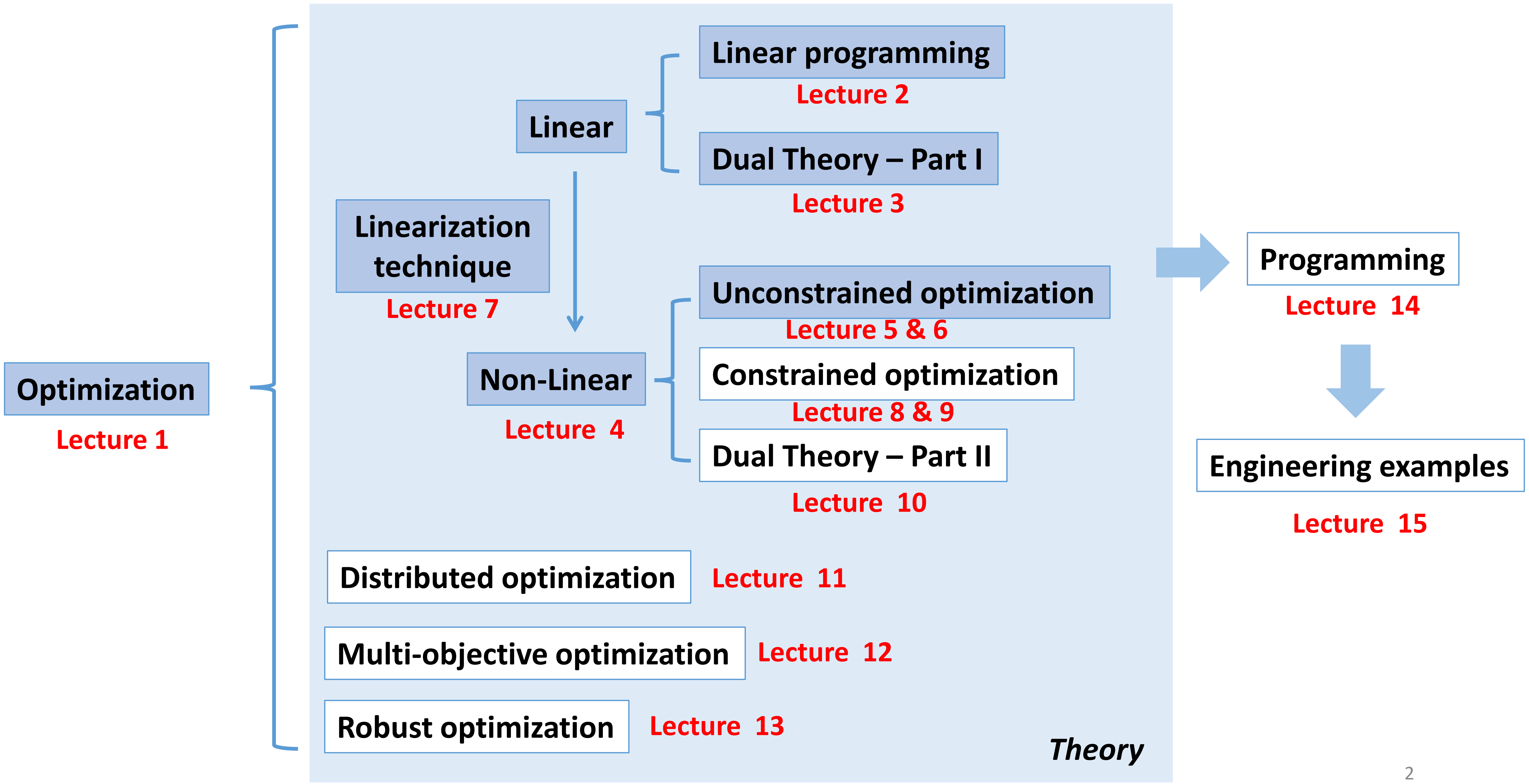
MAEG4070 Engineering Optimization

Lecture 7 Linearization Techniques

Yue Chen
MAE, CUHK
email: yuechen@mae.cuhk.edu.hk
Oct 5, 2022

Content of this course (tentative)



Overview

Despite the various algorithms we have learned to solve nonlinear optimization problems, they can be computationally inefficient with a growing number of variables; may only reach a local optimum.

In today's lecture, we will introduce a new kind of optimization problem – mixed integer linear programming (**MILP**); and try to solve the nonlinear optimization by turning it into MILPs via linearization techniques.

$$
\begin{aligned}
\min_{x,z} \quad & c^T x + d^T z \\
\text{s.t.} \quad & Ax + Bz \leq b \\
& x \in \mathbb{R}^n, z \in \{0,1\}^m \ (\text{or } z \in \mathbb{Z}^n)
\end{aligned}
$$

Typical MILP problems

Knapsack problem

- The weight capacity of the knapsack is c
- Weight of each item is $w_k, \forall k = 1, \dots, K$
- Value of each item is $v_k, \forall k = 1, \dots, K$
- We aim to maximize the total value

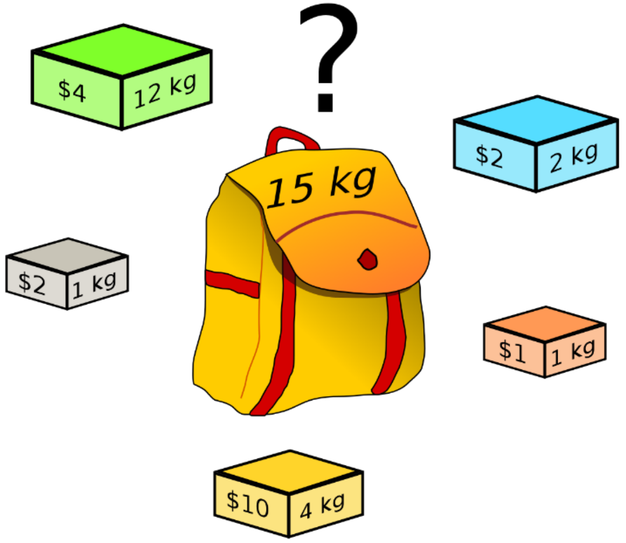
$$\max_{x_k, \forall k=1,\dots,K} \sum_{k=1}^{K} v_k x_k$$

$$\text{s.t.} \sum_{k=1}^{K} w_k x_k \leq c$$

$$x_k \in \{0, 1\}, \forall k = 1, \dots, K$$

$x_k = 1$ means item k is included

$x_k = 0$ means item k is not included



Solution:

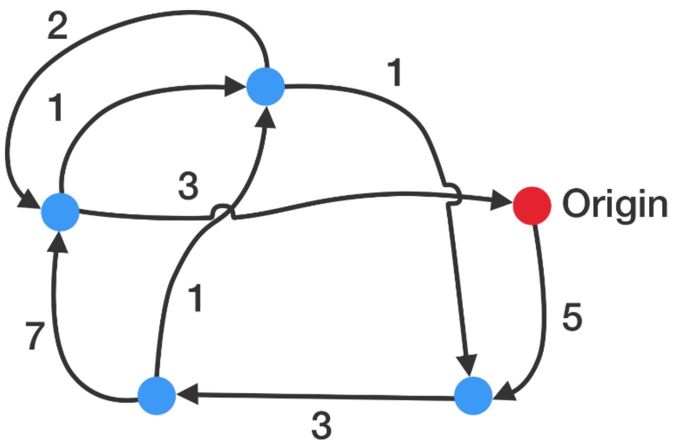
- Try all possible situations
- Dynamic programming
- heuristic

Pictures from Google image

Typical MILP problems

Traveling Salesman problem

- There are n cities, and the salesman want to find the shortest route to visit every city once and returns to the origin.
- The distance between city i and city j is d_{ij}, and $d_{ii} = \infty$.



$x_{ij} = 1$ if the salesman travels from city i to city j.

To ensure each city is visited only once, there is only one way in and only one way out of the city.

$$\sum_{j=1}^{n} x_{ij} = 1, \forall i = 1, ..., n$$

$$\sum_{i=1}^{n} x_{ij} = 1, \forall j = 1, ..., n$$

Pictures from Google image

Typical MILP problems

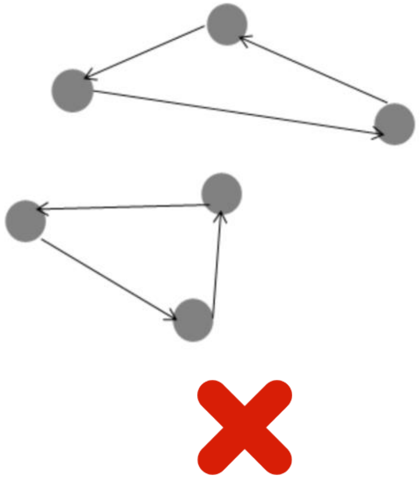
However..
It is possible that there are loops.

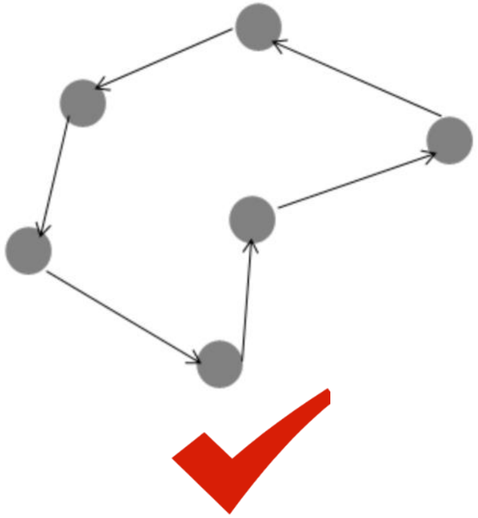
To eliminate the subtours, Dantzig, Fulkerson and Johnson proposed the DFJ formulation in 1954.

$$\sum_{i,j \in S} x_{ij} \leq |S| - 1, \forall S \subset \{1, \dots, n\}, 1 < |S| < n$$

For example, let $n = 10$, $S = \{2,3,4\}$, $|S| = 3$
The subtour elimination constraint is

$$x_{23} + x_{24} + x_{34} + x_{32} + x_{42} + x_{43} \leq 2$$





Pictures from Google image

Typical MILP problems

The problem can be modeled as

$$\min_{x_{ij},\forall i,j} \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}x_{ij}$$
$$\sum_{j=1}^{n} x_{ij} = 1, \forall i = 1, ..., n$$
$$\sum_{i=1}^{n} x_{ij} = 1, \forall j = 1, ..., n$$
$$\sum_{i,j\in S} x_{ij} \leq |S| - 1, \forall S \subset \{1, ..., n\}, 1 < |S| < n$$
$$x_{ij} \in \{0, 1\}, \forall i, j$$

Another formulation

$$\min_{x_{ij},\forall i,j} \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}x_{ij}$$
$$\sum_{j=1}^{n} x_{ij} = 1, \forall i = 1, ..., n$$
$$\sum_{i=1}^{n} x_{ij} = 1, \forall j = 1, ..., n$$
$$u_i - u_j + nx_{ij} \leq n - 1, 1 < i \neq j \leq n$$
$$x_{ij} \in \{0, 1\}, \forall i, j, u_i \in \mathbb{R}$$

$x_{ij} = 1$ if the salesman travels from city i to city j.

Solution Methods

To solve the MILP, we can

- Enumeration
  - ✓ The number of feasible solutions are finite
  - ✓ High computational burden under high dimension

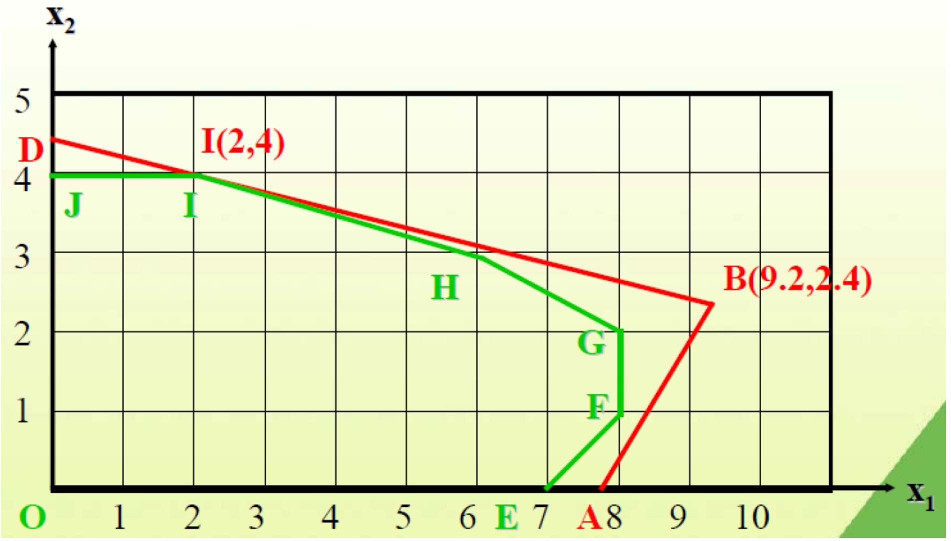
- Relaxation and rounding

- Branch and bound
  - ✓ Relax integrality requirement
  - ✓ Enumeration on non-integer solutions
  - ✓ Cut branches without an optimal solution
  - ✓ Used by solvers: CPLEX, Gurobi

Solution Methods

We compare the performance of different methods using the following example. Let's consider the optimization:

$$
\begin{aligned}
\max_{x_1,x_2}\ & 3x_1 + 13x_2 \\
\text{s.t. } & 2x_1 + 9x_2 \le 40 \\
& 11x_1 - 8x_2 \le 82 \\
& x_1, x_2 \in \mathbb{Z}_+
\end{aligned}
$$



	Relaxed LP	Rounding	Nearest feasible	Exact Solution
Optimal point	(9.2, 2.4)	(9, 2)	(8, 2)	(2, 4)
Optimal value	58.8	infeasible	50	58

Pictures from Google image

Linearization techniques-1

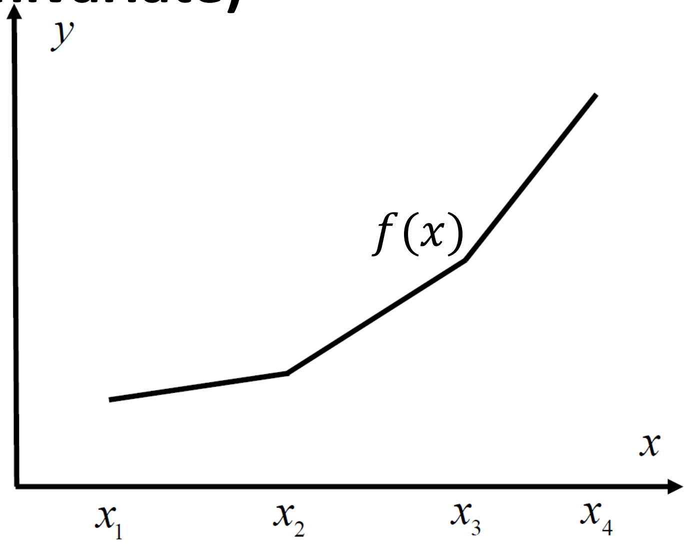
Minimizing a convex piecewise linear function (univariate)

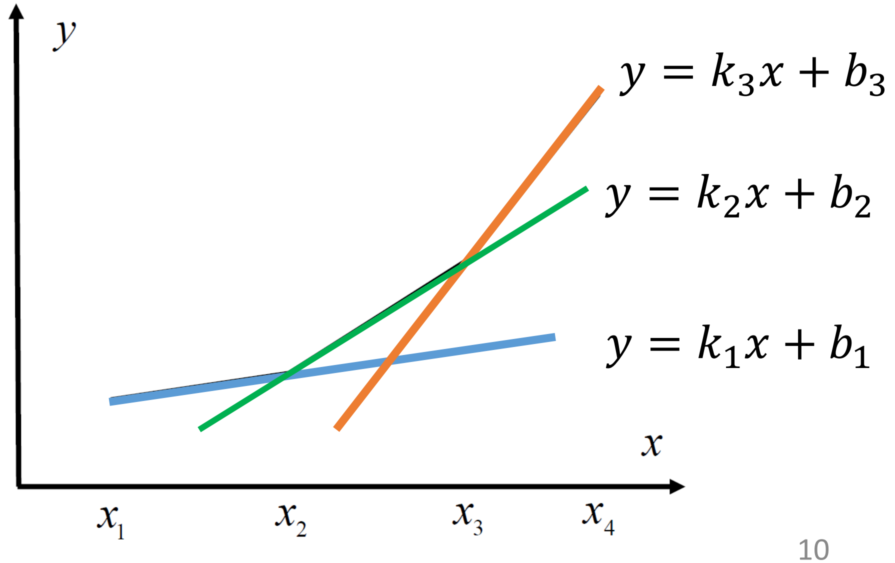
$$\min_{x} \; f(x)$$
$$\text{s.t. } x_1 \le x \le x_4$$

where

$$f(x) = \begin{cases} k_1 x + b_1, & x \in [x_1, x_2] \\ k_2 x + b_2, & x \in [x_2, x_3] \\ k_3 x + b_3, & x \in [x_3, x_4] \end{cases}$$

$$\min_{x,\sigma} \; \sigma$$
$$\text{s.t. } \sigma \ge k_1 x + b_1$$
$$\sigma \ge k_2 x + b_2$$
$$\sigma \ge k_3 x + b_3$$
$$x_1 \le x \le x_4$$





Linearization techniques-1

Minimizing a convex piecewise linear function (univariate)

Another equivalent form

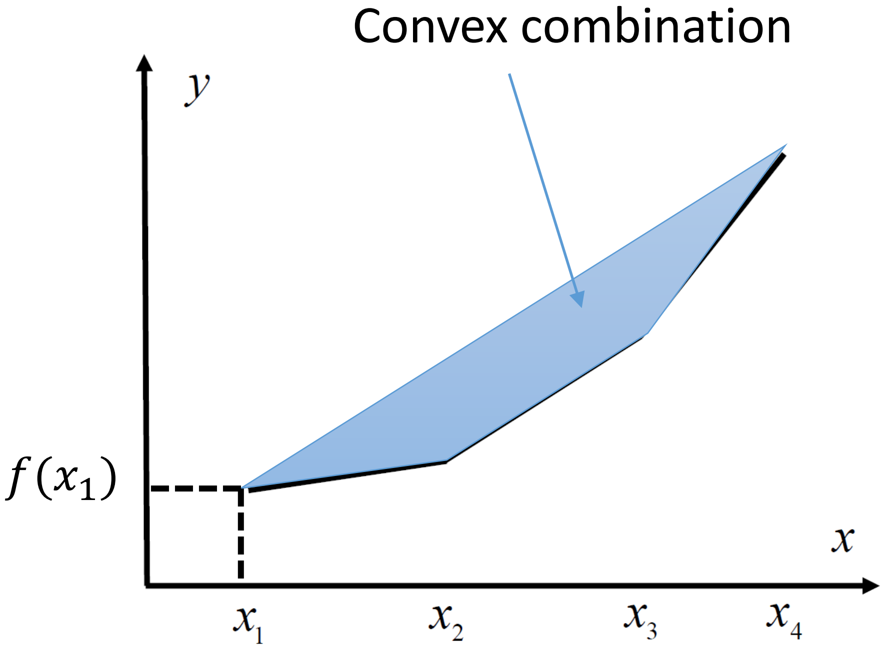
$$\min_{x,y,\lambda} \; y$$

$$\text{s.t. } x = \sum_{n=1}^{N} \lambda_n x_n$$

$$y = \sum_{n=1}^{N} \lambda_n f(x_n)$$

$$0 \leq \lambda_n \leq 1, \forall n = 1, \ldots, N$$

$$\sum_{n=1}^{N} \lambda_n = 1$$

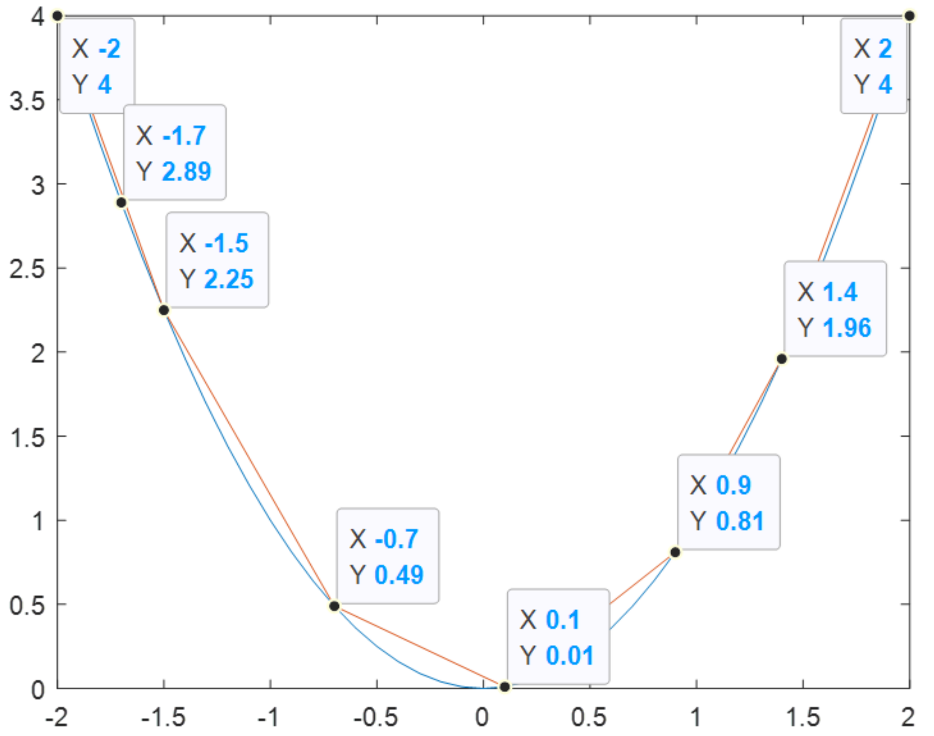


Example

For the function $f(x) = x^2$
Take some sample points

$$x = -2, -1.5, -0.7, 0.1, 0.9, 1.4, 2$$

We can use a piecewise linear function to approximate it, then minimizing $f(x)$ is equivalent to



$$\begin{aligned}
\min_{x,y,\sigma} \ & y \\
\text{s.t. } & x = -2\sigma_1 - 1.5\sigma_2 - 0.7\sigma_3 + 0.1\sigma_4 + 0.9\sigma_5 + 1.4\sigma_6 + 2\sigma_7 \\
& y = 4\sigma_1 + 2.25\sigma_2 + 0.49\sigma_3 + 0.01\sigma_4 + 0.81\sigma_5 + 1.96\sigma_6 + 4\sigma_7 \\
& \sigma_k \geq 0, \forall k = 1, \dots, 7 \\
& \sum_{k=1}^{7} \sigma_k = 1
\end{aligned}$$

$$x^* = 0.1, y^* = 0.01$$

More segments, more accurate
But more time-consuming

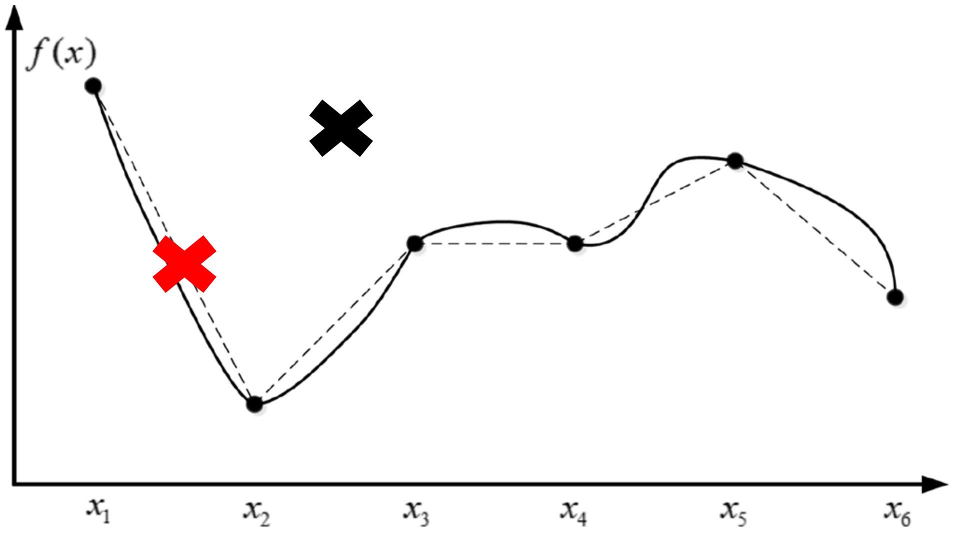
Linearization techniques-2

When the function appear in constraints

Representing a piecewise linear function (univariate)

We can approximate a nonlinear function by a piecewise linear function as in the Fig.

How to represent this piecewise linear function in a MILP form?



$$x = \sum_{n=1}^{N} \lambda_n x_n$$

$$y = \sum_{n=1}^{N} \lambda_n f(x_n)$$

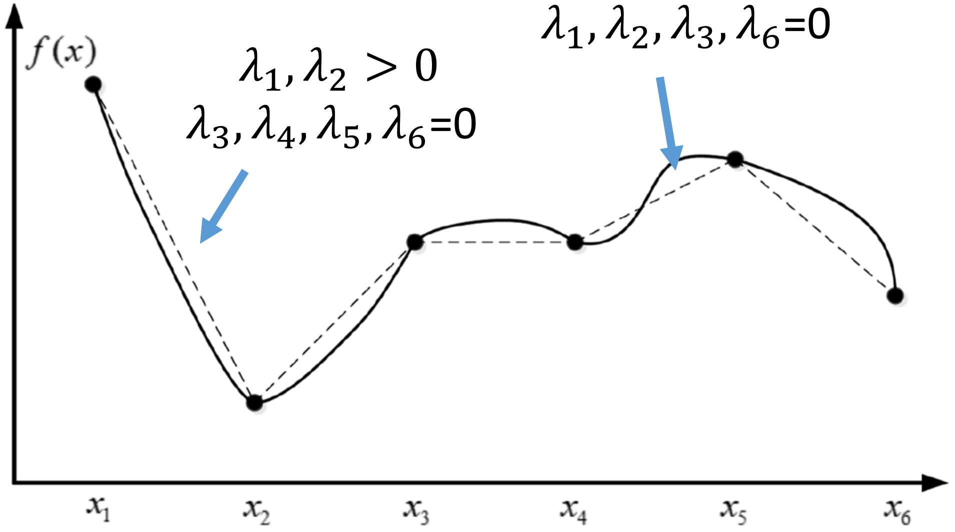
$$0 \le \lambda_n \le 1, \forall n = 1, ..., N; \sum_{n=1}^{N} \lambda_n = 1$$

?

How to tackle this issue?

Linearization techniques-2

Representing a piecewise linear function (univariate)



Special-ordered set of Type 2 (SOS2)

- An ordered set of non-negative variables, of which at most two consecutive elements can take strictly positive values, and the remaining ones equals to zero.

$$x = \sum_{n=1}^{N} \lambda_n x_n$$

$$y = \sum_{n=1}^{N} \lambda_n f(x_n)$$

$$0 \leq \lambda_n \leq 1, \forall n = 1, ..., N; \sum_{n=1}^{N} \lambda_n = 1$$

λ is S0S2

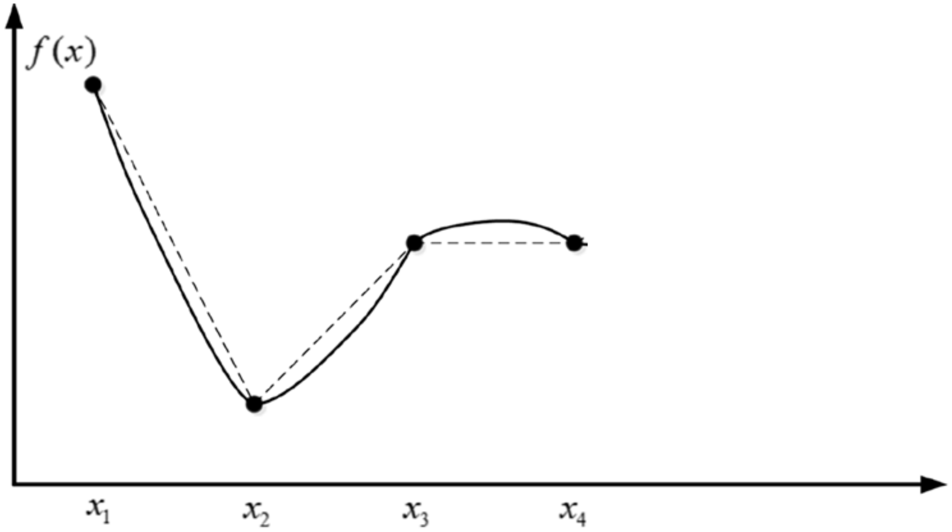
$$\lambda_1 \leq \theta_1$$
$$\lambda_2 \leq \theta_1 + \theta_2$$
$$\lambda_3 \leq \theta_2 + \theta_3$$
$$\ldots$$
$$\lambda_{N-1} \leq \theta_{N-2} + \theta_{N-1}$$
$$\lambda_N \leq \theta_{N-1}$$

$$\lambda_i \geq 0, n = 1, ..., N; \sum_{n=1}^{N} \lambda_n = 1$$

$$\theta_s \in \{0, 1\}, s = 1, ..., (N-1); \sum_{n=1}^{N-1} \theta_s = 1$$

Example

$$\lambda_1 \le \theta_1$$
$$\lambda_2 \le \theta_1 + \theta_2$$
$$\lambda_3 \le \theta_2 + \theta_3$$
$$\lambda_4 \le \theta_3$$
$$\theta \in \{0,1\}, \sum_{i=1}^{4} \theta_i = 1$$



If $\theta_1 = 1$, then we have $\lambda_1 \le 1, \lambda_2 \le 1, \lambda_3 \le 0, \lambda_4 \le 0$
Since $0 \le \lambda_i \le 1$, we have $\lambda_3 = 0, \lambda_4 = 0$

$$x = x_1\lambda_1 + x_2\lambda_2$$
$$y = y_1\lambda_1 + y_2\lambda_2$$

Line segment between (x_1, y_1) and (x_2, y_2)

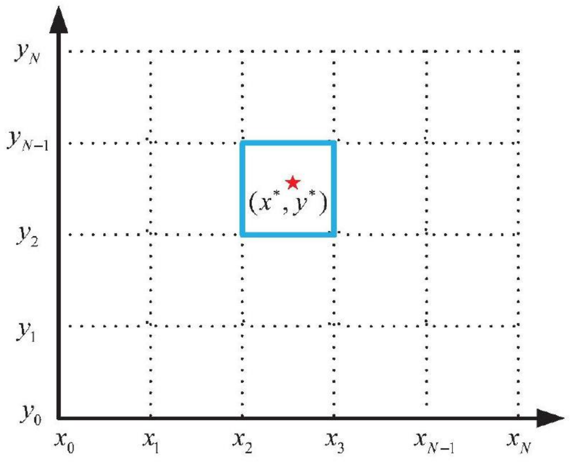
Linearization techniques-2

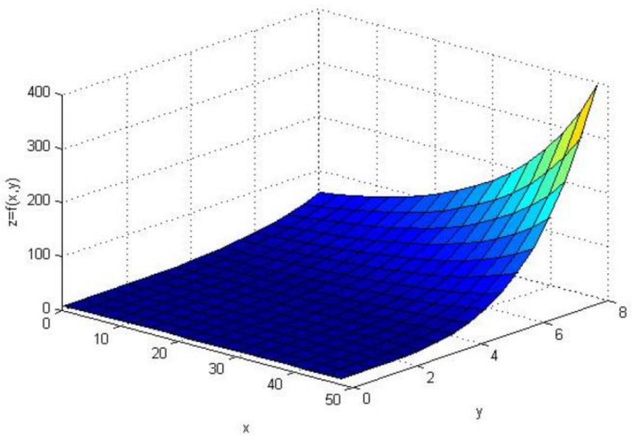
Representing a piecewise linear function (bivariate)

$$x = \sum_{m=1}^{M} \sum_{n=1}^{N} \lambda_{mn} x_n; \lambda^n = \sum_{m=1}^{M} \lambda_{mn} \text{ is S0S2}$$

$$y = \sum_{m=1}^{M} \sum_{n=1}^{N} \lambda_{mn} y_m; \lambda^m = \sum_{n=1}^{N} \lambda_{mn} \text{ is S0S2}$$

$$f(x,y) = \sum_{m=1}^{M} \sum_{n=1}^{N} \lambda_{mn} f_{mn}, \lambda_{mn} \geq 0, \forall m, \forall n; \sum_{m=1}^{M} \sum_{n=1}^{N} \lambda_{mn} = 1$$





Linearization techniques-3

Linearize the product of two binary variables

How about $z = 3xy$?

Consider $z = xy, x, y \in \{0,1\}$
It can be linearized by

$$0 \le z \le y$$
$$0 \le x - z \le 1 - y$$

Proof of equivalence:

1. If $y = 0$, then the first inequality becomes $z = 0$ and the second $0 \le x \le 1$. Meanwhile, we have $z = xy = 0$.
2. If $x = 0$, then we have $0 \le z \le y$ and $0 \le -z \le 1 - y$, therefore, $z = 0$ and $0 \le y \le 1$.
3. If $x = 1$ and $y = 1$, then we have $0 \le z \le 1$ and $0 \le 1 - z \le 0$, thus, $z = 1$.

Linearization techniques-4

Linearize the product of a binary and a continuous variable

Consider $z = xy, x \in [x_l, x_u], y \in \{0,1\}$

How about $z = 3xy$?

It can be linearized by

$$x_l y \leq z \leq x_u y$$
$$x_l(1-y) \leq x - z \leq x_u(1-y)$$

Proof of equivalence:

1. If $y = 0$, then the first inequality becomes $z = 0$ and the second $x_l \leq x \leq x_u$. Meanwhile, we have $z = xy = 0$.
2. If $y = 1$, then the second inequality becomes $x = z$ and the first $x_l \leq x = z \leq x_u$. Meanwhile, we have $z = xy = x$.

Linearization techniques-5

Linearize monomial of binary variables

Consider $z = x_1 x_2 \dots x_N, x_n \in \{0,1\}, \forall n = 1, \dots, N$

It is equivalent to

$$z \in \{0,1\}$$

$$z \leq \frac{x_1 + \cdots + x_N}{N}$$

$$z \geq \frac{x_1 + \cdots + x_N - n + 1}{N}$$

- If one of the $x_n = 0$, then the first inequality removes $z = 1$ from the feasible region, and the second inequality is redundant.
- If all $x_n = 1$, then the second inequality removes $z = 0$ from the feasible region, and the first inequality is redundant.

Linearization techniques-6

Complementary and slackness condition in KKT condition (will learn in lecture 8)

Consider condition $0 \leq x \perp y \geq 0$
It is equivalent to $x, y \geq 0, xy = 0$
And can be linearized by

$$0 \leq x \leq Mz$$
$$0 \leq y \leq M(1 - z)$$
$$z \in \{0,1\}^n$$

Proof of equivalence:

1. If $x = 0, y > 0$, then let $z = 0$
2. If $x > 0, y = 0$, then let $z = 1$
3. If $x = 0, y = 0$, then let $z = 0$ or $z = 1$

Remark: *M* can be chosen as the upper bound of the values of x, y; called big-M method in literature.

Still remember the complementary and slackness condition?

Primal problem

$$\min_x \; c^T x$$
$$\text{s.t.} \; Ax \geq b$$
$$x \geq 0$$

Dual problem

$$\max_\lambda \; b^T \lambda$$
$$A^T \lambda \leq c$$
$$\lambda \geq 0$$

Complementary and slackness: Suppose x^*, λ^* are the primal and dual optimal solutions, respectively. Then, we have

$$a_n^T x^* > b \quad \Rightarrow \quad \lambda_n^* = 0$$
$$\lambda_n^* > 0 \quad \Rightarrow \quad a_n^T x^* = b$$

$$0 \leq \lambda \perp (Ax - b) \geq 0$$

How to linearize this constraint?

Linearization techniques-7

Minimum values

Consider $y = \min\{x_1, \dots, x_n\}, x_i \in [x_i^l, x_i^u]$
Let $L = \min\{x_1^l, \dots, x_n^l\}$. It can be represented as

$$x_i^l \le x_i \le x_i^u, \forall i$$
$$y \le x_i, \forall i$$
$$x_i - (x_i^u - L)(1 - z_i) \le y, \forall i$$
$$z_i \in \{0,1\}, \sum_{i=1}^{n} z_i = 1$$

Proof of equivalence:

- Only one $z_i = 1$ and others =0.
- If $z_i = 1$, we have $x_i^l \le x_i \le x_i^u, y \le x_i, x_i \le y$
- If $z_i = 0$, we have $x_i^l \le x_i \le x_i^u, y \le x_i, x_i - y \le x_i^u - L$

For example, if $z_1 = 1$
$y = x_1$
$y \le x_2, \dots, y \le x_n$

Linearization techniques-8

Maximum values

Consider $y = \max\{x_1, \ldots, x_n\}, x_i \in \left[x_i^l, x_i^u\right]$
Let $U = \max\{x_1^u, \ldots, x_n^u\}$. It can be represented as

$$x_i^l \le x_i \le x_i^u, \forall i$$
$$y \ge x_i, \forall i$$
$$x_i + \left(U - x_i^l\right)(1 - z_i) \ge y, \forall i$$
$$z_i \in \{0,1\}, \sum_{i=1}^{n} z_i = 1$$

Proof of equivalence:

- Only one $z_i = 1$ and others =0.
- If $z_i = 1$, we have $x_i^l \le x_i \le x_i^u, y \ge x_i, x_i \ge y$
- If $z_i = 0$, we have $x_i^l \le x_i \le x_i^u, y \ge x_i, y - x_i \le U - x_i^l$

Linearization techniques-9

Absolute values

$$y = |x|, x \in \mathbb{R}, |x| \leq x^u$$

It can be represented as

$$\begin{aligned} 0 \leq y - x \leq 2x^u z, &\qquad x^u(1-z) \geq x \\ 0 \leq y + x \leq 2x^u(1-z), &\qquad -x^u z \leq x \\ -x^u \leq x \leq x^u, &\qquad z \in \{0,1\} \end{aligned}$$

$z = 0$

$$\begin{aligned} 0 \leq y - x \leq 0, &\qquad x^u \geq x \\ 0 \leq y + x \leq 2x^u, &\qquad 0 \leq x \\ -x^u \leq x \leq x^u & \end{aligned}$$

$z = 1$

$$\begin{aligned} 0 \leq y - x \leq 2x^u, &\qquad 0 \geq x \\ 0 \leq y + x \leq 0, &\qquad -x^u \leq x \\ -x^u \leq x \leq x^u & \end{aligned}$$

Thanks！