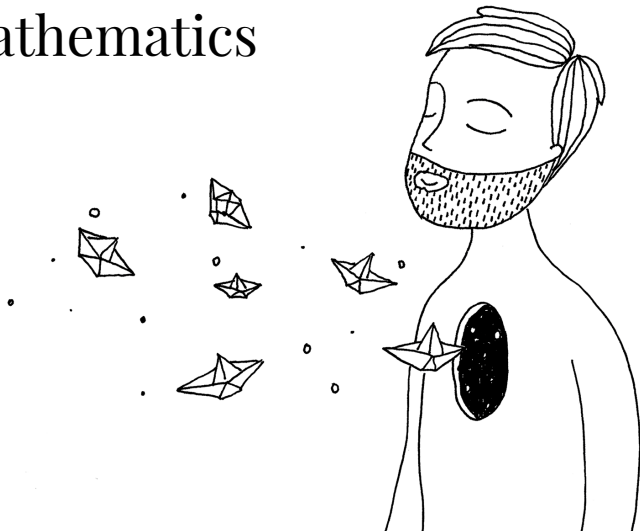


# 4509 – Bridging Mathematics

Optimization

PAULO FAGANDINI



Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , and let  $S \subseteq \mathbb{R}^n$ , the problem is:

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

Or finding  $x^* \in S$  such that

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

## Definition

If  $x \in S$ ,  $x$  is said to be **feasible** for the optimization problem  $\min_{x \in S} f(x)$ .

Note that to minimize  $f(x)$  is equivalent to maximize  $-f(x)$ .

## Definition

$x_0$  is a **local minimum** of  $f$  in  $S$  if, there is an open ball  $B(x_0, r)$  such that for any  $x \in B(x_0, r) \cap S$  it holds that  $f(x_0) \leq f(x)$ .

$x^*$  is a **global minimum** of  $f$  in  $S$  if for any  $x \in S$  it holds that  $f(x^*) \leq f(x)$ .

## Definition

Let the  $\arg \min_{x \in S} f(x)$  represent the set of all global solutions to the problem of minimizing  $f$  with  $x \in S$ .

**Note:** For a maximization problem you have equivalently local maximum, global maximum, and  $\arg \max_{x \in S} f(x)$ .

## Conjecture

*Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a continuous function. Let  $S \in \mathbb{R}^n$  be compact. Then the problem*

$$\max_{x \in S} f(x)$$

*has at least one solution, or equivalently  $\arg \min_{x \in S} f(x) \neq \emptyset$ .*

## Conjecture

*If  $S_1 \subseteq S_2$ , then  $\min(f, S_1) \geq \min(f, S_2)$ .*

Prove the first one, 10 min.

Proof.

- $f$  is continuous in  $S$ , and  $S$  is compact, so  $f(S)$  is compact.

Prove the first one, 10 min.

Proof.

- $f$  is continuous in  $S$ , and  $S$  is compact, so  $f(S)$  is compact.
- $f(S)$  compact means it is bounded, and furthermore closed, so the supremum (lowest upper bound) is in the set.

Prove the first one, 10 min.

Proof.

- $f$  is continuous in  $S$ , and  $S$  is compact, so  $f(S)$  is compact.
- $f(S)$  compact means it is bounded, and furthermore closed, so the supremum (lowest upper bound) is in the set.
- $\exists s \in S$  such that  $f(s) = \sup f(S)$ .

Prove the first one, 10 min.

Proof.

- $f$  is continuous in  $S$ , and  $S$  is compact, so  $f(S)$  is compact.
- $f(S)$  compact means it is bounded, and furthermore closed, so the supremum (lowest upper bound) is in the set.
- $\exists s \in S$  such that  $f(s) = \sup f(S)$ .

Maybe it can be shown with contradiction, choosing an  $x$ , then see if it is the maximizer, done, if it is not, it is because you know of another that gives you a higher value in  $f(x)$ , so you move to that one. Then you can build a sequence, which you now it has a subsequence that converges in  $S$  ( $S$  is compact), and therefore the limit must be the maximizer, and therefore  $f(x_n)$  converges to the maximum of the function as well.





Now the second one 5 min.

Now the second one 5 min.

Proof.

Trivial. By contradiction. Let  $s_1 \in S \leq s_2 \in S$ . As  $s_i \in S_1 \Rightarrow s_1 \in S_2$ , because  $S_1 \subseteq S_2$ , then  $s_2$  cannot be a minimizer.



These are the necessary conditions for optimality for the unconstrained problem ( $S = \mathbb{R}^n$ ), when  $f$  is at least twice differentiable:

## Conjecture

$x_0$  is a local optimum of  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  if:

1. *First Order Condition:*  $\nabla f(x_0) = 0$ .
2. *Second Order Conditions:*
  - 2.1  $H(f, x_0) \geq 0$  (Hessian positive semi definite) then  $x_0$  is a local minimum.
  - 2.2  $H(f, x_0) \leq 0$  (Hessian negative semi definite) then  $x_0$  is a local maximum.

# Example

Let  $f(x, y) = x^2 + y^2$ ,

1. FOC:  $\frac{\partial f}{\partial x}(x^*) = 0$  and  $\frac{\partial f}{\partial y}(x^*) = 0$

1.1  $f_x = 2x$ , so  $f_x = 0 \Rightarrow x^* = 0$

1.2  $f_y = 2y$ , so  $f_y = 0 \Rightarrow y^* = 0$

2. SOC:

2.1 Hessian:

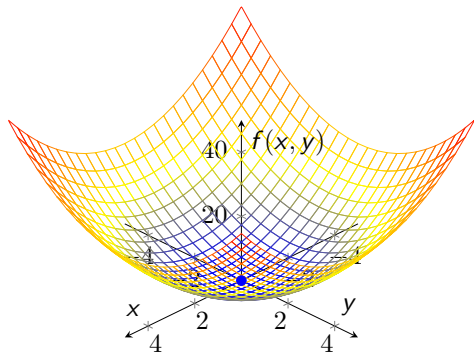
$$\mathcal{H} = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

2.2  $\det(\mathcal{H} - \lambda I) = (2 - \lambda)^2$ , so  $\det(\mathcal{H} - \lambda I) = 0 \Rightarrow \lambda = 2$

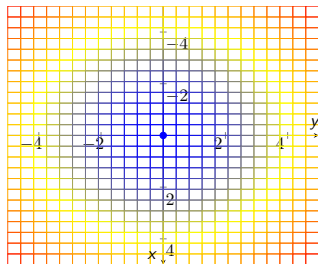
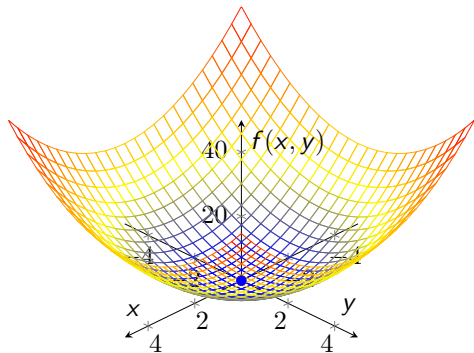
2.3 All e.v. are strictly positive, so  $\mathcal{H}$  is positive definite.

2.4 Finally  $(0, 0)$  is a minimum.

# Example



# Example



# Example

Let  $f(x, y) = x^2 - y^2$ ,

1. FOC:  $\frac{\partial f}{\partial x}(x^*) = 0$  and  $\frac{\partial f}{\partial y}(x^*) = 0$

1.1  $f_x = 2x$ , so  $f_x = 0 \Rightarrow x^* = 0$

1.2  $f_y = -2y$ , so  $f_y = 0 \Rightarrow y^* = 0$

2. SOC:

2.1 Hessian:

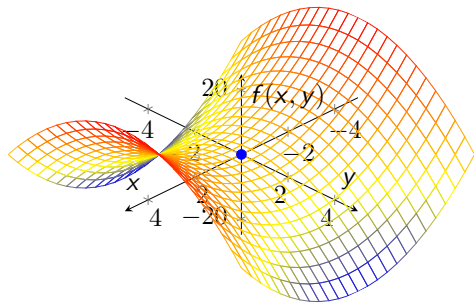
$$\mathcal{H} = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

2.2  $\det(\mathcal{H} - \lambda I) = -(4 - \lambda^2)$ , so  $\det(\mathcal{H} - \lambda I) = 0 \Rightarrow \lambda = \pm 2$

2.3 The e.v.s are not strictly positive or negative, so we cannot say anything about the positiveness of  $\mathcal{H}$ .

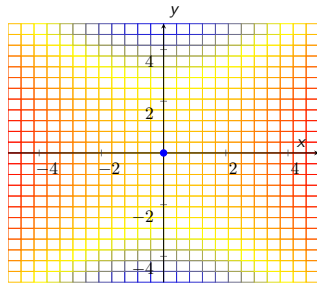
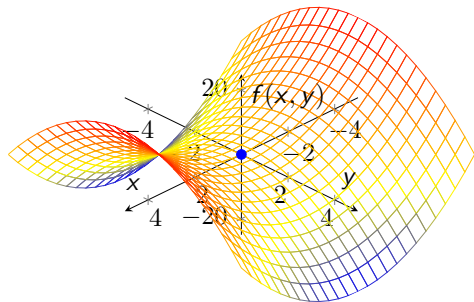
2.4 Finally we cannot say that  $(0, 0)$  is a minimum or a maximum.

# Example





# Example



# Equality Constraints

Consider now the constrained problem, and assume that  $S \subseteq \mathbb{R}^n$  can be described as a set of equations that  $x \in \mathbb{R}^n$  must satisfy, say  $h_i(x) = 0$ .

$$S = \{x \in \mathbb{R}^n | h_i(x) = 0, \quad i = 1, \dots, m\}$$

The problem is now

$$\begin{array}{ll} \min_{x \in \mathbb{R}^n} & f(x) \\ \text{s.t.} & h_i(x) = 0, \quad i = 1, \dots, m \end{array}$$

From now on, to simplify notation, we will use  $\min_x$  instead of  $\min_{x \in \mathbb{R}^n}$ .

## Definition (Mangasarian-Fromowitz constraint qualification)

The feasible point  $x^* \in \mathbb{R}^n$  is said to be **regular** if the set of gradients  $\nabla h_i(x^*)$  for  $i = 1, \dots, m$  is l.i.

If this is not satisfied, the solution you find might not be an optimum.

## Definition

Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , continuous and differentiable.

Consider the following optimization problem:

$$\begin{array}{ll} \min_x & f(x) \\ \text{s.t.} & h_i(x) = 0, \quad i = 1, \dots, n \end{array}$$

Define the function  $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  such that:

$$\mathcal{L}(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) = f(x) + \sum_{i=1}^m \lambda_i h_i(x)$$

$\mathcal{L}$  is called the **Lagrangian**, and  $\lambda_i$  for  $i = 1, \dots, m$  are called the **Lagrange multipliers**.

## Theorem

*Let  $x^*$  to be a local minimum of  $f$ , such that  $h_i(x^*) = 0$  for  $i = 1, \dots, m$ . Also let  $x^*$  be regular. Then, there is a vector  $\lambda = (\lambda_1, \dots, \lambda_m)^t \in \mathbb{R}^m$  such that:*

$$\begin{aligned} \frac{\partial f(x^*)}{\partial x_i} + \sum_{j=1}^m \lambda_j \frac{\partial h_j(x^*)}{\partial x_i} &= 0, \quad i = 1, \dots, n \\ \frac{\partial f(x^*)}{\partial \lambda_j} &= 0, \quad j = 1, \dots, m \end{aligned}$$

## Theorem (Second Order Conditions)

*Let  $x^*$  be a local minimum for  $f$ , satisfying  $h_j(x^*) = 0$  for every  $j = 1, \dots, m$ . Assume further that  $x^*$  is regular. Consider  $\lambda \in \mathbb{R}^m$  the vector of Lagrange multipliers of the problem, then the matrix*

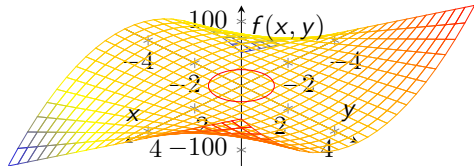
$$\mathcal{H} = H(f, x^*) + \sum_{j=1}^m \lambda_j H(h_j, x^*)$$

*is positive semi definite in the set  $M := \{y \in \mathbb{R}^n \mid \nabla h_j(x^*) \cdot y = 0, \forall j = 1, \dots, m\}$*

# Example

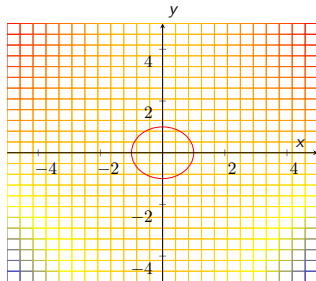
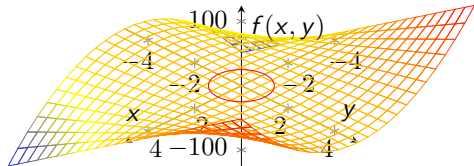
Let  $f(x, y) = x^2y$ . Maximize  $f(x, y)$  such that  $x^2 + y^2 = 1$ .

# Example

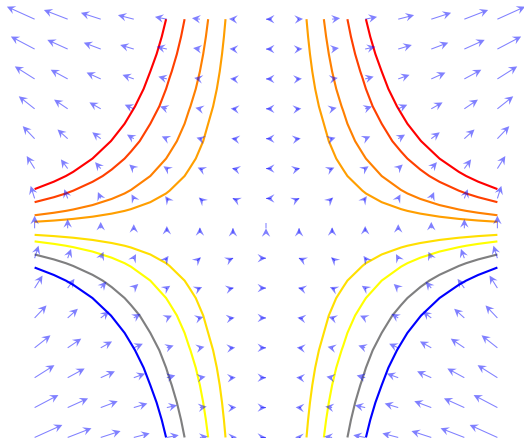




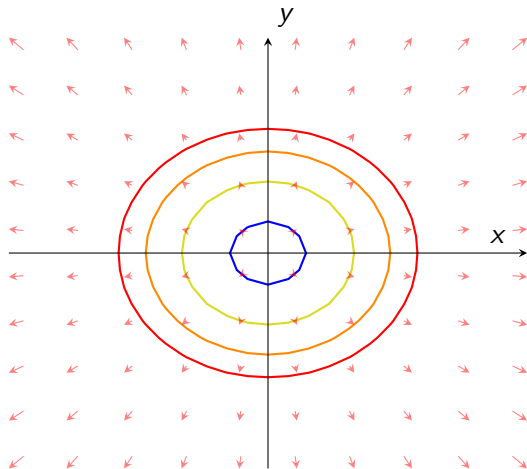
# Example



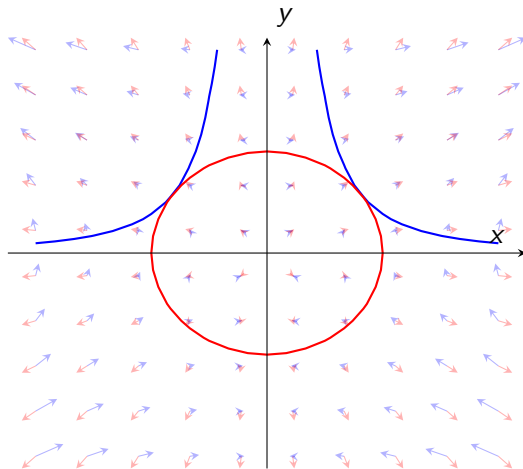
# Example



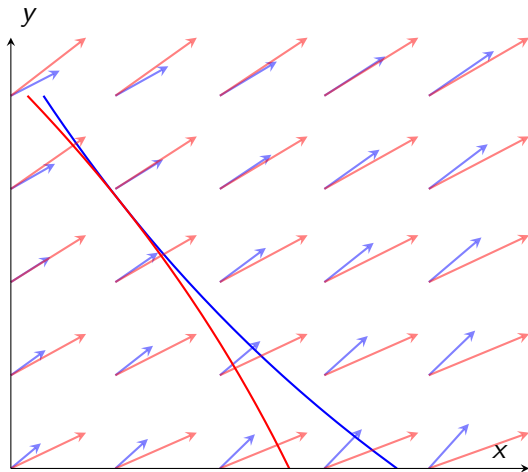
# Example



# Example



# Example



# Example

The Lagrangian is:

$$\mathcal{L} = x^2y + \lambda(x^2 + y^2 - 1)$$

# Inequality Constraints

Assume now that  $S \subseteq \mathbb{R}^n$  can be described as a set of equations and inequalities that  $x \in \mathbb{R}^n$  must satisfy, say  $h_i(x) = 0$  and  $g_j(x) \leq 0$ .

$$S = \{x \in \mathbb{R}^n | h_i(x) = 0, i = 1, \dots, m\} \cap \{x \in \mathbb{R}^n | g_j(x) \leq 0, j = 1, \dots, p\}$$

The problem is now

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & h_i(x) = 0, \quad i = 1, \dots, m \\ & g_j(x) \leq 0, \quad j = 1, \dots, p \end{aligned}$$

## Definition

Let  $x^* \in \mathbb{R}^n$  be such that  $h_i(x^*) = 0, i = 1, \dots, m$  and  $g_j(x^*) \leq 0, j = 1, \dots, p$ .  $x^*$  is called **regular** for the constraints if the set of gradients

$$\{\nabla h_i(x^*), \nabla g_j(x^*), i = 1, \dots, m, \quad j \in J_A\}$$

is *l.i.*, where  $J_A \subseteq \{1, \dots, p\}$  represents the active constraints in  $x^*$ .



## Definition

Let  $x^*$  be a solution to the problem in the previous slide. The inequality constraint  $g_k(x^*)$  is called **active**, if  $g_k(x^*) = 0$ . Otherwise it is considered **slack**.

If you knew ex-ante which constraints are active, you can use the Lagrange method to find the solution.

## Theorem (Karush-Kuhn-Tucker)

Let  $x^*$  be a local minimum for the problem:

$$\begin{array}{ll}\min_x & f(x) \\ \text{s.t.} & h_i(x) = 0, i = 1, \dots, m \\ & g_j(x) \leq 0, j = 1, \dots, p\end{array}$$

Such that  $x^*$  is regular for the constraints, then there are multipliers  $\lambda_i, i = 1, \dots, m$  and  $\mu_j, j = 1, \dots, p$  such that:

1.  $\mu_j \geq 0$  for  $j = 1, \dots, p$ .
2.  $\nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla h_i(x^*) + \sum_{j=1}^p \mu_j \nabla g_j(x^*) = 0$ .
3.  $\sum_{j=1}^p \mu_j g_j(x^*) = 0$

## Theorem (Second Order Conditions)

*Let  $x^*$  be a local minimum of  $f$  that satisfies  $h_i(x^*) = 0, i = 1, \dots, m$ ,  $g_j(x^*) \leq 0, j = 1, \dots, p$ . Assume further than  $x^*$  is regular for the constraints. Then the matrix*

$$\mathcal{H} = H(f, x^*) + \sum_{i=1}^m \lambda_i H(h_i, x^*) + \sum_{j=1}^p \mu_j H(g_j, x^*)$$

*Is positive semi definite in the set that is orthogonal to the active constraints:*

$$M := \{y \in \mathbb{R}^n \mid \nabla h_j(x^*) \cdot y = 0, \forall j = 1, \dots, m\} \cap \{y \in \mathbb{R}^n \mid \nabla g_k(x^*) \cdot y = 0, k \in J_A\}$$

## Theorem (Envelope's Theorem)

*Consider the following optimization problem,*

$$\begin{aligned} \min_x \quad & f(x, \alpha) \\ \text{s.t.} \quad & h_j(x, \alpha) = 0, j = 1, \dots, m \end{aligned}$$

*Where  $\alpha = (\alpha_1, \dots, \alpha_l) \in \mathbb{R}^l$  are parameters of the problem. Consider further that all the functions ( $f$ ,  $h$ s) are continuously differentiable. Let  $x(\alpha)$  a solution and  $\min(f)(\alpha) = f(x(\alpha), \alpha)$  the minimum value taken by  $f$ . Then,  $\forall k = 1, \dots, l$  holds that:*

$$\frac{d \min(f)(\alpha)}{d\alpha_k} = \frac{\partial f(x(\alpha), \alpha)}{\partial \alpha_k} + \sum_{j=1}^m \lambda_j \frac{\partial h_j(x(\alpha), \alpha)}{\partial \alpha_k}$$

*Where  $\lambda_j$  is the multiplier of the optimality conditions.*

## Theorem (Berge's Maximum)

*Consider the following optimization problem*

$$\begin{aligned} \max_x \quad & f(x, \alpha) \\ \text{s.t.} \quad & g_j(x, \alpha) \leq 0, j = 1, \dots, p \end{aligned}$$

*Assume that for  $\alpha^*$ , the solution is  $x^* = x(\alpha^*)$ . Then, if  $f$  and the  $g$ s are continuous in  $(x^*, \alpha^*)$ , and the set defined by the inequality constraints is compact, then the function  $\max(f)(\alpha) = f(x^*(\alpha), \alpha)$  is continuous in  $\alpha^*$ . Furthermore, if the solution  $x^*(\alpha)$  is unique, then,  $x^*(\alpha)$  is also continuous.*