PRO 5961 Métodos de Otimização Não Linear

Celma de Oliveira Ribeiro

Aula 7 - 2023

Departamento de Engenharia de Produção
Universidade de São Paulo

The problem

Given

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

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

g_i e h_i constraints

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

$\boxed{\text{Constrained optimization is a rather more difficult subject !!}}$

Constrained problems

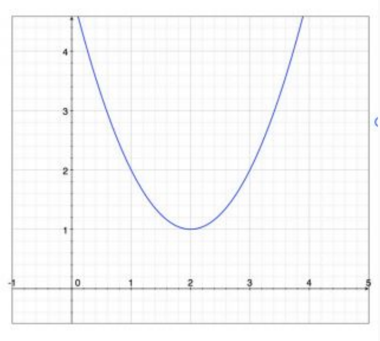
Usually we write

$$g(x) = \begin{bmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_m(x) \end{bmatrix} \text{ and } h(x) = \begin{bmatrix} h_1(x) \\ h_2(x) \\ \vdots \\ h_l(x) \end{bmatrix}$$

Constrained problems

Example: Univariate Constrained Optimization

$$\min_{x \in \mathbb{R}} (x-2)^2 - 1$$

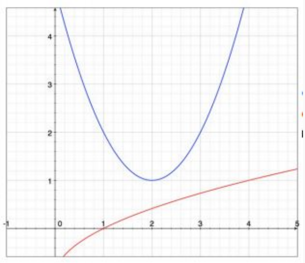


Constrained problems

Example: Univariate Constrained Optimization

$$\min_{x\in\mathbb{R}} (x-2)^2 - 1$$

$$s.t \quad \begin{array}{c} \sqrt{x} \leq 1 \\ x \geq 0 \end{array}$$



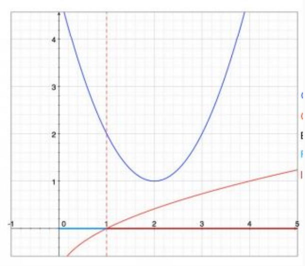
Note: $\sqrt{x} \leq 1 \iff \sqrt{x} - 1 \leq 0 \quad x \geq 0 \iff -x \leq 0$

Constrained problems

Example: Univariate Constrained Optimization

$$\min_{x \in \mathbb{R}} (x-2)^2 - 1$$

$$s.t \quad \begin{array}{c} \sqrt{x} \leq 1 \\ x \geq 0 \end{array}$$



Feasible set : $0 \leq x \leq 1$

Constrained problems

Some problems can be easily solved:

Example 1 - linear programming

$$
\begin{array}{llllllll}
\text{min} & 5x_1 & +2x_2 & +3x_3 & -x_4 & +x_5 & \\
\text{s.a} & x_1 & +2x_2 & +2x_3 & & & \leq 8 \\
 & 3x_1 & +4x_2 & +x_3 & & +x_5 & =7 \\
 & x_j \geq 0 & \forall j & & & &
\end{array}
$$

$$
\begin{array}{ll}
min & c^t x \\
s.a & Ax = b \\
 & Dx \leq d \\
 & Mx \geq f \\
 & x \geq 0
\end{array}
$$

Constrained problems

Some problems can be easily solved:

Example 2 - Quadratic programming

$$\begin{array}{lllllll}
min & x_1^2 + 9x_2^2 - 3x_3^2 & & & & \\
s.a & x_1 & +2x_2 & +2x_3 & & \leq 8 \\
 & 3x_1 & +4x_2 & +x_3 & +x_5 & = 7 \\
 & x_j \geq 0 & \forall j & & &
\end{array}$$

$$\begin{array}{ll}
min & \frac{1}{2}x^t Q x + c^t x \\
s.a & Ax = b_e \\
 & Dx \leq b_i
\end{array}$$

Some issues

- Constrained optimization is rather more difficult than unconstrained optimization

Constrained problems

Some issues

- Constrained optimization is rather more difficult than unconstrained optimization
- Algorithms are generally more complicated

Constrained problems

Some issues

- Constrained optimization is rather more difficult than unconstrained optimization
- Algorithms are generally more complicated
- A possible approach is to **transform** the original problem in order to use unconstrained optimization techniques.

Constrained problems

Some issues

- Constrained optimization is rather more difficult than unconstrained optimization
- Algorithms are generally more complicated
- A possible approach is to **transform** the original problem in order to use unconstrained optimization techniques.

$\boxed{\text{How to transform the problems?}}$

Scaling Variables

Scaling by variable transformation

- Cannot be described precisely in general terms
- Converts the variables from units that typically reflect the physical nature of the problem to units that display certain desirable properties during the minimization process.
- The variables of the scaled problem should be of similar magnitude and of order unity in the region of interest.
- If typical values of all the variables are known, a problem can be transformed so that the variables are all of the same order of magnitude

Example

Var	Interpretation	Units	Typical value
x_1	Gas flow	lb/hr	11000
x_2	Water flow	lb/hr	1675
x_3	Sterm thermal resistance	$(BTU/(hrft^2\ {}^0F))^{-1}$	100
x_4	Waste build-up	$(BTU/(hrft^2\ {}^0F))^{-1}$	6×10^{-4}
x_5	Gas-side radiation	$BTU/(hrft^2\ {}^0R^4)$	5.4×10^{-10}

Use linear transformations of the variables $x = Dy$ with x_i the original variables, y_i the transformed variables, and D is a constant diagonal matrix.

For instance, d_1 could be set to 1.1×10^4.

Scaling variables

advantages and disadvantages

Consider linear scaling:

- Some accuracy may be lost.
  Suppose $x_i \in [200.1242, 200.1806]$ and $y_i = x_i/200.1242$. Then
  $y_i \in [1.0, 1.000282]$ (suppose seven digit representation)

Scaling variables

advantages and disadvantages

Consider linear scaling:

- Some accuracy may be lost.
  Suppose $x_i \in [200.1242, 200.1806]$ and $y_i = x_i/200.1242$. Then $y_i \in [1.0, 1.000282]$ (suppose seven digit representation)
- The magnitude of a variable may vary substantially during the minimization. What might be a good scaling at one point may prove harmful at another.

advantages and disadvantages

Consider linear scaling:

- Some accuracy may be lost.
  Suppose $x_i \in [200.1242, 200.1806]$ and $y_i = x_i/200.1242$. Then $y_i \in [1.0, 1.000282]$ (suppose seven digit representation)
- The magnitude of a variable may vary substantially during the minimization. What might be a good scaling at one point may prove harmful at another.
- if a realistic range of values for a variable is known, try to use the information
  Example: if $x_i \in [a_i, b_i]$, consider $y_i = \frac{2x_i}{b_i - a_i} - \frac{b_i + a_i}{b_i - a_i}$ This transformation guarantees that $y_i \in [-1, 1]$

Be careful....

- The interval specifying the range of values for a given variable must be a realistic one.

Scaling variables

Be careful....

- The interval specifying the range of values for a given variable must be a realistic one.
- When the variables are scaled by a linear transformations, the derivatives of the objective function are also scaled.

Scaling variables

Be careful....

- The interval specifying the range of values for a given variable must be a realistic one.
- When the variables are scaled by a linear transformations, the derivatives of the objective function are also scaled.
- Even a mild scaling such as $x_j = 10y_j$ may have a substantial effect on the Hessian, and significantly alter the convergence rate of an optimization algorithm.

Transformations of problems - Constraints

1. **Scaling**

Pre-conditioning

- Appropriate for linear constraints

1. **Scaling**

Pre-conditioning

- Appropriate for linear constraints
- Generalize the idea of scaling of variables

1. **Scaling**

Pre-conditioning

- Appropriate for linear constraints
- Generalize the idea of scaling of variables
- Multiply the coeficient matrix and the right hand side vector by a suitably matrix M

1. **Scaling**

Pre-conditioning

- Appropriate for linear constraints
- Generalize the idea of scaling of variables
- Multiply the coeficient matrix and the right hand side vector by a suitably matrix M
- Matrix M
  - does not change the set of feasible points
  - makes it easier to find feasible points

2. Slack Variables

Let $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^l$. Consider the problems:

P1 $\min_{x\in\mathbb{R}^n} \{f(x) \,|\, g(x) \leq 0 \quad h(x) = 0\}$

P2 $\min_{x\in\mathbb{R}^n, w\in\mathbb{R}^m} \{f(x) \,|\, g(x) + w = 0 \quad h(x) = 0 \quad w \geq 0\}$

Then we have:

2. Slack Variables

Let $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^l$. Consider the problems:

P1 $\min_{x \in \mathbb{R}^n} \{f(x) \,|\, g(x) \leq 0 \quad h(x) = 0\}$

P2 $\min_{x \in \mathbb{R}^n, w \in \mathbb{R}^m} \{f(x) \,|\, g(x) + w = 0 \quad h(x) = 0 \quad w \geq 0\}$

Then we have:

(i) Problem (P1) has a minimum if and only if Problem (P2) has a minimum

2. Slack Variables

Let $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^l$. Consider the problems:

P1 $\min_{x \in \mathbb{R}^n} \{ f(x) \,|\, g(x) \leq 0 \quad h(x) = 0 \}$

P2 $\min_{x \in \mathbb{R}^n, w \in \mathbb{R}^m} \{ f(x) \,|\, g(x) + w = 0 \quad h(x) = 0 \quad w \geq 0 \}$

Then we have:

(i) Problem (P1) has a minimum if and only if Problem (P2) has a minimum

(ii) If one of the problems has a minimum,then the minima are equal

2. Slack Variables

Let $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^l$. Consider the problems:

P1 $\min_{x \in \mathbb{R}^n} \{ f(x) \,|\, g(x) \le 0 \quad h(x) = 0 \}$

P2 $\min_{x \in \mathbb{R}^n, w \in \mathbb{R}^m} \{ f(x) \,|\, g(x) + w = 0 \quad h(x) = 0 \quad w \ge 0 \}$

Then we have:

(i) Problem (P1) has a minimum if and only if Problem (P2) has a minimum

(ii) If one of the problems has a minimum,then the minima are equal

(iii) To each minimizer x^* of (P1) there corresponds a minimizex^*r $\begin{bmatrix} x^* \\ w^* \end{bmatrix}$ of (P2) and vice-versa

Entrega aula - **Slack Variables**

Coloque os problemas abaixo na forma

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & g(x) = 0 \end{array}$$

a)

$$\begin{array}{llllll} \max & 3x_1 & + 2x_2 & + 7x_3 & \\ \text{s.a} & 2x_1 & + 3x_2 & & \leq 42 \\ & 2x_1 & & - x_3 & \leq 18 \\ & 3x_1 & - x_2 & + 4\,x_3 & \geq 24 \\ & x_1 \geq 0 & x_2 \leq 0 & x_3 \in \mathbb{R} & \end{array}$$

Entrega aula - Slack Variables

b)

$$
\begin{array}{llllll}
\min & x_1 \times x_2 \times x_3 & & & \\
\text{s.a} & 2x_1 & +\ 3x_2 & & \leq 42 \\
 & x_1^4 & +x_2^4 & & = x_3 \\
 & 3x_1 & \text{-}\ x_2 & +\ 4\ x_3 & \geq 24
\end{array}
$$

3. Changing the functional form

Consider $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$,$b \in \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^l$, $d \in \mathbb{R}^l$. Let $\omega : \mathbb{R} \to \mathbb{R}$ and $\lambda : \mathbb{R} \to \mathbb{R}$ be strictly monotonically increasing and continuous on $\mathbb{R}^1$. Consider the problems:

P1 $\min_{x \in \mathbb{R}^n} \{f(x) \,|g(x) \leq b \quad h(x) = d\,\}$

P2 $\min_{x \in \mathbb{R}^n} \{f(x) \,|\Lambda(x) \leq \beta \quad \Omega(x) = \theta\,\}$ with

$$\Lambda_i(x) = \lambda(g_i(x)) \qquad \beta_i = \lambda(b_i)$$
$$\Omega_i(x) = \omega(g_i(x)) \qquad \theta_i = \omega(b_i)$$

Then we have:

3. Changing the functional form

Consider $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$,$b \in \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^l$, $d \in \mathbb{R}^l$. Let $\omega : \mathbb{R} \to \mathbb{R}$ and $\lambda : \mathbb{R} \to \mathbb{R}$ be strictly monotonically increasing and continuous on $\mathbb{R}^1$. Consider the problems:

P1 $\min_{x \in \mathbb{R}^n} \{f(x) \,|g(x) \leq b \quad h(x) = d\}$

P2 $\min_{x \in \mathbb{R}^n} \{f(x) \,|\Lambda(x) \leq \beta \quad \Omega(x) = \theta\}$ with

$$\Lambda_i(x) = \lambda(g_i(x)) \qquad \beta_i = \lambda(b_i)$$
$$\Omega_i(x) = \omega(g_i(x)) \qquad \theta_i = \omega(b_i)$$

Then we have:

(i) Problem (P1) has a minimum if and only if Problem (P2) has a minimum

3. Changing the functional form

Consider $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$,$b \in \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^l$, $d \in \mathbb{R}^l$. Let $\omega : \mathbb{R} \to \mathbb{R}$ and $\lambda : \mathbb{R} \to \mathbb{R}$ be strictly monotonically increasing and continuous on $\mathbb{R}^1$. Consider the problems:

P1 $\min_{x \in \mathbb{R}^n} \{f(x) \,|g(x) \leq b \quad h(x) = d\}$

P2 $\min_{x \in \mathbb{R}^n} \{f(x) \,|\Lambda(x) \leq \beta \quad \Omega(x) = \theta\}$ with

$$\begin{array}{ll} \Lambda_i(x) = \lambda(g_i(x)) & \beta_i = \lambda(b_i) \\ \Omega_i(x) = \omega(g_i(x)) & \theta_i = \omega(b_i) \end{array}$$

Then we have:

(i) Problem (P1) has a minimum if and only if Problem (P2) has a minimum

(ii) If one of the problems has a minimum,then the minima are equal and they have the same minimizers

Example

Consider problem P1 with:

$f : \mathbb{R}^3 \to \mathbb{R},\ f(x_1, x_2, x_3) = 2a^{x_1} b^{x_2}$

$g : \mathbb{R}^3 \to \mathbb{R}^2\ g(x_1, x_2, x_3) = \begin{bmatrix} x_1 + x_2 + x_3 \\ x_1 \times x_2 \end{bmatrix}$

$b = \begin{bmatrix} 13 \\ 1 \end{bmatrix}$

Example

Consider problem P1 with:

$f : \mathbb{R}^3 \to \mathbb{R},\ f(x_1, x_2, x_3) = 2a^{x_1} b^{x_2}$

$g : \mathbb{R}^3 \to \mathbb{R}^2\ g(x_1, x_2, x_3) = \begin{bmatrix} x_1 + x_2 + x_3 \\ x_1 \times x_2 \end{bmatrix}$

$b = \begin{bmatrix} 13 \\ 1 \end{bmatrix}$

Possible solution: consider the logarithm in the objective function

4. Altering the feasible region

Transformations of problems - Constraints

If possible, the objective function shall be non-linear and the constraints linear.

Consider

$$
\begin{array}{llll}
\min & x_1 & + 2x_2 & \\
\text{s.a} & x_1 \times x_2 & & = 2 \\
 & x_1 & & \geq 1 \\
 & x_2 & & \geq 1
\end{array}
$$

4. Altering the feasible region

Transformations of problems - Constraints

If possible, the objective function shall be non-linear and the constraints linear.

Consider

$$
\begin{array}{lll}
\min & x_1 + 2x_2 & \\
\text{s.a} & x_1 \times x_2 & = 2 \\
 & x_1 & \geq 1 \\
 & x_2 & \geq 1
\end{array}
$$

Let $x_2 = 2/x_1$, you obtain:

$$
\begin{array}{lll}
\min & x_1 + 4/x_1 & \\
\text{s.a} & x_1 & \geq 1 \\
 & x_1 & \leq 2
\end{array}
$$

4. Altering the feasible region

Transformations of problems - Constraints

Let $\mathbb{S}_1 \subseteq \mathbb{S} \subseteq \mathbb{S}_2 \subseteq \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ Consider the problems:

P1 $\min_{x \in \mathbb{S}_1} f(x)$

P2 $\min_{x \in \mathbb{S}} f(x)$

P3 $\min_{x \in \mathbb{S}_2} f(x)$

and assume that they all have minima and minimizers. Then:

4. Altering the feasible region

Transformations of problems - Constraints

Let $\mathbb{S}_1 \subseteq \mathbb{S} \subseteq \mathbb{S}_2 \subseteq \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ Consider the problems:

P1 $\min_{x \in \mathbb{S}_1} f(x)$

P2 $\min_{x \in \mathbb{S}} f(x)$

P3 $\min_{x \in \mathbb{S}_2} f(x)$

and assume that they all have minima and minimizers. Then:

(i) $\min_{x \in \mathbb{S}_1} f(x) \geq \min_{x \in \mathbb{S}} f(x) \geq \min_{x \in \mathbb{S}_2} f(x)$

4. Altering the feasible region

Transformations of problems - Constraints

Let $\mathbb{S}_1 \subseteq \mathbb{S} \subseteq \mathbb{S}_2 \subseteq \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ Consider the problems:

P1 $\min_{x \in \mathbb{S}_1} f(x)$

P2 $\min_{x \in \mathbb{S}} f(x)$

P3 $\min_{x \in \mathbb{S}_2} f(x)$

and assume that they all have minima and minimizers. Then:

(i) $\min_{x \in \mathbb{S}_1} f(x) \geq \min_{x \in \mathbb{S}} f(x) \geq \min_{x \in \mathbb{S}_2} f(x)$

(ii) If $x^* \in \arg\min_{x \in \mathbb{S}_2} f(x)$ and $x^* \in \mathbb{S}$, then

- $\min_{x \in \mathbb{S}} f(x) = \min_{x \in \mathbb{S}_2} f(x)$
- $\arg\min_{x \in \mathbb{S}} f(x) = \left(\arg\min_{x \in \mathbb{S}_2} f(x)\right) \cap \mathbb{S}$

Relaxations

$\boxed{\text{How to solve a problem?}}$

$\boxed{\text{Consider a problem with some objective but a larger feasible set.}}$

If the solution of the *relaxed* problem lies in $\mathbb{S}$ a solution has been found!

Transformations of problems - Constraints

Relaxations

How to solve a problem?

Consider a problem with some objective but a larger feasible set.

If the solution of the *relaxed* problem lies in $\mathbb{S}$ a solution has been found!

Some nice cases:

- $\mathbb{S}_2$ is convex while $\mathbb{S}$ is not
- $\mathbb{S}_2$ temporarily ignores some constraints

Relaxations

$\boxed{\text{How to solve a problem?}}$

Alternative 1

Consider a problem with same objective but a larger feasible set.

If the solution of the *relaxed* problem lies in $\mathbb{S}$ a solution has been found! Some nice cases:

- $\mathbb{S}_2$ is convex while $\mathbb{S}$ is not
- $\mathbb{S}_2$ temporarily ignores some constraints

Relaxations

$\boxed{\text{How to solve a problem?}}$

Alternative 1

Consider a problem with same objective but a larger feasible set.

If the solution of the *relaxed* problem lies in $\mathbb{S}$ a solution has been found! Some nice cases:

- $\mathbb{S}_2$ is convex while $\mathbb{S}$ is not
- $\mathbb{S}_2$ temporarily ignores some constraints

Alternative 2

If the optimizer is known to lie in a subset of $\mathbb{S}$, confine the search to this subset

1. **Adding terms**

Main idea adding terms that depend on the constraints:

- do not consider the constraints explicitly
- make the constraints easier to deal with

1. Adding terms

Main idea adding terms that depend on the constraints:

- do not consider the constraints explicitly
- make the constraints easier to deal with

Alternatives

- Penalty functions: the function is large for values of the decision variables that violate the constraints

Transformations of problems - Objective function

1. **Adding terms**

Main idea adding terms that depend on the constraints:

- do not consider the constraints explicitly
- make the constraints easier to deal with

Alternatives

- Penalty functions: the function is large for values of the decision variables that violate the constraints
- Barrier functions: build a barrier to violating constraints. These methods are generally applicable only to inequality constrained optimization problems