Bender's Decomposition

Prasangsha Ganguly

August 27, 2021

Consider the following problem,

$$Max\ Z = c^T x + h^T y \tag{1}$$
$$S.T.\ \ Ax + Gy \leq b \tag{2}$$
$$x \in \chi \cap \mathbb{Z}^n \tag{3}$$
$$y \geq 0 \tag{4}$$

In this problem, we have two types of decision variables:

- x belongs to a feasible region of integer values, like χ looks like a polyhedron, and we are interested in the integer values in the region χ. So, in other words, the values of the variable x can take only integer solutions inside the region χ.
- The other variable y have continuous values.

In this context, we call the variable set x is the set of **complicating variables**. What we mean by the complicating variables is that, since the other set of variable y is continuous, if I remove the x variables from our problem, then we end up with a problem with continuous variables which is a linear program and easy to solve. That is, without x, the resulting problem becomes easy.

$**$ The main idea of Bender's decomposition is that, suppose that the variable x is fixed, some oracle say, are going to provide the solution of the complicating variable x are going to be $\bar{x}$ say. If that is the case, then the resulting problem is a linear program (LP). The objective of the LP depends on the value of x that we are assigning as $\bar{x}$.

$$LP(x): \quad Max\ Z_{LP}(\bar{x}) = h^T y \tag{5}$$
$$S.T.\ \ Gy \leq b - A\bar{x} \tag{6}$$
$$y \geq 0 \tag{7}$$

This problem is easy to solve.

The dual of this problem is $D(x)$. As the primal problem $LP(x)$ depends on x, the dual will also depend on x. Let the dual variables corresponding to constraint (6) be u.

The dual problem is given by,

$$D(x): \quad MinZ_(\bar{x}) = u^T(b - A\bar{x}) \tag{8}$$

$$S.T. \;\; u^TG \geq h^T \tag{9}$$

$$u \geq 0 \tag{10}$$

Now, let us consider the original problem.

$$Max\ Z = c^Tx + h^Ty \tag{11}$$

$$S.T. \;\; Ax + Gy \leq b \tag{12}$$

$$x \in \chi \cap \mathbb{Z}^n \tag{13}$$

$$y \geq 0 \tag{14}$$

We can rewrite this problem as,

$$Max\ Z = c^Tx + \gamma(x) \tag{15}$$

$$\gamma(x) = Max\ h^Ty \tag{16}$$

$$S.T., \;\; Gy \leq b - Ax \tag{17}$$

$$y \geq 0 \tag{18}$$

$$x \in \chi \cap \mathbb{Z}^n \tag{19}$$

So, we have an optimization problem inside another optimization problem. Essentially, what we are saying is that, as our overall problem is a maximization problem, so this problem is same as, maximizing only c^Tx + a new variable γ, which is essentially the optimal value of another maximization problem involving only y. Now, the inner problem is a linear problem and we can consider the dual problem. At the optimal solution, the primal objective is same as the dual objective.

So,

$$Max\ Z = c^Tx + \gamma(x) \tag{20}$$

$$\gamma(x) = Min\ u^T(b - A\bar{x}) \tag{21}$$

$$S.T., \;\; u^TG \geq h^T \tag{22}$$

$$u \geq 0 \tag{23}$$

$$x \in \chi \cap \mathbb{Z}^n \tag{24}$$

Important:

- We need to make sure that, the y variables are continuous. If the y variables are integers, then this mechanism won't work. Because in that case, the dual of the inner problem $D(x)$ is not well defined. Because, we don't have a well defined definition of an integer problem.

- We are also going to assume that the original problem is bounded.

$$Max\ Z = c^Tx + h^Ty \tag{25}$$

$$S.T. \;\; Ax + Gy \leq b \tag{26}$$

$$x \in \chi \cap \mathbb{Z}^n \tag{27}$$

$$y \geq 0 \tag{28}$$

That is, objective of this problem will never go to $+\infty$. This implies that $LP(x)$ is also bounded. Hence, the dual of $LP(x)$, which is $D(x)$ is feasible.

Hence, if we are going to use Bender's decomposition, then we must show these two conditions hold.

Now, we can use $D(x)$ to characterize the solutions of $LP(x)$. The reason we are doing this transformation is because this dual problem is going to help us create a reformulation of the original problem. Essentially, we are going to take this dual problem and we are going to replace this by the extreme points and extreme rays that come from Minkowski's representation theorem.

Let us consider the $D(x)$,

$$D(x): \quad MinZ_(\bar{x}) = u^T(b - A\bar{x}) \tag{29}$$

$$S.T. \;\; u^T G \geq h^T \tag{30}$$

$$u \geq 0 \tag{31}$$

We have already considered that as because the original full problem is bounded, $LP(x)$ is bounded and hence this $D(x)$ is feasible. Now let us consider the w^i for $i = 1, ..., k$ are the extreme points of $D(x)$, and v^j for $j = 1, ..., l$ are the set of extreme rays of $D(x)$. Then, $D(x)$ is equivalent to,

$$Min \;\; Z_d(\bar{x}) = (b - A\bar{x})(\sum_{i=1}^{k} \lambda_i w^i + \sum_{j=1}^{l} \mu_j v^j) \tag{32}$$

$$S.T., \;\; \sum_{i=1}^{k} \lambda_i = 1 \tag{33}$$

$$\lambda_i, \mu_j \geq 0 \tag{34}$$

Now, if consider $LP(x)$,, we note that the problem depends on the value of $\bar{x}$ that we provide in. Depending on the value of $\bar{x}$ that we provide, the value of $LP(x)$ changes.

$*$ It is important to note that, for some specific valid value of x, $LP(x)$ can be infeasible. For example consider the uncapacitated facility location problem below,

$$Max \;\; Z = -\sum_{i \in F} f_i x_i + \sum_{i \in F} \sum_{j \in N} c_{ij} y_{ij} \tag{35}$$

$$\sum_{i \in F} y_{ij} = 1 \;\; \forall j \in N \tag{36}$$

$$-x_i + y_{ij} \leq 0 \;\; \forall i \in F, j \in N \tag{37}$$

$$x_i \in \{0, 1\} \;\; \forall i \in F \tag{38}$$

$$y_{ij} \geq 0 \;\; \forall i \in F, j \in N \tag{39}$$

As we know, there are two decision variables, the binary variables x_i determine out of all the candidate locations which are to be opened for facility; the continuous variable y_{ij}

depicts the fraction of demand of the node j satisfied by the facility i. We assume that the demand of a node may be satisfied by multiple facilities. The objective tries to maximize the profit obtained by satisfying the demand − the cost of opening the facilities. The constr. (36) says, all the demands need to be satisfied; constr. (37) says, a demand from facility i can only be satisfied if the facility is open. Now, if we use $x = 0$, i.e., don't open any facility, then it is valid in terms of feasibility of x variable. Also, that is a good solution to start with. Because, considering $x = 0$, will help the objective. However, if we set $x = 0$, then all the y are forced to 0 through the constraint (37). Then the constraint (36) can't be satisfied anymore. So, the problem becomes infeasible.

So, the takeaway is, for some value of x, the linear problem $LP(x)$ may become infeasible. So, we have to make sure that we do something with our x variables such that $LP(x)$ is feasible. Now, as our assumption $LP(x)$ is always bounded, so the dual of it, $D(x)$ is always feasible. If $LP(x)$ is infeasible for some x, then $D(x)$ is unbounded. Lets consider the extreme point and extreme direction representation of the $D(x)$:

$$Min \;\; Z_d(\bar{x}) = (b - A\bar{x})^T(\sum_{i=1}^{k} \lambda_i w_i + \sum_{j=1} \mu_j v^j) \tag{40}$$

$$S.T., \;\; \sum_{i=1}^{k} \lambda_i = 1 \tag{41}$$

$$\lambda_i, \mu_j \geq 0 \tag{42}$$

If this problem is unbounded, then we must have the extreme direction with the condition,

$$(b - A\bar{x})^T v^j < 0.$$

If this occurs, then we can have μ large enough to minimize Z_d to $-\infty$ which makes $D(x)$ as unbounded. Thus, we want to make sure, no matter what values of x you give to me, the condition $(b - A\bar{x})^T v^j < 0$ is not satisfied for all the extreme rays of the dual problem $D(x)$. Which will mean the primal problem $LP(x)$ is feasible. If we look at the dual problem, then we see that $\bar{x}$ or the provided values of x are present only in the objective of the dual problem. So, the dual problem doesn't change with the value of the x. But the primal LP problem depends on the values of the x as the $\bar{x}$ is in the constraints. So, independent of the value of the x, I can extract the extreme rays and extreme points of the dual problem. So, the first conclusion we have is that, we want to guarantee in the optimization problem that no matter which value that I give for x, we want the following constraint to hold:

$$(b - A\bar{x})v^j \geq 0 \;\; \forall j \in [1, l] \tag{43}$$

This constraint guarantees that $LP(x)$ is always feasible. These constraints are called **Bender's Feasibility Cut**.

Now, suppose that by adding the cuts in the dual problem, for every single value of the x, that I feed into my sub-problem, satisfies all of these constraints. So, we know that the dual problem is going to be bounded. So, in the problem,

$$Min\ \ Z_d(\bar{x}) = (b - A\bar{x})^T(\sum_{i=1}^{k} \lambda_i w_i + \sum_{j=1} \mu_j v^j) \tag{44}$$

$$S.T.,\ \ \sum_{i=1}^{k} \lambda_i = 1 \tag{45}$$

$$\lambda_i, \mu_j \geq 0 \tag{46}$$

μ_j will be 0. This is because, as we are minimizing the objective and $(b - A\bar{x})v^j \geq 0$, so, $\mu_j = 0$ to make the objective minimum. So, the dual problem consists only of the extreme points and not the extreme directions. So, the dual problem can be solved by looking at the extreme points and picking up the best one. So, consider our reformulation,

$$Max\ Z = c^T x + \gamma(x) \tag{47}$$

$$\gamma(x) = Max\ h^T y \tag{48}$$

$$S.T.,\ \ Gy \leq b - Ax \tag{49}$$

$$y \geq 0 \tag{50}$$

$$x \in \chi \cap \mathbb{Z}^n \tag{51}$$

Now, what is $\gamma(x)$? As we have seen in the last point, $\gamma(x)$ can be identified by just looking at the extreme points w_i's of the dual problem. So, we can write,

$$\gamma_x = Min_{i=1,...,k}\ \{(b - Ax)w^i\}$$

As in the final problem, we are maximizing $\gamma(x)$ in $Max\ Z = c^T x + \gamma(x)$, so, $\gamma(x) \leq (b - AX)w^i\ \ \forall i = 1, ..., k$.

Let us call this η, where $\eta = \gamma(x) \leq (b - AX)w^i\ \ \forall i = 1, ..., k$. This is called **Bender's Optimality Cut**.

So, the whole formulation is,

$$Max\ Z = c^T x + \eta \tag{52}$$

$$(b - Ax)v^j \geq 0\ \ \forall j \in [1, l] \tag{53}$$

$$\eta \leq (b - AX)w^i\ \ \forall i = 1, ..., k \tag{54}$$

$$x \in \chi \cap \mathbb{Z}^n \tag{55}$$

This is the Bender's reformulation.