geq 3$ such planes
2. find all $\binom{n}{2}$ pairs of these planes³, determining the (candidates of) boundary edges through \mathbf{p}_i – if $n > 3$, then at most n of them actually are part of the boundary
3. for each edge (edge candidate), determine \mathbf{v} such that $\mathbf{p}_i + t \mathbf{v}$ represents traversing this edge (edge candidate) away from \mathbf{p}_i (in n D we can use the SVD to find \mathbf{v} , in 3D the cross-product is easier)
4. ignore all edge candidates outside of the boundary as well as edges that would lead to decreasing the objective function when moving away from \mathbf{p}_i
5. if the set of remaining edges is empty, then \mathbf{p}_i is optimal, otherwise, we continue along one of the remaining edges to \mathbf{p}_{i+1} (for ex., along the one that has the greatest effect on increasing f)
6. if we’ve found an edge on the boundary, represented by \mathbf{v} , that does not lead to a decreasing of the objective function, we need to find the maximally-large value of t , leading to \mathbf{p}_{i+1} ; we check for all planes with slack, whether moving along \mathbf{v} leads to decreasing their slack – if so, we determine the value of t for each such plane that would decrease the corresponding slack to 0. Of all those values of t , we keep the smallest; then $\mathbf{p}_{i+1} = \mathbf{p}_i + t \mathbf{v}$

Let’s do this for the above specified example:

² given non-negative slack t_k , the right-hand side gets smaller ($b_k - t_k$) – leaving the plane into the direction of the normal, however, would increase the right-hand side according to $\mathbf{n}^\top (\mathbf{p} + t \mathbf{n}) = \mathbf{n}^\top \mathbf{p} + t \mathbf{n}^\top \mathbf{n} = \mathbf{n}^\top \mathbf{p} + t |\mathbf{n}|^2$ with $t > 0$

³ we assume that no two planes overlap with each other, i.e., that no two plane equations are multiples of each other

1.9.1 Initial guess \mathbf{p}_0

We start the iteration with $\mathbf{p}_0 = \mathbf{0} = (0, 0, 0)^\top$. The non-zero entries of $\mathbf{t}_0 = (0 \ 0 \ 0 \mid 8 \ 2 \ 4 \ 8 \ 6 \mid 0)^\top$, not counting $t_{09} = f(\mathbf{p}_0) = \mathbf{c}^\top \mathbf{0} = 0$, correspond to the three x_i . This means that there are three boundary planes that contain \mathbf{p}_0 and their equations are given by $\mathbf{A}_{1,2,3} \dots \mathbf{x} = b_{1,2,3}$:

$$\begin{pmatrix} \mathbf{a}_1^\top \\ \mathbf{a}_2^\top \\ \mathbf{a}_3^\top \end{pmatrix} \mathbf{x} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1^\top \\ x_2^\top \\ x_3^\top \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Considering the three possible pairs of these three planes, $\{(1, 2), (1, 3), (2, 3)\}$, we find three corresponding edge vectors $\{\mathbf{v}_{1,2}, \mathbf{v}_{1,3}, \mathbf{v}_{2,3}\}$ – since each one of them needs to lie in both of the defining planes, they accordingly are perpendicular to both of the according plane normals: $\mathbf{v}_{1,2} = (0, 0, 1)^\top \perp \mathbf{a}_1, \mathbf{a}_2$; $\mathbf{v}_{1,3} = (0, 1, 0)^\top \perp \mathbf{a}_1, \mathbf{a}_3$; and $\mathbf{v}_{2,3} = (1, 0, 0)^\top \perp \mathbf{a}_2, \mathbf{a}_3$; The only way for these vectors to possibly leave the boundary right away, would be if the lead to decreasing the zero slack of the non-involved plane – note that, by construction, the zero-slack of the two defining planes remains zero when moving along the corresponding edge vectors. This is easily checked by computing the inner product of $\mathbf{v}_{i,j}$ and \mathbf{a}_k with $k \neq i, j$ – as long as this inner product is non-positive, the edge does not leave the boundary:

$$\begin{aligned} \mathbf{v}_{1,2}^\top \mathbf{a}_3 &= (0, 0, 1) (0, 0, -1)^\top = -1 < 0 \\ \mathbf{v}_{1,3}^\top \mathbf{a}_2 &= (0, 1, 0) (0, -1, 0)^\top = -1 < 0 \\ \mathbf{v}_{2,3}^\top \mathbf{a}_1 &= (1, 0, 0) (-1, 0, 0)^\top = -1 < 0 \end{aligned}$$

We also check the prospective effect on the objective function, given that we would move along any of these three edge candidates $\mathbf{v}_{i,j}$ – if the inner product of $\mathbf{v}_{i,j}$ and \mathbf{c} is non-negative, then increasing t along $\mathbf{p}_0 + t \mathbf{v}_{i,j}$ would not lead to decreasing the objective function:

$$\begin{aligned} \mathbf{v}_{1,2}^\top \mathbf{c} &= (0, 0, 1) (1, 1, 1)^\top = 1 > 0 \\ \mathbf{v}_{1,3}^\top \mathbf{c} &= (0, 1, 0) (1, 1, 1)^\top = 1 > 0 \\ \mathbf{v}_{2,3}^\top \mathbf{c} &= (1, 0, 0) (1, 1, 1)^\top = 1 > 0 \end{aligned}$$

Given the above, we could follow each one of these three edge vectors, staying on the boundary and increasing the objective function – to proceed we choose one of them. Since all of the $\mathbf{v}_{i,j}^\top \mathbf{c}$ are equal, we simply choose the second one, $\mathbf{v}_{1,3} = (0, 1, 0)^\top$ – why not? :-). To find t , we check how moving along $\mathbf{v}_{1,3}$ impacts the non-zero slack in \mathbf{t} (in \mathbf{p}_0 , the five slack-values of $t_{4,5,\dots,8}$ are positive). We compute the inner products of $\mathbf{v}_{1,3}$ and $\mathbf{a}_{4,5,\dots,8}$ – if this inner product is positive for a plane k , then moving along $\mathbf{v}_{1,3}$ leads to decreasing the slack corresponding to plane k :

$$\begin{aligned} \mathbf{v}_{1,3}^\top \mathbf{a}_4 &= (0, 1, 0) (0, -4, 1)^\top = -4 < 0 \\ \mathbf{v}_{1,3}^\top \mathbf{a}_5 &= (0, 1, 0) (1, 0, -1)^\top = 0 \\ \mathbf{v}_{1,3}^\top \mathbf{a}_6 &= (0, 1, 0) (1, 2, 0)^\top = 2 > 0 \rightarrow \text{moving along } \mathbf{v}_{1,3} \text{ leads to decreasing } t_6 \\ \mathbf{v}_{1,3}^\top \mathbf{a}_7 &= (0, 1, 0) (1, 4, 1)^\top = 4 > 0 \rightarrow \text{moving along } \mathbf{v}_{1,3} \text{ leads to decreasing } t_7 \\ \mathbf{v}_{1,3}^\top \mathbf{a}_8 &= (0, 1, 0) (0, 0, 1)^\top = 0 \end{aligned}$$

For the two planes in question, with indices 6 and 7, we check which value of t would decrease “their” slack to 0:

$$\begin{aligned} t_6(t) &= b_6 - \mathbf{a}_6^\top (\mathbf{p}_0 + t \mathbf{v}_{1,3}) = (b_6 - \mathbf{a}_6^\top \mathbf{p}_0) - t \mathbf{a}_6^\top \mathbf{v}_{1,3} = 4 - 2t \rightarrow \text{decreases to 0 for } t = 2 \\ t_7(t) &= b_7 - \mathbf{a}_7^\top (\mathbf{p}_0 + t \mathbf{v}_{1,3}) = (b_7 - \mathbf{a}_7^\top \mathbf{p}_0) - t \mathbf{a}_7^\top \mathbf{v}_{1,3} = 8 - 4t \rightarrow \text{decreases to 0 for } t = 2 \end{aligned}$$

The farthest we can move along $\mathbf{p}_0 + t \mathbf{v}_{1,3}$ (without leaving the boundary) is thus given for $t = 2$, bringing us to the next-best basic feasible solution $\mathbf{p}_1 = \mathbf{p}_0 + 2 (0, 1, 0)^\top = (0, 2, 0)^\top$ with the (increased) objective function $f(\mathbf{p}_1) = 2$. Filling in the updated slack, we get $\mathbf{t}_1 = \mathbf{b} - \mathbf{A} \mathbf{p}_1 = (0 \ 2 \ 0 \mid 16 \ 2 \ 0 \ 0 \ 6 \mid 2)^\top$. Note that we’ve reached a special basic feasible solution in the sense that four values of slack (including the x_i) are equal to 0! This makes the next iteration non-generic / special.

1.9.2 Special case \mathbf{p}_1 : intersection of four boundary planes

Following our generalized approach, we first identify the planes that contain \mathbf{p}_1 – in fact, we already know which planes these are: planes 1 and 3 (row-index of \mathbf{A} and \mathbf{b}) defined the edge $\mathbf{v}_{1,3}$ that we came along to \mathbf{p}_1 (from \mathbf{p}_0); then we saw that the t -value of 2 brought two additional slack-variables down to 0, i.e., 6 and 7 (again by their row-indices).

Having four planes in \mathbf{p}_1 means that we can find – all in all – $\binom{4}{2} = 6$ pairs of planes that determine edge candidates: $\{(1, 3), (1, 6), (1, 7), (3, 6), (3, 7), (6, 7)\}$. The first of them, $(1, 3)$, would bring us right back to \mathbf{p}_0 – that’s not what we’ll do, of course. Using the cross-product (since we’re in 3D) and normalization, we get the following six edge candidates – each one normal to the two defining planes:

$$(\mathbf{v}_{1,3}, \mathbf{v}_{1,6}, \mathbf{v}_{1,7}, \mathbf{v}_{3,6}, \mathbf{v}_{3,7}, \mathbf{v}_{6,7}) \approx \begin{pmatrix} 0 & 0 & 0 & 0.8944 & -0.9701 & 0.6667 \\ -1 & 0 & -0.2425 & -0.4472 & 0.2425 & -0.3333 \\ 0 & -1 & 0.9701 & 0 & 0 & 0.6667 \end{pmatrix}$$

We check the inner products (for each edge candidate with the two other, non-defining planes) to see, which of these edge candidates stay on the boundary:

$$\begin{pmatrix} \mathbf{a}_6^\top \\ \mathbf{a}_7^\top \end{pmatrix} \mathbf{v}_{1,3} = \begin{pmatrix} -2 \\ -4 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{a}_3^\top \\ \mathbf{a}_7^\top \end{pmatrix} \mathbf{v}_{1,3} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{a}_3^\top \\ \mathbf{a}_6^\top \end{pmatrix} \mathbf{v}_{1,3} = \begin{pmatrix} -0.9701 \\ -0.4850 \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{a}_1^\top \\ \mathbf{a}_7^\top \end{pmatrix} \mathbf{v}_{1,3} = \begin{pmatrix} -0.8944 \\ -0.8944 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{a}_1^\top \\ \mathbf{a}_6^\top \end{pmatrix} \mathbf{v}_{1,3} = \begin{pmatrix} 0.9701 \\ -0.4851 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{a}_1^\top \\ \mathbf{a}_3^\top \end{pmatrix} \mathbf{v}_{1,3} = \begin{pmatrix} -0.6667 \\ -0.6667 \end{pmatrix}$$

This shows that the edge candidates that are given by plane-pairs $\{(1, 6), (3, 7)\}$ necessarily leave the boundary (even when flipped 180°) and thus must be discarded. Next, we check the prospective impact on the objective function:

$$\mathbf{c}^\top (\mathbf{v}_{1,3}, \mathbf{v}_{1,7}, \mathbf{v}_{3,6}, \mathbf{v}_{6,7}) = (-1, 0.7276, 0.4472, 1)$$

Obviously, “walking” back along $\mathbf{v}_{1,3}$ would decrease the objective function, again (as mentioned above) – traversing the other 3 edges $\{\mathbf{v}_{1,7}, \mathbf{v}_{3,6}, \mathbf{v}_{6,7}\}$, however, would lead to increasing the objective function. Since $\mathbf{v}_{6,7}$ promises the steepest increase, we follow that edge to \mathbf{p}_2 . Check all planes with slack, we get:

$$\begin{aligned} \mathbf{v}_{6,7}^\top \mathbf{a}_2 &= (2/3, -1/3, 2/3) (0, -1, 0)^\top = 1/3 > 0 \rightarrow \text{moving along } \mathbf{v}_{6,7} \text{ leads to decreasing } t_2 \\ \mathbf{v}_{6,7}^\top \mathbf{a}_4 &= (2/3, -1/3, 2/3) (0, -4, 1)^\top = 2 > 0 \rightarrow \text{moving along } \mathbf{v}_{6,7} \text{ leads to decreasing } t_4 \\ \mathbf{v}_{6,7}^\top \mathbf{a}_5 &= (2/3, -1/3, 2/3) (1, 0, -1)^\top = 0 \\ \mathbf{v}_{6,7}^\top \mathbf{a}_8 &= (2/3, -1/3, 2/3) (0, 0, 1)^\top = 2/3 > 0 \rightarrow \text{moving along } \mathbf{v}_{6,7} \text{ leads to decreasing } t_8 \end{aligned}$$

For the planes with decreasing slack, let’s see how far we could “go”:

$$\begin{aligned} t_2(t) &= b_2 - \mathbf{a}_2^\top (\mathbf{p}_1 + t \mathbf{v}_{6,7}) = (b_2 - \mathbf{a}_2^\top \mathbf{p}_1) - t \mathbf{a}_2^\top \mathbf{v}_{6,7} = 0 - (-2) - (1/3)t \rightarrow \text{decreases to 0 for } t = 6 \\ t_4(t) &= b_4 - \mathbf{a}_4^\top (\mathbf{p}_1 + t \mathbf{v}_{6,7}) = (b_4 - \mathbf{a}_4^\top \mathbf{p}_1) - t \mathbf{a}_4^\top \mathbf{v}_{6,7} = 8 - (-8) - 2t \rightarrow \text{decreases to 0 for } t = 8 \\ t_8(t) &= b_8 - \mathbf{a}_8^\top (\mathbf{p}_1 + t \mathbf{v}_{6,7}) = (b_8 - \mathbf{a}_8^\top \mathbf{p}_1) - t \mathbf{a}_8^\top \mathbf{v}_{6,7} = 6 - (2/3)t \rightarrow \text{decreases to 0 for } t = 9 \end{aligned}$$

Accordingly, we get $t = 6$ and $\mathbf{p}_2 = \mathbf{p}_1 + 6(2/3, -1/3, 2/3)^\top = (4, 0, 4)^\top$ with $f(\mathbf{p}_2) = 8$ (further increased). Filling in all slack, we get $\mathbf{t}_2 = (4 \ 0 \ 4 \mid 4 \ 2 \ 0 \ 0 \ 2 \mid 8)^\top$.

1.9.3 Regular case \mathbf{p}_2 , kind of ...

Arriving in \mathbf{p}_2 , we find ourselves in a regular case of a basic feasible solution, where three intersecting planes have 0 slack: 2, 6, and 7 (by their row-indices). The three possible pairs of these three planes determine three edge candidates: $(0, 0, -1)^\top$, $(-1/\sqrt{2}, 0, 1/\sqrt{2})^\top$, $(-2/3, 1/3, -2/3)^\top$. They all three stay on the boundary (good), but two of them (the first and the third) would lead to decreasing the objective function (in the last case that’s expected, since that’s the edge that we came along from \mathbf{p}_1).

The remaining edge, $(-1/\sqrt{2}, 0, 1/\sqrt{2})^\top$, is normal to \mathbf{c} , which means that moving along it does not change the objective function. We thus could stop here (at \mathbf{p}_2) and claim that we’ve found an optimum, knowing, however, that there are more such optimal solutions (actually infinitely many).