

Continuous optimization
ENT 305

Elise Grosjean

Ensta-Paris
Institut Polytechnique de Paris
2025

And to sum up the courses ...

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

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

And to sum up the courses ...

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

	Find a local solution
No constraints	Gradient Descent
Affine constraints	Penalty methods
Non-linear constraints	

- 1 Reminders
 - KKT
 - Sensitivity analysis
- 2 Penalty methods
 - Quadratic penalization
 - Augmented Lagrangian
- 3 Projected gradient method
 - Projection
 - Method
 - Combination with penalty methods

KKT

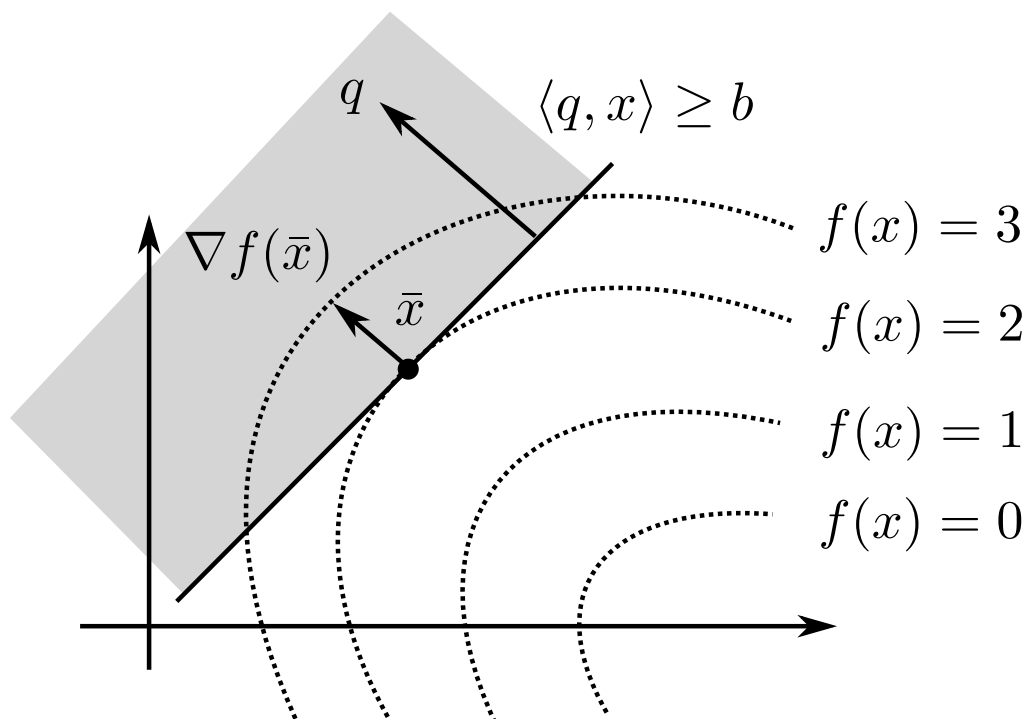
- We investigate in this section the problem

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

- 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 $g_j(x) < 0$.

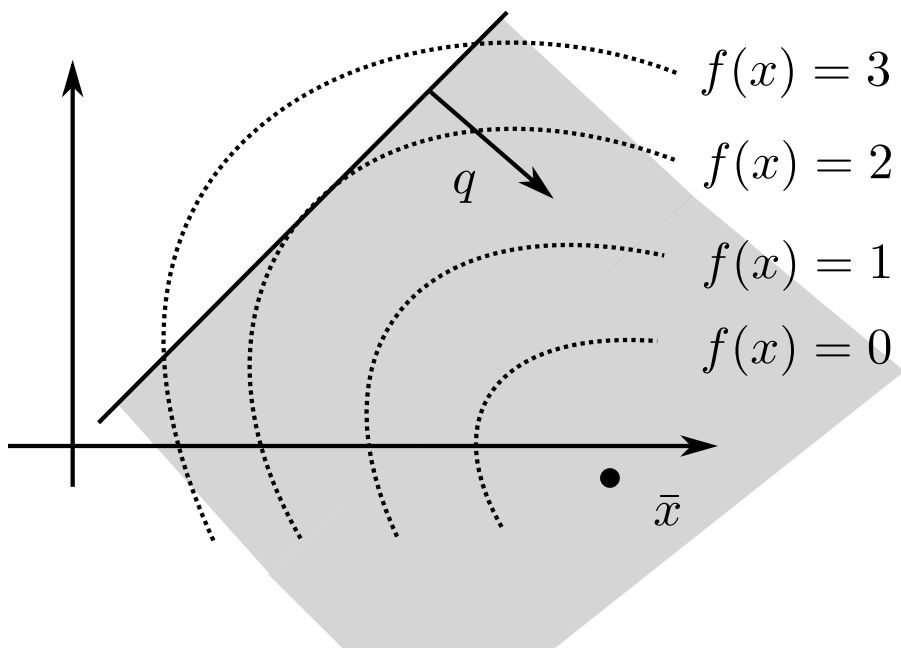
Linear constraints

$q = \nabla g$ **Example Case of one active inequality constraint:**



Linear constraints

$q = \nabla g$ **Example** Case of **one inactive inequality constraint**:



Linear constraints

- Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ and let $h: \mathbb{R}^n \rightarrow \mathbb{R}^{m_1}$ and $g: \mathbb{R}^n \rightarrow \mathbb{R}^{m_2}$ be C^2 .
- Let the **Lagrangian** $L: \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \rightarrow \mathbb{R}$ 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**.

KKT conditions

Let $\bar{x}$ be feasible for (P) .

The **Karush-Kuhn-Tucker (KKT) conditions** at $\bar{x}$ are the existence of

$$(\mu, \lambda) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$$

such that:

- 1 ?
- 2 ?
- 3 ?
- 4 ?

KKT conditions

Let $\bar{x}$ be feasible for (P) .

The **Karush-Kuhn-Tucker (KKT) conditions** at $\bar{x}$ are the existence of

$$(\mu, \lambda) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$$

such that:

- | | |
|---|---|
| 1 | ? |
| 2 | ? |
| 3 | ? |
| 4 | ? |

KKT conditions

Let $\bar{x}$ be feasible for (P) .

The **Karush-Kuhn-Tucker (KKT) conditions** at $\bar{x}$ are the existence of

$$(\mu, \lambda) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$$

such that:

1 Stationarity:

$$\nabla_x L(\bar{x}, \mu, \lambda) = 0.$$

2 Dual feasibility:

$$\lambda_j \geq 0, \quad \forall j \in \mathcal{I}.$$

3 Complementarity:

$$\lambda_j g_j(\bar{x}) = 0, \quad \forall j \in \mathcal{I}.$$

4 Primal feasibility:

$$h_i(\bar{x}) = 0, \quad \forall i, \quad g_j(\bar{x}) \leq 0, \quad \forall j.$$

KKT conditions

Let $\bar{x}$ be feasible for (P) .

The **Karush-Kuhn-Tucker (KKT) conditions** at $\bar{x}$ are the existence of

$$(\mu, \lambda) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$$

such that:

1 Stationarity:

$$\nabla_x L(\bar{x}, \mu, \lambda) = 0.$$

2 Dual feasibility:

$$\lambda_j \geq 0, \quad \forall j \in \mathcal{I}.$$

3 Complementarity:

$$\lambda_j g_j(\bar{x}) = 0, \quad \forall j \in \mathcal{I}.$$

4 Primal feasibility:

$$h_i(\bar{x}) = 0, \quad \forall i, \quad g_j(\bar{x}) \leq 0, \quad \forall j.$$

KKT conditions

Let $\bar{x}$ be feasible for (P) .

The **Karush-Kuhn-Tucker (KKT) conditions** at $\bar{x}$ are the existence of

$$(\mu, \lambda) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$$

such that:

1 Stationarity:

$$\nabla_x L(\bar{x}, \mu, \lambda) = 0.$$

2 Dual feasibility:

$$\lambda_j \geq 0, \quad \forall j \in \mathcal{I}.$$

3 Complementarity:

$$\lambda_j g_j(\bar{x}) = 0, \quad \forall j \in \mathcal{I}.$$

4 Primal feasibility:

$$h_i(\bar{x}) = 0, \quad \forall i, \quad g_j(\bar{x}) \leq 0, \quad \forall j.$$

KKT conditions

Let $\bar{x}$ be feasible for (P) .

The **Karush-Kuhn-Tucker (KKT) conditions** at $\bar{x}$ are the existence of

$$(\mu, \lambda) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$$

such that:

1 Stationarity:

$$\nabla_x L(\bar{x}, \mu, \lambda) = 0.$$

2 Dual feasibility:

$$\lambda_j \geq 0, \quad \forall j \in \mathcal{I}.$$

3 Complementarity:

$$\lambda_j g_j(\bar{x}) = 0, \quad \forall j \in \mathcal{I}.$$

4 Primal feasibility:

$$h_i(\bar{x}) = 0, \quad \forall i, \quad g_j(\bar{x}) \leq 0, \quad \forall j.$$

Lagrangian formulation

Main idea of the Lagrangian formulation:

Consider the primal problem (P):

$$\inf_{x \in \mathbb{R}^n} f(x) \text{ s.t. } g(x) \leq 0$$

Let p^* be the optimal cost of (P), and L the associated Lagrangian formulation. Then $p^* = \inf_{x \in \mathbb{R}^n} \sup_{\lambda \in \mathbb{R}_+^m} L(x, \lambda)$. Indeed,

$$\begin{aligned} \sup_{\lambda \in \mathbb{R}_+^m} L(x, \lambda) &= \sup_{\lambda \in \mathbb{R}_+^m} \left(f(x) + \sum_{j=1}^m \lambda_j g_j(x) \right), \\ &= \begin{cases} f(x) & \text{if } g_j(x) \leq 0, \forall j = 1, \dots, m \\ \infty & \text{otherwise (if } x \text{ not feasible and } g_j(x) > 0), \end{cases} \end{aligned}$$

Exercise

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

- 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:

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

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:

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

Exercise

2. It holds:

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

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

We also have

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

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

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. Prove that D^2f is positive semidefinite. 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 $D^2f(x, y)$ 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,

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

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.

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

Example 4.

Consider the problem

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

The unique feasible point is $\bar{x} = 0$, hence it is 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 $\nabla g_1(0) = 0$.

KKT

We consider the problem

$$\min_{x,y} f(x,y) = (x-1)^2 + 2(y-1)^2$$

subject to

$$h(x,y) = x + y - 1 = 0, \quad g(x,y) = -x \leq 0 \text{ (i.e. } x \geq 0 \text{)}.$$

Exercise.

- 1 Write the Lagrangian
- 2 Write the KKT conditions
- 3 Find $(\bar{x}, \bar{y}, \bar{\lambda}, \bar{\mu})$ satisfying the KKT conditions
- 4 Are the LICQ conditions satisfied?
- 5 Is the solution global?

KKT exercise: solution (affine constraints)

Problem: $f(x, y) = (x - 1)^2 + 2(y - 1)^2$, $x + y = 1$, $x \geq 0$.

$$L = f + \mu(x + y - 1) - \lambda x, \quad \lambda \geq 0.$$

The KKT conditions are

$$\begin{cases} 2(x - 1) + \mu - \lambda = 0, \\ 4(y - 1) + \mu = 0, \\ x + y = 1, \quad x \geq 0, \\ \lambda \geq 0, \quad \lambda x = 0. \end{cases}$$

The unique candidate is

$$(\bar{x}, \bar{y}) = (1/3, 2/3), \quad \bar{\mu} = 4/3, \quad \bar{\lambda} = 0.$$

Since $\bar{x} > 0$, the inequality is inactive. The equality constraint has nonzero gradient $(1, 1)$: **LICQ holds**. Strict convexity of f and convexity of the feasible set give **the unique global minimizer**.

KKT

We consider the problem

$$\min_{x,y} f(x,y) = x$$

subject to

$$g(x,y) = y^2 - x^3 \leq 0.$$

Exercise.

- 1 Write the Lagrangian
- 2 Find $(\bar{x}, \bar{y})$ the global solution
- 3 Are the LICQ conditions satisfied at that point?
- 4 What about KKT conditions?

KKT exercise: solution (nonlinear constraint)

Problem: $\min_{(x,y) \in \mathbb{R}^2} x$ subject to $y^2 - x^3 \leq 0$.

$$L(x, y, \lambda) = x + \lambda(y^2 - x^3), \quad \lambda \geq 0.$$

Feasibility implies $x^3 \geq y^2 \geq 0$, hence $x \geq 0$. The origin is feasible:

(0, 0) is a global minimizer. At (0, 0),

$$\nabla f = (1, 0), \quad \nabla g = (-3x^2, 2y)|_{(0,0)} = (0, 0).$$

Thus **LICQ fails**. Stationarity would require

$$(1, 0) + \lambda(0, 0) = (0, 0),$$

The minimizer does not satisfy KKT.

Non-linear constraints

Exercise.

Consider:

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

- 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).

Sensitivity analysis

- Consider the family of optimization problems

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

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 changes the optimal cost by approximately $-\lambda_i \delta y_i$ (or $-\mu_i \delta y_i$ for an equality constraint).

Sensitivity analysis: active vs. inactive constraints

Consider $Ax \leq b$, with multiplier $\lambda \geq 0$.

Inactive constraint

$$Ax < b$$

Changing the constraint has no local impact.

For a sufficiently small perturbation δb :

$$Ax < b + \delta b.$$

By complementarity:

$$\lambda = 0$$

$$V(b + \delta b) \simeq V(b).$$

Active constraint

$$Ax = b$$

Changing the constraint may affect the optimal cost.

If $V \in \mathcal{C}^1$: $V'(b) = -\lambda$

$$V(b + \delta b) \simeq V(b) - \lambda \delta b.$$

- $\delta b > 0$: relaxing the constraint may reduce the cost.
- $\delta b < 0$: tightening the constraint may increase the cost.

Sensitivity analysis

Exercise.

A company rents a machine for d days and uses it to produce two goods, which cannot be produced simultaneously. Let x_1 and x_2 denote the production times allocated to each good. The resulting profits (in k€) are:

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

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 rental cost of the machine is 70€/day. Should the company rent the machine for a longer period?

Sensitivity analysis

1. Problem:

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

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

Sensitivity analysis

3. Let $\bar{x}$ be a solution. Let $\mu \in \mathbb{R}$ 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:

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

- 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: $[\mu] = [\lambda_1] = [\lambda_2] = \text{k€}/\text{day}$.

Sensitivity analysis

4. Suppose that μ, λ satisfy the KKT conditions. By complementarity, $\lambda_1 = \lambda_2 = 0$. The stationarity condition holds with

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

The sign condition holds because the inequality constraints are inactive. Lagrangian:

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

The point $\bar{x}$ is feasible and satisfies the KKT conditions. The constraints are affine and the cost function is convex.

Therefore, the KKT conditions are sufficient, and $\bar{x}$ is a global solution.

Sensitivity analysis

5. Should the company rent the machine longer?

The optimal production cost satisfies

$$V'(d) = -\mu(d).$$

At $d = 10$, we have $\mu(10) = 0.04$, hence

$$V(10 + \delta d) \simeq V(10) - 0.04 \delta d.$$

The rental cost is 0.07 k€/day. The total optimized cost is therefore

$$C(d) = V(d) + 0.07d.$$

Thus,

$$C'(10) = -\mu(10) + 0.07 = -0.04 + 0.07 = 0.03.$$

Conclusion: An extra rental day generates approximately 40€ in additional production profit, but costs 70€.

The company should not extend the rental period. 

- 1 Reminders
 - KKT
 - Sensitivity analysis
- 2 Penalty methods
 - Quadratic penalization
 - Augmented Lagrangian
- 3 Projected gradient method
 - Projection
 - Method
 - Combination with penalty methods

Motivation: KKT conditions may fail

Example 1: a nonlinear inequality constraint.

$$\min_{(x,y) \in \mathbb{R}^2} f(x,y) = x, \quad g(x,y) = y^2 - x^3 \leq 0.$$

Since $y^2 \leq x^3$, every feasible point satisfies $x \geq 0$. Thus $(0,0)$ is the **unique global minimizer**.

The Lagrangian is $L(x,y,\lambda) = x + \lambda(y^2 - x^3)$. Stationarity at $(0,0)$ would imply

$$\nabla_{(x,y)} L(0,0,\lambda) = \begin{pmatrix} 1 - 3\lambda \cdot 0^2 \\ 2\lambda \cdot 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \neq 0.$$

- No multiplier $\lambda \geq 0$ satisfies the KKT equations.
- LICQ fails: $\nabla g(0,0) = (0,0)$.

KKT are not always necessary without a constraint qualification.