Solving Mixed-Integer Optimization Problems using Benders Decomposition

Shuvomoy Das Gupta

September 17, 2024

This is a blog on how to solve mixed-integer optimization problems using Benders decomposition algorithm, due to Jacques F. Benders. The formulation presented here closely follows [Garfinkel and Nemhauser, 1972, Section 4.8].

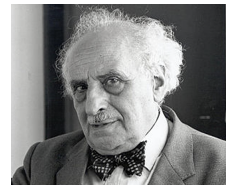

Figure 1: Jacques F. Benders (1925-2017).

Contents

Things to recall.

We recall the following facts that we will use to develop the algorithm.

- **Fact 1: Strong duality in linear programming problems (LPs).** Strong duality does not hold in LPs if and only if both primal and dual problems are infeasible.

- **Fact 2: Extreme point of a polyhedron.** Consider the polyhedron $P = \{x \mid Ax \succeq b\}$ where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ with $m \geq n$. A point $\tilde{x} \in P$ is an extreme point of P if $\mathrm{rank}(A^{\tilde{x}}_{\mathrm{active}}) = n$.

- **Fact 3: Extreme ray of a recesion cone and a polyhedron.** Consider the recesion cone $C = \{x \mid Ax \succeq 0\}$, where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ with $m \geq n$. A point $\tilde{r} \in C$ is an extreme ray of C if $\mathrm{rank}(A^{\tilde{r}}_{\mathrm{active}}) = n-1$. An extreme ray of C is also called the the extreme ray of the underlying polyhedron $\{x \mid Ax \succeq b\}$.

Notation.

- The symbol $\preceq$ or $\succeq$ corresponds to element-wise greater than or smaller than type inequlaity.
- Consider the polyhedron $P = \{x \in \mathbb{R}^4 \mid Ax \succeq b\}$ where

$$A = \begin{bmatrix} a_1^\top \\ a_2^\top \\ \cdots \\ a_5^\top \end{bmatrix} \in \mathbb{R}^{5\times 4}, b = \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_5 \end{bmatrix} \in \mathbb{R}^5.$$

Consider a feasible point $\tilde{x} \in P$, where $a_i^\top \tilde{x} = b_i$ for $i \in \{1,3,5\}$ and $a_i^\top \tilde{x} > b_i$ for $i \in \{2,4\}$. At $\tilde{x}$, we have

$$A^{\tilde{x}}_{\mathrm{active}} = \begin{bmatrix} a_1^\top \\ a_3^\top \\ a_5^\top \end{bmatrix} \in \mathbb{R}^{3\times 4}, A^{\tilde{x}}_{\mathrm{nonactive}} = \begin{bmatrix} a_2^\top \\ a_4^\top \end{bmatrix} \in \mathbb{R}^{2\times 4}.$$

- **Fact 4: Optimal cost of an LP.** Consider the LP

$$\left(\begin{array}{ll} \underset{x \in \mathbb{R}^d}{\text{minimize}} & c^\top x \\ \text{subject to} & Ax \succeq b. \end{array}\right)$$

  where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ with the feasible set $\{x \mid Ax \succeq b\}$ having at least one extreme point. Then,

  - [Bertsimas and Tsitsiklis, 1997, Theorem 4.14] The polyhedron $\{x \mid Ax \succeq b\}$ can have only a finite number of extreme points and extreme rays.
  - [Bertsimas and Tsitsiklis, 1997, Theorem 4.14] The optimal cost of this problem is $-\infty$ if and only if some extreme ray r of $\{x \mid Ax \succeq b\}$ satisfies $c^\top r < 0$. In other words, the optimal cost of this problem is finite if and only if for every extreme ray r of $\{x \mid Ax \succeq b\}$, we have $c^\top r \geq 0$.
  - [Bertsimas and Tsitsiklis, 1997, Theorem 2.7] If the optimal cost of this problem is finite, then there exists an optimal solution which is an extreme point of $\{x \mid Ax \succeq b\}$.

Mixed-integer optimization problem to solve.

We want to solve the following mixed-integer optimization problem (MIP):

$$p^\star = \left(\begin{array}{ll} \underset{x \in \mathbb{Z}^n, v \in \mathbb{R}^p}{\text{maximize}} & c_1^\top x + c_2^\top v \\ \text{subject to} & A_1 x + A_2 v \preceq b, \\ & x \succeq 0,\, v \succeq 0, \end{array}\right) \qquad (\mathcal{P})$$

where $A_1 \in \mathbb{R}^{m \times n}, A_2 \in \mathbb{R}^{m \times p}, b \in \mathbb{R}^m, c_1 \in \mathbb{R}^n, c_2 \in \mathbb{R}^p$. We will assume that $p^\star$ is finite and denote the optimal solution by $x^\star, v^\star$. Note that we are not saying anything about the boundedness of the feasible set.

Problem $(\mathcal{P})$ for fixed $x' \in \mathbb{Z}_+^n$.

Let us consider some $x' \in \mathbb{Z}_+^n$, i.e., some fixed integer-valued $x' \succeq 0$, and define:

$$\begin{aligned} p^\star(x') &= \left(\begin{array}{ll} \underset{v \in \mathbb{R}^p}{\text{maximize}} & c_1^\top x' + c_2^\top v \\ \text{subject to} & A_1 x' + A_2 v \preceq b, \\ & v \succeq 0. \end{array}\right) \\ &= c_1^\top x' + \left(\begin{array}{ll} \underset{v \in \mathbb{R}^p}{\text{maximize}} & c_2^\top v \\ \text{subject to} & A_2 v \preceq b - A_1 x', \\ & v \succeq 0. \end{array}\right) \end{aligned}$$

$$
\begin{aligned}
&= c_1^\top x' - \underbrace{\left(\begin{array}{ll} \underset{v \in \mathbb{R}^p}{\text{minimize}} & -c_2^\top v \\ \text{subject to} & A_2 v \preceq b - A_1 x', \quad \triangleright\, \texttt{d.v.}\ u \succeq 0 \\ & -v \preceq 0. \quad \triangleright\, \texttt{d.v.}\ \lambda \succeq 0 \end{array}\right)}_{p^\star_{\text{sub}}(x')} \\
&= c_1^\top x' - p^\star_{\text{sub}}(x'). \qquad (\mathcal{P}(x'))
\end{aligned}
$$

Any $x \in \mathbb{Z}^n_+$ such that $p^\star(x)$ is finite, is called an *admissible solution*. As $p^\star$ is finite, we can write

$$
p^\star = \max_{\substack{x \in \mathbb{Z}^n_+, \\ x:\text{admissible}}} p^\star(x). \tag{1}
$$

Clearly the optimal $x^\star$ to $(\mathcal{P})$ is an admissible solution.

Dual of the minimization problem associated with $p^\star_{sub}(x')$.

We will follow the style of [Boyd and Vandenberghe, 2004, Chapter 5]. The Lagrangian of the minimization problem associated with $p^\star_{\text{sub}}(x')$ is:

$$
\begin{aligned}
\mathcal{L}(v, u, \lambda) &= -c_2^\top v + u^\top \left(A_2 v - (b - A_1 x')\right) - v^\top \lambda \\
&= \left(-c_2 + A_2^\top u - \lambda\right)^\top v - u^\top (b - A_1 x'),
\end{aligned}
$$

with the dual function:

$$
g(u, \lambda) = \inf_v \mathcal{L}(v, u, \lambda) = \begin{cases} -u^\top (b - A_1 x'), & \text{if } -c_2 + A_2^\top u - \lambda = 0, \\ -\infty, & \text{else.} \end{cases}
$$

As a result, the dual problem of the minimization problem associated with $p^\star_{\text{sub}}(x')$ is:

$$
\begin{aligned}
d^\star_{\text{sub}}(x') &= \left(\begin{array}{ll} \underset{u \in \mathbb{R}^m, \lambda \in \mathbb{R}^p}{\text{maximize}} & -(b - A_1 x')^\top u \\ \text{subject to} & -c_2 + A_2^\top u - \lambda = 0, \\ & u \succeq 0,\ \lambda \succeq 0. \end{array}\right) \\
&= \left(\begin{array}{ll} \underset{u \in \mathbb{R}^m, \lambda \in \mathbb{R}^p}{\text{maximize}} & -(b - A_1 x')^\top u \\ \text{subject to} & -c_2 + A_2^\top u = \lambda \succeq 0, \\ & u \succeq 0. \end{array}\right) \\
&= \left(\begin{array}{ll} \underset{u \in \mathbb{R}^m}{\text{maximize}} & -(b - A_1 x')^\top u \\ \text{subject to} & A_2^\top u \succeq c_2, \\ & u \succeq 0. \end{array}\right)
\end{aligned}
$$

Now, due to weak duality, we always have $d^\star_{\text{sub}}(x') \leq p^\star_{\text{sub}}(x')$. For notational convenience, define:

$$d^\star(x') = c_1^\top x' + \begin{pmatrix} \underset{u\in\mathbb{R}^m}{\text{minimize}} & (b - A_1 x')^\top u \\ \text{subject to} & A_2^\top u \succeq c_2, \\ & u \succeq 0. \end{pmatrix} \qquad (\mathcal{D}(x'))$$

So, the relationship between $(\mathcal{P}(x'))$ and $(\mathcal{D}(x'))$ is as follows.

$$\begin{aligned} p^\star(x') &= c_1^\top x' - p^\star_{\text{sub}}(x') \\ &\leq c_1^\top x' - d^\star_{\text{sub}}(x') \\ &= c_1^\top x' - \begin{pmatrix} \underset{u\in\mathbb{R}^m}{\text{maximize}} & -(b - A_1 x')^\top u \\ \text{subject to} & A_2^\top u \succeq c_2 \\ & u \succeq 0. \end{pmatrix} \\ &= c_1^\top x' + \begin{pmatrix} \underset{u\in\mathbb{R}^m}{\text{minimize}} & (b - A_1 x')^\top u \\ \text{subject to} & A_2^\top u \succeq c_2, \\ & u \succeq 0. \end{pmatrix} \\ &= d^\star(x') \end{aligned} \qquad (2)$$

Relationship between $(\mathcal{P})$, $(\mathcal{P}(x'))$ and $(\mathcal{D}(x'))$.

We can have two possible relationships[1] between $(\mathcal{P})$, $(\mathcal{P}(x'))$ and $(\mathcal{D}(x'))$.

- **Case 1.** $(\mathcal{D}(x'))$ **is unbounded for some** $x' \in \mathbb{Z}^n_+$**.** If $(\mathcal{D}(x'))$ is unbounded for some $x' \in \mathbb{Z}^n_+$, i.e.,

  $$\exists_{x'\in\mathbb{Z}^n_+}\ d^\star(x') = -\infty,$$

  then strong duality will hold due to Fact 1, and we will have

  $$\exists_{x'\in\mathbb{Z}^n_+}\ p^\star(x') = -\infty,$$

  which means that $(\mathcal{P}(x'))$ will be infeasible for that $x' \in \mathbb{Z}^n_+$. This can happen.

- **Case 2.** $(\mathcal{D}(x'))$ **has a finite optimal solution for some** $x' \in \mathbb{Z}^n_+$**.** If $(\mathcal{D}(x'))$ has a finite optimal solution for some $x' \in \mathbb{Z}^n_+$, then strong duality will hold for that x'.

So, combining both cases above, we have for every $x' \in \mathbb{Z}^n_+$,

$$p^\star \geq d^\star(x') = p^\star(x'), \qquad (3)$$

where the inequality follows from (1).

[1] Note that, the case that $(\mathcal{D}(x'))$ *does not have a feasible solution for some* $x' \in \mathbb{Z}^n_+$ cannot happen, which we can see as follows. Suppose $(\mathcal{D}(x'))$ does not have a feasible solution some $x' \in \mathbb{Z}^n_+$, i.e.,

$$\exists_{x'\in\mathbb{Z}^n_+}\ d^\star(x') = \infty.$$

This means that the feasible set of $(\mathcal{D}(x'))$, which is the polyhedron $S \triangleq \{u \mid A_2^\top u \succeq c_2, u \succeq 0\} = \emptyset$ (which does not contain x'), is empty. This also means that, $d^\star(x') = \infty$ for some $x' \in \mathbb{Z}^n_+$ also implies $d^\star(\tilde{x}) = \infty$ for every $\tilde{x}$, be it real or integer valued.

Recalling Fact 1, in Case 1, for every $\tilde{x}$, $(\mathcal{P}(\tilde{x}))$ will be either infeasible, i.e., $p^\star(\tilde{x}) = \infty$ or unbounded, i.e., $p^\star(\tilde{x}) = -\infty$ [Bertsimas and Tsitsiklis, 1997, Table 4.2, Page 151]. In other words, $(\mathcal{P})$ itself will be infeasible or unbounded. This is a contradiction to our assumption that $p^\star$ is finite for some $(x^\star, v^\star)$.

Representation of $(\mathcal{P})$ via extreme points of S.

Let us denote the extreme points and extreme rays of the polyhedron S by E and R, respectively.

Using Fact 4, from 1, we can construct the following integer linear program, which we then will use to construct the decomposition algorithm:

$$
\begin{aligned}
p^\star &= \left(\underset{\substack{x \in \mathbb{Z}^n_+ \\ x:\text{admissible}}}{\text{maximize}}\ p^\star(x) \right) \\
&= \left(\underset{\substack{x \in \mathbb{Z}^n_+ \\ x:\text{admissible}}}{\text{maximize}} \left\{ c_1^\top x + \begin{pmatrix} \underset{u \in \mathbb{R}^m}{\text{minimize}} & (b - A_1 x)^\top u \\ \text{subject to} & A_2^\top u \succeq c_2, \\ & u \succeq 0. \end{pmatrix} \right\} \right) \\
&= \left(\underset{\substack{x \in \mathbb{Z}^n_+ \\ \forall_{r \in R}\, (b - A_1 x)^\top r \geq 0}}{\text{maximize}} \left\{ c_1^\top x + \begin{pmatrix} \underset{u \in \mathbb{R}^m}{\text{minimize}} & (b - A_1 x)^\top u \\ \text{subject to} & A_2^\top u \succeq c_2, \\ & u \succeq 0. \end{pmatrix} \right\} \right) \\
&= \left(\underset{\substack{x \in \mathbb{Z}^n_+ \\ \forall_{r \in R}\, (b - A_1 x)^\top r \geq 0}}{\text{maximize}} \left\{ c_1^\top x + \begin{pmatrix} \underset{u \in \mathbb{R}^m}{\text{minimize}} & (b - A_1 x)^\top u \\ \text{subject to} & A_2^\top u \succeq c_2, \\ & u \succeq 0. \end{pmatrix} \right\} \right) \\
&= \left(\underset{\substack{x \in \mathbb{Z}^n_+ \\ \forall_{r \in R}\, (b - A_1 x)^\top r \geq 0}}{\text{maximize}} \left\{ c_1^\top x + \min_{u \in E} \left((b - A_1 x)^\top u \right) \right\} \right) \\
&= \begin{pmatrix} \underset{x}{\text{maximize}} & c_1^\top x + \min_{u \in E} \left((b - A_1 x)^\top u \right) \\ \text{subject to} & \forall_{r \in R}\, (b - A_1 x)^\top r \geq 0, \\ & x \in \mathbb{Z}^n_+. \end{pmatrix} \\
&= \begin{pmatrix} \underset{x,t}{\text{maximize}} & t \\ \text{subject to} & t \leq \min_{u \in E} \left(c_1^\top x + (b - A_1 x)^\top u \right), \\ & \forall_{r \in R}\, (b - A_1 x)^\top r \geq 0, \\ & x \in \mathbb{Z}^n_+. \end{pmatrix} \\
&= \begin{pmatrix} \underset{x,t}{\text{maximize}} & t \\ \text{subject to} & \forall_{u \in E}\, t \leq c_1^\top x + (b - A_1 x)^\top u, \\ & \forall_{r \in R}\, (b - A_1 x)^\top r \geq 0, \\ & x \in \mathbb{Z}^n_+. \end{pmatrix} \qquad (\mathcal{P}^{\text{enum}})
\end{aligned}
$$

▷ because x being admissible is the same as $\forall_{r \in R}\, (b - A_1 x)^\top r \geq 0$ due to Fact 4.

▷ because we are enforcing the admissibility, an optimal solution to the inner LP can be found just by searching over E (due to Fact 4).

One may think that now we can find the optimal solution via enumeration in $(\mathcal{P}^{\text{enum}})$. Of course, doing that is impractical, as we need to know all the extreme points and extreme rays of S, which can be an astronomically large number. However, we can use $(\mathcal{P}^{\text{enum}})$ to construct a decomposition algorithm as follows.

Setup for the decomposition algorithm.

- **Notation.** At k-th iteration ($k \in \{1, 2, 3, \ldots\}$) of the decomposition algorithm, the subset of E is denoted by $E(k)$ and the subset R is denoted by $R(k)$.
- **Master problem.** At iteration k, denote the *master* problem by

$$p^\star\left(E(k), R(k)\right) = \left(\begin{array}{ll} \underset{x,t}{\text{maximize}} & t \\ \text{subject to} & \forall_{u \in E(k)}\, t \leq c_1^\top x + (b - A_1 x)^\top u, \\ & \forall_{r \in R(k)}\, (b - A_1 x)^\top r \geq 0, \\ & x \in \mathbb{Z}_+^n, \end{array}\right) \tag{$\mathcal{M}(k)$}$$

with its solution denoted by $x^{(k)}, t^{(k)}$. As ($\mathcal{M}(k)$) is a relaxation of ($\mathcal{P}^{\text{enum}}$), we have

$$p^\star \leq p^\star\left(E(k), R(k)\right). \tag{4}$$

- **Subproblem.** Denote the *sub*problem by

$$d^\star(x^{(k)}) = c_1^\top x^{(k)} + \left(\begin{array}{ll} \underset{u \in \mathbb{R}^m}{\text{minimize}} & (b - A_1 x^{(k)})^\top u \\ \text{subject to} & A_2^\top u \succeq c_2, \\ & u \succeq 0, \end{array}\right) \tag{$\mathcal{S}(k)$}$$

  with its solution denoted by $u^{(k)}$.

Decomposition algorithm.

- **Step 1: Initialization.** Here $k = 1$. If the $E(1)$ or $R(1)$ is nonempty, then go to step 2. If $E(1) = R(1) = \emptyset$, let $d^\star(k)$ be arbitrarily large and $x^{(1)}$ be any non-negative integer vector and go to Step 3.
- **Step 2: Solve the master problem.** Solve ($\mathcal{M}(k)$) with the current value of $E(k)$ and $R(k)$. There are two possibilities [2]:
  - If $p^\star\left(E(k), R(k)\right) = \infty$, then then ($\mathcal{M}(k)$) is not bounded above. Set $t^{(k)}$ to be an arbitrarily large number and compute an $x^{(k)}$ that is a feasible solution to ($\mathcal{M}(k)$) for the fixed value of $t^{(k)}$. Proceed to step 3. (Note that this case would not happen, if $x \in \{0, 1\}^n$.)
  - If $p^\star\left(E(k), R(k)\right)$ is finite, then we collect $x^{(k)}, t^{(k)}$ and proceed to step 3.

Note that for both the cases, we must have

$$p^\star \leq p^\star\left(E(k), R(k)\right) \leq p^\star\left(E(k-1), R(k-1)\right),$$

as $\mathcal{M}(k-1)$ is a relaxation of $\mathcal{M}(k)$ and $\mathcal{M}(k)$ is a relaxation of ($\mathcal{P}$) (or ($\mathcal{P}^{\text{enum}}$)).

[2] Note that $p^\star(k) = -\infty$ cannot happen. If $p^\star(k) = -\infty$, then ($\mathcal{M}(k)$) is infeasible, and because ($\mathcal{M}(k)$) is a relaxation of ($\mathcal{P}^{\text{enum}}$), we must have ($\mathcal{P}^{\text{enum}}$) infeasible, which violates our assumption that $p^\star$ is finite.

- **Step 3: Solve the subproblem.** Solve ($\mathcal{S}(k)$) for the current value of $x^{(k)}, t^{(k)}$. There are two possibilities again:
  - If $d^\star(x^{(k)}) = -\infty$, then we have found one extreme point $u^{(k)} \in E$ and one extreme ray $r^{(k)} \in R$ such that $d^\star(x^{(k)})$ can decrease as much as we vary $\theta \geq 0$ in $u^{(k)} + \theta r^{(k)}$. We set $E(k+1) := E(k) \cup \{u^{(k)}\}$ and $R(k+1) := R(k) \cup \{r^{(k)}\}$, i.e., we add the following constraints lazily to ($\mathcal{M}(k)$):

$$t \leq c_1^\top x + (b - A_1 x)^\top u^{(k)},$$
$$(b - A_1 x')^\top r^{(k)} \geq 0,$$

    Then we set $k := k+1$ and go to Step 2.
  - If $d^\star(x^{(k)})$ finite, then we have found one optimal solution $u^{(k)} \in E$, i.e. $u^{(k)}$is an extreme point of S. Using (3) and (4) we have:

$$d^\star(x^{(k)}) \leq p^\star \leq p^\star\left(E(k), R(k)\right). \tag{5}$$

    In this $d^\star(x^{(k)})$ finite case, one of two things can happen:
    * If $d^\star(x^{(k)}) = (E(k), R(k))$, then from (5) we must have, $p^\star = d^\star(x^{(k)}) = (E(k), R(k))$. So we have found the optimal value and $x^\star := x^{(k)}$ is an optimal solution to ($\mathcal{P}$). By solving $\left(\mathcal{P}(x^{(k)})\right)$, we can compute an optimal $v^\star := v^{(k)}$ to ($\mathcal{P}$).
    * If $d^\star(x^{(k)}) < (E(k), R(k))$, then we $E(k+1) := E(k) \cup \{u^{(k)}\}$, i.e, we add the following constraints lazily to ($\mathcal{M}(k)$)

$$t \leq c_1^\top x + (b - A_1 x)^\top u^{(k)}.$$

      Then we set $k := k+1$ and proceed to Step 2.

In worst case, the number of iterations of this algorithm is bounded by $|E| + |R|$.

Benefits of the algorithm.

When $\mathcal{S}(k)$ has an optimal solution with finite objective value, after step 3, we have a feasible solution to ($\mathcal{P}$) along with a measure of suboptimality $p^\star\left(E(k), R(k)\right) - d^\star(x^{(k)})$ from (5). So, in practice, even if we terminate early, we can have a good enough feasible solution with small suboptimality gap.

References

Dimitris Bertsimas and John N Tsitsiklis. *Introduction to linear optimization*, volume 6. Athena Scientific Belmont, MA, 1997.

Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

R. Garfinkel and G. Nemhauser. *Integer Programming*. John Wiley and Sons, New York, 1972.