Transformation Techniques in Optimization

Hailiang Zhao
Zhejiang University

February 25, 2024

Outline

Outline

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:

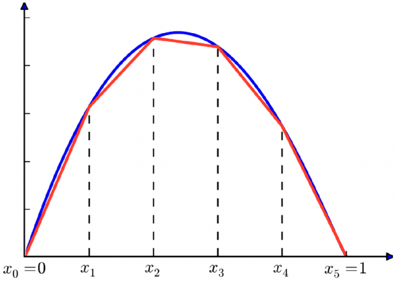

Figure 1: Piecewise linear approximation.

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 kth-order Taylor polynomial.

First-Order Taylor Polynomial Example

An approximation of $f(x) = e^x$ at $(0, f(0))$:

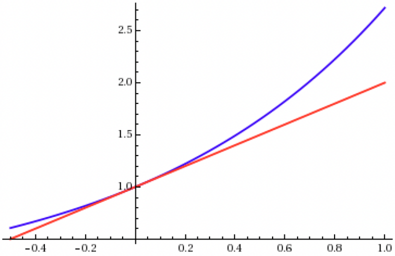

Figure 2: First-order Taylor polynomial.

Outline

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:

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

We also need to add some new constraints:

$$z_{ij} \leq x_i, \forall i, j, \tag{2}$$
$$z_{ij} \leq y_j, \forall i, j, \tag{3}$$
$$z_{ij} \geq x_i + y_j - 1, \forall i, j, \tag{4}$$
$$z_{ij} \in \{0, 1\}, \forall i, j. \tag{5}$$

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

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

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

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

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

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

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:

$$z_{ij} \leq y_j, \forall i, j, \tag{10}$$

$$z_{ij} \leq \mu_j \cdot x_i, \forall i, j, \tag{11}$$

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

$$z_{ij} \geq 0, \forall i, j. \tag{13}$$

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:

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

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

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

$$y_1^2 - y_2^2 := x_1 \cdot x_2. \tag{16}$$

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:

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

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

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_2 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:

$$l_1 \cdot x_2 \leq z \leq u_1 \cdot x_2. \tag{19}$$

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:

$$z \geq x_i, \forall i, \tag{20}$$
$$z \leq x_i + m \cdot y_i, \forall i, \tag{21}$$
$$\textstyle\sum_i y_i \leq n - 1, \tag{22}$$
$$y_i \in \{0, 1\}, \forall i, \tag{23}$$

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

Minimum Operators

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

$$z \leq x_i, \forall i, \tag{24}$$

$$z \geq x_i - m \cdot y_i, \forall i. \tag{25}$$

In this case, (25) and (22) are used to ensure that for only one i, z has to be greater then 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

$$f(x) \leq z, \tag{26}$$
$$-f(x) \leq z. \tag{27}$$

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:

$$\min_{x_i, y_j} \sum_i |x_i| \tag{28}$$

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

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

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:

$$
\begin{aligned}
& \min_{x_i, y_j} \sum_i |x_i^+ - x_i^-| && (31)\\
s.t. \quad & x_i^+ - x_i^- + \sum_j a_{ij} y_j = b_i, \forall i \in [m], && (32)\\
& x_i = x_i^+ - x_i^-, \forall i \in [m], && (33)\\
& x_i^+, x_i^- \geq 0, \forall i \in [m], && (34)\\
& x_i, y_j \in \mathbb{R}, \forall i \in [m], j \in [n]. && (35)
\end{aligned}
$$

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:

$$\min_{x_i, y_j} \max_i |x_i| \tag{36}$$

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

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

x_i is the deviation for the ith observation b_i and y_j is the jth 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:

$$
\begin{aligned}
& \min_{x_i, y_j} x && (39)\\
s.t. \quad & x \geq b_i - \textstyle\sum_j a_{ij} y_j, \forall i, && (40)\\
& x \geq -(b_i - \textstyle\sum_j a_{ij} y_j), \forall i, && (41)\\
& x \geq 0, && (42)\\
& y_j \in \mathbb{R}, \forall j. && (43)
\end{aligned}
$$

Floor and Ceiling Functions

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

$$y \le f(x) < y + 1, \tag{44}$$
$$y \in \mathbb{Z}. \tag{45}$$

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

$$y - 1 < f(x) \le y, \tag{46}$$
$$y \in \mathbb{Z}. \tag{47}$$

Multiple Breakpoint Function

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

$$f(x) = \begin{cases} a_1 x + b_1 & c_0 \leq x \leq c_1, \\ a_2 x + b_2 & c_1 \leq x \leq c_2, \\ \vdots & \vdots \\ a_n x + b_n & c_{n-1} \leq x \leq c_n. \end{cases} \tag{48}$$

Tasi's Method

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

$$f(x) = \sum_i t_i \cdot (a_i x + b_i) \tag{49}$$

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

$$\sum_i t_i = 1 \text{ and } t_i \in \{0, 1\}. \tag{51}$$

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:

$$
\begin{aligned}
f(x) &= \textstyle\sum_i z_i && (52)\\
s.t. \quad \textstyle\sum_i c_{i-1} t_i &\le x \le \textstyle\sum_i c_i t_i, && (53)\\
\textstyle\sum_i t_i &= 1 \text{ and } t_i \in \{0,1\}, \forall i, && (54)\\
g_i(x) - (1 - t_i)m &\le z_i, \forall i, && (55)\\
g_i(x) + (1 - t_i)m &\ge z_i, \forall i, && (56)\\
-t_i m &\le z_i \le t_i m, \forall i, && (57)\\
z_i &\in \mathbb{R}, && (58)
\end{aligned}
$$

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:

$$f(x) = \sum_i t_i(a_i x_i + b_i) \tag{59}$$

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

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

$$x_i \in \mathbb{R}, \forall i. \tag{62}$$

Multiple Breakpoint Function

Mirzapour's Method (Cont'd)

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

$$f(x) = \begin{cases} a_x + b_1 & x \leq c_1, \\ a_2 x + b_2 & c_1 < x \leq c_2, \\ \vdots & \vdots \\ a_n x + b_n & c_{n-1} < x. \end{cases} \tag{63}$$

We only need to substitute (60) with

$$(c_{i-1} + \tfrac{1}{m}) t_i \leq x_i \leq c_i t_i, \forall i \in \{2, ..., n-1\}, \tag{64}$$

$$x_1 \leq c_1 t_1, \tag{65}$$

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

Outline

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.