CS/ISyE 719:Benders Decomposition for Solving Two-stage Stochastic Optimization Models

Jim Luedtke

Department of Industrial and Systems Engineering
University of Wisconsin-Madison

September 15, 2016

Module Overview

1. Review some linear programming terms and theory
2. Benders decomposition for problems with relatively complete recourse
3. Extension removing that assumption
4. Multi-cut vs. Single-cut versions
5. Advanced Benders cut selection

Next module: Improved methods via regularization

Some definitions

Definition: Polyhedron

A polyhedron is a set that can be written as $P = \{x \in \mathbb{R}^n : Ax \leq b\}$.

Definition: Extreme point (vertex)

$x \in P$ is an extreme point (vertex) of P if there does not exist $x^1, x^2 \in P$ and $\lambda \in (0,1)$ such that $x^1 \neq x^2$ and $x = \lambda x^1 + (1-\lambda)x^2$

Some definitions (2)

Definition: Ray and Extreme Ray

- $r \in \mathbb{R}^n$ is a ray of P if $x + \lambda r \in P$ for all $x \in P$ and $\lambda \geq 0$
- $r \in \mathbb{R}^n$ is an extreme ray of P if r is a ray of P, and there does not exist rays r^1 and r^2 of P and $\lambda \in (0, 1)$ such that $r^1 \neq \alpha r^2$ for any $\alpha \geq 0$ and $r = \lambda r^1 + (1 - \lambda) r^2$.

Characterization of rays of a polyhedron

Definition

The recession cone of a polyhedron $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ is the set

$$\operatorname{recc}(P) := \{r \in \mathbb{R}^n : Ar \leq 0\}$$

Theorem

A vector $r \in \mathbb{R}^n$ is a ray of a polyhedon $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ if and only if $r \in \operatorname{recc}(P)$.

Fundamental Theorem

Minkowski's Theorem

Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a polyhedron with nonempty set of extreme points $V(P) = \{v^1, v^2, \ldots, v^t\}$ and extreme rays $R(P) = \{r^1, r^2, \ldots, r^p\}$. Then,

$$P = \left\{ \begin{array}{ll} x \in \mathbb{R}^n : & x = \sum_{j=1}^{t} \lambda_j x^j + \sum_{j=1}^{p} \mu_j r^j, \\ & \sum_{j=1}^{t} \lambda_j = 1, \\ & \lambda \in \mathbb{R}^t_+, \mu \in \mathbb{R}^p_+ \end{array} \right\}.$$

- Provides an alternative representation of a polyhedron.

A few more LP facts

Theorem

Suppose $P \neq \emptyset$. The linear program $\max\{c^\top x : x \in P\}$ is unbounded if and only if there exists an extreme ray $r \in R(P)$ with $c^\top r > 0$.

Theorem

If a a linear program has an optimal solution, then it has an extreme point optimal solution.

And finally...

Strong Duality Theorem

Consider a primal and dual pair of linear programs. Exactly one of the following holds:

- Both the primal and the dual have an optimal solution, and their objective values are equal.
- The primal problem is unbounded, and the dual problem is infeasible.
- The dual problem is unbounded, and the primal problem is infeasible.
- Both the prial and dual problems are infeasible.

Recall – Two-Stage Stochastic LP with Recourse

$$\begin{aligned} z^{SP} = \min\ & c^\top x + \textstyle\sum_{k=1}^{K} p_k Q_k(x) \\ \text{s.t.}\ & Ax = b \\ & x \in \mathbb{R}^{n_1}_+ \end{aligned}$$

where for $k = 1, \ldots, K$

$$\begin{aligned} Q_k(x) \stackrel{\text{def}}{=} Q(x, \xi^k) = \min\ & q_k^\top y \\ \text{s.t.}\ & W_k y = h_k - T_k x \\ & y \in \mathbb{R}^{n_2}_+ \end{aligned}$$

How to solve this problem?

- Previously: Extensive form $\Rightarrow$ Large-scale LP
- This Module: Benders decomposition $\Rightarrow$ Exploit structure

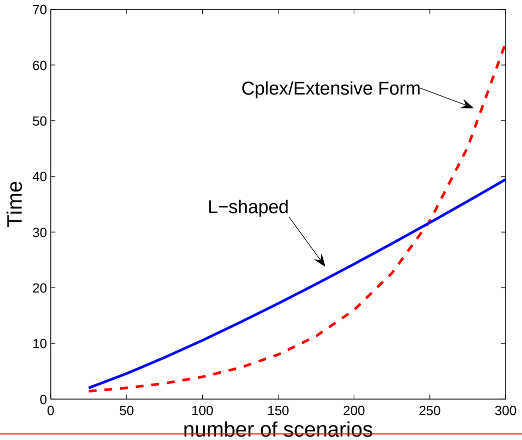
Review: Extensive Form

Easiest way to solve a two-stage stochastic program is to build and solve the extensive form

$$
\begin{array}{ccccccccccc}
c^\top x & + & p_1 q_1^\top y_1 & + & p_2 q_2^\top y_2 & + & \cdots & + & p_K q_K^\top y_K & & \\
Ax & & & & & & & & & = & b \\
T_1 x & + & W_1 y_1 & & & & & & & = & h_1 \\
T_2 x & & & + & W_2 y_2 & & & & & = & h_2 \\
\vdots & & & & & & \ddots & & & & \vdots \\
T_K x & & & & & & & + & W_K y_K & = & h_K \\
x \in X & & y_1 \geq 0 & & y_2 \geq 0 & & & & y_K \geq 0 & &
\end{array}
$$

- This is sometimes called the deterministic equivalent, but I prefer the term extensive form

Small SP's are Easy!



- Using barrier/interior point method is usually faster than simplex methods for solving extensive form LPs.

The Upshot

- If it is too large to solve directly, exploit the structure.
- Fix the first stage variables x, then the problem decomposes by scenario

$$
\begin{array}{ccccccccccc}
c^\top x & + & p_1 q_1^\top y_1 & + & p_2 q_2^\top y_2 & + & \cdots & + & p_K q_K^\top y_K & & \\
Ax & & & & & & & & & = & b \\
T_1 x & + & W_1 y_1 & & & & & & & = & h_1 \\
T_2 x & & & + & W_2 y_2 & & & & & = & h_2 \\
\vdots & & & & & & \ddots & & & & \vdots \\
T_K x & & & & & & & + & W_K y_K & = & h_K \\
x \in X & & y_1 \geq 0 & & y_2 \geq 0 & & & & y_K \geq 0 & &
\end{array}
$$

Key Idea

- Benders Decomposition: Characterize the solution of a scenario linear program as a function of first stage solution x

Recall: Good Modeling Assumption

Relatively Complete Recourse

For every feasible first-stage feasible solution ($Ax = b, x \in \mathbb{R}^{n_1}_+$) and every scenario k there exists a feasible recourse decision y_k ($W_k y_k = h_k - T_k x, y_k \in \mathbb{R}^{n_2}_+$).

Simplifies Benders decomposition algorithm

- We'll first present the algorithm under this assumption
- Then relax the assumption

Assumption is very important when using sample average approximation

- Relatively complete recourse should hold for all possible realizations of random outcomes
- Otherwise, solution to SAA problem may not even be feasible to original problem

Scenario Value Function

Primal

$$Q_k(x) = \min_y \; q_k^\top y$$
$$\text{s.t. } W_k y = h_k - T_k x$$
$$y \in \mathbb{R}_+^{n_2}$$

Dual

$$\max_\pi \; \pi^\top (h_k - T_k x)$$
$$\text{s.t. } \pi^\top W_k \le q_k$$

Assumption

$\Pi_k \stackrel{\text{def}}{=} \{\pi : \pi^\top W_k \le q_k\} \neq \emptyset$.

If this assumption is violated:

- $\Rightarrow$ For every x, second-stage primal problem is either infeasible or unbounded
- $\Rightarrow$ Stochastic LP is either infeasible or unbounded

Scenario Value Function

Primal

$$
\begin{aligned}
Q_k(x) = \min_y \ & q_k^\top y \\
\text{s.t.} \ & W_k y = h_k - T_k x \\
& y \in \mathbb{R}_+^{n_2}
\end{aligned}
$$

Dual

$$
\begin{aligned}
\max_\pi \ & \pi^\top (h_k - T_k x) \\
\text{s.t.} \ & \pi^\top W_k \le q_k
\end{aligned}
$$

Assumption

$\Pi_k \stackrel{\text{def}}{=} \{\pi : \pi^\top W_k \le q_k\} \neq \emptyset.$

Relatively complete recourse $\Rightarrow$ Primal has a feasible solution

- Combine the two $\Rightarrow$ Strong duality applies: Primal and dual each have an optimal solution

Using Strong Duality

Assuming $\Pi_k \neq \emptyset$ and relatively complete recourse:

$$\begin{aligned} Q_k(x) &= \min_y \{q_k^\top y : W_k y = h_k - T_k x, y \in \mathbb{R}_+^{n_2}\} \\ &= \max_\pi \{\pi^\top (h_k - T_k x) : \pi^\top W_k \leq q_k\} \\ &= \max \{(\pi^k)^\top (h_k - T_k x) : \pi^k \in \mathrm{V}(\Pi_k)\} \end{aligned}$$

where $\mathrm{V}(\Pi_k)$ is the finite set of extreme points of Π_k.

Structure of $Q_k(\cdot)$

$Q_k(\cdot)$ is a piecewise-linear convex function.

Benders Reformulation

$$z^{SP} = \min_x \; c^\top x + \sum_{k=1}^K p_k Q_k(x) = \min_{x,\theta} \; c^\top x + \sum_{k=1}^K p_k \theta_k$$

$$\text{s.t. } Ax = b, x \in \mathbb{R}_+^{n_1} \qquad \text{s.t. } Ax = b, x \in \mathbb{R}_+^{n_1}$$

$$\theta_k \geq Q_k(x), \; k = 1, \ldots, K$$

Observe:

$$\begin{aligned}
\theta_k \geq Q_k(x) &\Leftrightarrow \theta_k \geq \max_\pi \{\pi^\top (h_k - T_k x) : \pi^\top W_k \leq q_k\} \\
&\Leftrightarrow \theta_k \geq \max\{(\pi^k)^\top (h_k - T_k x) : \pi^k \in \mathrm{V}(\Pi_k)\} \\
&\Leftrightarrow \theta_k \geq (\pi^k)^\top (h_k - T_k x), \quad \pi^k \in \mathrm{V}(\Pi_k)
\end{aligned}$$

Benders Reformulation

$$
\begin{aligned}
z^{SP} = \min_{x,\theta}\ & c^\top x + \textstyle\sum_{k=1}^{K} p_k \theta_k \\
\text{s.t. }\ & Ax = b, x \in \mathbb{R}^{n_1}_+ \\
& \theta_k \geq (\pi^k)^\top (h_k - T_k x), \quad k = 1, \ldots, K, \pi^k \in \mathrm{V}(\Pi_k)
\end{aligned}
$$

Explicit linear program formulation

- Many fewer variables than deterministic equivalent form (just 1 per scenario)
- Potentially HUGE number of constraints
- Theoretically solvable by ellipsoid algorithm
- Practically solvable by a cutting plane algorithm $\Rightarrow$ Benders decomposition/L-shaped algorithm

Benders Decomposition (L-Shaped Method)

At iteration t of Benders decomposition, a Master Problem (MP) is solved:

$$
\begin{aligned}
(\mathrm{MP})_t : z_t \stackrel{\text{def}}{=} \min_{\theta,x}\ & c^\top x + \sum_{k=1}^{K} p_k \theta_k \\
\text{s.t. }\ & Ax = b, x \in \mathbb{R}_+^{n_1} \\
& \theta_k \geq (\pi^k)^\top (h_k - T_k x),\ k = 1, \ldots, K,\ \pi^k \in \hat{V}^{k,t}
\end{aligned}
$$

where $\hat{V}^{k,t} \subseteq \mathrm{V}(\Pi_k)$ for $k = 1, \ldots, K$

- $|\hat{V}^{k,t}| \leq t$, so constraints of $(\mathrm{MP})_t$ are a (small) subset of constraints of Benders reformulation
- $\Rightarrow z_t \leq z^{SP}$
- Given an optimal solution $(\hat{x}^t, \hat{\theta}^t)$ to $(\mathrm{MP})_t$, we must determine if any of the excluded constraints are violated

Benders Subproblems

Let $(\hat{x}^t, \hat{\theta}^t)$ be an optimal solution to $(\mathrm{MP})_t$:

- For each k, are the following constraints all satisfied?

$$\hat{\theta}_k^t \geq (\pi^k)^\top (h_k - T_k \hat{x}^t), \quad \pi^k \in \mathrm{V}(\Pi_k)$$

- True if and only if:

$$\hat{\theta}_k^t \geq \max\{\pi^\top (h_k - T_k \hat{x}^t) : \pi \in \Pi_k\} = Q_k(\hat{x}^t)$$

Thus, for each k, we solve the second-stage subproblem:

$$\begin{aligned} Q_k(\hat{x}^t) &= \max_{\pi} \{\pi^\top (h_k - T_k \hat{x}^t) : \pi^\top W_k = q_k\} \\ &= \min_{y} \{q_k^\top y : W_k y = h_k - T_k x, y \in \mathbb{R}_+^{n_2}\} \end{aligned}$$

Benders Subproblems

For each k, we solve the second-stage subproblem:

$$\begin{aligned} Q_k(\hat{x}^t) &= \max_{\pi}\{\pi^\top(h_k - T_k\hat{x}^t) : \pi^\top W_k = q_k\} \\ &= \min_{y}\{q_k^\top y : W_k y = h_k - T_k x, y \in \mathbb{R}_+^{n_2}\} \end{aligned}$$

If $\hat{\theta}_k^t \geq Q_k(\hat{x}^t)$: Do nothing!

- All the constraints for scenario k are satisfied

If $\hat{\theta}_k^t < Q_k(\hat{x}^t)$:

- Let $\hat{\pi}^k$ be an extreme point optimal dual solution
- The constraint $\theta_k \geq (\hat{\pi}^k)^\top(h_k - T_k x)$ is violated by $(\hat{x}^t, \hat{\theta}^t)$
- Add to the master problem: $\hat{V}^k \leftarrow \hat{V}^k \cup \{\hat{\pi}^k\}$

Upper bounds

- At each iteration, master problem objective value provides a lower bound on z^{SP}
- Because $\hat{x}^t$ is a feasible solution, we also obtain an upper bound after having solved the scenario subproblems:

$$z^{SP} \leq c^\top \hat{x}^t + \sum_{k=1}^{K} p_k Q_k(\hat{x}^t)$$

- Thus, we can terminate the algorithm when these bounds are equal, or "close enough" (e.g., within ϵ)

Recap: Benders Decomposition

Assume for simplicity: $\{x \in \mathbb{R}^{n_1}_{+} : Ax = b\}$ is bounded

Initialization:

- Let $\hat{x}^0$ be an optimal solution to $\min\{c^\top x : Ax = b, x \in \mathbb{R}^{n_1}_{+}\}$
- For each scenario k, solve scenario subproblem $Q_k(\hat{x}^0)$, let $\hat{\pi}^k$ be an optimal dual solution, and set $\hat{V}^k = \{\hat{\pi}^k\}$.

For $t = 1, 2, \ldots$

1. Solve Master Problem to obtain solution $(\hat{x}^t, \hat{\theta}^t)$ with objective value $\hat{z}_t$
2. For each $k = 1, \ldots, K$:
    - Solve scenario subproblem to evaluate $Q_k(\hat{x}^t)$, and let $\hat{\pi}^k$ be an optimal dual solution
    - If $\hat{\theta}^t_k < Q_k(\hat{x}^t)$: Set $\hat{V}^k \leftarrow \hat{V}^k \cup \{\hat{\pi}^k\}$
3. If $\hat{z}_t \geq c^\top \hat{x}^t + \sum_{k=1}^{K} p_k Q_k(\hat{x}^t) - \epsilon$: Break.

Convergence/correctness

When algorithm finds no violated cuts, it has found optimal solution

- At every iteration, $\hat{z}_t \leq z^{SP}$.
- No cuts $\Rightarrow$

$$\hat{z}_t = c^\top \hat{x}^t + \sum_{k=1}^{K} p_k \hat{\theta}_k^t \geq c^\top \hat{x}^t + \sum_{k=1}^{K} p_k Q_k(\hat{x}^t) \geq z^{SP}$$

The algorithm terminates finitely

- Only finitely many extreme point dual solutions
- Worst-case (would be terrible!), algorithm enumerates all extreme point dual solutions, and no more violated cuts could be found

No guarantee on rate of convgerence (similar to simplex)

- In practice, often converges "pretty fast", but occasionally disastrous

Implementation Notes

Must use a cut violation tolerance, e.g., $\delta \approx 10e^{-6}$:

- Numerically, conclude $\hat{\theta}_k^t < Q_k(\hat{x}^t)$ only if $\hat{\theta}_k^t < Q_k(\hat{x}^t) - \delta$
- δ should match the feasibility tolerance of LP solver
- Otherwise, you may add a "cut" that LP solver does not think is violated $\Rightarrow$ infinite loop

Not necessary to solve all scenario subproblems in every iteration

- But only obtain an upper bound (and hence can consider terminating) when all are solved
- Focus on "useful" scenarios

Scenario subproblems can be solved in parallel

- Parallel scalability eventually limited by master problem

Implementation Notes (cont'd)

Benders decomposition is a cutting plane algorithm for solving LPs. Which algorithm should we use to solve the master LP? Why?

We should use the dual simplex algorithm!

- Dual solution remain feasible after adding cuts
- So the algorithm can be initialized with that dual solution/basis

A First Example

$$\min x_1 + x_2$$

subject to

$$\begin{aligned} \xi_1 x_1 + x_2 &\geq 7 \\ \xi_2 x_1 + x_2 &\geq 4 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

- $\xi = (\xi_1, \xi_2) \in \Xi = \{(1, 1/3), (5/2, 2/3), (4, 1)\}$
- Each outcome has $p_k = \frac{1}{3}$

Huh?

- This problem doesn't make sense!

Recourse Formulation

$$
\begin{aligned}
\min\ & x_1 + x_2 + \textstyle\sum_{k=1}^{3} p_k Q_k(x) \\
\text{s.t.}\ & x_1, x_2 \geq 0
\end{aligned}
$$

where

$$
\begin{aligned}
Q_k(x) = \min\ & y_1 + y_2 \\
& \xi_{1k} x_1 + x_2 + y_1 \geq 7 \\
& \xi_{2k} x_1 + x_2 + y_2 \geq 4 \\
& y_1, y_2 \geq 0
\end{aligned}
$$

- $(\xi_{1k}, \xi_{2k}) \in \{(1, 1/3), (5/2, 2/3), (4, 1)\}$
- Each outcome has $p_k = \frac{1}{3}$

Benders for Stochastic MIP

Let $X = \{x \in \mathbb{R}^{n_1}_+ \times \mathbb{Z}^{p_1}_+ : Ax = b\}$ be a mixed-integer set

- But assume all recourse variables are continuous!

Benders algorithm is exactly the same, except master problem is updated:

$$(\text{MP})_t : z_t \stackrel{\text{def}}{=} \min_{\theta,x} \ c^\top x + \sum_{k=1}^{K} p_k \theta_k$$

$$\text{s.t. } x \in X$$

$$\theta_k \geq (\pi^k)^\top (h_k - T_k x), \ k = 1, \dots, K, \ \pi^k \in \hat{V}^{k,t}$$

where $\hat{V}^{k,t} \subseteq \mathrm{V}(\Pi_k)$ for $k = 1, \dots, K$

- Master problems is now a mixed-integer linear program

Example: Facility location

Example data:

- Three possible facilities and four customers
- Fixed costs: $f = [120, 100, 90]$
- Capacity: $C = [26, 25, 18]$
- Two equally likely scenarios: $d^1 = [12, 8, 6, 11]$, $d^2 = [8, 11, 7, 6]$
- Penalty for unmet demand: $\lambda_j = 20$

Iteration 1: Master problem (no θ variable yet)

$$\min\ 120x_1 + 100x_2 + 90x_3$$
$$\text{s.t. } x_i \in \{0, 1\}, i = 1, 2, 3$$

Optimal solution: $\hat{x} = (0, 0, 0)$
Optimal value (lower bound on SMIP): 0

Example: Iteration 1

Subproblems with $\hat{x} = (0, 0, 0)$:

$$\begin{array}{rl} \min & \sum_{i=1}^{3} \sum_{j=1}^{4} c_{ij} y_{ij} + \sum_{j=1}^{4} 30 z_j \\ \text{s.t.} & \sum_{i=1}^{4} y_{ij} + z_j = d_j^1, \ \forall j \\ & \sum_{j=1}^{4} y_{ij} \leq 26 \cdot 0 \\ & \sum_{j=1}^{4} y_{ij} \leq 25 \cdot 0 \\ & \sum_{j=1}^{4} y_{ij} \leq 18 \cdot 0 \\ & y_{ij} \geq 0, z_j \geq 0 \end{array}$$	$$\begin{array}{rl} \min & \sum_{i=1}^{3} \sum_{j=1}^{4} c_{ij} y_{ij} + \sum_{j=1}^{4} 30 z_j \\ \text{s.t.} & \sum_{i=1}^{4} y_{ij} + z_j = d_j^2, \ \forall j \\ & \sum_{j=1}^{4} y_{ij} \leq 26 \cdot 0 \\ & \sum_{j=1}^{4} y_{ij} \leq 25 \cdot 0 \\ & \sum_{j=1}^{4} y_{ij} \leq 18 \cdot 0 \\ & y_{ij} \geq 0, z_j \geq 0 \end{array}$$

Yields Benders cut:

$$\theta_1 \geq 1140 - 728x_1 - 675x_2 - 468x_3$$

Yields Benders cut:

$$\theta_2 \geq 990 - 728x_1 - 675x_2 - 468x_3$$

Upper bound: $\sum_i f_i \hat{x}_i + \sum_k p_k Q_k(\hat{x}) = 0 + 1/2(1140 + 990) = 1065$

Example: Iteration 2

Updated master problem

$$
\boxed{\begin{array}{ll}
\min & 120x_1 + 100x_2 + 90x_3 + 1/2(\theta_1 + \theta_2) \\
\text{s.t.} & \theta_1 \geq 1140 - 728x_1 - 675x_2 - 468x_3 \\
 & \theta_2 \geq 990 - 728x_1 - 675x_2 - 468x_3 \\
 & x_i \in \{0, 1\}, i = 1, 2, 3
\end{array}}
$$

Optimal solution: $\hat{x} = (0, 1, 1)$, $\hat{\theta} = (0, 0)$
Optimal value (lower bound on SMIP): 190

Example: Iteration 2

Subproblems with $\hat{x} = (0, 1, 1)$:

$$\begin{array}{ll} \min & \sum_{i=1}^{3} \sum_{j=1}^{4} c_{ij} y_{ij} + \sum_{j=1}^{4} 30 z_j \\ \text{s.t.} & \sum_{i=1}^{4} y_{ij} + z_j = d_j^1, \ \forall j \\ & \sum_{j=1}^{4} y_{ij} \leq 26 \cdot 0 \\ & \sum_{j=1}^{4} y_{ij} \leq 25 \cdot 1 \\ & \sum_{j=1}^{4} y_{ij} \leq 18 \cdot 1 \\ & y_{ij} \geq 0, z_j \geq 0 \end{array}$$	$$\begin{array}{ll} \min & \sum_{i=1}^{3} \sum_{j=1}^{4} c_{ij} y_{ij} + \sum_{j=1}^{4} 30 z_j \\ \text{s.t.} & \sum_{i=1}^{4} y_{ij} + z_j = d_j^2, \ \forall j \\ & \sum_{j=1}^{4} y_{ij} \leq 26 \cdot 0 \\ & \sum_{j=1}^{4} y_{ij} \leq 25 \cdot 1 \\ & \sum_{j=1}^{4} y_{ij} \leq 18 \cdot 1 \\ & y_{ij} \geq 0, z_j \geq 0 \end{array}$$

Yields Benders cut:

$$\theta_1 \geq 200 - 130x_1 - 18x_3$$

Yields Benders cut:

$$\theta_2 \geq 142 - 104x_1$$

Upper bound: $\sum_i f_i \hat{x}_i + \sum_k p_k Q_k(\hat{x}) = 190 + 1/2(182 + 142) = 352$

Example: Iteration 3

Updated master problem

$$
\boxed{
\begin{array}{ll}
\min & 120x_1 + 100x_2 + 90x_3 + 1/2(\theta_1 + \theta_2) \\
\text{s.t.} & \theta_1 \geq 1140 - 728x_1 - 675x_2 - 468x_3 \\
 & \theta_1 \geq 200 - 130x_1 - 18x_3 \\
 & \theta_2 \geq 990 - 728x_1 - 675x_2 - 468x_3 \\
 & \theta_2 \geq 142 - 104x_1 \\
 & x_i \in \{0,1\}, i = 1,2,3
\end{array}
}
$$

Optimal solution: $\hat{x} = (1,0,1)$, $\hat{\theta} = (52,38)$
Optimal value (lower bound on SMIP): 255

Example: Iteration 3

Subproblems with $\hat{x} = (1, 0, 1)$:

$\begin{array}{rl} \min & \sum_{i=1}^{3} \sum_{j=1}^{4} c_{ij} y_{ij} + \sum_{j=1}^{4} 30 z_j \\ \text{s.t.} & \sum_{i=1}^{4} y_{ij} + z_j = d_j^1, \ \forall j \\ & \sum_{j=1}^{4} y_{ij} \leq 26 \cdot 1 \\ & \sum_{j=1}^{4} y_{ij} \leq 25 \cdot 0 \\ & \sum_{j=1}^{4} y_{ij} \leq 18 \cdot 1 \\ & y_{ij} \geq 0, z_j \geq 0 \end{array}$	$\begin{array}{rl} \min & \sum_{i=1}^{3} \sum_{j=1}^{4} c_{ij} y_{ij} + \sum_{j=1}^{4} 30 z_j \\ \text{s.t.} & \sum_{i=1}^{4} y_{ij} + z_j = d_j^2, \ \forall j \\ & \sum_{j=1}^{4} y_{ij} \leq 26 \cdot 1 \\ & \sum_{j=1}^{4} y_{ij} \leq 25 \cdot 0 \\ & \sum_{j=1}^{4} y_{ij} \leq 18 \cdot 1 \\ & y_{ij} \geq 0, z_j \geq 0 \end{array}$

Yields Benders cut:

$$\theta_1 \geq 237 - 26x_1 - 125x_2$$

Yields Benders cut:

$$\theta_2 \geq 208 - 26x_1 - 125x_2$$

Upper bound: $\sum_i f_i \hat{x}_i + \sum_k p_k Q_k(\hat{x}) = 210 + 1/2(211 + 182) = 406.5$

Example: Iteration 5 (skipped one!)

Updated master problem

$$
\begin{array}{ll}
\min & 120x_1 + 100x_2 + 90x_3 + 1/2(\theta_1 + \theta_2) \\
\text{s.t.} & \theta_1 \geq 1140 - 728x_1 - 675x_2 - 468x_3 \\
& \theta_1 \geq 200 - 130x_1 - 18x_3 \\
& \theta_1 \geq 237 - 26x_1 - 125x_2 \\
& \theta_1 \geq 141 - 36x_3 \\
& \theta_2 \geq 990 - 728x_1 - 675x_2 - 468x_3 \\
& \theta_2 \geq 142 - 104x_1 \\
& \theta_2 \geq 208 - 26x_1 - 125x_2 \\
& \theta_2 \geq 124 - 36x_3 \\
& x_i \in \{0,1\}, i = 1,2,3
\end{array}
$$

Optimal solution: $\hat{x} = (0,1,1)$, $\hat{\theta} = (182, 142)$
Optimal value (lower bound on SMIP): 352

- Matches upper bound from Iteration 2 $\Rightarrow$ Optimal
- Subproblems yield no violated cuts

Without Relatively Complete Recourse

Recall the second-stage scenario subproblem:

$$Q_k(x) = \min_y \{q_k^\top y : W_k y = h_k - T_k x, y \in \mathbb{R}_+^{n_2}\}$$

Without relatively complete recourse, this problem might be infeasible for some x:

- Yields implicit constraints on x
- Let $C_k = \{x \in \mathbb{R}^{n_1} : \exists y \in \mathbb{R}_+^{n_2} \text{ s.t. } W_k y = h_k - T_k x\}$

Need to make these implicit constraints explicit in two-stage formulation:

$$\begin{aligned} \min\ & c^\top x + \textstyle\sum_{k=1}^K p_k Q_k(x) \\ \text{s.t. }\ & Ax = b, x \in \mathbb{R}_+^{n_1} \\ & x \in C_k,\ k = 1, \ldots, K \end{aligned}$$

Characterizing implicit constraints

For what x is this problem feasible?

$$Q_k(x) = \min_y \{q_k^\top y : W_k y = h_k - T_k x, y \in \mathbb{R}_+^{n_2}\}$$

Recall the dual:

$$\max_\pi \{\pi^\top (h_k - T_k x) : \pi^\top W_k \leq q_k\}$$

- Dual assumed to be feasible: Primal feasible $\Leftrightarrow$ Dual bounded
- Dual bounded $\Leftrightarrow$ $(r^k)^\top (h_k - T_k x) \leq 0$ for every extreme ray $r^k \in \mathrm{R}(\Pi_k)$ of dual feasible region Π_k
- Thus, subproblem for scenario k is feasible if and only if

$$x \in C_k \stackrel{\text{def}}{=} \{x \in \mathbb{R}^{n_1} : (r^k)^\top (h_k - T_k x) \leq 0, r^k \in \mathrm{R}(\Pi_k)\}$$

Modified Benders Reformulation

$$
\begin{aligned}
\min_{x,\theta}\ & c^\top x + \textstyle\sum_{k=1}^{K} p_k \theta_k \\
\text{s.t. }\ & Ax = b, x \in \mathbb{R}_+^{n_1} \\
& \theta_k \geq (\pi^k)^\top (h_k - T_k x), \quad k = 1, \ldots, K, \pi^k \in \mathrm{V}(\Pi_k) \\
& (r^k)^\top (h_k - T_k x) \leq 0, \quad k = 1, \ldots, K, r^k \in \mathrm{R}(\Pi_k)
\end{aligned}
$$

Again can solve with a cutting plane algorithm

Solving Modified Benders Reformulation

Master problem:

- Replace full set of extreme points $\mathrm{V}(\Pi_k)$ with a subset $\hat{V}^k$
- Replace full set of extreme rays $\mathrm{R}(\Pi_k)$ with a subset $\hat{R}^k$

Given a master solution $(\hat{x}^t, \hat{\theta}^t)$ solve each scenario k subproblem:

- If subproblem k feasible and $\hat{\theta}_k^t < Q_k(\hat{x}^t)$:
  - Add "optimality cut": $\hat{V}^k \leftarrow \hat{V}^k \cup \{\hat{\pi}^k\}$
- If subproblem k is infeasible:
  - Simplex algorithm yields a dual extreme ray $\hat{r}^k$ with $(\hat{r}^k)^\top(h_k - T_k\hat{x}^t) > 0$
  - Add "feasibility cut": $\hat{R}^k \leftarrow \hat{R}^k \cup \{\hat{r}^k\}$

Single-cut vs. Multi-cut Benders

Benders algorithm we have seen is referred to as multi-cut version

- Cuts are used to approximate value function of each scenario
- Many cuts may be added in each iteration

Alternative: Single-cut implementation

- Define $\mathcal{Q}(x) = \sum_{k=1}^{K} Q_k(x)$
- Basis of Benders reformulation: A single variable Θ
- Any scenario subproblem infeasible $\Rightarrow \mathcal{Q}(x)$ not defined: Enforce all feasibility cuts as in multi-cut Benders

$$
\begin{aligned}
\min_{x,\Theta} \ & c^\top x + \Theta \\
\text{s.t. } & Ax = b, x \in \mathbb{R}_+^{n_1} \\
& (r^k)^\top (h_k - T_k x) \le 0, \quad k = 1, \ldots, K, r^k \in \mathrm{R}(\Pi_k) \\
& \Theta \ge \mathcal{Q}(x)
\end{aligned}
$$

Need explicit reformulation of last constraint...

Single-cut Benders decomposition

Assuming all scenario subproblems are feasible for some x:

$$
\begin{aligned}
\mathcal{Q}(x) &= \sum_{k=1}^{K} p_k Q_k(x) \\
&= \sum_{k=1}^{K} p_k \min_{y} \{ q_k^\top y : W_k y = h_k - T_k x, y \in \mathbb{R}_+^{n_2} \} \\
&= \sum_{k=1}^{K} p_k \max_{\pi} \{ \pi^\top (h_k - T_k x) : \pi^\top W_k = q_k \} \\
&= \sum_{k=1}^{K} p_k \max \{ (\pi^k)^\top (h_k - T_k x) : \pi^k \in \mathrm{V}(\Pi_k) \}
\end{aligned}
$$

Single-cut Benders decomposition

Assuming all scenario subproblems are feasible for some x:

$$\mathcal{Q}(x) = \sum_{k=1}^{K} p_k \max\{(\pi^k)^\top (h_k - T_k x) : \pi^k \in \mathrm{V}(\Pi_k)\}$$

Therefore $\Theta \geq \mathcal{Q}(x)$ if and only if:

$$\Theta \geq \sum_{k=1}^{K} p_k (\pi^k)^\top (h_k - T_k x), \quad (\pi^1, \ldots, \pi^K) \in \mathrm{V}(\Pi_1) \times \cdots \times \mathrm{V}(\Pi_K)$$

- HUGE number of constraints
- But for cutting plane algorithm only need to be able to efficiently find a violated constraint

Single-cut Benders master problem

Master problem after adding t optimality cuts

$$\begin{aligned}
\min_{x,\Theta}\ & c^\top x + \Theta \\
\text{s.t. } & Ax = b, x \in \mathbb{R}^{n_1}_+ \\
& \Theta \geq d_i - c_i^\top x, & i = 1, \ldots, t \\
& (r^k)^\top (h_k - T_k x) \leq 0, & k = 1, \ldots, K, r^k \in \hat{R}^k
\end{aligned}$$

The constraints $\Theta \geq d_i - c_i^\top x$, $i = 1, \ldots, t$ are just a compact way of writing a subset of these constraints:

$$\Theta \geq \sum_{k=1}^{K} p_k (\pi^k)^\top (h_k - T_k x), \quad (\pi^1, \ldots, \pi^K) \in \mathrm{V}(\Pi_1) \times \cdots \times \mathrm{V}(\Pi_K)$$

Single-cut Benders master problem

Let $(\hat{x}, \hat{\Theta})$ be a master problem optimal solution

- Solve all scenario subproblems for this x
- If any one of them is infeasible: Add feasibility cut(s)
- Else if $\hat{\Theta} \geq \sum_{k=1}^{K} Q_k(\hat{x})$: Solution is feasible, hence optimal
- Else:
  - Let $\hat{\pi}^k$ be optimal extreme point dual solution, $k = 1, \ldots, K$
  - Add following cut which is violated by $(\hat{x}, \hat{\Theta})$:

$$\Theta \geq \sum_{k=1}^{K} p_k (\hat{\pi}^k)^\top (h_k - T_k x) \stackrel{\text{def}}{=} d_{t+1} - c_{t+1}^\top x$$

where $d_{t+1} = \sum_{k=1}^{K} p_k (\hat{\pi}^k)^\top h_k$ and $c_{t+1} = \sum_{k=1}^{K} (\hat{\pi}^k)^\top T_k$

Single-cut vs. Multi-cut Benders

Single-cut Benders

- Fewer variables and optimality cuts $\Rightarrow$ Master problem typically solves faster
- Must solve all subproblems to generate an optimality cut

Multi-cut Benders

- Many cuts added per iteration $\Rightarrow$ Typically converges in many fewer interations
- Master problem may become large and slow to solve
- Do not need to solve all subproblems every iteration

Cut Selection (in Multi-cut Benders)

Given $(\hat{x}, \hat{\theta})$ optimal to master problem, key step is finding a violated cut:

- Optimality cut: Dual solution $\hat{\pi}^k$ with $\hat{\theta}_k < (\hat{\pi}^k)^\top (h_k - T_k \hat{x})$
- Feasibility cut: Dual extreme ray $\hat{r}^k$ with $(\hat{r}^k)^\top (h_k - T_k \hat{x}) > 0$

Benders method provides a recipe for finding some violated constraint whenever one exists

- But, there may be many different violated constraints
- "Good" choice can lead to faster convergence

Alternative Cut Generation Problem

Proposed by Fischetti, Salvagnin, and Zanette (2010):

$$
\begin{aligned}
\hat{v}_k = \max\ & \pi^\top (h_k - T_k \hat{x}) - \hat{\theta}_k \pi_0 \\
\text{s.t. }\ & \pi^\top W_k \le q_k \pi_0 \\
& \|\pi\|_1 + \pi_0 \le 1 \\
& \pi_0 \ge 0
\end{aligned}
$$

Theorem

- If $\hat{v}_k = 0$, then $\hat{\theta}_k \ge Q_k(\hat{x})$.
- If $\hat{v}_k > 0$, then either there exists $\hat{\pi}^k \in \Pi_k$ with

  $$\hat{\theta}_k < (\hat{\pi}^k)^\top (h_k - T_k \hat{x})$$

  or there exists a ray $\hat{r}^k$ of Π_k with

  $$(\hat{r}^k)^\top (h_k - T_k \hat{x}) > 0.$$

Alternative Cut Generation Problem

$$\begin{aligned} \hat{v}_k = \max\ & \pi^\top (h_k - T_k \hat{x}) - \hat{\theta}_k \pi_0 \\ \text{s.t. }\ & \pi^\top W_k \leq q_k \pi_0 \\ & \|\pi\|_1 + \pi_0 \leq 1 \\ & \pi_0 \geq 0 \end{aligned}$$

Conclusion from Theorem:

- We can solve this problem for each scenario to generate either feasibility or optimality cuts

Convergence?

- Theorem does not guarantee $\hat{\pi}^k$ or $\hat{r}^k$ are extreme point/ray
- No problem: There are finitely many extreme point solutions to the cut generation linear program