Benders Decomposition

The principles of classical Bender’s decomposition
(where the subproblem is a linear program)

Motivating Problem

Solving optimization problems monolithically become intractable as the number of variables and constraints increases.

Benders decomposition addresses this problem by solving the original problem in stages, where the problem solved in each stage consists of fewer constraints and variables than the original problem.

Stages

A first-stage master problem is solved for a subset of variables. This is also called the relaxed master problem.

A second-stage subproblem given the values of the subset of variables determined in the first-stage determines the optimal values of the remaining variables.

Stages

If the subproblem at the second stage is found to be infeasible, then one or more constraints are generated and added to the first stage master problem. These added constraints are known as as Bender’s cuts.The new master problem is then resolved and the process repeats.

Bender's feasibility cuts

Each Bender's feasibility cut excludes the infeasible assignment of the master variables that prompted the cut, and possibly assignments that can be no better. Bender's feasibility cuts arise from identifying extreme rays of the feasible region of the dual of a subproblem. They arise when the solution of the dual of a subproblem is unbounded.

Bender's optimality cuts

Each Bender's optimality cut excludes the non optimal assignments of the master variables that prompted the cut, and possibly assignments that can be no better. Bender's feasibility cuts arise from identifying extreme points of the dual of a subproblem. They arise when the solution of the dual of a subproblem does not agree with the solution of the master problem

Termination

Each Bender’s feasibility cut enforces an extreme ray of the feasible polytope of the full problem solution. Each Bender’s feasibility cut enforces an extreme point. Since the feasible region of the solution has a finite number of extreme points and rays Bender’s decomposition is guaranteed to terminate.

Termination

Suppose the master problem is solved at iteration k with objective value V. If the “best” objective value for any subproblem at any iteration iteration i where $0<i<k+1$ is equal to V, then the algorithm terminates at iteration k with an optimal solution.

Complicating Variables

The essence of Benders’ Decomposition lies in determining which variables to fix, such as to simplify the resulting subproblem. This decision will often require specific knowledge of the problem at hand as well as known ways to solve similar problems quickly.

Complicating Variables

An example of a complicating variable set selection is to choose all binary or integer variables in a MILP problem.
It must be the case that the subproblem solving for the non-complicating variables be convex.
An LP is an example of such a convex subproblem.

Complicating Variables

Suppose one is able to identify a minimal subset of the variables containing all the “complicating variables”. Where complicating variables are variables which if fixed allow the resulting problem to be solved much more easily than the original problem.

In that case the master problem in Benders decomposition will contain these complicating variables and the subproblem will contain the remaining variables.

Pseudo-Code

Algorithm 1 BENDERSDECOMPOSITION()

Choose $\bar{y}$ in original problem
$\bar{z} \leftarrow -\infty$
$k \leftarrow 0$
while (sub-problem dual has feasible solution $\beta \geq \bar{z}$) **do**
    Derive lower bound function $\beta_{\bar{y}}(y)$ with $\beta_{\bar{y}}(\bar{y}) = \beta$
    $k \leftarrow k + 1$
    $y^k \leftarrow \bar{y}$
    Add $z \geq \beta_{\bar{y}}(y)$ to master problem
    if (master problem is infeasible) **then**
        Stop. The original problem is infeasible.
    else
        Let $(\bar{z}, \bar{y})$ be the optimal value and solution to the master problem.
return $(\bar{z}, \bar{y})$

Original Problem

$$\begin{aligned} \underset{x,y}{\text{minimize}} \quad & c^T x + d^T y \\ \text{subject to} \quad & Ax + By \geq b, \\ & y \geq z, \\ & x \geq 0, \\ & y \in \{0,1\}^q \end{aligned}$$

$x \in \mathbb{R}^p$

$y \in \mathbb{R}^q$

$c \in \mathbb{R}^p$

$d \in \mathbb{R}^q$

$z \in \mathbb{R}^q$

$b \in \mathbb{R}^m$

$A \in \mathbb{R}^{m \times p}$

$B \in \mathbb{R}^{m \times q}$

where x,y,b,c,f are vectors

Original Problem

$$\begin{aligned} \underset{x,y}{\text{minimize}} \quad & c^T x + d^T y \\ \text{subject to} \quad & Ax + By \geq b, \\ & y \geq z, \\ & x \geq 0, \\ & y \in \{0,1\}^q \end{aligned}$$

$x \in \mathbb{R}^p$
$y \in \mathbb{R}^q$
$c \in \mathbb{R}^p$
$d \in \mathbb{R}^q$
$z \in \mathbb{R}^q$
$b \in \mathbb{R}^m$
$A \in \mathbb{R}^{m \times p}$
$B \in \mathbb{R}^{m \times q}$

Suppose that the y variables are complicating variables. When they are fixed the problem becomes much easier to solve.

Rewriting the problem

The original problem can be rewritten as a theoretical master problem and a subproblem.

$$\underset{y}{\text{minimize}} \quad V^*(y)$$

Where $V^*(y)$ is defined as the optimal objective value of the sub-problem defined by fixing y.

$$\begin{aligned} V^*(y) = \underset{x}{\text{minimize}} \quad & c^T x + d^T y \\ \text{subject to} \quad & Ax \geq b - By, \\ & x \geq 0 \end{aligned}$$

Master problem (MP)

The master problem can be rewritten practically by introducing variable v:

$$\begin{aligned} \underset{y}{\text{minimize}} \quad & v \\ \text{subject to} \quad & \\ & v \geq V^*(y) \\ & y \geq f, \\ & y \in \{0,1\}^q \end{aligned}$$

Where $V^*(y)$ is defined as the optimal objective value of the sub-problem defined by fixing y.

Notice that all variables in the the master problem are <u>complicating variables</u>.

Also, all constraints in the master problem are constraints involving <u>complicating variables</u>.

Subproblem

$$\begin{aligned} V^*(y) = \underset{x}{\text{minimize}} \quad & c^T x + d^T y \\ \text{subject to} \quad & Ax \geq b - By, \\ & x \geq 0 \end{aligned}$$

Notice that the subproblem contains all constraints that relate only to the non conflicting variables.

Subproblem (SP)

Here the subproblem is a linear program when the values of y are fixed. Notice that If the subproblem is unbounded for some vector $\bar{y}$, then the original problem itself is unbounded.

If the subproblem is bounded, then because the problem is a linear program we can find the optimal value of the subproblem by solving its dual (by strong duality).

Doing so will give us the $V^*(\bar{y})$ value for any given $\bar{y}$.

Subproblem (SP)

Let’s write the subproblem as:

$$\begin{aligned} V^* = \underset{x}{\text{minimize}} \quad & c^T x + d^T \bar{y} \\ \text{subject to} \quad & Ax \geq b - B\bar{y}, \\ & x \geq 0 \end{aligned}$$

Where ȳ indicates a fixed y vector. i.e consider ȳ to contain constants.

Then: Because $\bar{y}$ was fixed, $d^T \bar{y}$ and $B\bar{y}$ must be constant!

Subproblem (SP)

So, we can rewrite the subproblem:

$$
\begin{aligned}
V^* - d^T\bar{y} = \quad \underset{x}{\text{minimize}} \quad & c^T x \\
\text{subject to} \quad & Ax \geq b - B\bar{y}, \\
& x \geq 0
\end{aligned}
$$

Dualize (SP) Dual Subproblem (DSP)

Let λ represent the dual variables in the subproblem.

$$
\begin{aligned}
W^* = \quad & \underset{\lambda}{\text{maximize}} \quad (b - B\bar{y})^T\lambda \\
& \text{subject to} \quad A^T\lambda \leq c, \\
& \qquad\qquad\qquad \lambda \geq 0
\end{aligned}
$$

Define λ^* as the values of λ when the dual sub-problem is solved.

Then $W^* = (b - B\bar{y})^T\lambda^*$

SP and DSP

Because the sub-problem is convex $V^* - d^T \bar{y} = W^*$ by strong duality.

$$V^* - d^T \bar{y} = (b - B\bar{y})^T \lambda^*$$

Benders' Decomposition

As part of Benders' decomposition we will be solving the primal and dual subproblems iteratively. Generating these subproblems using constant $\bar{y}$ vectors resulting from the solution of a master problem. The master problem will in turn be modified as at each iteration Benders' cut constraints are added.

Notice

That y is not present in any of the constraints of a dual subproblem. It is present in the objective.

This means that the feasible region of the dual subproblem is the same in all iterations of Benders' decomposition !!!

Notice

It also means that if a solution vector λ is found for the dual subproblem at iteration k, then that solution vector λ will be feasible for all subsequent iterations.

Notice

However, because the $\bar{y}$ used at iteration $j>k$ may not be the same as the $\bar{y}$ used at iteration k and because $\bar{y}$ is present in the objective… the solution vector λ may not be optimal for the dual subproblem at iteration j.

Bender’s decomposition

In Bender's decomposition we will solving the subproblem and dual subproblem iteratively.

Let us define y^k = the $\bar{y}$ defining the sub-problem solved in the k'th iteration of Benders' decomposition.

Bender's decomposition

Then let $V^*[y^k]$ = the optimal value of the objective function of the sub-problem solved in the k'th iteration of Benders' decomposition.

$$
\begin{aligned}
= \underset{x}{\text{minimize}} \quad & c^T x + d^T y^k \\
\text{subject to} \quad & Ax \geq b - By^k, \\
& x \geq 0
\end{aligned}
$$

Let $\lambda^*[y^k]$ = the vector of dual variables λ necessary to produce $V^*[y^k]$

Bender's decomposition

Then, at the k'th iteration of Benders' decomposition:

$$V^*[y^k] - d^T y^k = (b - By^k)^T \lambda^*[y^k]$$

So..

$$\forall k$$

$$V^*[y^k] - d^T y^k = (b - By^k)^T \lambda^*[y^k]$$

Bender's decomposition

$$V^*(y) - d^T y \geq (b - By)^T \lambda$$

Where:

$$\begin{aligned} V^*(y) = \underset{x}{\text{minimize}} \quad & c^T x + d^T y \\ \text{subject to} \quad & Ax \geq b - By, \\ & x \geq 0 \end{aligned}$$

Notice the $\geq$ rather than $\leq$

this is because the dual subproblem is maximization.

Notice that in the above we can use the λ's generated at the subproblem dual at ANY iteration of Bender's decomposition subproblem.

Bender’s optimality cuts

$B_k(y) =$ The Bender’s cut generated at the k’th iteration from variables x.

$$B_k(y) := \{v \geq (b - By)^T \lambda^*[y^k] + d^T y\}$$

Where y is variable in the master problem,

$\lambda^*[y^k]$ is a constant in the master problem,

and v is a new variable that will be introduced into the master problem.

Cuts of this form are called Bender’s optimality cuts. This is because the cut off non optimal portions of the feasible region of the dual subproblem. Why they do that can be see from examining the second type of cut.

Bender's feasibility cuts

Suppose the dual subproblem is unbounded for some ȳ. Then:

$$\infty = \underset{\lambda}{\text{maximize}} \quad (b - B\bar{y})^T \lambda$$
$$\text{subject to} \quad A^T \lambda \leq c,$$
$$\lambda \geq 0$$

How can we constrain y in the master problem to prevent ȳ being used to generate a subproblem again? Introduce this constraint into the Master problem: $B_k(y) := \{(b - By)^T \lambda^*[y^k] \leq 0\}$

Bender’s cuts

- Bender’s feasibility cuts ensure that the next dual subproblem problem that is generated will not be unbounded, at least not in the same way.
- If a subproblem or dual subproblem is found to be infeasible then the original problem itself must be infeasible.
- So, the only case left is that the solutions generated by are subproblem are suboptimal in the master problem. And this is the case that Bender’s optimality cuts address.

End

? Questions ?