Assignment 3 — Benders' Decomposition

Koen Moors and Jeroen van Riel

June 2023

1 – Facility location problem

First, we reformulate the problem in the standard form (P)

$$\begin{aligned} \min\ & c_x^\top x + c_y^\top y \\ & Ax = b, \\ & Bx + Dy = d, \\ & x \in \{0,1\}^F, y \geq 0. \end{aligned}$$

Note that the current problem does not have first-stage "$Ax = b$" constraints. We order the y_{ij} variables as $y_{11}, y_{12}, \dots, y_{1n}, y_{21}, y_{22}, \dots$. For the first m constraints

$$\sum_{j=1}^{n} y_{ij} \geq 1,\ i \in C,$$

the first m rows of D are given by

$$\begin{aligned} D_{ij} &= 1,\ \text{ for } i \in C,\ (i-1)n < j \leq in, \\ D_{mn+j,j} &= -1,\ \text{ for } j \in F, \end{aligned}$$

and zero otherwise. The '-1's correspond to m slack variables. The first m rows of B are zero and the first m entries of d are all 1.

The remaining mn constraints

$$y_{ij} - x_j \leq 0,\ i \in C, j \in F,$$

are encoded by the last mn rows of D, which are given by

$$\begin{aligned} D_{mn+(i-1)n+j,(i-1)n+j} &= 1,\ \text{ for } i \in C, j \in F, \\ D_{mn+(i-1)n+j,mn+m+(i-1)n+j} &= 1,\ \text{ for } i \in C, j \in F, \end{aligned}$$

which are just two $mn \times mn$ identity matrix, the last one corresponding to mn slack variables. The corresponding entries in B are given by

$$B_{mn+(i-1)n+j,j} = -1,\ \text{ for } i \in C, j \in F,$$

and the last mn entries of d are zero. Finally, the cost vectors are simply given as $(c_x)_j = f_j$ and $(c_y)_{ij} = c_{ij}$ (similarly flattened and the last $m + mn$ entries corresponding to the slack variables are simply zero).

2 – Solution for the uncapacitated case

After fixing variables x, the problem P is given by

$$\begin{aligned} \min\ & c_y^\top y \\ & Dy = d - Bx \\ & y \geq 0, \end{aligned}$$

and it's dual is simply given by

$$\max\ (d - Bx)^\top \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{1a}$$

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} D \leq c_y^\top, \tag{1b}$$

with $\alpha \in \mathbb{R}^n$ and $\beta \in \mathbb{R}^{mn}$. Note that the entries in D corresponding to the $m + mn$ slack variables require that

$$\begin{aligned} \alpha &\geq 0, \\ \beta &\leq 0. \end{aligned}$$

From the structure of B, D and d, it is not difficult to see that (1) is equivalent to

$$\begin{aligned} \max\ & \sum_{i=1}^{m} \alpha_i + \sum_{j \in O(\bar{x})} \beta_{ij} \\ & \alpha_i + \beta_{ij} \leq c_{ij} \quad \text{for all } i \in C, j \in F, \\ & \alpha \geq 0, \beta \leq 0. \end{aligned}$$

Consider a candidate dual solution, given by

$$\begin{aligned} \alpha_i &= \min_{j \in O(\bar{x})} c_{ij}, \\ \beta_{ij} &= \begin{cases} 0, & j \in O(\bar{x}), \\ -\max\{\alpha_i - c_{ij}, 0\}, & j \in C(\bar{x}). \end{cases} \end{aligned}$$

This solution is feasible, because for $j \in O(\bar{x})$, we have

$$\alpha_i + \beta_{ij} = \min_{j \in O(\bar{x})} c_{ij} \leq c_{ij}$$

and for $j \in C(\bar{x})$, we have

$$\alpha_i + \beta_{ij} = \alpha_i - \max\{\alpha_i - c_{ij}, 0\} = \alpha_i + \min\{c_{ij} - \alpha_i, 0\} = \min\{c_{ij}, \alpha_i\} \leq c_{ij}.$$

Note that the objective value for the given solution is simply $\sum_{i=1}^{m} \alpha_i$. We now argue that this objective value is optimal. Suppose that we fullfill every customer's demand from the cheapest among the open facilities, i.e., we set

$$y_{ij} = \begin{cases} 1 & \text{if } j = \operatorname{argmin}_{j \in O(\bar{x})} c_{ij}, \\ 0 & \text{otherwise}, \end{cases}$$

then clearly

$$c_y^\top y = \sum_{i=1}^{m} \min_{j \in O(\bar{x})} c_{ij} = \sum_{i=1}^{m} \alpha_i,$$

which provides a lower bound on the optimal objective value of the primal problem. However, by strong duality, it is also an upper bound on the objective value of the dual. Since the proposed solution actually attains this bound, it must be optimal. Hence, we conclude that it is always optimal to serve each customer from the cheapest open facility in the uncapacitated variant, which makes perfect sense intuitively.

3 – Polyhedral classifiers

Part 2

The "no good" cut is defined as

$$\sum_{(y,i):\bar{s}_{yi}=0} s_{yi} + \sum_{(y,i):\bar{s}_{yi}=1} (1 - s_{yi}) \geq 1. \tag{2}$$

Suppose $s \in \{0,1\}^{Y \times \{1,\ldots,K\}}$ is such that (2) is violated. Both summations are non-negative, because $s_{yi} \in \{0,1\}$, so the inequality can only be violated when both summations are zero. Hence,

$$\sum_{(y,i):\bar{s}_{yi}=0} s_{yi} = 0$$

implies $s_{yi} = 0$ for all $(y,i) : \bar{s}_{yi} = 0$ and

$$\sum_{(y,i):\bar{s}_{yi}=1} (1 - s_{yi}) = 0$$

implies $s_{yi} = 1$ for all $(y,i) : \bar{s}_{yi} = 1$, so $s = \bar{s}$. Therefore, we conclude that $\bar{s}$ is the only binary point that violates (2)

Part 3 + 4

Let $s \in \{0,1\}^{Y \times \{1,\ldots,K\}}$ be a feasible assignment. We want to show that the following cut holds

$$\sum_{(y,i):\bar{s}_{yi}=1} (1 - s_{yi}) \geq 1. \tag{3}$$

Assume $\bar{s}$ is such that the remaining subproblem is infeasible, otherwise we have already found an optimal solution. Then we know there exist $w \in Y$, $k \in \{1,\ldots,K\}$ such that the inequality

$$\sum_{j=1}^{n} a_{kj} y_j \geq b_k + \varepsilon - M(1 - \bar{s}_{wk})$$

is violated. This implies that $\bar{s}_{wk} = 1$. However, since s is feasible, we have

$$\sum_{j=1}^{n} a_{kj} y_j \geq b_k + \varepsilon - M(1 - s_{wk})$$

and therefore $s_{wk} = 0$. Hence

$$\sum_{(y,i):\bar{s}_{yi}=1} (1 - s_{yi}) \geq (1 - s_{wk}) = 1.$$

The advantage of using 3 as a strengthened cut is that is is very easy to obtain from the optimal assignment $\bar{s}$. However, after implementation we see that the cut is still relatively weak and the computational time is very long.

However, note that we also have $s_{wk} = 0$ and hence $1 - s_{wk} \geq 1$ would also be a feasible cut. This cut is stronger than (3). However, to know w and k you would need to find a violated inequality first, which takes some work.