

# Transformation Techniques in Optimization

Hailiang Zhao  
Zhejiang University

February 24, 2024

# Outline

## Introduction

## Transformation Techniques

- Multiplication of Binary Variables

- Multiplication of Binary and Continuous Variables

- Multiplication of Two Continuous Variables

- Maximum and Minimum Operators

- Absolute Value Function

- Floor and Ceiling Functions

- Multiple Breakpoint Function

## References

# Outline

## Introduction

## Transformation Techniques

- Multiplication of Binary Variables

- Multiplication of Binary and Continuous Variables

- Multiplication of Two Continuous Variables

- Maximum and Minimum Operators

- Absolute Value Function

- Floor and Ceiling Functions

- Multiple Breakpoint Function

## References

# Introduction

Many optimization problems have been formulated in the non-linear programming form. Finding a global optimum for them in acceptable computational time is challenging.

Linear programming (LP) forms of the optimization models are often recommended rather than solving integer or non-linear forms. Two ways to solve the non-linear optimizations ——

1. **Transformations** The non-linear equations or functions are replaced by an *exact* equivalent LP formulation
2. **Linear Approximations** Find the equivalent of a non-linear function with *the least deviation* around the point of interest or separate straight-line segments

# Transformations

Transformation into the LP model generally requires particular manipulations and substitutions in the original non-linear model along with the implementation of valid inequalities.

After solving the modified problem, the optimal values of the initial decision variables can be easily determined by reversing the transformation.

# Linear Approximations

Linear approximation of a function is an approximation (an affine function) that relies on a set of linear segments for calculation purposes. ▶ Piecewise or first-order methods

## Piecewise Example

Divide the curve and using linear interpolations between the points:

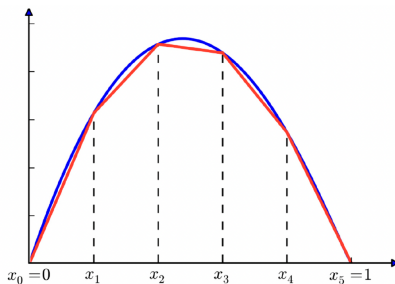


Figure 1: Piecewise linear approximation.

# Linear Approximations

Taylor's theorem approximates the output of a function  $f(x)$  around a given point by providing a  $k$ -times differentiable function and a polynomial of degree  $k$ , which is known as the  $k$ th-order Taylor polynomial.

## First-Order Taylor Polynomial Example

Divide the curve and using linear interpolations between the points:

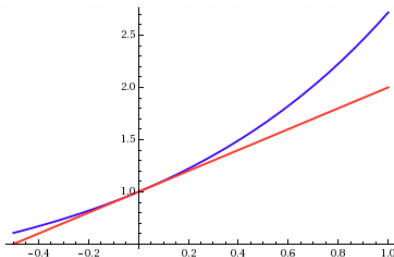


Figure 2: First-order Taylor polynomial.

# Outline

## Introduction

## Transformation Techniques

- Multiplication of Binary Variables

- Multiplication of Binary and Continuous Variables

- Multiplication of Two Continuous Variables

- Maximum and Minimum Operators

- Absolute Value Function

- Floor and Ceiling Functions

- Multiple Breakpoint Function

## References



## Multiplication of Binary Variables

Consider two binary variables  $x_i$  ( $i \in [m]$ ) and  $y_j$  ( $j \in [n]$ ). To linearize the term  $x_i \cdot y_j$ , we replace it with an additional binary variable:

$$z_{ij} := x_i \cdot y_j, \forall i \in [m], j \in [n]. \quad (1)$$

We also need to add some new constraints:

$$z_{ij} \leq x_i, \forall i, j, \quad (2)$$

$$z_{ij} \leq y_j, \forall i, j, \quad (3)$$

$$z_{ij} \geq x_i + y_j - 1, \forall i, j, \quad (4)$$

$$z_{ij} \in \{0, 1\}, \forall i, j. \quad (5)$$

## Multiplication of Binary Variables

It's easy to verify the correctness of the transformation with the following value table:

$x$	$y$	$x \cdot y$	Constraints	ImPLY
0	0	0	$z \leq 0$ $z \leq 0$ $z \geq -1$ $z \in \{0,1\}$	$z = 0$
0	1	0	$z \leq 0$ $z \leq 1$ $z \geq 0$ $z \in \{0,1\}$	$z = 0$
1	0	0	$z \leq 1$ $z \leq 0$ $z \geq 0$ $z \in \{0,1\}$	$z = 0$
1	1	1	$z \leq 1$ $z \leq 1$ $z \geq 1$ $z \in \{0,1\}$	$z = 1$

When binary variables have power ( $x_i^p$ ), w.l.o.g., one can omit the power of  $p$  ( $x_i := x_i^p$ ) and apply the same technique.

## Multiplication of Binary Variables

The extension to products of more than two variables is straightforward. In general, the multiplication of binary variables  $x_{i_k}^p$  ( $k \in [K], i_k \in I_k \in [m_k]$ ) for  $K \geq 2$  with different powers  $p$  can be linearized by replacing it with a new variable

$$z_j := \prod_{k=1}^K x_{i_k}^p, \quad (6)$$

where  $j = (i_1, \dots, i_K)$ . Additional variables:

$$z_j \leq x_{i_k}, \forall k, \forall i_k \in I_k, \forall j \in \cup_{k=1}^K I_k, \quad (7)$$

$$z_j \geq \sum_{k=1}^K x_{i_k} - (K - 1), \forall k, \forall i_k \in I_k, \forall j \in \cup_{k=1}^K I_k, \quad (8)$$

$$z_j \in \{0, 1\}, \forall j \in \cup_{k=1}^K I_k. \quad (9)$$

## Multiplication of Binary and Continuous Variables

Let  $x_i$  be a binary variable ( $i \in [m]$ ) and  $y_j$  be a continuous variable for which  $0 \leq y_j \leq u_j$  holds ( $j \in [n]$ ). To linearize the bilinear term  $x_i \cdot y_j$ , we replace it with the auxiliary variable  $z_{ij}$ . Additional variables:

$$z_{ij} \leq y_j, \forall i, j, \quad (10)$$

$$z_{ij} \leq u_j \cdot x_i, \forall i, j, \quad (11)$$

$$z_{ij} \geq y_j + u_j \cdot (x_i - 1), \forall i, j, \quad (12)$$

$$z_{ij} \geq 0, \forall i, j. \quad (13)$$

$x$	$y$	$x \cdot y$	Constraints	Imply
0	$m : 0 \leq m \leq u$	0	$z \leq m$ $z \leq 0$ $z \geq m - u$ $z \geq 0$	$z = 0$
1	$m : 0 \leq m \leq u$	$m$	$z \leq m$ $z \leq u$ $z \geq m$ $z \geq 0$	$z = m$

## Multiplication of Two Continuous Variables

Linearization of multiplication of continuous variables can be complex. Below provides a hint for bounded variables.

We assume that term  $x_1 \cdot x_2$  must be converted. First of all, we define two new continuous variables  $y_1$  and  $y_2$  as follows:

$$y_1 := \frac{1}{2}(x_1 + x_2), \quad (14)$$

$$y_2 := \frac{1}{2}(x_1 - x_2). \quad (15)$$

Then  $x_1 \cdot x_2$  can be replaced with a separate function:

$$y_1^2 - y_2^2 := x_1 \cdot x_2. \quad (16)$$

## Multiplication of Two Continuous Variables

Note that  $y_1^2 - y_2^2$  can be linearized with piecewise approximation. We can eliminate the non-linear function at the cost of having to approximate the objective.

If  $l_1 \leq x_1 \leq u_1$  and  $l_2 \leq x_2 \leq u_2$ , then the lower and upper bounds on  $y_1$  and  $y_2$  are:

$$\frac{1}{2}(l_1 + l_2) \leq y_1 \leq \frac{1}{2}(u_1 + u_2), \quad (17)$$

$$\frac{1}{2}(l_1 - u_2) \leq y_2 \leq \frac{1}{2}(u_1 - l_2). \quad (18)$$

## Multiplication of Two Continuous Variables

If

- ▶ one of the variables is not referenced in any other term except in the products of the above form,
- ▶ and the lower bounds  $l_1$  and  $l_2$  are non-negative,

there is a simpler way.

Suppose  $x_1$  is not used in any other terms. We can substitute  $x_1 \cdot x_2$  with a single variable  $z$  with the following additional constraints:

$$l_1 \cdot x_2 \leq z \leq u_1 \cdot x_2. \quad (19)$$

Once the resulting mathematical formulation is solved in terms of  $z$  and  $x_2$ , it is required to calculate  $x_1 = \frac{z}{x_2}$  whenever  $x_2 > 0$ .  $x_1$  is undetermined when  $x_2 = 0$  since the extra constraints on  $z$  ensure that  $l_1 \leq x_1 \leq u_1$  only when  $x_2 > 0$ .

## Maximum Operators

Assume there is a general non-linear structure in the form of  $\max_i \{x_i\}$ , where  $i \in [n]$ . It can be transformed into  $z := \max_i \{x_i\}$  with the following additional constraints:

$$z \geq x_i, \forall i, \quad (20)$$

$$z \leq x_i + m \cdot y_i, \forall i, \quad (21)$$

$$\sum_i y_i \leq n - 1, \quad (22)$$

$$y_i \in \{0, 1\}, \forall i, \quad (23)$$

where  $m$  is a sufficiently large number. (21) and (22) are used to ensure that for only one  $i$ ,  $z$  has to be less than or equal to  $x_i$  (preventing  $z$  from being  $\infty$ ).



## Minimum Operators

For the term  $\min_i \{x_i\}$ , (20) and (21) are replaced by:

$$z \leq x_i, \forall i, \quad (24)$$

$$z \geq x_i - m \cdot y_i, \forall i. \quad (25)$$

In this case, (25) and (22) are used to ensure that for only one  $i$ ,  $z$  has to be greater than or equal to  $x_i$ .

# Absolute Value Function

## Absolute Value in Constraints

For  $|f(x)| \leq z$  where  $f(x)$  is linear, we can have it replaced by

$$f(x) \leq z, \quad (26)$$

$$-f(x) \leq z. \quad (27)$$

The same logic can be applied for  $|f(x)| \geq z$  and  $|f(x)| + g(x) \leq z$  (or  $\geq$ ).

# Absolute Value Function

## Absolute Value in the Objective Function

If the objective is

$$\max_{x,y} -|f(x)| + g(y)$$

or

$$\min_{x,y} |f(x)| + g(y),$$

we can substitute  $|f(x)|$  by  $z$  and add two extra constraints  $f(x) \leq z$  and  $-f(x) \leq z$ .

# Absolute Value Function

## Minimizing the Sum of Absolute Deviations

The problem is:

$$\min_{x_i, y_j} \sum_i |x_i| \quad (28)$$

$$s.t. \quad x_i + \sum_j a_{ij} y_j = b_i, \forall i \in [m], \quad (29)$$

$$x_i, y_j \in \mathbb{R}, \forall i \in [m], j \in [n]. \quad (30)$$

# Absolute Value Function

## Minimizing the Sum of Absolute Deviations (Cont'd)

To linearize it, we replace  $x_i$  with  $x_i^+ - x_i^-$  (where the two variables are non-negative). The problem is transformed into:

$$\min_{x_i, y_j} \sum_i |x_i^+ - x_i^-| \quad (31)$$

$$s.t. \quad x_i^+ - x_i^- + \sum_j a_{ij} y_j = b_i, \forall i \in [m], \quad (32)$$

$$x_i = x_i^+ - x_i^-, \forall i \in [m], \quad (33)$$

$$x_i^+, x_i^- \geq 0, \forall i \in [m], \quad (34)$$

$$x_i, y_j \in \mathbb{R}, \forall i \in [m], j \in [n]. \quad (35)$$

At the optimal solution, it can be proven that  $x_i^+ \cdot x_i^- = 0$ . Therefore, the model is reformulated to a linear programming form, as (31) replaced by  $\min_{x_i, y_j} \sum_i (x_i^+ + x_i^-)$ .

# Absolute Value Function

## Minimizing the Maximum of Absolute Values

The problem is:

$$\min_{x_i, y_j} \max_i |x_i| \quad (36)$$

$$s.t. \quad x_i + \sum_j a_{ij} \cdot y_j = b_i, \forall i \in [m], \quad (37)$$

$$x_i, y_j \in \mathbb{R}, \forall i \in [m], j \in [n]. \quad (38)$$

$x_i$  is the deviation for the  $i$ th observation  $b_i$  and  $y_j$  is the  $j$ th variable in the linear equation.

# Absolute Value Function

## Minimizing the Maximum of Absolute Values (Cont'd)

We can use  $x$  to substitute  $\max_i |x_i|$ , and the problem is re-formulated as:

$$\min_{x_i, y_j} x \quad (39)$$

$$s.t. \quad x \geq b_i - \sum_j a_{ij} y_j, \forall i, \quad (40)$$

$$x \geq -(b_i - \sum_j a_{ij} y_j), \forall i, \quad (41)$$

$$x \geq 0, \quad (42)$$

$$y_j \in \mathbb{R}, \forall j. \quad (43)$$

## Floor and Ceiling Functions

For  $\lfloor f(x) \rfloor$ , we can replace it by  $y$  and adding the following constraints:

$$y \leq f(x) < y + 1, \quad (44)$$

$$y \in \mathbb{Z}. \quad (45)$$

For  $\lceil f(x) \rceil$ , we can replace it by  $y$  and adding the following constraints:

$$y - 1 < f(x) \leq y, \quad (46)$$

$$y \in \mathbb{Z}. \quad (47)$$



## Multiple Breakpoint Function

Suppose there is a general continuous multiple breakpoint function that can be defined as follows:

$$f(x) = \begin{cases} a_1x + b_1 & c_0 \leq x \leq c_1, \\ a_2x + b_2 & c_1 \leq x \leq c_2, \\ \vdots & \vdots \\ a_nx + b_n & c_{n-1} \leq x \leq c_n. \end{cases} \quad (48)$$

### Tasi's Method

Firstly, we can simplify the formulation into the following one:

$$f(x) = \sum_i t_i \cdot (a_i x + b_i) \quad (49)$$

$$s.t. \quad \sum_i c_{i-1} t_i \leq x \leq \sum_i c_i t_i, \quad (50)$$

$$\sum_i t_i = 1 \text{ and } t_i \in \{0, 1\}. \quad (51)$$

# Multiple Breakpoint Function

## Tasi's Method (Cont'd)

We define  $g_i(x) = a_i x + b_i$ . Then, we replace  $t_i g_i(x)$  with  $z_i$ , the problem is then transformed into:

$$f(x) = \sum_i z_i \quad (52)$$

$$s.t. \quad \sum_i c_{i-1} t_i \leq x \leq \sum_i c_i t_i, \quad (53)$$

$$\sum_i t_i = 1 \text{ and } t_i \in \{0, 1\}, \forall i, \quad (54)$$

$$g_i(x) - (1 - t_i)m \leq z_i, \forall i, \quad (55)$$

$$g_i(x) + (1 - t_i)m \geq z_i, \forall i, \quad (56)$$

$$-t_i m \leq z_i \leq t_i m, \forall i, \quad (57)$$

$$z_i \in \mathbb{R}, \quad (58)$$

where  $m$  is a sufficiently large number.

# Multiple Breakpoint Function

## Mirzapour's Method

$f(x)$  can also be linearized by introducing some binary variables  $t_i$  and also converting variable  $x$  to  $n$  independent variables  $x_i$ , where  $x = \sum_i x_i$ .

The problem is then transformed into:

$$f(x) = \sum_i t_i(a_i x_i + b_i) \quad (59)$$

$$\text{s.t. } c_{i-1}t_i \leq x_i \leq c_i t_i, \forall i, \quad (60)$$

$$\sum_i t_i = 1 \text{ and } t_i \in \{0, 1\}, \forall i, \quad (61)$$

$$x_i \in \mathbb{R}, \forall i. \quad (62)$$

# Multiple Breakpoint Function

## Mirzapour's Method (Cont'd)

If  $f(x)$  is dis-continuous:

$$f(x) = \begin{cases} a_1x + b_1 & x \leq c_1, \\ a_2x + b_2 & c_1 < x \leq c_2, \\ \vdots & \vdots \\ a_nx + b_n & c_{n-1} < x. \end{cases} \quad (63)$$

We only need to substitute (60) with

$$(c_{i-1} + \frac{1}{m})t_i \leq x_i \leq c_i t_i, \forall i \in \{2, \dots, n-1\}, \quad (64)$$

$$x_1 \leq c_1 t_1, \quad (65)$$

$$(c_{n-1} + \frac{1}{m})t_n \leq x_n. \quad (66)$$

# Outline

## Introduction

## Transformation Techniques

- Multiplication of Binary Variables

- Multiplication of Binary and Continuous Variables

- Multiplication of Two Continuous Variables

- Maximum and Minimum Operators

- Absolute Value Function

- Floor and Ceiling Functions

- Multiple Breakpoint Function

## References

# References

1. Asghari M, Fathollahi-Fard A M, Mirzapour Al-E-Hashem S M J, et al. [Transformation and linearization techniques in optimization: A state-of-the-art survey\[J\]](#). Mathematics, 2022, 10(2): 283.
2. Pshenichnyj B N. [The linearization method for constrained optimization\[M\]](#). Springer Science & Business Media, 2012.