Benders Decomposition

Yuping Huang, Dr. Qipeng Phil Zheng

Department of Industrial and Management Systems Engineering
West Virginia University

IENG 593G Nonlinear Programming, Spring 2012

Benders Decomposition

Benders Decomposition:

1. A solution method for solving certain large-scale linear optimization problems.
2. In each iteration, **new constraints** added to the problem and then make it progress towards a solution. ("row generation")

Consider the following problem:

$$\begin{aligned} \min \quad & \mathbf{c}^\mathbf{T}\mathbf{x} + \mathbf{d}^\mathbf{T}\mathbf{y} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} + \mathbf{D}\mathbf{y} \geq \mathbf{b} \\ & \mathbf{x} \geq 0, \mathbf{y} \geq 0 \end{aligned}$$

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

Benders Decomposition

The original problem [OP]:

$$\begin{aligned} \min \quad & \mathbf{c}^{\mathbf{T}}\mathbf{x} + \mathbf{f}(\mathbf{y}) \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} + \mathbf{D}\mathbf{y} \geq \mathbf{b} \\ & \mathbf{x} \geq 0, \mathbf{y} \in \mathbb{Y} \end{aligned}$$

For a fixed value of y, the OP is given by

$$\begin{aligned} \min \quad & \mathbf{c}^{\mathbf{T}}\mathbf{x} + \mathbf{f}(\hat{\mathbf{y}}) \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} - \mathbf{D}\hat{\mathbf{y}} \\ & \mathbf{x} \geq 0 \end{aligned} \quad \Longrightarrow \quad \mathbf{f}(\hat{\mathbf{y}}) + \left\{ \begin{aligned} \min \quad & \mathbf{c}^{\mathbf{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} - \mathbf{D}\hat{\mathbf{y}} \\ & \mathbf{x} \geq 0 \end{aligned} \right\}$$

The resulting model to solve becomes:

$$\begin{aligned} [\mathbf{SP}]: \quad \min \quad & \mathbf{c}^{\mathbf{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} - \mathbf{D}\hat{\mathbf{y}} \qquad (u) \\ & \mathbf{x} \geq 0 \end{aligned}$$

Benders Decomposition

Let u be the dual variable associated with $\mathbf{Ax} \geq \mathbf{b} - \mathbf{D\hat{y}}$. The dual subproblem is given by

$$[\mathbf{DSP}]: \quad \max \quad (\mathbf{b} - \mathbf{D\hat{y}})^T \mathbf{u} \tag{1}$$

$$\text{s.t.} \quad \mathbf{A}^T \mathbf{u} \leq \mathbf{c} \tag{2}$$

$$\mathbf{u} \geq 0 \tag{3}$$

Note:

1. Only LP are considered in this model. By strong duality, the optimal objective value of primal problem is equal to that of dual problems.
2. The feasible region of the dual formulation does not depend on the value of y. Assuming that the feasible region defined by (2) and (3) is not empty, we can find all extreme points based on the fixed $\hat{y}$.

Benders Decomposition

To find an extreme point $\mathbf{u}_j$ that maximizes the value of objective function $(\mathbf{b} - \mathbf{D}\hat{\mathbf{y}})^T \mathbf{u}_j$, let $z = \max\{(\mathbf{b} - \mathbf{D}\hat{\mathbf{y}})^T \mathbf{u}_j : j = 1, 2, \ldots, Q\}$. The **DSP** can be reformulated as follows:

$$
\begin{aligned}
\min \quad & z \\
\text{s.t.} \quad & z \geq (\mathbf{b} - \mathbf{D}\hat{\mathbf{y}})^T \mathbf{u}_j \quad \text{for } j = 1, 2, \ldots, Q \\
& z \text{ unrestricted}
\end{aligned}
$$

The reformulation of the **OP** in terms of z and y-variables can be obtained by replacing $\mathbf{c}^\mathbf{T}\mathbf{x}$ with z:

$$
\begin{aligned}
[\mathbf{MP}]: \quad \min \quad & \mathbf{f}(\mathbf{y}) + z \\
\text{s.t.} \quad & z \geq (\mathbf{b} - \mathbf{D}\mathbf{y})^T \hat{\mathbf{u}}_j \quad \text{for } j = 1, 2, \ldots, Q \\
& \mathbf{y} \in \mathbb{Y}, z \text{ unrestricted}
\end{aligned}
$$

Benders Decomposition

When it comes to a large size problem, it's impractical to enumerate all extreme points of X in the subproblem.

Let $k < Q$, the **relaxed MP** with less constraints is given by

$$
\begin{aligned}
[\mathbf{RMP}]: \quad \min \quad & \mathbf{f}(\mathbf{y}) + z \\
\text{s.t.} \quad & z \geq (\mathbf{b} - \mathbf{D}\mathbf{y})^T \hat{\mathbf{u}}_j, \ \text{for } j = 1, 2, \ldots, k \\
& \mathbf{y} \in \mathbb{Y}, z \text{ unrestricted}
\end{aligned}
$$

Let $(\bar{y}, \bar{z})$ denote an optimal solution to RMP.

- † If $(\bar{y}, \bar{z})$ is also feasible to MP, then it is optimal to the original problem.
- † In order to check this optimality condition, we equivalently check if the inequality $(\mathbf{b} - \mathbf{D}\bar{\mathbf{y}})^T \mathbf{u}_j - \bar{z} \leq 0$, for $j = 1, 2, \ldots, Q$ holds true.
- † If the current solution of RMP, $(\bar{y}, \bar{z})$, violates some constraints in MP, this means we get $\bar{z} < (\mathbf{b} - \mathbf{D}\bar{\mathbf{y}})^T \mathbf{u}_{k+1}$.
- † Therefore, we impose **Benders' Cut** to RMP

$$z \geq (\mathbf{b} - \mathbf{D}\mathbf{y})^T \mathbf{u}_{k+1}$$

Benders Decomposition Algorithm

- **Initialization**: Let $\hat{\mathbf{y}}$:= initial feasible solution, only solve for the function of y to get the initial LB and then fix y to solve for UB.
- **Step 1**: Solve the RMP, $\min_y\{f(\mathbf{y}) + z | cuts, \mathbf{y} \in Y, z \text{ unrestricted}\}$
  If RMP is infeasible, then stop; else LB := $f(\hat{\mathbf{y}}) + \hat{z}$.
- **Step 2**: Solve the SP, $\max_u\{\mathbf{f}(\hat{\mathbf{y}}) + (\mathbf{b} - \mathbf{D}\hat{\mathbf{y}})^T\mathbf{u} | \mathbf{A}^T\mathbf{u} \leq \mathbf{c}, \mathbf{u} \geq 0\}$, get extreme point $\hat{u}$.
  Get UB := $\mathbf{f}(\hat{\mathbf{y}}) + (\mathbf{b} - \mathbf{D}\hat{\mathbf{y}})^T\hat{\mathbf{u}}$;
  Add cut $z \geq (\mathbf{b} - \mathbf{D}\mathbf{y})^T\hat{\mathbf{u}}$ to RMP.
- If $UB - LB = 0$ or $UB - LB < \epsilon$, the current solution is optimal and stop.
  If $UB - LB > 0$ or $UB - LB > \epsilon$, perform next iteration and go to Step 1.

Example

Solve the following problem by using Benders decomposition.

$$\begin{aligned} \min \quad & 2x_1 + 3x_2 + 2y_1 \\ \text{s.t.} \quad & x_1 + 2x_2 + y_1 \geq 3 \\ & 2x_1 - x_2 + 3y_1 \geq 4 \\ & \mathbf{x} \geq 0, \ \mathbf{y} \geq 0 \end{aligned}$$

$$\mathbf{c}^T = [2\ 3]^T, \mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

Example

Iteration 1:

Step 1: Let $\mathbf{u} = (0,0)^T$. From RMP,

$$\begin{aligned} \min \quad & 2y_1 + z \\ \text{s.t.} \quad & z \geq \left(\begin{bmatrix} 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \end{bmatrix} y_1 \right)^T \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ & y_1 \geq 0, z \text{ unrestricted} \end{aligned}$$

$(\bar{y}, \bar{z}) = (0,0)$, LB = 0.

Step 2:

$$\begin{aligned} \max \quad & 2y_1 + \left(\begin{bmatrix} 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \end{bmatrix} [0] \right)^T \mathbf{u} \\ \text{s.t.} \quad & \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}^T \mathbf{u} \leq \begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ & \mathbf{u} \geq 0 \end{aligned}$$

There exists 4 extreme points in the feasible region of DSP.

$$\mathbf{u}^1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{u}^2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{u}^3 = \begin{bmatrix} 1.5 \\ 0 \end{bmatrix}, \mathbf{u}^4 = \begin{bmatrix} 1.6 \\ 0.2 \end{bmatrix}$$

Example

The dual optimal solution: $\mathbf{u} = [1.6, 0.2]^T$. UB $= 5.6$.

$UB > LB$, add the cut $z \geq (\mathbf{b} - \mathbf{D}\mathbf{y})^T \begin{bmatrix} 1.6 \\ 0.2 \end{bmatrix} = 5.6 - 2.2y_1$ to RMP.

Iteration 2:

$$\begin{aligned} \min \quad & 2y_1 + z \\ \text{s.t.} \quad & z \geq 0 \\ & z \geq 5.6 - 2.2y_1 \\ & y_1 \geq 0, z \text{ unrestricted} \end{aligned}$$

$(\bar{y}, \bar{z}) = (2.545, 0)$, LB = 5.091.

$$\begin{aligned} \max \quad & 2(2.545) + \left(\begin{bmatrix} 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \end{bmatrix} [2.545] \right)^T \mathbf{u} \\ \text{s.t.} \quad & \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}^T \mathbf{u} \leq \begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ & \mathbf{u} \geq 0 \end{aligned}$$

The dual optimal solution is $\mathbf{u} = [1.5\ 0]^T$ and UB $= 5.772$.

$UB > LB$, add the cut $z \geq = 4.5 - 1.5y_1$ to RMP.

Example

Iteration 3:

$$\begin{aligned} \min \quad & 2y_1 + z \\ \text{s.t.} \quad & z \geq 0 \\ & z \geq 5.6 - 2.2y_1 \\ & z \geq 4.5 - 1.5y_1 \\ & y_1 \geq 0, z \text{ unrestricted} \end{aligned}$$

$(\bar{y}, \bar{z}) = (1.571, 2.143)$, LB = 5.286.

$$\begin{aligned} \max \quad & 2(2.545) + \left(\begin{bmatrix} 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \end{bmatrix} [2.545] \right)^T \mathbf{u} \\ \text{s.t.} \quad & \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}^T \mathbf{u} \leq \begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ & \mathbf{u} \geq 0 \end{aligned}$$

The dual optimal solution is $\mathbf{u} = [1.6, 0.2]^T$ and $UB = 5.286$.

The process has converged because $UB = LB$. The optimal solution of OP is $(x_1, x_2, y_1) = (0, 0.714, 1.571)$ and the optimal value is 5.286.

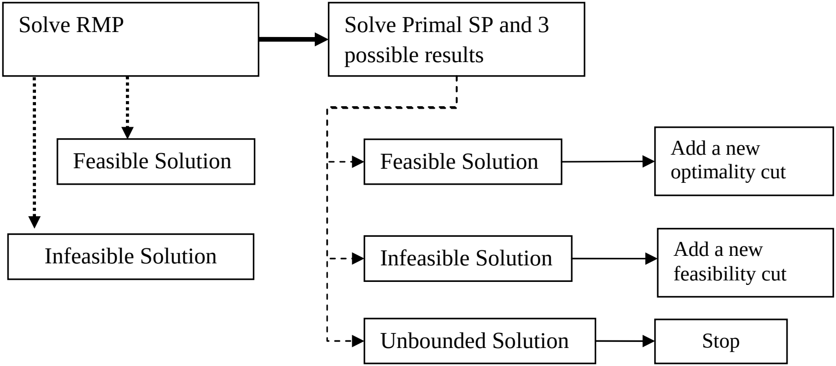
Benders Decomposition

- ▷ Benders decomposition is a cutting plane method due to adding a constraint at each iteration.
- ▷ It is outer approximation method.
- ▷ It reduces search region by adding linear constraints while preserving the original feasible region.

In most application the master problem (MP) is

- an MILP
- a nonlinear programming problem (NLP), or
- a mixed integer/nonlinear programming problem (MINLP).

Special Cases for Subproblem



Special Cases for Subproblem

1. If both primal and dual subproblem have finite optimal solutions, two optimal solutions are converged, then the optimal solution is obtained; else add a new Benders optimality cut to RMP.
2. If the primal subproblem generates infeasible solutions, the dual subproblem also has unbounded solutions, then add a new Benders feasibility cut to RMP.
3. If the primal subproblem has unbounded solution, which means the dual subproblem is infeasible and RMP is unbounded, the objective value will go to $-\infty$. Stop the process.

Special Cases for Subproblem

If the primal subproblem is infeasible and the dual subproblem is infeasible too, what can we do???

Special Cases for Subproblem

Consider the infeasible primal subproblem:

$$\begin{aligned} \min \quad & \mathbf{c}^{\mathbf{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} - \mathbf{D}\hat{\mathbf{y}} \\ & \mathbf{x} \geq 0 \end{aligned}$$

$$\begin{aligned} \min \quad & \mathbf{c}^{T}\mathbf{x} + Mw \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} + \mathbf{e}\mathbf{w} \geq \mathbf{b} - \mathbf{D}\hat{\mathbf{y}} \\ & \mathbf{x} \geq 0, \mathbf{w} \geq 0 \end{aligned}$$

where M is a large value, w is an artificial variable and **e** is Euclidean matrix. In this way, we can make the infeasible inequality system become feasible since the inequality will hold after adding ew to the inequality system. Then the objective function should be penalized by adding a very large value. So the modified subproblem can be solved by using the previous rules.

Special Cases for Subproblem

We can think in another way:

Primal Subproblem:

$$\begin{aligned} \min \quad & \mathbf{0} \cdot \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} - \mathbf{D}\hat{\mathbf{y}} \\ & \mathbf{x} \geq 0 \end{aligned}$$

Dual Subproblem:

$$\begin{aligned} \max \quad & (\mathbf{b} - \mathbf{D}\hat{\mathbf{y}})^T \mathbf{u} \\ \text{s.t.} \quad & \mathbf{A}^T \mathbf{u} \leq \mathbf{0} \\ & \mathbf{u} \geq 0 \end{aligned}$$

- The dual subproblem cannot be infeasible. Since the trivial solution $\mathbf{u} = 0$ is always feasible.
- If the dual problem has an unbounded solution, then there exists a feasible region that is unbounded along an extreme direction (extreme ray). Let $\lambda > 0$ and $\hat{\mathbf{u}} \in C$, we still have $\lambda\hat{\mathbf{u}} \in C$. So $(\mathbf{b} - \mathbf{D}\hat{\mathbf{y}})^T \lambda\mathbf{u} \to +\infty$. To avoid the unbounded case, we have to make $(\mathbf{b} - \mathbf{D}\hat{\mathbf{y}})^T \mathbf{u} \leq 0$ such that $(\mathbf{b} - \mathbf{D}\hat{\mathbf{y}})^T \lambda\mathbf{u} \to -\infty$. This is attributed to the feasible case. Therefore the expression is $(\mathbf{b} - \mathbf{D}\hat{\mathbf{y}})^T \mathbf{u} \leq 0$.
- The cuts of this type are referred as ”Benders feasibility cuts”.

Special Cases for Subproblem

The new relaxed master problem is given by:

$$\begin{aligned}
\textbf{[RMP]}: \quad \min \quad & \mathbf{f}(\mathbf{y}) + z \\
\text{s.t.} \quad & z \geq (\mathbf{b} - \mathbf{D}\mathbf{y})^T \mathbf{u}_j^p \quad \text{for } j = 1, 2, \ldots, Q \\
& 0 \geq (\mathbf{b} - \mathbf{D}\mathbf{y})^T \mathbf{u}_i^r \quad \text{for } i = 1, 2, \ldots, N \\
& \mathbf{y} \in \mathbb{Y}, z \text{ unrestricted}
\end{aligned}$$

where $\mathbf{u}_j^p$ are extreme points of the feasible region in DSP,

and $\mathbf{u}_i^r$ are extreme rays of the feasible region.

Convergence

- In LP, suppose that the set Y is closed and bounded and that $f(\mathbf{y})$ and $\mathbf{Dy}$ are both continuous on S. We can terminate the computation in a finite number of iterations with an optimal solution.
- The reason is the finite number of constraints generated in the RMP; that is, a finite number of extreme points and directions are in any polyhedron. Besides, since the feasible region is convex, stalling doesn't occur in the LP problem.
- For the integer programming case, convergence can be established based on the assumption that either Y is finite or that the set of dual multipliers is finite. But for the general NLP, these assumptions do not apply. For nonlinear cases, please see Generalized Benders Decomposition.