

Comparing the strength of alternative formulations (more abstractly)

Recall the following formulation for the uncapacitated facility location problem:

$$\min \sum_{i=1}^n f_i y_i + \sum_{i=1}^n \sum_{j=1}^m c_{ij} x_{ij} \quad (1)$$

$$x_{ij} \leq y_i, \quad \forall i, j \quad (2)$$

$$\sum_{i=1}^n x_{ij} = 1, \quad \forall j \quad (3)$$

$$x_{ij} \geq 0, \quad \forall i, j \quad (4)$$

$$0 \leq y_i \leq 1, \quad \forall i \quad (5)$$

$$y_i \text{ integer}, \quad \forall i. \quad (6)$$

Here, y_i represents the decision to open facility i , and x_{ij} represents the proportion of customer j 's demand that is fulfilled by facility i . For simplicity, we write $\forall i$ instead of the more tedious $\forall i \in \{1, \dots, n\}$. Similarly, $\forall j$ means for all $j \in \{1, \dots, m\}$.

An alternative formulation replaces the nm coupling inequalities (2) with only n constraints:

$$\sum_{j=1}^m x_{ij} \leq m y_i, \quad \forall i. \quad (7)$$

Note that this formulation is correct as well; both formulations model the same feasible region.

Which formulation is stronger? Is it strictly better?

To formalize our discussion, let P_1 denote the initial LP relaxation's feasible region, i.e., P_1 is the set of $(x, y) \in \mathbb{R}^{n \times m} \times \mathbb{R}^n$ that satisfies constraints (2), (3), (4), (5).

Similarly, let P_2 be the set of all $(x, y) \in \mathbb{R}^{n \times m} \times \mathbb{R}^n$ satisfying (7), (3), (4), (5).

First question: is $P_1 \subseteq P_2$? In other words, is P_1 at least as strong as P_2 ?

Perhaps the formulations are equally strong?

Second question: is $P_2 \subseteq P_1$?