2023_1019_0148_21

1 Preliminaries

1.1 Standard Form of LP

Any unbounded variable x can be replaced by a pair of non-negative variables (u, v).

$$\begin{cases} x = u - v \\ u \geq 0 \\ v \geq 0 \end{cases} \tag{1}$$

Any in-equality constraint can be converted into an equality constraint, by introducing an additional assistant non-negative variable x'.

$$\begin{aligned} \mathbf{a}_i^T \mathbf{x} \geq b_i \quad & \Leftrightarrow \quad \begin{cases} \mathbf{a}_i^T \mathbf{x} - x' & = b_i \\ x' & \geq 0 \end{cases} \\ & \Leftrightarrow \quad \begin{cases} \begin{bmatrix} \mathbf{a}_i^T & (-1) \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ x' \end{bmatrix} & = b_i \\ x' & \geq 0 \end{cases} \end{aligned} \tag{2}$$

So we can safely represent any Linear Programming (LP) problem in its standard form, with only non-negative variables and only equality constraints.

1.1.1 Example

The LP with equalities and in-equalities

$$\begin{aligned} \text{Minimize} \quad & \mathbf{c}^T \color{red}{\mathbf{x}} \\ \text{s.t.} \quad & \begin{cases} \mathbf{A}^{(e)} \color{red}{\mathbf{x}} & = \mathbf{b}^{(e)} & \in \mathbb{R}^{m^{(e)} \times 1} \\ \mathbf{A}^{(i)} \color{red}{\mathbf{x}} & \geq \mathbf{b}^{(i)} & \in \mathbb{R}^{m^{(i)} \times 1} \\ \color{red}{\mathbf{x}} & \geq \mathbf{0} & \in \mathbb{R}^{n \times 1} \end{cases} \end{aligned} \tag{3}$$

can be transformed to a standard form without in-equalities

$$\begin{aligned} \text{Minimize} \quad & \mathbf{c}^T \color{red}{\mathbf{x}} \\ \text{s.t.} \quad & \begin{cases} \mathbf{A}^{(e)} \color{red}{\mathbf{x}} & = \mathbf{b}^{(e)} & \in \mathbb{R}^{m^{(e)} \times 1} \\ \mathbf{A}^{(i)} \color{red}{\mathbf{x}} - \mathbf{y} & = \mathbf{b}^{(i)} & \in \mathbb{R}^{m^{(i)} \times 1} \\ \color{red}{\mathbf{x}} & \geq \mathbf{0} & \in \mathbb{R}^{n \times 1} \\ \color{red}{\mathbf{y}} & \geq \mathbf{0} & \in \mathbb{R}^{m^{(i)} \times 1} \end{cases} \end{aligned} \tag{4}$$

1.2 In-equality Form of LP

Conversely, any equality constraint can also be converted into in-equality constraints, by using a pair of "opposite" constraints.

$$\begin{aligned} \mathbf{a}_i^T \mathbf{x} = b_i \quad & \Leftrightarrow \quad \begin{cases} \mathbf{a}_i^T \mathbf{x} \geq b_i \\ \mathbf{a}_i^T \mathbf{x} \leq b_i \end{cases} \\ & \Leftrightarrow \quad \begin{cases} +\mathbf{a}_i^T \mathbf{x} \geq +b_i \\ -\mathbf{a}_i^T \mathbf{x} \geq -b_i \end{cases} \\ & \Leftrightarrow \quad \left\{ \begin{bmatrix} \mathbf{a}_i^T \\ -\mathbf{a}_i^T \end{bmatrix} \mathbf{x} \geq \begin{bmatrix} +b_i \\ -b_i \end{bmatrix} \right. \end{aligned} \tag{5}$$

1.2.1 Example

The LP with equalities and in-equalities Eq.(3) can also be transformed to an in-equality form without equalities

$$\begin{aligned} \text{Minimize} \quad & \mathbf{c}^T \color{red}{\mathbf{x}} \\ \text{s.t.} \quad & \begin{cases} +\mathbf{A}^{(e)} \color{red}{\mathbf{x}} & \geq +\mathbf{b}^{(e)} & \in \mathbb{R}^{m^{(e)} \times 1} \\ -\mathbf{A}^{(e)} \color{red}{\mathbf{x}} & \geq -\mathbf{b}^{(e)} & \in \mathbb{R}^{m^{(e)} \times 1} \\ \mathbf{A}^{(i)} \color{red}{\mathbf{x}} & \geq \mathbf{b}^{(i)} & \in \mathbb{R}^{m^{(i)} \times 1} \\ \color{red}{\mathbf{x}} & \geq \mathbf{0} & \in \mathbb{R}^{n \times 1} \end{cases} \end{aligned} \tag{6}$$

2 LP and Duality

2.1 Definition of Dual

Assume $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^{m\times 1}$, $\mathbf{c} \in \mathbb{R}^{n\times 1}$. The LP problem **(in in-equality form)**

$$\begin{aligned} &\text{Minimize} \quad \mathbf{c}^T\mathbf{x} \\ &\text{s.t.} \quad \begin{cases} \mathbf{A}\mathbf{x} & \geq \mathbf{b} \quad \in \mathbb{R}^{m\times 1} \\ \mathbf{x} & \geq \mathbf{0} \quad \in \mathbb{R}^{n\times 1} \end{cases} \end{aligned} \tag{7}$$

has a dual problem defined as

$$\begin{aligned} &\text{Maximize} \quad \mathbf{b}^T\lambda \\ &\text{s.t.} \quad \begin{cases} \mathbf{A}^T\lambda & \leq \mathbf{c} \quad \in \mathbb{R}^{n\times 1} \\ \lambda & \geq \mathbf{0} \quad \in \mathbb{R}^{m\times 1} \end{cases} \end{aligned} \tag{8}$$

and we can conclude that the dual derivation process is simple:

- $(\mathbf{A} \Leftrightarrow \mathbf{A}^T)$ coefficient matrix is transposed
- $(\mathbf{b} \Leftrightarrow \mathbf{c})$ cost and constraint vectors are interchanged
- $(\leq \Leftrightarrow \geq)$ constraint in-equalities are reversed
- (maximizing $\Leftrightarrow$ minimizing) optimization direction is reversed

and also note that the sizes/lengths of matrices/vectors are changed accordingly.

Following this definition, the LP problem **(in standard form)**

$$\begin{aligned} &\text{Minimize} \quad \mathbf{c}^T\mathbf{x} \\ &\text{s.t.} \quad \begin{cases} \mathbf{A}\mathbf{x} & = \mathbf{b} \quad \in \mathbb{R}^{m\times 1} \\ \mathbf{x} & \geq \mathbf{0} \quad \in \mathbb{R}^{n\times 1} \end{cases} \end{aligned} \tag{9}$$

which can be converted into in-equality form as

$$\begin{aligned} &\text{Minimize} \quad \mathbf{c}^T\mathbf{x} \\ &\text{s.t.} \quad \begin{cases} \begin{bmatrix} +\mathbf{A} \\ -\mathbf{A} \end{bmatrix}\mathbf{x} & \geq \begin{bmatrix} +\mathbf{b} \\ -\mathbf{b} \end{bmatrix} \quad \in \mathbb{R}^{(2m)\times 1} \\ \mathbf{x} & \geq \mathbf{0} \quad \in \mathbb{R}^{n\times 1} \end{cases} \end{aligned} \tag{10}$$

should have a dual problem as

$$\begin{aligned} &\text{Maximize} \quad \begin{bmatrix} +\mathbf{b}^T & -\mathbf{b}^T \end{bmatrix}\lambda \\ &\text{s.t.} \quad \begin{cases} \begin{bmatrix} +\mathbf{A}^T & -\mathbf{A}^T \end{bmatrix}\lambda & \leq \mathbf{c} \quad \in \mathbb{R}^{n\times 1} \\ \lambda & \geq \mathbf{0} \quad \in \mathbb{R}^{(2m)\times 1} \end{cases} \end{aligned} \tag{11}$$

Partitioning $\lambda \in \mathbb{R}^{(2m)\times 1}$ as $\begin{bmatrix} \lambda_1 & \in \mathbb{R}^{m\times 1} \\ \lambda_2 & \in \mathbb{R}^{m\times 1} \end{bmatrix}$ and we have

$$\begin{aligned} &\text{Maximize} \quad \mathbf{b}^T(\lambda_1 - \lambda_2) \\ &\text{s.t.} \quad \begin{cases} \mathbf{A}^T(\lambda_1 - \lambda_2) & \leq \mathbf{c} \quad \in \mathbb{R}^{n\times 1} \\ \lambda_1 & \geq \mathbf{0} \quad \in \mathbb{R}^{(m)\times 1} \\ \lambda_2 & \geq \mathbf{0} \quad \in \mathbb{R}^{(m)\times 1} \end{cases} \end{aligned} \tag{12}$$

and we can safely combine the free non-negative variables pairs $\lambda' = (\lambda_1 - \lambda_2)$, resulting in

$$\begin{aligned} &\text{Maximize} \quad \mathbf{b}^T\lambda' \\ &\text{s.t.} \quad \begin{cases} \mathbf{A}^T\lambda' & \leq \mathbf{c} \quad \in \mathbb{R}^{n\times 1} \\ \lambda' & \geq -\infty \quad \in \mathbb{R}^{m\times 1} \\ & \text{(variables not necessarily non-negative)} \end{cases} \end{aligned} \tag{13}$$

2.2 The Duality Theorem

So what is the relationship between Eq.(7) and its dual Eq.(8) defined in section 2.1 ?

We will explore , based on the dual of standard form, Eq.(9) and Eq.(13).

2.2.1 Weak Duality Lemma

If Eq.(9) and Eq.(13) are both feasible, then any feasible solution $\mathbf{x}$ for Eq.(9) as well as any feasible solution λ' for Eq.(13) satisfies:

$$\begin{aligned} &\mathbf{b}^T\lambda' \\ =&(\mathbf{A}\mathbf{x})^T\lambda' \quad \text{by} \quad \mathbf{A}\mathbf{x} = \mathbf{b} \quad \text{in Eq.(9)} \\ =&\mathbf{x}^T(\mathbf{A}^T\lambda') \\ \leq&\mathbf{x}^T\mathbf{c} \quad \text{by} \quad \begin{cases} \mathbf{A}^T\lambda' \leq \mathbf{c} & \text{in Eq.(13)} \\ \mathbf{x} \geq \mathbf{0} & \text{in Eq.(9)} \end{cases} \\ =&\mathbf{c}^T\mathbf{x} \end{aligned} \tag{14}$$

meaning that,

- the objective of **any feasible** solution of a standard form LP as Eq.(9) should always be larger than or equal to the objective of **any feasible** solution of its dual as Eq.(13).

and therefore,

- If either problem has an unbounded optimal objective, the other problem will be in-feasible.

- If both Eq.(9) and Eq.(13) are feasible, the objective of **the optimal** solution of a standard form LP as Eq.(9) should always be larger than or equal to the objective of **the optimal** solution of its dual as Eq.(13).
    - If we come across a pair of feasible solutions $(\mathbf{x}, \lambda')$ where the equality holds, then they should be (one of) the corresponding optimal solutions of each problem.

(Note that Eq.(9) seeks for minimum, and Eq.(13) seeks for maximum.)

2.2.2 Strong Duality Theorem

However, the above comparison is weak, and these questions currently remain:

- If either problem is in-feasible, should the other problem have an unbounded optimal objective?
- If either problem has a finite optimal objective, should the other problem also have a finite optimal objective?
- If both Eq.(9) and Eq.(13) are feasible, in which case the two problems have the equal optimal objective?

TODO

3 Benders Decomposition

3.1 To decompose the problem: a bi-level viewpoint

It may be helpful to consider the Benders decomposition (BD) algorithm as a bi-level nested loop, where the outer layer loop aims at optimizing some part of the set of decision variables, while the inner layer loop aims at optimizing the remaining part according to a given (partial) solution from the outer layer. In the context of BD, the outer layer is called a master problem (MP), while the inner is called a sub-problem (SP).

In this section, the "part of variables" for outer and inner layer (i.e., the MP and the SP) are denoted by $\mathbf{x}$ and $\mathbf{y}$ respectively.

For a given MILP (or even more generalized ?) model

$$\begin{aligned} \text{Minimize} \quad & f_1(\mathbf{x}) + f_2(\mathbf{x}, \mathbf{y}) \\ \text{s.t.} \quad & \begin{cases} (\mathbf{x}, \mathbf{y}) \in S \\ \mathbf{x} \in D_x \\ \mathbf{y} \in D_y \end{cases} \end{aligned} \tag{15}$$

where S defines the feasibility space (of decision variables $\mathbf{x}$ and $\mathbf{y}$) limited by the constraints, we can decompose it into a master problem (MP) as Eq.(16) and a sub-problem (SP) as Eq.(17).

$$\begin{aligned} & \text{MP:} \\ \text{Minimize} \quad & f_1(\mathbf{x}) + q \\ \text{s.t.} \quad & \begin{cases} (\mathbf{x}, q) \in S' \\ \mathbf{x} \in D_x \end{cases} \end{aligned} \tag{16}$$

SP with given $\mathbf{x} = \bar{\mathbf{x}_k}$ at the k-th outer iteration:

$$\begin{aligned} \text{Minimize} \quad & f_2(\bar{\mathbf{x}_k}, \mathbf{y}) \\ \text{s.t.} \quad & \begin{cases} (\bar{\mathbf{x}_k}, \mathbf{y}) \in S \\ \mathbf{y} \in D_y \end{cases} \end{aligned} \tag{17}$$

where an additional assistant variable q is introduced into the master problem, to embed the "information" about the partial objective from the sub-problem. These "information" will be implemented in the newly defined feasibility space S' by some techniques (which are the core difficulties of Benders decomposition methods, and will be explained in later sections).

3.2 How the two layers interact

In many bi-level meta-heuristics, the inner layer is responsible to optimize the sub-problem, providing clue for the outer layer to evaluate the partial solution it is focusing on. However, the results from the inner layer can also serve as other roles.

In BD, the results from the inner layer are used to generate constraints (named "cuts") for the outer layer MP. Note that, for a model that already completely expresses a real-world application, we can still add "redundant" constraints, and hopefully they may reduce the search space of variables. These additional constraints are referred to as "speed-up" constraints. Speed-up constraints can be added before starting optimization. However, it is also possible to dynamically add speed-up constraints during optimization, if they cannot be obtained in advance. The procedure of BD can be regarded as dynamically adding speed-up constraints to the master problem according to information found by the inner layer. Specifically, the feasibility space S' in Eq.(16) is dynamically changed, embedded with more and more information from the inner layer results.

So, how is the information expressed? How does the generated constraints look like?

3.3 Techniques to generate constraints for MP

Given an outer-layer partial solution $\mathbf{x} = \bar{\mathbf{x}_k}$, the inner layer has returned a (possibly approximately) optimal solution $\mathbf{y} = \mathbf{y}_k^*$, with the corresponding objective value $f_2(\bar{\mathbf{x}_k}, \mathbf{y}_k^*)$, of the SP. So what is the form of the generated speed-up constraints for MP according to $\mathbf{y}_k^*$ and $f_2(\bar{\mathbf{x}_k}, \mathbf{y}_k^*)$?

Let's first look at a simple example. The constraint

$$q \begin{cases} = f_2(\bar{\mathbf{x}_k}, \mathbf{y}_k^*), & \mathbf{x} = \bar{\mathbf{x}_k}, \\ \text{(unconstrained)}, & \mathbf{x} \neq \bar{\mathbf{x}_k}, \end{cases} \tag{18}$$

provides precise objective for MP when $\mathbf{x} = \bar{\mathbf{x}_k}$, and otherwise does nothing. This successfully embeds the information from inner-layer result into the outer-layer MP model. However, the result of the model is useless with this kind of constraints. The $\mathbf{x}$ will be assigned to some value not yet listed in any of these constraints, so the q can happily escape from all of these constraints and get a value of $-\infty$, resulting the objective function being minimized to $(-\infty + f_1(\mathbf{x}))$. There is one exception, i.e., if all possible values of $\mathbf{x}$ are enumerated by outer layer, solved by inner layer, and contribute a constraint. However, enumerating all feasible solutions is computationally impossible in most applications. (Otherwise we do not need to research on optimization methods!)

3.3.1 for linear-programming-style SP

3.3.2 for more general SP

3.3.3 No-good cuts for more general SP

$$ \tag{19}$$