Continuous optimization ENT305A

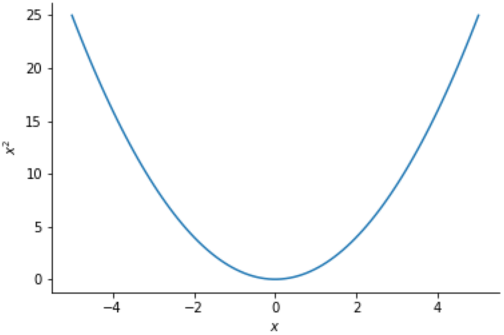
Elise Grosjean

Ensta-Paris
Institut Polytechnique de Paris
September 2024

What is an optimization problem?

$$f : x \to x^2, x \in [-5, 5]$$

$\mathcal{D} = K = \mathbb{R}$



What is an optimization problem?

Notation.

Let $\bar{B}(\bar{x}, \varepsilon)$ denote the closed ball of center $\bar{x}$ and radius ε.

Definition.

A feasible point $\bar{x}$ is a local solution to (P) if and only if there exists $\varepsilon > 0$ such that $\bar{x}$ is a **global** solution to the following **localized** problem:

$$\inf_{x \in \mathbb{R}^n} f(x), \quad x \in K \cap \bar{B}(\bar{x}, \varepsilon).$$

What is an optimization problem?

Constraints.
Most of the time, the feasible set K is described by

$$K = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} h_i(x) = 0, & \forall i \in \mathcal{E} \\ g_j(x) \leq 0, & \forall j \in \mathcal{I} \end{array} \right\},$$

where $h\colon \mathbb{R}^n \to \mathbb{R}^{m_1}$, $g : \mathbb{R}^n \to \mathbb{R}^{m_2}$.

We call the expressions

- $h_i(x) = 0$: **equality constraint**
- $g_j(x) \leq 0$: **inequality constraint**.

Existence of a solution

Theorem 1 (existence of extreme value (Weierstrass))

Assume the following:

- *?*
- *?*

Then the optimization problem (P) *has (at least) one solution.*

Existence of a solution

Theorem 2 (existence of extreme value (Weierstrass))

Assume the following:

- *K is* **non-empty** *and* **compact** *(i.e. closed and bounded)*
- *f is* **continuous** *on K.*

Then the optimization problem (P) has (at least) one solution.

Remarks. If $K = \{x \in \mathbb{R}^n \mid h_i(x) = 0, \forall i \in \mathcal{E}, g_j(x) \leq 0, \forall j \in \mathcal{I}\}$, where h_i, g_j are continuous, then K is closed. In practical exercises, it is not necessary to justify the continuity of h_i or g_j.

Optimality conditions

Let us fix a continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ for the whole section. Let us consider

$$\inf_{x \in \mathbb{R}^n} f(x) \qquad (P)$$

The function f is said to be **stationary** at $x \in \mathbb{R}^n$ if $\nabla f(x) = 0$.

Theorem 3 (Necessary optimality condition)

Let $\bar{x} \in \mathbb{R}^n$ be a **local solution** *of (P). Then, f is* **stationary** *at $\bar{x}$.*

Remark. Stationarity is only a necessary condition!

Optimality conditions

Theorem 4

Assume that f *is twice continuously differentiable. Let* $\bar{x}$ *be a* **stationary** *point.*

- **Necessary** *condition.*
  If $\bar{x}$ *is a* **local solution** *of* (P), *then* $D^2 f(\bar{x})$ *is* **positive semi-definite**, *that is to say,*

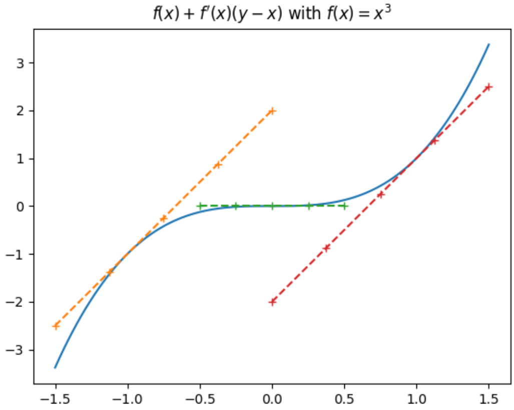
$$\langle h, D^2 f(\bar{x}) h \rangle \geq 0, \quad \text{for all } h \in \mathbb{R}^n.$$

- **Sufficient** *condition.*
  If $D^2 f(\bar{x})$ *is* **positive definite**, *that is to say if*

$$\langle h, D^2 f(\bar{x}) h \rangle > 0, \quad \text{for all } h \in \mathbb{R}^n \setminus \{0\} \ ,$$

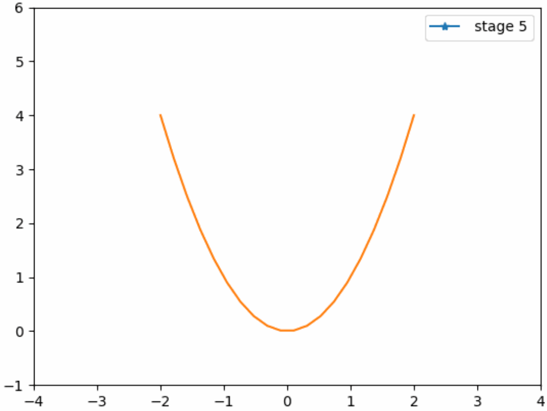
then $\bar{x}$ *is a* **local solution** *of* (P).

illustration



illustration

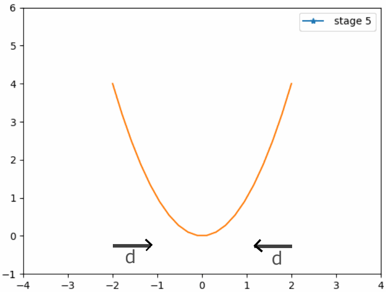
Descent gradient with $\alpha = 0.25$



illustration

Descent gradient

$x_{k+1} = x_k + \alpha d$



Taylor:

$$f(x + td) \simeq f(x) + td \ f'(x),$$

$$\lim_{t \to 0} \frac{f(x+td) - f(x)}{t} = d \ f'(x)$$

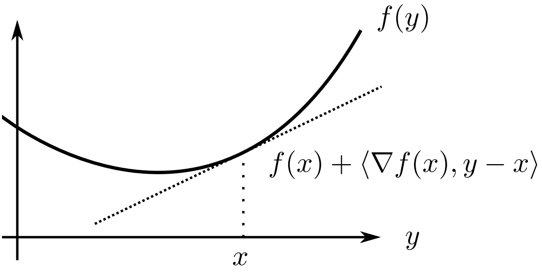
Optimality conditions

Theorem 5

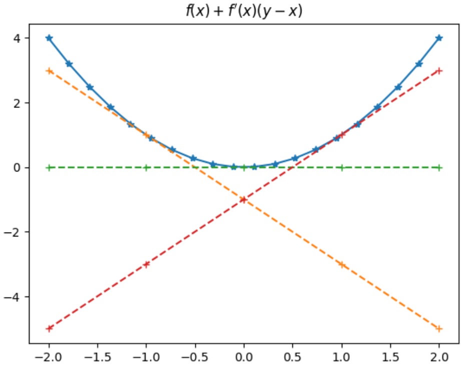
- *The function f is convex if and only if*

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle,$$

for all x and $y \in \mathbb{R}^n$.



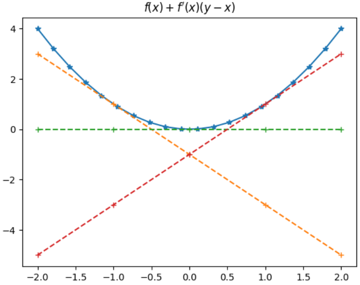
Optimality conditions



Optimality conditions

Theorem 6

- *The function f is convex if and only if f is twice differentiable, and $D^2 f(x)$ is symmetric* **positive semi-definite** *for all $x \in \mathbb{R}^n$.*



Optimality conditions

Theorem 7

Assume that f is **convex**. *Let $\bar{x}$ be a* **stationary** *point of f. Then it is a* **global solution** *of (P).*

Proof. For all $x \in \mathbb{R}^n$, we have

$$f(x) \geq f(\bar{x}) + \langle \nabla f(\bar{x}), x - \bar{x} \rangle = f(\bar{x}).$$

Linear equality constraints

- We investigate in this section the problem

$$\inf_{x \in \mathbb{R}^n} f(x), \quad \text{s.t.} \begin{cases} h_i(x) = 0, & \forall i \in \mathcal{E} \\ g_j(x) \leq 0, & \forall j \in \mathcal{I}. \end{cases} \qquad (P)$$

- Let $x \in \mathbb{R}^n$ be feasible. Let $j \in \mathcal{I}$. We say that
  - the inequality constraint j is **active** if $g_j(x) = 0$
  - the inequality constraint j is **inactive** if $g_j(x) < 0$.

Linear constraints

- Let $f\colon \mathbb{R}^n \to \mathbb{R}$ and let $h\colon \mathbb{R}^n \to \mathbb{R}^{m_1}$ and $g\colon \mathbb{R}^n \to \mathbb{R}^{m_2}$ be two continuously differentiable functions.

- Let the **Lagrangian** $L\colon \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \to \mathbb{R}$ be defined by

$$\begin{aligned} L(x,\mu,\lambda) &= f(x) + \langle \mu, h(x)\rangle + \langle \lambda, g(x)\rangle \\ &= f(x) + \sum_{i=1}^{m_1} \mu_i h_i(x) + \sum_{j=1}^{m_2} \lambda_j g_j(x). \end{aligned}$$

The variables μ, λ are referred to as **dual variables**.

Linear equality constraints

Theorem 8

Assume that h and g are affine, that it to say, there exists $A \in \mathbb{R}^{m_2 \times n}$ *and* $b \in \mathbb{R}^m_2$ *such that*

$$g(x) = Ax + b.$$

Let $\bar{x}$ *be a* **local solution** *to* (P).

Then there exists $(\mu, \lambda) \in \mathbb{R}^{m1} \times \mathbb{R}^{m_2}$ *such that the following three conditions, referred to as Karush-Kuhn-Tucker (KKT) conditions, are satisfied:*

1. **Stationarity** *condition: ?*
2. **Sign** *condition: ?*
3. **Complementarity** *condition: ?*

Linear equality constraints

Theorem 9

Assume that h and g are affine, that it to say, there exists $A \in \mathbb{R}^{m_2 \times n}$ and $b \in \mathbb{R}_2^m$ such that

$$g(x) = Ax + b.$$

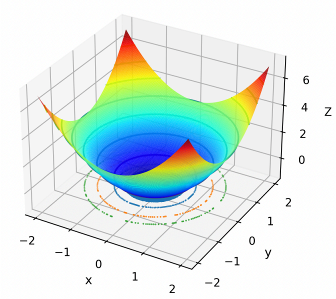
Let $\bar{x}$ be a **local solution** *to (P).*

Then there exists $(\mu, \lambda) \in \mathbb{R}^{m1} \times \mathbb{R}^{m_2}$ such that the following three conditions, referred to as Karush-Kuhn-Tucker (KKT) conditions, are satisfied:

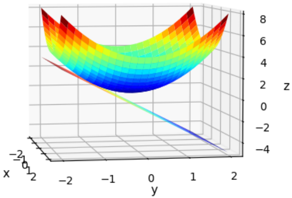
1. **Stationarity** *condition:* $\nabla_x L(\bar{x}, \mu, \lambda) = 0$.
2. **Sign** *condition: for all $j \in \mathcal{I}$,* $\lambda_j \geq 0$.
3. **Complementarity** *condition: for all $j \in \mathcal{I}$,* $g_j(\bar{x}) < 0 \Longrightarrow \lambda_j = 0$.

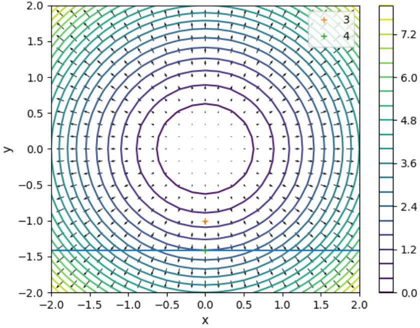
Linear constraints

Illustration.

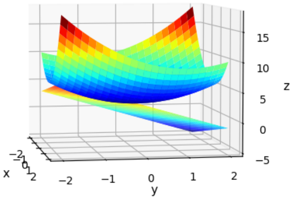


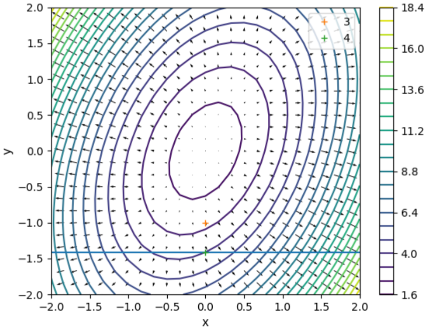
KKT stationarity



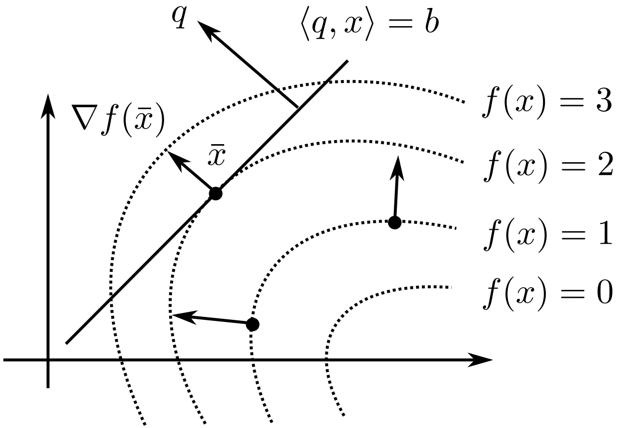


KKT stationarity



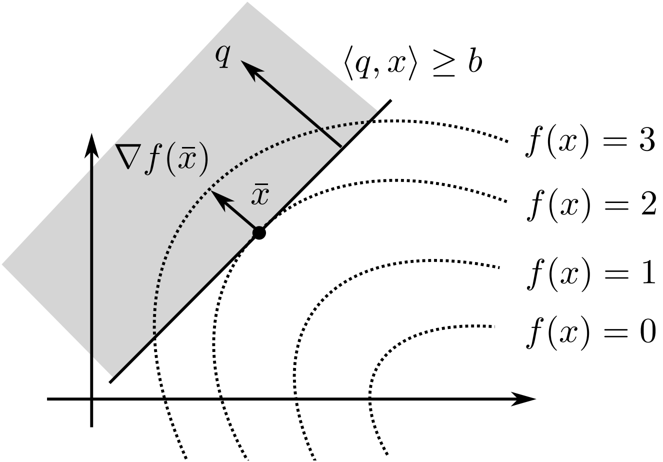


Linear constraints



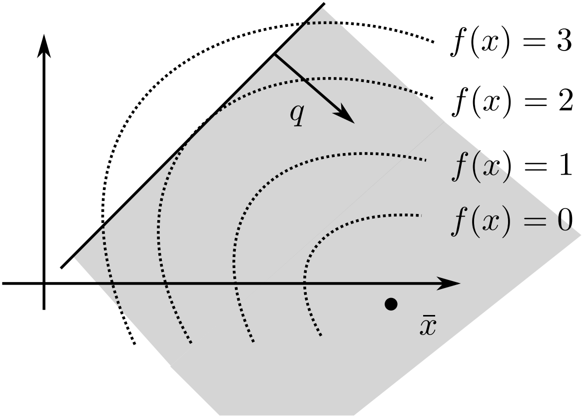
Linear constraints

Example 2(a). Case of **one (active) inequality equality constraint**:



Linear constraints

Example 2(b). Case of **one (inactive) inequality equality constraint**:



Non-linear constraints

Definition 10

Let $\bar{x}$ be a feasible point. Let the set of **active inequality constraints** $\mathcal{I}_0(\bar{x})$ be defined by

$$\mathcal{I}_0(\bar{x}) = \{j \in \mathcal{I} \,|\, g_j(\bar{x}) = 0\}.$$

We say that the **Linear Independence Qualification Condition (LICQ)** holds at $\bar{x}$, if the following set of vectors is linearly indepedent:

$$\{\nabla h_i(\bar{x})\}_{i \in \mathcal{E}} \cup \{\nabla g_j(\bar{x})\}_{j \in \mathcal{I}_0(\bar{x})}$$

Non-linear constraints

Non-linear constraints

Theorem 11

Let $\bar{x}$ be a **local solution** *to (P). Assume that the LICQ holds at $\bar{x}$. Then there exists a unique (μ, λ) such that the* **KKT conditions** *are satisfied.*

Remarks.

- At a numerical level, a solution that does not satisfy the LICQ is hard to compute.

Non-linear constraints

Example 4.

Consider the problem

$$\inf_{x \in \mathbb{R}} x, \quad \text{subject to: } x^2 \leq 0.$$

Unique feasible point: $\bar{x} = 0$, thus the solution.

Lagrangian:

$$L(x, \lambda) = x + \lambda x^2.$$

At zero:

$$\nabla_x L(0, \lambda) = 1 + 2\lambda\bar{x} = 1 \neq 0.$$

The LICQ is not satisfied, since $\nabla g_1(0) = 0$.

Non-linear constraints

Theorem 12

Assume that

- f *is* **convex**
- *for all* $i \in \mathcal{E}$, *the map* $x \mapsto h_i(x)$ *is* **affine**
- *for all* $j \in \mathcal{I}$, *the map* $x \mapsto g_j(x)$ *is* **convex**.

Then any feasible point $\bar{x}$ *satisfying the* **KKT conditions** *is a* **global solution** *to the problem.*

Remark. The result holds whether the LICQ holds or not at $\bar{x}$.

Exercise

Exercise. Consider the function $f\colon (x,y) \in \mathbb{R}^2 \mapsto \exp(x+y^2) + y + x^2$.

1. Prove that f is coercive. *Indication: Use* $\exp(z) \geq 1 + z$
2. Compute $\nabla f(x,y)$ and $\nabla^2 f(x,y)$.
3. We recall that a symmetric matrix of size 2 of the form $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ is positive semidefinite if and only if $a + c \geq 0$ and $ac - b^2 \geq 0$. Using this fact, prove that f is convex.
4. We consider the following problem:

$$\inf_{(x,y)\in\mathbb{R}^2} f(x,y), \quad \text{subject to: } \begin{cases} -x - y \leq 0 \\ -x - 2 \leq 0. \end{cases} \qquad (\mathcal{P})$$

   Verify that $(0,0)$ is feasible and satisfies the KKT conditions.
5. Is the point $(0,0)$ a global solution to problem $(\mathcal{P})$?

Exercise

Solution.

1. We use the inequality: $\exp(z) \geq 1 + z$, which yields:

$$
\begin{aligned}
f(x, y) &\geq x + y^2 + y + x^2 \\
&= \frac{1}{2}(x^2 + y^2) + \frac{1}{2}(x^2 + 2x + 1) + \frac{1}{2}(y^2 + 2y + 1) - 1 \\
&= \frac{1}{2}\|(x, y)\|^2 + \frac{1}{2}(x + 1)^2 + \frac{1}{2}(y + 1)^2 - 1 \underset{\|(x,y)\| \to \infty}{\longrightarrow} \infty.
\end{aligned}
$$

Exercise

2. It holds:

$$\frac{\partial f}{\partial x} = \exp(x + y^2) + 2x, \qquad \frac{\partial f}{\partial y} = 2y\exp(x + y^2) + 1.$$

Therefore, $\nabla f(x, y) = \begin{pmatrix} \exp(x + y^2) + 2x \\ 2y\exp(x + y^2) + 1 \end{pmatrix}$.

We also have

$$\frac{\partial^2 f}{\partial x^2} = \exp(x + y^2) + 2, \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 2y\exp(x + y^2),$$

$$\frac{\partial^2 f}{\partial y^2} = 2\exp(x + y^2) + 4y^2\exp(x + y^2).$$

Thus, $D^2 f(x, y) = \begin{pmatrix} \exp(x + y^2) + 2 & 2y\exp(x + y^2) \\ 2y\exp(x + y^2) & (2 + 4y^2)\exp(x + y^2) \end{pmatrix}$.

Exercise

3. Proof of positive definiteness of $D^2 f$. It holds:

$$a + c = (3 + 4y^2)\exp(x + y^2) + 2 \geq 0$$

and

$$ac - b^2 = 2\exp(2x + 2y^2) + 4(1 + 2y^2)\exp(x + y^2) \geq 0.$$

It follows that $D^2 f(x, y)$ is positive semidefinite, for all (x, y). Therefore f is a convex function.

Exercise

4. Feasibility of $(0,0)$: we have $0+0 \geq 0$ and $0+2>0$.
   KKT conditions. Lagrangian:

$$L(x,y,\lambda_1,\lambda_2) = \exp(x+y^2) + y + x^2 - \lambda_1(x+y) - \lambda_2(x+2).$$

Therefore,

$$\frac{\partial L}{\partial x}(0,0,\lambda_1,\lambda_2) = 1-\lambda_1-\lambda_2, \qquad \frac{\partial L}{\partial y}(0,0) = 1-\lambda_1.$$

Taking $\lambda_1 = 1$ and $\lambda_2 = 0$, we have:

1. Stationarity: $\frac{\partial L}{\partial x}(0,0,1,0) = \frac{\partial L}{\partial y}(0,0,1,0) = 0$.
2. Sign condition: $\lambda_1 \geq 0$, $\lambda_2 \geq 0$.
3. Complementarity: the second constraint is inactive and the corresponding Lagrange multiplier is null.

Exercise

5. We have the following:
   - The cost function is convex.
   - The functions $-(x+y)$ and $-(x+2)$ are convex.
   - The point $(0,0)$ is feasible and satisfies the KKT conditions.

   Therefore $(0,0)$ is a global solution.

Non-linear constraints

Exercise.

Consider:

$$\inf_{x\in\mathbb{R}^2} f(x) := -x_1 - x_2, \quad \text{s.t.} \begin{cases} g_1(x) = & x_1^2 + 2x_2^2 - 3 & \leq 0 \\ g_2(x) = & x_1 - 1 & \leq 0. \end{cases}$$

- Show that $\bar{x} = (1, 1)$ is feasible
- Verify that the LICQ and the KKT conditions hold at $\bar{x} = (1, 1)$.
- Prove that $\bar{x} = (1, 1)$ is a global solution.

Non-linear constraints

Verification of the LICQ.

$$\nabla g_1(\bar{x}) = \begin{pmatrix} 2\bar{x}_1 \\ 4\bar{x}_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \quad \text{and} \quad \nabla g_2(\bar{x}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

We have: $\mathcal{E} = \emptyset$, $\mathcal{I}_0(\bar{x}) = \{1, 2\}$. The vectors $\nabla g_1(\bar{x})$ and $\nabla g_2(\bar{x})$ are linearly independent, since

$$\det \begin{pmatrix} 2 & 4 \\ 1 & 0 \end{pmatrix} = -4 \neq 0.$$

Thus the LICQ is satisfied at $\bar{x}$.

Non-linear constraints

KKT conditions.

- Lagrangian:
  $L(x, \lambda) = (-x_1 - x_2) + \lambda_1(x_1^2 + 2x_2^2 - 3) + \lambda_2(x_1 - 1)$.
- Stationarity condition:

$$\begin{pmatrix} -1 \\ -1 \end{pmatrix} + \begin{pmatrix} 2\bar{x}_1 \\ 4\bar{x}_2 \end{pmatrix} \lambda_1 + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \lambda_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

  It is satisfied at $\bar{x}$ with $\lambda_1 = 1/4 \geq 0$ and $\lambda_2 = 1/2 \geq 0$.
- The sign condition is satisfied.
- The complementarity condition is satisfied (all inequality constraints are active).

1 Reminders
- Optimization problem
- Existence of a solution

2 Optimality conditions for constrained problems
- Linear constraints
- Non-linear constraints
- Sensitivity analysis

Sensitivity analysis

- Consider the family of optimization problems

$$\inf_{x\in\mathbb{R}^n} f(x), \quad \text{s.t.} \begin{cases} h_i(x) = y_i, & \forall i \in \mathcal{E}, \\ g_j(x) \leq y_j, & \forall j \in \mathcal{I}, \end{cases} \qquad (P(y))$$

parametrized by the vector $y \in \mathbb{R}^m$.

- Let the **value function** V be defined by

$$V(y) = \text{val}(P(y)).$$

- A variation δy_i in the i-th constraint generates a variation of the optimal cost of $\lambda_i \delta y_i$.

Sensitivity analysis

Exercise.

A company decides to rent an engine over d days. The engine can be used to produce two different objects. The two objects are not produced simultaneously. Let x_1 and x_2 denote the times dedicated to the production of each object. The resulting benefits (in k€) are given by:

$$\frac{x_1}{1+x_1} \quad \text{and} \quad \frac{x_2}{4+x_2}.$$

Sensitivity analysis

1. Formulate the problem as a minimization problem.
2. Justify the existence of a solution.
3. Write the KKT conditions. What is the unit of the dual variable?
4. Verify that $\bar{x} = (4, 6)$ satisfies the KKT conditions for $d = 10$ days. Is it a global solution to the problem?
5. The renting cost of the engine is 70€/day. Is it of interest for the company to rent the engine for a longer time?

Sensitivity analysis

1. Problem:

$$\inf_{x \in \mathbb{R}^2} -\frac{x_1}{1+x_1} - \frac{x_2}{4+x_2}, \quad \text{s.t.} \begin{cases} x_1 + x_2 = & d \\ -x_1 \leq & 0 \\ -x_2 \leq & 0 \end{cases}$$

2. The feasible set is obviously compact and non-empty and the cost function is continuous. Therefore, there exists a solution.

Sensitivity analysis

3. Let $\bar{x}$ be a solution. Let $\mu \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}^2$ be the associated Lagrange multipliers. Lagrangian:

$$L(x, \mu, \lambda) = -\frac{x_1}{1 + x_1} - \frac{x_2}{4 + x_2} + \mu(x_1 + x_2 - d) - \lambda_1 x_1 - \lambda_2 x_2.$$

KKT conditions:

- Stationarity:

$$-\frac{1}{(1 + \bar{x}_1)^2} + \mu - \lambda_1 = 0, \qquad -\frac{4}{(4 + \bar{x}_2)^2} + \mu - \lambda_2 = 0.$$

- Sign condition: $\lambda_1 \geq 0$, $\lambda_2 \geq 0$.
- Complementarity: $\bar{x}_1 > 0 \Rightarrow \lambda_1 = 0$, $\bar{x}_2 > 0 \Rightarrow \lambda_2 = 0$.

- Units: $[\mu] = [\lambda_1] = [\lambda_2] =$ k€/day.

Sensitivity analysis

4. Let μ, λ be such that the KKT conditions hold true. By complementarity condition, we necessarily have $\lambda_1 = \lambda_2 = 0$. The stationarity condition holds true with

$$\mu = \frac{1}{(1+\bar{x}_1)^2} = \frac{4}{(4+\bar{x}_2)^2} = \frac{1}{25} = 0.04.$$

The sign condition trivially holds true since the inequality constraints are inactive. Lagrangian:

$$L(x,\mu,\lambda) = -\frac{x_1}{1+x_1} - \frac{x_2}{4+x_2} + 0.04(x_1 + x_2 - d).$$

If $x_1 + x_2 > d$, the cost associated to constraints is increased, otherwise decreased (company rents the engine for the $d - x_1 - x_2$ remaining days).

The point $\bar{x}$ is feasible and satisfies the KKT conditions. We have affine constraints and a convex cost function, therefore, the KKT conditions are sufficient. The point $\bar{x}$ is a global solution.

Sensitivity analysis

5. d is fixed.
   Increasing the renting time of y days will generate a variation of cost of μy (approximately), that is, an augmentation of the benefit of 40€/day (less the renting price). It corresponds to the benefit that the company can have from another firm for renting the engine. Thus, the cost will corresponds to:

$$c(\overline{x}, \mu, \lambda) = -\frac{4}{1+4} - \frac{6}{4+6} - 0.04y + 0.07y = -1.4 + 0.03y.$$

   It would be of interest for the company to reduce the renting time.

And to sum up the courses ...

	Necessary conditions	Sufficient conditions
Abstract formulation (exist.)		if K compact, $f \in C^0(K)$ then at least one solution
		if K closed, $f \in C^0(K)$, coercive then at least one solution

	Necessary conditions	Sufficient conditions
No constraints $K = \mathbb{R}^d$ (opt.)	if $\overline{x}$ local sol., $f \in C^2(K)$ then, $D^2 f(\overline{x})$ is positive semi-def.	if $f \in C^2(K)$, $\nabla f(\overline{x}) = 0$, $D^2 f(\overline{x})$ positive def. then $\overline{x}$ local sol.
Affine constraints	$\overline{x}$ local sol. then KKT	f convex, then KKT=global sol.
Non-linear constraints	$\overline{x}$ local sol., LICQ then KKT	f convex, h affine, g convex, then KKT=global sol.