Mathematical analysis 3
Chapter 4 : Extrema of multivariable functions

ensia
The National School of
Artificial Intelligence
المدرسة الوطنية للذكاء الاصطناعي

2023/2024

Course outline

Absolute maximum and minimum values

Let D be a subset of $\mathbb{R}^m$ and let $p \in D$.

Definition (Absolute maximum and minimum values)

A function f defined on D has

- *an* ***absolute minimum*** *at p if*

$$\forall u \in D : \quad f(p) \leq f(u).$$

- *an* ***absolute maximum*** *at p if*

$$\forall u \in D : \quad f(p) \geq f(u).$$

- *an* ***absolute extremum*** *at p if it has an absolute minimum or an absolute maximum at p.*

Local maximum and minimum values

Definition (Local maximum and minimum values)

A function f defined on an open set D has

- *a **local minimum** at p if $\exists V(p) \subset D$ such that*

$$\forall u \in V(p) \cap D : \quad f(p) \leq f(u),$$

- *a **local maximum** at p if $\exists V(p) \subset D$ such that*

$$\forall u \in V(p) \cap D : \quad f(p) \geq f(u).$$

- *a **local extremum** at p if it has a local minimum or a local maximum at p.*

Some remarks

Remarks

- A local extremum may not be absolute.
- A function f has a global extremum at p over D if $f(u) - f(p)$ does not change sign for all $u \in D$, and a local extremum at p over an open set D if there exists $V(p) \subset D$ such that $f(u) - f(p)$ does not change sign for all $u \in V(p) \cap D$.

Course outline

Optimality necessary condition of first-order

Definition (Critical point)

Let D be an open set in $\mathbb{R}^m$. A point $p \in D$ is called a ***critical point*** *of a differentiable function $f : D \to \mathbb{R}$ if $\nabla f(p) = 0$.*

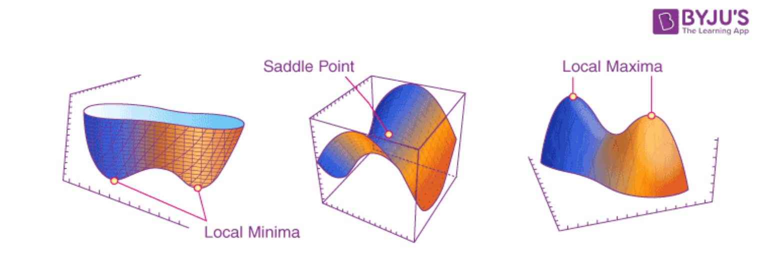
The following result gives a necessary condition for a point to be an extremum:

Theorem (Necessary condition of first-order optimality)

Let D be an open set in $\mathbb{R}^m$ and $f : D \to \mathbb{R}$ be a C^1 function. Then, if f has a local extremum at a point $p \in D$, it is ***necessarily*** *a critical point.*

Remark. The converse is false.

Nature of critical points



To determine the nature of critical points, we have two methods:

- By definition,
- Using second-order partial derivatives.

Otherwise, we use the Taylor formula around the critical point to a high enough order.

Study of extrema using the definition: Examples

Example. Let $f(x,y)=x^2+y^2+xy$. We have $f\in C^1(\mathbb{R}^2)$. The critical points of f: We have

$$\nabla f(x,y)=\begin{pmatrix}\frac{\partial f}{\partial x}\\ \frac{\partial f}{\partial y}\end{pmatrix}=\begin{pmatrix}2x+y\\ 2y+x\end{pmatrix}.$$

Thus, (x,y) is a critical point of f if and only if

$$\begin{cases}2x+y=0,\\ 2y+x=0,\end{cases}$$

which implies

$$\begin{cases}x=0,\\ y=0.\end{cases}$$

Nature of the critical points: We have

$$f(x,y)-f(0,0)=(x+\frac{1}{2}y)^2+\frac{3}{4}y^2>0 \quad \forall (x,y)\in\mathbb{R}^2\setminus\{(0,0)\}.$$

Therefore, f has a unique extremum at $(0,0)$ which is a global minimum.

Study of extrema using the definition: Examples

Example. Let $f(x,y) = x^2 + y^2 - 2x - 4y$. We have $f \in C^1(\mathbb{R}^2)$. The critical points of f: We have

$$\nabla f(x,y) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x-2 \\ 2y-4 \end{pmatrix}.$$

Thus, (x,y) is a critical point of f if and only if

$$\begin{cases} 2x-2=0, \\ 2y-4=0, \end{cases}$$

which implies

$$\begin{cases} x=1, \\ y=2. \end{cases}$$

Study of extrema using the definition: Examples

Nature of the critical points: We use the change of variables $(x,y) = (h+1, k+2)$ Then,

$$\begin{aligned} f(x,y) - f(1,2) &= f((h+1, k+2)) - f(1,2) \\ &= (h+1)^2 + (k+2)^2 - 2(h+1) - 4(k+2) + 5 \\ &= h^2 + k^2 > 0. \end{aligned}$$

Therefore, f has a unique extremum at $(1,2)$ which is a global minimum.

Study of extrema using the definition: Examples

Let $f(x,y)=x^3+y^2$. We have $f\in C^1(\mathbb{R}^2)$. The critical points of f: We have

$$\nabla f(x,y)=\begin{pmatrix}\frac{\partial f}{\partial x}\\ \frac{\partial f}{\partial y}\end{pmatrix}=\begin{pmatrix}3x^2\\ 2y\end{pmatrix}.$$

Thus, (x,y) is a critical point of f if and only if

$$\begin{cases}3x^2=0,\\ 2y=0,\end{cases}$$

which implies

$$\begin{cases}x=0,\\ y=0.\end{cases}$$

Study of extrema using the definition: Examples

Nature of the critical points: We observe that

$$f(-1,0) = -1 < f(0,0) < 1 = f(1,0),$$

thus f does not have a global extremum at $(0,0)$.
Furthermore,

$$f(x,0) - f(0,0) = x^3 < 0 \text{ if } x < 0$$
$$f(x,0) - f(0,0) = x^3 > 0 \text{ if } x > 0.$$

Therefore, $f(x,y) - f(0,0)$ changes sign in any neighborhood of $(0,0)$.
We conclude that f does not even have a local extremum at $(0,0)$.

Optimality necessary condition of second-order

We will generalize to functions of several variables what is known for extrema of real functions of one variable, namely if $f : \mathbb{R} \to \mathbb{R}$ is C^2 in the neighborhood of a critical point p (i.e., if $f'(p) = 0$), then:

- if $f''(p) > 0$, f has a local minimum at p,
- if $f''(p) < 0$, f has a local maximum at p,
- if $f''(p) = 0$, we cannot conclude, further calculations are needed (for example, Taylor expansion of order greater than 2).

For functions of several variables, if f is C^2 in the neighborhood of a critical point p, we will focus on a "certain notion of positivity of the Hessian matrix.

Notions of "positive definite" and "negative definite" matrix

Definition

A symmetric matrix $A \in M_m(\mathbb{R})$ *is:*

- ***positive definite*** *if:*

$$\forall X \in \mathbb{R}^m \setminus \{0\}, X^T A X > 0$$

- ***negative definite*** *if* $-A$ *is positive definite i.e.:*

$$\forall X \in \mathbb{R}^m \setminus \{0\}, X^T A X < 0.$$

- ***indefinite*** *if* A *is neither positive definite nor negative definite.*

Notions of "positive definite" and "negative definite" matrix

The following rule is very practical for determining if a matrix is positive definite:

Proposition (Sylvester's criterion)

A matrix A is positive definite if and only if $\forall k \in \{1, \ldots, n\}, \det(\Delta_k) > 0$, where $\det(\Delta_k)$ is the k-th leading principal minor.

Example.

The matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ is positive definite.

Proposition

A matrix A is positive definite if and only if $sp(A) \subset \mathbb{R}_+^*$.

Notions of "positive definite" and "negative definite" matrix

Example: Let $f(x,y) = x^2 + y^2 + 2y^3$. We have

$$\nabla^2(f)(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 2+12y \end{pmatrix}.$$

In particular,

$$\nabla^2(f)(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

is a positive definite matrix.

$$\nabla^2(f)(x,-1) = \begin{pmatrix} 2 & 0 \\ 0 & -10 \end{pmatrix}$$

is an invertible and indefinite matrix.

$$\nabla^2(f)\left(0,-\frac{1}{6}\right) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

is a non-invertible matrix.

Fundamental theorem

Theorem

Let D be an open set in $\mathbb{R}^m$ and $f : D \subseteq \mathbb{R}^m \to \mathbb{R}$ be a function of class C^2 with p a critical point of f (i.e., $\nabla f(p) = 0$). Then,

- *if $Hess(f)(p)$ is positive definite, then f has a strict local minimum at p,*
- *if $Hess(f)(p)$ is negative definite, then f has a strict local maximum at p,*
- *if $Hess(f)(p)$ is invertible and indefinite, then f has neither a local maximum nor a local minimum at p (i.e., p is a saddle point),*
- *if $Hess(f)(p)$ is non-invertible, no conclusion. In this case, the critical point p is said to be degenerate (another method is needed to determine its nature, i.e., to determine the sign of $f(p+h) - f(p)$ for h very close to 0 in $\mathbb{R}^m$).*

Some examples

Example.

Let f be the function defined on $\mathbb{R}^3$ by $f(x,y,z) = x^2 + y^2 + z^2$. The only critical point of f is $p = (0,0,0)$. Using the Hessian of f, we deduce that f has a strict local minimum at $(0,0,0)$.

Example.

Let f be the function defined on $\mathbb{R}^3$ by $f(x,y,z) = x^3 - 3x + y^2 - 2y + z^2$.

- *The function f has 2 critical points at $(1,1,0)$ and $(-1,1,0)$.*
- *We deduce that since $Hess(f)((1,1,0))$ is positive definite, f has a strict local minimum at $(1,1,0)$.*
- *Also, since $Hess(f)((-1,1,0))$ is invertible and indefinite, f has neither a local maximum nor a local minimum at $(-1,1,0)$ (it is a saddle point).*

Special case of functions with two variables:

For a function of class C^2, let's define

$$c = \frac{\partial f}{\partial x}(\alpha,\beta), \quad d = \frac{\partial f}{\partial y}(\alpha,\beta), \quad \nabla f(\alpha,\beta) = (c,d)^t$$

$$r = \frac{\partial^2 f}{\partial x^2}(\alpha,\beta), \quad s = \frac{\partial^2 f}{\partial x \partial y}(\alpha,\beta), \quad t = \frac{\partial^2 f}{\partial y^2}(\alpha,\beta).$$

Theorem (Monge's theorem)

Let D be an open set in $\mathbb{R}^2$, $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ be a function of class C^2, and $p = (\alpha,\beta) \in D$. Then, (α,β) is a critical point if and only if $c = d = 0$. Furthermore,

- *If $rt - s^2 > 0$ and $r > 0$, then f has a strict local minimum at p.*
- *If $rt - s^2 > 0$ and $r < 0$, then f has a strict local maximum at p.*
- *If $rt - s^2 < 0$, then f does not have an extremum at p (saddle point).*
- *If $rt - s^2 = 0$ (the critical point p is degenerate), no conclusion.*

Course outline

1. Introduction to optimization of multivariable functions
2. Optimization without constraints
3. Optimization with constraints
   - Direct method
   - Lagrange multipliers

Introduction

What is the problematic?

Determine the local extrema of a function f in a set A, where

$$A = \{x \in D \mid g_1(x) = 0, \ldots, g_k(x) = 0\},$$

D is an open subset of $\mathbb{R}^m$, f and $g_1, \ldots, g_k : D \to \mathbb{R}$, $(k \leq m-1)$.

Introduction

Definition

We say that a function f *has a maximum or minimum at* $p \in A$ *under the constraints* $g_1(x) = 0, \dots, g_k(x) = 0$*, if the restriction* $f|_A$ *has a maximum (resp. minimum) at this point* p*, that is,*

$$g_1(p) = \ldots = g_k(p) = 0,$$

$$f(p) = \max_{x \in D, g_1(x) = \ldots = g_k(x) = 0} f(x)$$

or

$$f(p) = \min_{x \in D, g_1(x) = \ldots = g_k(x) = 0} f(x).$$

Introduction

How to do?

To solve this problem, two methods are proposed.

Introduction

How to do?

To solve this problem, two methods are proposed.

1. **Direct method**,
2. **Lagrange multipliers.**

Course outline

Direct method

The case $n = 2$

In this case, let $A = f(x,y) \in U = \{(x,y) \mid g(x,y) = 0\}$, the direct method consists of solving the explicit equation $g(x,y) = 0$ and expressing y as a function of x or vice versa, then substituting into $f(x,y)$. Finally, one must find the extrema of this new function, which is a function of one variable, and go back to f.

Direct method

Example: Find the extrema (if they exist) of f:

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

subject to the constraint $x - 2y = -1$.
Answer: $D_f = \mathbb{R}^2$, f is C^1 on $\mathbb{R}^2$ because it's a polynomial.
We need to determine the extrema of f under the constraint $h(x,y) = x - 2y + 1 = 0$.
We have $x - 2y = -1 \Rightarrow x = 2y - 1$. The problem is reduced to finding the extrema of the function g:

$$g(y) = f(2y-1, y) = y^2 - 2y + 1$$

Direct method

Let's study the variations of g: $g'(y) = 2y - 2$, $g''(y) = 2$, $g'(y) = 0 \Rightarrow y = 1$:

y	$-\infty$	1	$+\infty$
$g'(y)$	$-$	0	$+$
g	$\searrow$	minimum at $y = 1$	$\nearrow$

So, g has a minimum at $y_0 = 1$ (it gives a global minimum). Going back to $x_0 = 2y_0 - 1 = 1$, we obtain that $(1, 1; f(1, 1))$ is a minimum under constraint for f.

Direct method

The case $n = 3$

• If $A = f(x,y,z) \in U = \{g_1(x,y,z) = 0, g_2(x,y,z) = 0\}$, the method consists of solving the system of equations

$$\begin{cases} g_1(x,y,z) = 0 \\ g_2(x,y,z) = 0 \end{cases}$$

and expressing 2 variables in terms of the third, then substituting into $f(x,y,z)$. Finally, one must find the extrema of this new function, which is a function of one variable, and go back to f.

• If $A = f(x,y,z) \in U = \{g(x,y,z) = 0\}$, the method consists of solving the equation $g(x,y,z) = 0$ and expressing 1 variable in terms of the other two, then substituting into $f(x,y,z)$. Finally, one must find the free extrema of this new function, which is a function of two variables, and go back to f.

Direct method

Example: Find the extrema (if they exist) of f:

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

subject to the constraints:

$$A = \{(x,y,z) \in \mathbb{R}^3 \mid x+y+z = 3, z = 1\}$$

Answer: $D_f = \mathbb{R}^3$, f is C^∞ on $\mathbb{R}^3$ because it's a polynomial.
We need to determine the extrema of f under the constraints φ_1 and φ_2

$$\varphi_1(x,y,z) = x+y+z-3, \quad \varphi_2(x,y,z) = z-1.$$

We have:

$$x+y+z-3 = 0, \quad z-1 = 0$$

Direct method

This implies:

$$y = 2 - x, \quad z = 1$$

The problem reduces to finding the extrema of the function g:

$$g(x) = f(x, 2-x, 1) = x^2 + (2-x)^2 + 1 = 2x^2 - 4x + 5$$

Let's study the variations of g: $g'(y) = 4x - 4$, $g'(x) = 0 \Rightarrow x = 1$:

x	$-\infty$	1	$+\infty$
$g'(x)$	$-$	0	$+$
g	$\searrow$	minimum at $x = 1$	$\nearrow$

So, g has a minimum at $x_0 = 1$. Going back to $y_0 = 2 - x_0 = 1$ and $z_0 = 1$, we obtain that $(1, 1, 1; f(1, 1, 1))$ is a minimum for f.

Course outline

Theorem (Lagrange multipliers)

Let f be a real-valued function on U and let a be an element of A. Suppose that f and $\varphi_1, \varphi_2, \ldots, \varphi_p$ are C^1 functions.
If:

- $(a, f(a))$ *is an extremum of* f *under the constraints* $\varphi_1, \varphi_2, \ldots, \varphi_p$,
- *The vectors* $\nabla\varphi_1(a), \nabla\varphi_2(a), \ldots, \nabla\varphi_p(a)$ *are linearly independent,*

Then, there exist p real numbers $\lambda_1, \lambda_2, \ldots, \lambda_p$ and an auxiliary function $F = F(X) = f(X) + \lambda_1\varphi_1(X) + \lambda_2\varphi_2(X) + \ldots + \lambda_p\varphi_p(X)$, such that $a_1, a_2, \ldots, a_n$ and $\lambda_1, \lambda_2, \ldots, \lambda_p$ are solutions of the system:

$$\begin{cases} \frac{\partial F}{\partial x_1}(x_1, \ldots, x_n) = 0 \\ \vdots \\ \frac{\partial F}{\partial x_n}(x_1, \ldots, x_n) = 0 \\ \varphi_1(x_1, \ldots, x_n) = 0 \\ \vdots \\ \varphi_p(x_1, \ldots, x_n) = 0 \end{cases}$$

Lagrange multipliers

Remarks

1) The points $(a_1, a_2, \ldots, a_n)$ resulting from the solutions of (S) are called critical points of f under the constraints $\varphi_1, \varphi_2, \ldots, \varphi_p$.

2) The working method consists of:

- First determining the "doubtful" points such that the vectors: $\nabla\varphi_1(a), \nabla\varphi_2(a), \ldots, \nabla\varphi_p(a)$ are linearly dependent, then rejecting those that do not satisfy the constraints and testing the remaining ones (using the definition method, without forgetting to apply the constraints).
- Then determining the critical points of f - using the Lagrange multipliers - and testing them (using the definition method, without forgetting to apply the constraints).

Lagrange multipliers

Example.
Find the extrema of f:

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

under the constraint φ:

$$\varphi(x,y) = x - 2y + 1$$

using the method of Lagrange multipliers.
Answer. Let $D_f = D_\varphi = \mathbb{R}^2$: f, φ are C^1 functions $\mathbb{R}^2$ as they are polynomials. Search for doubtful points: We search for points (x,y) where the vector $\nabla\varphi$ is linearly dependent, i.e., we search for (x,y) such that $\nabla\varphi(x,y) = (0,0)$.

Lagrange multipliers

We then solve the system:

$$\begin{cases} \frac{\partial\varphi}{\partial x}(x,y) = 0 \\ \frac{\partial\varphi}{\partial y}(x,y) = 0 \end{cases} \longrightarrow \begin{cases} 1 = 0 \\ 2 = 0 \end{cases}$$

In this case, the doubtful points are the critical points of φ. This system has no solution, so there are no doubtful points. Search for critical points: Using Lagrange multipliers, consider the auxiliary function (the Lagrangian):

$$F(x,y) = f(x,y) + \lambda\varphi(x,y) = -x + y^2 + \lambda(x - 2y + 1)$$

If (x,y) gives an extremum of f under the constraint $\varphi(x,y) = 0$, then there exists $\lambda \in \mathbb{R}$ such that $\nabla F(x,y) = (0,0)$.

Lagrange multipliers

Solve the system: (S)

$$\begin{cases} \frac{\partial F}{\partial x}(x,y) = 0 \\ \frac{\partial F}{\partial y}(x,y) = 0 \\ \varphi(x,y) = 0 \end{cases}$$

$$-1 + \lambda = 0 \qquad (1)$$

$$2y - 2\lambda = 0 \qquad (2)$$

$$x - 2y + 1 = 0 \qquad (3)$$

From (1) and (2), we get $\lambda = 1$ and $y = \lambda = 1$. From (3), we get $x = 2y - 1$, so $x = 1$, and thus $(1,1)$ is the only solution to the system (S). We obtain a single critical point: $M = (1,1)$.

Lagrange multipliers

Testing M: To do this, we use the definition and see the sign of $f(x,y) - f(M) = (x,y)^2$.

$$f(1+h_1, 1+h_2) - f(1,1) = -(1+h_1) + (1+h_2)^2 + 1 - 1$$

with

$$(1+h_1) - 2(1+h_2) + 1 = 0$$

i.e.,

$$f(1+h_1, 1+h_2) - f(1,1) = -h_1 + h_2^2 + 2h_2, \quad \text{with} \quad h_1 - 2h_2 = 0$$

Then

$$f(1+h_1, 1+h_2) - f(1,1) = h_2^2 - 0 = h_2^2$$

Therefore, $(1,1)$ is a constrained minimum for f.

Lagrange multipliers

Theorem

Let f be a real-valued function on U and a an element of A: If A is a closed bounded set and f is continuous on A, then f attains its bounds, i.e., it has a maximum value (given by a constrained maximum) and a minimum value (given by a constrained minimum).

Lagrange multipliers

Example: Determine the maximum value of $f(x,y) = x^2 + 2y^2$ on the circle with equation $x^2 + y^2 = 1$:
Answer: $D_f = \mathbb{R}^2$; f is C^1 on $\mathbb{R}^2$ as it is a polynomial. The set $C((0,0);1)$ is a closed bounded set, so f attains its bounds.
We determine now the extrema of f under the constraint $\varphi = \varphi(x,y) = x^2 + y^2 - 1$, we have:
$D_\varphi = \mathbb{R}^2$; φ is C^1 on $\mathbb{R}^2$ as it is a polynomial.
Search for doubtful points: We search for (x,y) such that the vector $\nabla\varphi(x,y)$ is linearly dependent, i.e., we search for solutions of the system:

$$\begin{cases} \frac{\partial \varphi}{\partial x}(x,y) = 0 \\ \frac{\partial \varphi}{\partial y}(x,y) = 0 \\ 2x = 0 \\ 2y = 0 \end{cases}$$

Lagrange multipliers

In this case, the only solution is $(0,0)$, but this point does not satisfy the constraint, since $\varphi(0,0) \neq 0$, so this doubtful point is rejected (i.e., this doubtful point will not give a constrained extremum for f).

Search for critical points: Using Lagrange multipliers, consider the Lagrangian function:

$$F(x,y) = f(x,y) + \lambda\varphi(x,y) = x^2 + 2y^2 + \lambda(x^2 + y^2 - 1)$$

If (x,y) gives a constrained extremum of f under the constraint $\varphi(x,y) = 0$, then there exists $\lambda \in \mathbb{R}$ such that $\nabla F(x,y) = (0,0)$. First, solve the system (S):

$$\begin{cases} \frac{\partial F}{\partial x}(x,y) = 0 \\ \frac{\partial F}{\partial y}(x,y) = 0 \\ \varphi(x,y) = 0 \end{cases}$$

Lagrange multipliers

$$2x + 2\lambda x = 0 \qquad (1)$$
$$4y + 2\lambda y = 0 \qquad (2)$$
$$x^2 + y^2 = 1 \qquad (3)$$

There are four critical points: $(0,1), (0,-1), (1,0), (-1,0)$. Then, according to theorem 2, f attains its bounds, so it is unnecessary to perform the tests, it suffices to calculate:
$f(0,1) = 2, f(0,-1) = 2, f(1,0) = 1$, and $f(-1,0) = 1$. Conclusion: The maximum value of f on the circle with equation $x^2 + y^2 = 1$ is 2.