PRO 5970 Métodos de Otimização Não Linear

Basic concepts

Celma de Oliveira Ribeiro

2023

Departmento de Engenharia de Produção
Universidade de São Paulo

The Transportation Problem with Volume Discounts

Determine an optimal plan for shipping goods from *m* sources to *n* destinations, given supply and demand constraints.

- Assume shipping costs are linear on the volume

Shipping Costs, Supply, and Demand for Powerco

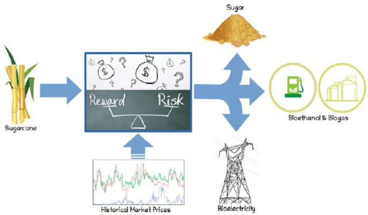
From	To: City 1	City 2	City 3	City 4	Supply (million kwh)
Plant 1	\$8	\$6	\$10	\$9	35
Plant 2	\$9	\$12	\$13	\$7	50
Plant 3	\$14	\$9	\$16	\$5	40
Demand (million kwh)	45	20	30	30	

- Assume the shipping costs are not be fixed. Volume discounts sometimes are available for large shipments

Entregar

Considere que a planta 1 oferece descontos de acordo com o volume. Há três faixas: $0 \leq x < 20$, $20 \leq x < 40$, $40 \leq x$ e assuma que a cada faixa ganha-se um desconto de 5%

Investment problem - Integrated sugar and ethanol plants



Part of the bagasse is used to cover internal needs in steam and electricity,

Investment problem - Integrated sugar and ethanol plants

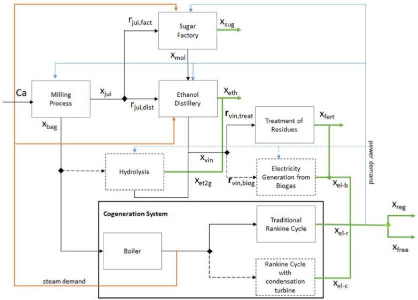


Fig. 1. Superstructure of the sugarcane mill process. Each Box represents a process in the sugarcane plant. A solid line indicates the flow of a resource or utility between two or more units. A technological routes subject to investment decision is shown in dotted lines.

Modelling the steam and power utility streams as part of the process superstructure would provide a way of describing these interdependencies, yet at the cost of increasing the model complexity significantly.

Investment model

Suppose one has the opportunity to invest in n assets. Their future returns are represented by random variables, $R_1, ..., R_n$, whose expected values and covariances are $\mu_i = E[Ri]$, $i = 1, ..., n$ and $\sigma_{ij} = Cov(Ri, Rj)$, $i, j = 1, ..., n$, respectively, estimated based on historical data. You want to find the portfolio of miinimum risk (risk is the variance of the portfolio)

Investment model

Suppose one has the opportunity to invest in n assets. Their future returns are represented by random variables, $R_1, ..., R_n$, whose expected values and covariances are $\mu_i = E[Ri]$, $i = 1, ..., n$ and $\sigma_{ij} = Cov(Ri, Rj)$, $i, j = 1, ..., n$, respectively, estimated based on historical data. You want to find the portfolio of miinimum risk (risk is the variance of the portfolio)

Some problems

	Peso	Volatilidade	Retorno Esperado
Bolsa	25,0%	21,09%	12,00%
Juros Longos	25,0%	11,35%	9,75%
Dólar	25,0%	15,68%	5,50%
Imobiliário	25,0%	6,46%	7,90%
Total	100,0%	6,2%	8,8%

Matriz de Correlação

	Bolsa	Juros Longos	Dólar	Imobiliário
Bolsa	1	0,3	-0,4	0,2
Juros Longos	0,3	1	-0,5	0,2
Dólar	-0,4	-0,5	1	-0,1
Imobiliário	0,2	0,2	-0,1	1

Some problems

Let x_i be the fraction of your wealth allocated to each asset

$$x_i \geq 0$$
$$\sum_{i=1}^{n} x_i = 1$$

- The return of the portfolio is a random variable $R(x) = \sum_{i=1}^{n} \mu_i x_i$
- The variance of the portfolio is:
  $\mathrm{var}(R(x)) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j = x^t \Sigma x$
  Σ is the covariance matrix

Usually the problem is modelled as

$$\min \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j$$

$$\begin{aligned} &\sum_{i=1}^{n} \mu_i x_i \geq R_0 \\ &\sum_{i=1}^{n} x_i = 1 \\ &x_i \geq 0 \qquad\qquad \mathrm{i} \in \{1, 2, \ldots n\} \end{aligned}$$

for a given R_0

Some problems

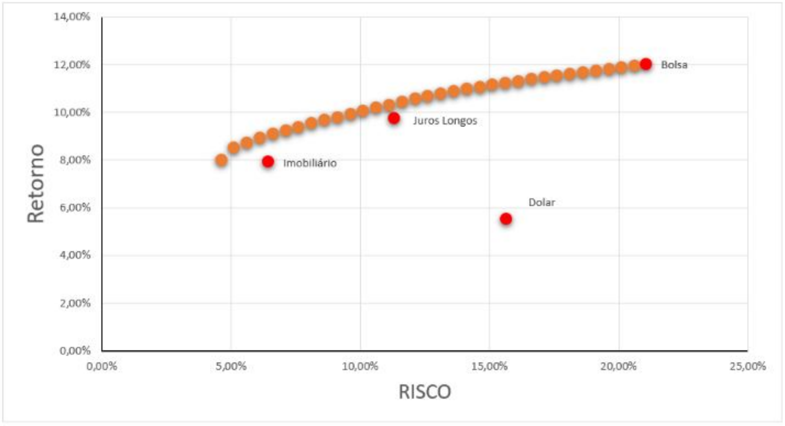


Figure 1: Solutions

Problem of interest

Objective

Minimize $f(x)$, $x \in \mathbb{S} \subset \mathbb{R}^n$

$\mathbb{S}$: feasible set

$f : \mathbb{R}^n \to \mathbb{R}$ - objective function

$$\min_{x \in \mathbb{S}} f(x)$$

Solution set

$$\arg \min_{x \in \mathbb{R}^n} \{f(x) | x \in \mathbb{S}\}$$

Feasibility

1. A point x is feasible for (P) if it satisfies all the constraints. For an unconstrained problem we will assume $x \in \mathbb{R}^n$

Feasibility

1. A point x is feasible for (P) if it satisfies all the constraints. For an unconstrained problem we will assume $x \in \mathbb{R}^n$
2. The set of all feasible points forms the feasible region, or feasible set. Let us denote it by $\mathcal{S}$.

Feasibility

1. A point x is feasible for (P) if it satisfies all the constraints. For an unconstrained problem we will assume $x \in \mathbb{R}^n$
2. The set of all feasible points forms the feasible region, or feasible set. Let us denote it by $\mathcal{S}$.
3. The goal of an optimization problem in minimization form, as above, is to find a feasible point x^* such that $f(x^*) \leq f(x)$ for any other feasible point x.

Types of problems

Three general forms of the feasible set

- Unconstrained
- Equality constrained
- Inequality constrained

Unconstrained Problem

General form

$$\min_{x \in \mathbb{S}} f(x)$$

$\mathbb{S}$ is an open set (usually, but not always, $\mathbb{S} = \mathbb{R}^n$).

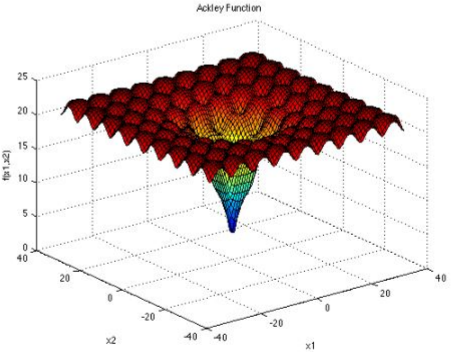
Example

$$\min_{x \in \mathbb{S}} |x|$$

$$\mathbb{S} = \mathbb{R}$$

.

Example 2



$$f(\mathbf{x}) = -a \exp\left(-b\sqrt{\frac{1}{d}\sum_{i=1}^{d} x_i^2}\right) - \exp\left(\frac{1}{d}\sum_{i=1}^{d}\cos(cx_i)\right) + a + \exp(1)$$

Example 3

$$min_{(\alpha_0,\alpha_1)} f(\alpha)$$

$$f(\alpha_0,\alpha_1) = \sum_{j=1}^{n} \left(Y_j - \alpha_0 - \alpha_1 X_j\right)^2$$

Example 4

$$f(\alpha) = \sum_{j=1}^{n} \left(Y_j - e^{(\alpha \times X_j)}\right)^2$$

Constrained problem

Given

$x \in \mathbb{R}^n$ - decision variables vector

$f : \mathbb{R}^n \to \mathbb{R}$ - objective function

g_i e h_i Constraints

$$
\begin{array}{lll}
 & \min \ f(x) & \\
\text{s.a} & g_i(x) \leq 0 & \text{i} \in \{1, 2, \ldots m\} \\
 & h_i(x) = 0 & \text{i} \in \{1, 2, \ldots l\}
\end{array}
$$

Find

$$\arg \min_{x \in \mathbb{R}^n} \{f(x) | x \in \mathbb{S}\} =$$

$$\arg \min_{x \in \mathbb{R}^n} \{f(x) | g_i(x) \leq 0 \quad h_i(x) = 0\}$$

Constrained problem

Examples

Example 1

$$\min_{x \in \mathbb{S}} |x|$$

$$\mathbb{S} = \{x \in \mathbb{R} | x \geq 7\}$$

.

Example 2

$$\begin{array}{llll} \max z = & 3x_1 & - x_2 & \\ \text{s.a} & -3x_1 & +3x_2 & \leq 6 \\ & -8x_1 & + 4x_2 & \leq 4 \\ & x_j \geq 0 & \forall j & \end{array}$$

- Plot the feasible set and the level curves of the objective function See: https://www.desmos.com/calculator?lang=pt-BR
- What happens if $f(x) = 3x_1 + x_2$

Example 3

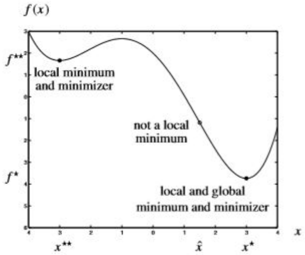
$$
\begin{array}{ll}
\min & (x_1 - 3)^2 + (x_2 - 2)^2 \\
\text{s.a} & x_1^2 - x_2 \leq 3 \\
& x_2 \leq 1 \\
& x_1 \geq 1
\end{array}
$$

Example 4

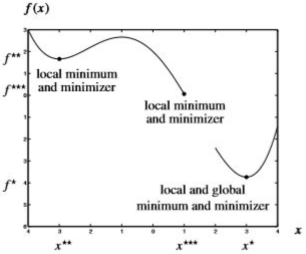
$$
\begin{array}{ll}
\min & (x_1 - 2)^2 + (x_2 - 1)^2 \\
\text{s.a} & x_1^2 - x_2 \leq 0 \\
& x_1 + x_2 \leq 2
\end{array}
$$

Exercice: Plot the feasible set and the level curves for both problems

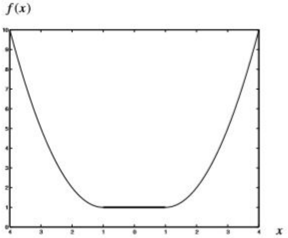
Types of problems



Types of problems



Multiple global minimizers

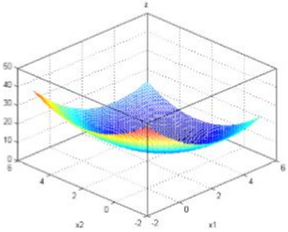


Types of problems

Example

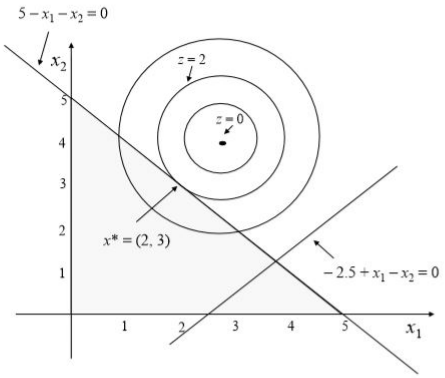
$$
\begin{array}{ll}
 & \min\ (x_1 - 3)^2 + (x_2 - 4)^2 \\
\text{s.a} & 5 - x_1 - x_2 \geq 0 \\
 & -2,5 + x_1 - x_2 \leq 0 \\
 & x_1 \geq 0 \ x_2 \geq 0
\end{array}
$$

Objective function



- Plot the feasible set
- Plot the level curves of the objective function

Objective function

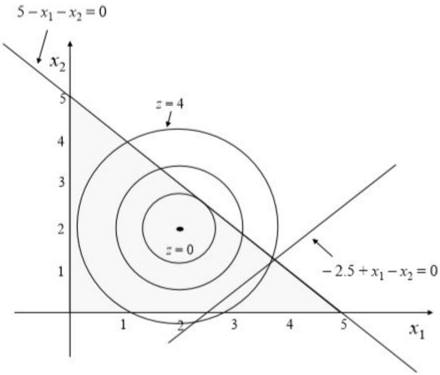


Types of problems

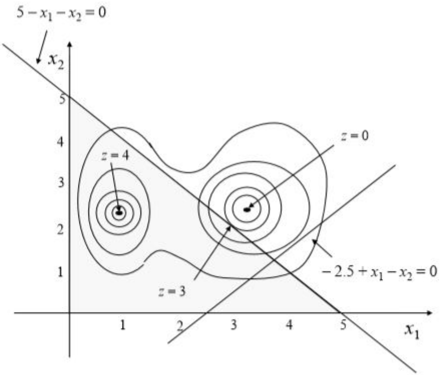
What if...

$$\begin{array}{ll} & \min\ (x_1-2)^2+(x_2-2)^2 \\ \text{s.a} & 5-x_1-x_2\geq 0 \\ & -2,5+x_1-x_2\leq 0 \\ & x_1\geq 0\ x_2\geq 0 \end{array}$$

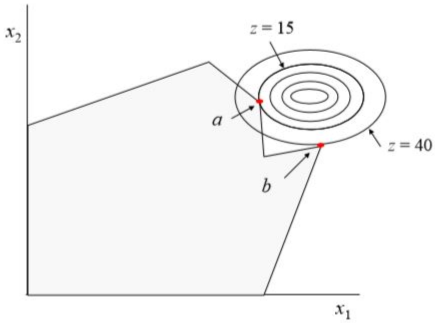
Objective function



Objective function



Constraints can affect the solution



Non linear models are much more difficult to solve

- It is hard to distinguish between local and global optimum
- Optimal are not restricted to extreme points
- Different starting points may lead to different final solutions
- It may be difficult to find a feasible starting point
- It is difficult to satisfy equality constraints (and to keep them satisfied)
- The use of solvers is far from a simple task

Some good news

- Relatively few algorithms implemented
- Solving non linear programs is difficult but not impossible.
- Looks for a simpler formulation
- Provide a good starting point
- Put resonable bounds on all variables

Identifying a solution

Global minimum

Vector x^* is a *global minimizer* if

$$f(x^*) \leq f(x) \quad \forall x \in \mathbb{S}$$

Local minimizer

A vector x^* is a *local minimizer* if there is a neighborhood of V of x^*, such that

$$f(x^*) \leq f(x) \quad \forall x \in V \cap \mathbb{S}$$

Optimization problems

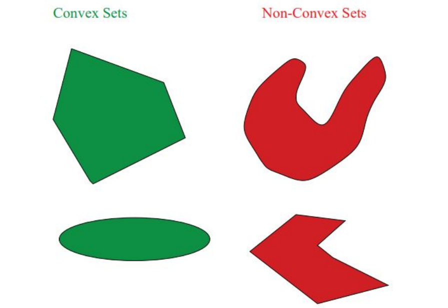
- according to the type of variable (continuous or discrete)
- according to constraints (constrained or unconstrained)
- optimal local versus global
- uncertainty in parameters (deterministic or stochastic)
- differentiability and convexity

Convex sets

A set $S \subset \mathbb{R}^n$ is *convex* if

$$\forall x, y \in S, \forall \lambda \in [0,1], \lambda x + (1-\lambda) y \in S$$

A set is convex if, given any two points in the set, the line segment connecting them lies entirely inside the set.



Analyze the following sets

1. $\mathbb{P} = \{x \in \mathbb{R} | -4 \leq x \leq 1\} \cup \{x \in \mathbb{R} | 2 \leq x \leq 4\}$
2. $\mathbb{W} = \left\{ x \in \mathbb{R} | \begin{array}{c} -4 \leq x \leq 1 \\ -1 \leq x \leq 4 \end{array} \right\}$
3. $\mathbb{M} = \{x \in \mathbb{R}^n | Ax \leq b\}$

$$\mathbb{M} = \{x \in \mathbb{R}^n | Ax \leq b\}$$

Let $w \in \mathbb{M}$, $y \in \mathbb{M}$ and $\lambda \in [0, 1]$. Then

- $Aw \leq b \Rightarrow \lambda A(w) \leq \lambda b$

$$\mathbb{M} = \{x \in \mathbb{R}^n | Ax \leq b\}$$

Let $w \in \mathbb{M}$, $y \in \mathbb{M}$ and $\lambda \in [0, 1]$. Then

- $Aw \leq b \Rightarrow \lambda A(w) \leq \lambda b$
- $Ay \leq b \Rightarrow (1 - \lambda) A(y) \leq (1 - \lambda) b$

$$\mathbb{M} = \{x \in \mathbb{R}^n | Ax \leq b\}$$

Let $w \in \mathbb{M}$, $y \in \mathbb{M}$ and $\lambda \in [0, 1]$. Then

- $Aw \leq b \Rightarrow \lambda A(w) \leq \lambda b$
- $Ay \leq b \Rightarrow (1-\lambda) A(y) \leq (1-\lambda) b$

Ento:

- $A(\lambda w + (1-\lambda) y) = \lambda Aw + (1-\lambda) A(y) \leq$

$$\mathbb{M} = \{x \in \mathbb{R}^n | Ax \leq b\}$$

Let $w \in \mathbb{M}$, $y \in \mathbb{M}$ and $\lambda \in [0,1]$. Then

- $Aw \leq b \Rightarrow \lambda A(w) \leq \lambda b$
- $Ay \leq b \Rightarrow (1-\lambda) A(y) \leq (1-\lambda) b$

Ento:

- $A(\lambda w + (1-\lambda) y) = \lambda Aw + (1-\lambda) A(y) \leq$
  $\leq \lambda b + (1-\lambda) b = b$

$$\mathbb{M} = \{x \in \mathbb{R}^n | Ax \leq b\}$$

Let $w \in \mathbb{M}$, $y \in \mathbb{M}$ and $\lambda \in [0, 1]$. Then

- $Aw \leq b \Rightarrow \lambda A(w) \leq \lambda b$
- $Ay \leq b \Rightarrow (1 - \lambda) A(y) \leq (1 - \lambda) b$

Ento:

- $A(\lambda w + (1 - \lambda) y) = \lambda Aw + (1 - \lambda) A(y) \leq$
  $\leq \lambda b + (1 - \lambda) b = b$

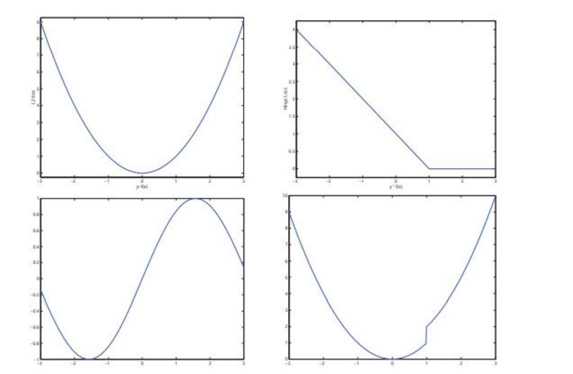
Convex functions

Let $S \subset \mathbb{R}^n$ be convex. A function $f : S \to \mathbb{R}$ is convex if

1. $$\forall x, y \in S, \forall \lambda \in [0,1],$$

   $$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$$

2. The line segment connecting two points $f(x_1)$ and $f(x_2)$ lies entirely on or above the function f .
3. The set of points lying on or above the function f is convex.



The top two figures are convex functions. The first function is strictly convex. Bottom figures are nonconvex functions.

Analyse if the following functions are convex.

1. $f(x) =< c, x > x \in \mathbb{R}^n \; c \in \mathbb{R}^n$
2. $f(X) = max\{X - X_0, 0\}$
3. $f(x) = \begin{cases} (x+5)^2 & \text{se } \; x \leq 0 \\ (x-5)^2 & \text{se } \; x > 0 \end{cases}$

   Note that $f(x) = \min\left\{(x+5)^2, (x-5)^2\right\}$
4. $f(x) = \max\left\{(x+5)^2, (x-5)^2\right\}$

Exercices

Let $f_l : \mathbb{R}^n \to \mathbb{R}, \quad l = 1, 2, \ldots, r$ e $f : \mathbb{R}^n \to \mathbb{R}$ given as:
$f(x) = \max_{l=1,2,\ldots,r} \{f_l(x)\}$

Verify: If f_l is convex $\forall l$ then f is convex

Theorem

Let S $\subset \mathbb{R}^n$ be convex and $f : S \to \mathbb{R}$ convex. Then

i. If f is convex in S, then there is at most one local minimum in S [1]

ii. If f is convex to S and has a local minimum in S, then the local minimum is also a global minimum.

iii. If f is strictly convex in S then it has at most one minimizer in S

If f is convex you just need to find a local minimum

Algorithms fall into three families:

- **Heuristics methods**
  - normally quick to execute but do not provide guarantees of optimality.
  - Include ant colony, particle swarm, and evolutionary algorithms
  - Some heuristics are stochastic in nature and have proof of convergence to an optimal solution (e.g Simulated annealing and multiple random starts).
  - No guarantee on the running time to reach optimality and there is no way to identify when one has reached an optimum point.
- **Approximate methods**
  - efficient algorithms that find approximate solutions to optimization problems
  - can provide a guarantee of the solution being at most ϵ away from the optimal solution.
- **Exact methods**
  - method of choice to solve an optimization problem to optimality.
  - The computational effort grows (at least) polynomially with the problem size

Important

1. Algorithms are usually iterative. In general you can only assure convergence to a **local** minimum

Important

1. Algorithms are usually iterative. In general you can only assure convergence to a **local** minimum
2. Iterative algorithms obtain points with decreasing values of the objective function at each step (or closer to satisfying constraints).

Important

1. Algorithms are usually iterative. In general you can only assure convergence to a **local** minimum
2. Iterative algorithms obtain points with decreasing values of the objective function at each step (or closer to satisfying constraints).
3. The choice of search directions, in general, aims to:
   a move from the current solution in a direction that decreases the objective function and maintains feasibility
   b move from the current solution towards the optimal (minimizer)

Definition

Let $\mathbb{S} \subset \mathbb{R}^n$ a convex set and $f : \mathbb{R}^n \to \mathbb{R}$ convex on $\mathbb{S}$, then the problem $min_{x \in \mathbb{S}} f(x)$ is called a convex problem or convex optimization problem.

Examples

Linear programming is always a convex problem

$$\begin{aligned} & \min\ c^t x \\ \text{s.t} \quad & Ax = b \\ & Dx \leq d \\ & x \geq 0 \end{aligned}$$

Quadratic programming is a convex problem iff the matrix Q is positive semidefinite

$$\begin{aligned} & \min\ x^t Q x + c^t x \\ \text{s.t} \quad & Ax = b \\ & Dx \leq d \\ & x \geq 0 \end{aligned}$$