UNIVERSITAT JAUME I

Bachelor in Computational Mathematics

Final Degree Project

Analysis of extrema problems

Author:
Elías MARTÍNEZ SORIANO

Academic tutor:
Sheldon Miriel Gil Dantas

Presentation date: __ of ________ of 2022
Academic year: 2021/2022

Abstract

This document gathers a systematic analysis of function extrema problems. To do so, a previous introduction to the theory is made to understand and develop the results and examples contained in later sections. Concepts, notations and important results about single and multi variable calculus and general topology are included. Afterwards, the project is fundamentally divided in two sections: a first one focused on analyzing the methods to classify critical points of a function and a second one concentrated in the behavior of the quantity of extrema in single and multi variable functions.

Keywords

Differential calculus, critical point, maximum, minimum, extrema, saddle point.

Contents

Chapter 1

Motivation and objectives

The study of critical points has been an extended topic of discussion in functional and mathematical analysis for a long time. The applications of local and global extrema are a big part of applied mathematics when it comes to optimization problems and in order to provide solutions or approximations to solutions of various equations.

A very useful approach to one of these goals is the variational method, in which the solutions of and equation are sought by finding the critical points of an appropriately chosen function. In other useful instances, the problem of finding solutions of some type of equation is equivalent to the problem of finding critical points of an appropriately chosen function. For example, solutions of the equation $10x + cos\ x - 4 + 3x^2e^{x^3}$ correspond to critical points of the function $f(x) = 5x^2 + sin\ x - 4x + e^{x^3}$. Therefore, if we minimize the function f we can find the solutions of the equation.

Since elementary calculus courses in high school and university, a familiar type of critical point is a local minimum or maximum. In order to find these, we can pick a initial point in the functions, then move to another point at which the function is lower, and repeat this process until the function can no longer be decreased or increased, respectively. It is important to notice that not every point has to be a maximum or a minimum and that not every function has to contain any of them whatsoever. The mentioned variational methods can be used to find critical points that are neither minimums nor maximums of functions.

Another useful feature of solving equations by looking for critical points is that it is sometimes possible to use the existence of one critical point together with other information about the function to prove that there is at least another critical point. However, when the critical point is unique, knowing if the local minimum or maximum behaves as absolute minimum or maximum one can not be assured.

It is in this project where we will approach in more depth these problems. First, we will study a different method from the commonly known Second Derivative Test that is taught in introductory calculus courses to classify critical points and how it differs from it on its applications to both single and multi variable functions. Secondly, we will study how the dimension of the space of the function we are working on can affect the absoluteness of unique critical points. Finally, it will be shown under what conditions we can affirm the existence of additional critical points given one of them exists in functions of two variables.

Chapter 2

Introduction to function extreme problems

In this section will be shown some previous concepts necessary to understand this project. Mainly, we will explain basic concepts and results about single and multi variable calculus and general topology. We will also establish the notation that is going to be used.

The intention of this chapter is to give and introduction and lay the foundations of the objects with which we are going to work in order to achieve the several objectives of the project, and therefore include the convenient statements and results without proving.

References: [10], [9], [8] and [7].

2.1 Notation

Definition 1. *On the n-dimensional Euclidean space* $\mathbb{R}^n$ *the intuitive notion of length of the vector* $x = (x_1, x_2, \ldots, x_n)$ *is captured by the formula*

$$\|\boldsymbol{x}\| := \sqrt{x_1^2 + \cdots + x_n^2}$$

Definition 2. *Let* $\cdot$ *denote the inner product operation in* $\mathbb{R}^n$*.*

Definition 3. *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be a real-valued function having all of its second order partial derivatives continuous in D, a neighborhood centered at c in* $\mathbb{R}^n$*:* $(\alpha \cdot \nabla)^2 f(c) = \sum_{i,j} \alpha_i \alpha_j \frac{\partial^2 f(c)}{\partial x_j \partial x_i}$*, where* α_i *is the ith component of* α *and* ∇ *is the del operator in* $\mathbb{R}^n$*.*

2.2 Concepts and results about single and multi variable functions

Definition 4. *Let c be a number in the domain D of a function* $f : \mathbb{R}^n \to \mathbb{R}$. *Then f(c) is the:*

- *absolute maximum value of* f *on* D *if* $f(c) \geq f(x)$ *for all* x *in* D.
- *absolute minimum value of* f *on* D *if* $f(c) \leq f(x)$ *for all* x *in* D.

An absolute maximum or minimum is sometimes called a global maximum or minimum. The maximum and minimum values f *of are called extreme values of* f.

Definition 5. *A function* $f : \mathbb{R}^n \to \mathbb{R}$ *has a*

- *local maximum at* c *if* $f(x) \leq f(c)$ *for all points* x *in some disk* D *with center* c. *The number* $f(c)$ *is called a local maximum value.*
- *local minimum at* c *if* $f(x) \geq f(c)$ *for all points* x *in some disk* D *with center* c. *The number* $f(c)$ *is called a local minimum value.*
- *saddle point at* c *if* $\forall\ 1 \leq i \leq n$, $\frac{\partial f(c)}{\partial x_i} = 0$ *but has neither a maximum nor a minimum value.*

If the inequalities hold for all points x *in the domain of* f, *then* f *has an absolute maximum or absolute minimum at c, respectively.*

Definition 6. *A critical number of a function* $f : \mathbb{R}^n \to \mathbb{R}$ *is a number c in the domain of* f *such that* $\frac{\partial f(c)}{\partial x_i} = 0\ \forall\ 1 \leq i \leq n$ *or either of the partial derivatives does not exist.*

Theorem 1. *(See p.160, [10])**(First Derivative Test)*** *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a continuous function and suppose that c is a critical point.*

- *If* f' *changes from positive to negative at c, then* f *has a local maximum at c.*
- *If* f' *changes from negative to positive at c, then* f *has a local minimum at c.*
- *If* f' *does not change sign at* f *(for example, if is positive on both sides of c or negative on both sides), then has no local maximum or minimum at c.*

Theorem 2. *If* $f : \mathbb{R}^n \to \mathbb{R}$ *has a local maximum or minimum at c and the first-order partial derivatives of* f *exist there, then* $\frac{\partial f(c)}{\partial x_i} = 0$ *for all* $1 \leq i \leq n$.

Theorem 3. *[10](p. 52, Chapter 1).* ***(The intermediate value theorem).*** *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is continuous on the close interval $[a, b]$ and let N be any number between $f(a)$ and $f(b)$, where $f(a) \neq f(b)$. Then there exists a number c in (a, b) such that $f(c) = N$.*

Theorem 4. ***(Rolle's Theorem).*** *Let $f : \mathbb{R} \to \mathbb{R}$ be a function that satisfies the following three hypotheses:*

(i) *f is continuous on the closed interval $[a, b]$.*

(ii) *f is differentiable on the open interval (a, b).*

(iii) *$f(a) = f(b)$*

Then there is a number c in (a, b) such that $f'(c) = 0$.

Theorem 5. ***(Mean Value Theorem).*** *Let $f : [a, b] \to \mathbb{R}$ be a continuous function on the closed interval $[a, b]$, and differentiable on the open interval (a, b), where $a < b$. Then there exists some c in (a, b) such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Theorem 6. ***(Fermat's Theorem).*** *If $f : \mathbb{R}^n \to \mathbb{R}$ has a local maximum or minimum at c, and if $\frac{\partial f(c)}{\partial x_i}$ exist for all $1 \leq i \leq n$, then $\frac{\partial f(c)}{\partial x_i} = 0$ for all $1 \leq i \leq n$.*

Fermat's Theorem can be rephrased combining Definition 6 with Theorem 6:

Corolary 6.1. *If f has a local maximum or minimum at c, then c is a critical number of f.*

Theorem 7. ***(Weierstrass' extreme value Theorem).*** *If a real-valued function $f : \mathbb{R}^n \to \mathbb{R}$ is continuous on the closed interval $[a, b]$, then f must attain a maximum and a minimum, each at least once. That is, there exist numbers c and d in [a,b] such that:*

$$f(c) \geq f(x) \geq f(d) \quad \forall x \in [a, b]$$

Definition 7. *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is a function taking as input a vector $x \in \mathbb{R}^n$ and outputting a scalar $f(x) \in \mathbb{R}$. If all second partial derivatives of f exist, then the Hessian matrix $\boldsymbol{H}$ of f is a square $n \times n$ matrix, defined and arranged as follows:*

$$\boldsymbol{H}_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Each of the coefficients of the matrix can be stated in the following way:

$$(\boldsymbol{H}_f)_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

Definition 8. *[3] A function $f : X \to Y$ between two topological spaces is proper or boundedly compact if the inverse image $f^{-1}(K)$ of every compact set $K \subset Y$ is compact in X.*

2.3 Topology concepts and results

Definition 9. *A metric space (X, d) is a non empty set X which over a function $d : X \times X \to \mathbb{R}$ is defined such that for all $x, y, z \in X$ it holds that:*

(i) *Non negativity:* $d(x, y) \geq 0$.

(ii) *Non degenerated:* $d(x, y) = 0$ *if and only if* $x = y$.

(iii) *Symmetry:* $d(x, y) = d(y, x)$.

(iv) *Triangle inequality:* $d(x, z) \leq d(x, y) + d(y, z)$.

This function is called metric or distance.

Definition 10. *A nonempty subset $A \subset (X, d)$ is said to be open if for any $x \in A$, there is $\delta > 0$ such that $B_\delta(x) \in A$. The empty set is open.*

Definition 11. *A subset $C \subset (X, d)$ is said to be closed if every sequence $(x_n)_{n=1}^{\infty} \subset C$ that converges, converges to an element of C. In symbols, if $\lim_{x\to\infty} x_n = x_0$.*

Definition 12. *Let X be a set. We say that a collection $\{U_i : i \in I\}$ of subsets of X is a cover of X when $X = \bigcup_{i\in I} U_i$.*

If X is a set and $\{U_i : i \in I\}$ is a cover of X, we say that $\{U_{i1}, \dots, U_{in}\}$ is a finite subcover of X when $X = U_{i_1} \cup \dots \cup U_{i_n} = \bigcup_{k=1}^{n} U_{i_k}$.

Theorem 8. ***(Heine-Borel Theorem).*** *A subset $K \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Chapter 3

A first derivative test for functions of several variables

When it comes to trying to analyze the critical points of a function, many first differential calculus courses in mathematics or engineering at university tend to teach the same following criterion:

Theorem 9. *(See p.369, [6]).* ***(Second Derivative Test).*** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real-valued function having all of its second order partial derivatives continuous in D, a neighborhood centered at c in $\mathbb{R}^n$, and let c be a critical point of f. Then at a the function f has*

(i) a local minimum at c if $(\alpha \cdot \nabla)^2 f(c) > 0$ for every unit vector α;

(ii) a local maximum at c if $(\alpha \cdot \nabla)^2 f(c) < 0$ for every unit vector α.

(iii) a saddle point at c if $(\alpha \cdot \nabla)^2 f(c)$ can change sign.

If $(\alpha \cdot \nabla)^2 f(c) = 0$, the test gives no information: f could have a local maximum or local minimum at c, or c could be a saddle point of f.

In two dimensions the conditions can be rewritten as:

(i) $f_{xx}(c) > 0, f_{xx}(c)f_{yy}(c) - (f_{xy}(c))^2 > 0$ (minimum);

(ii) $f_{xx}(c) < 0, f_{xx}(c)f_{yy}(c) - (f_{xy}(c))^2 > 0$ (maximum);

(iii) $|\mathbf{H}_f| = f_{xx}(c)f_{yy}(c) - (f_{xy}(c))^2 < 0$ (saddle point).

What we will now show is an alternative to the previous method, but one that only uses first derivatives. Consider the following theorem:

Theorem 10. *(See p.558-559,[4]). Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real-valued function which is continuous on a neighborhood D centered at a critical point c in $\mathbb{R}^n$ and differentiable on $D \setminus \{c\}$. Then*

(i) *f has a local maximum value at c if $(x-c) \cdot \nabla f(x) < 0 \ \forall \ x \in D \setminus \{c\}$;*

(ii) *f has a local minimum value at c if $(x-c) \cdot \nabla f(x) > 0 \ \forall \ x \in D \setminus \{c\}$;*

Proof. Since the proofs of both parts are analogous, only (i) will be proven. Suppose that c is a critical point of f and $(x-c) \cdot \nabla f(x) < 0$ on $D \setminus \{c\}$ and let $x \in D \setminus \{c\}$. Our aim is to find a local maximum value at c of f. Define $F : [0,1] \to \mathbb{R}$ such that $F(t) = f(c + t(x-c))$.

f is continuous on $[0,1]$ because it is composed of the f function, which is continuous and real valued. It is also differentiable along $(0,1)$:

$$F'(t) = (x-c) \cdot \nabla f(c + t(x-c))$$

Realize that conditions for the Mean Value Theorem [5] comply, therefore:

$$\exists t_0 \in (0,1) / F'(t_0) = \frac{F(1) - F(0)}{1 - 0} = F(1) - F(0)$$

which implies that

$$F'(t_0) = (x-c) \cdot \nabla f(c + t_0(x-c)) = f(c + 1(x-c)) - f(c + 0(x-c)) = f(x) - f(c)$$

i.e.

$$f(x) - f(c) = (x-c) \cdot \nabla f(c + t_0(x-c)) = \frac{1}{t_0}(t_0(x-c))\nabla f(c + t_0(x-c))$$

Now let $y = c + t_0(x-c)$, we have that $y - c = t_0(x-c)$.

Notice that $x - c \neq 0$, because $x \in D \setminus \{c\}$, and that $t_0 \in (0,1)$, therefore

$$0 < ||y-c|| = t_0||x-c|| < ||x-c|| < r$$

being r the radius of D, so $y \in D \setminus \{c\}$. The previous equation can be rewritten in the following way:

$$f(x) - f(c) = \frac{1}{t_0}(y - c) \cdot \nabla f(y)$$

Since it has been seen that $\frac{1}{t_0}(y-c) \cdot \nabla f(y) = \frac{1}{t_0}(t_0(x-c))\nabla f(c + t_0(x-c))$, following the initial assumption:

$$\frac{1}{t_0}(y - c) \cdot \nabla f(y) < 0$$

Therefore it holds that

$$f(x) - f(c) < 0 \Rightarrow f(x) < f(c) \forall x \in D \setminus \{c\}$$

which means that f has a local maximum value at c.

□

As an example to illustrate the functioning of this theorem, consider the following function:

Example 10.1. *[4]* $f(x, y) = x^2 + 2xy + 3y^2 + 2x + 10y + 9$

First, we find the possible critical points:

$$\begin{cases} \dfrac{\partial f(x,y)}{\partial x} = 2x + 2y + 2 \\ \dfrac{\partial f(x,y)}{\partial y} = 2x + 6y + 10 \end{cases}$$

Now we try to solve the following system with the equations above:

$$\begin{cases} 2x + 2y + 2 = 0 \\ 2x + 6y + 10 = 0 \end{cases}$$

Which results with the solution $(x, y) = (1, -2)$*, being the only critical point of the function.*

Applying the theorem that was just proved:

$$(x-1,y+2)\cdot\nabla f(x,y) = (x-1,y+2)\cdot(2x+2y+2,2x+6y+10) = 2(x-1)(x+y+1)+2(y+2)(x+3y+5)$$

Now, making a change of variable as the following:

$u = x-1$ *and* $v = y+2$*; the above expression becomes:*

$(u,v)\cdot\nabla f(u,v) = 2u(u+1+v-2+1)+2v(u+13v-6+5) = 2u^2+4vu+6v^2 = 2(u^2+2vu+3v^2) = 2\left((u+v)^2+2v^2\right) > 0\,\forall(u,v)\neq(0,0).$

Therefore $(x-1,y+2)\cdot\nabla f(x,y) > 0$ *for all* $(x,y)\neq(1,-2)$*, which means that* f *has a local minimum value at* $(1,-2)$*.*

We can visualize the function in the following 3D plot and its global minimum:

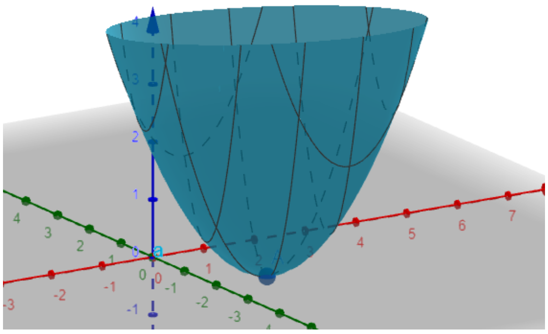

Figure 3.1: $f(x,y) = x^2+2xy+3y^2+2x+10y+9$.

Next, it will be shown that Theorem 10 is a more general method to classify critical points than Theorem 9, meaning that if a critical point can be classified using the latter, then it can also be classified using Theorem 10.

To do so, take the hypothesis of Theorem 9 and the result from (ii) of Theorem 10, respectively such as:

Proposition 10.1. *[4] Let* $f:\mathbb{R}^n\to\mathbb{R}$ *be a real-valued function having all of its second partial derivatives continuous in* D*, being* D *a neighborhood at* c *in* $\mathbb{R}^n$*, being* c *a critical point of* f*; suppose* $(\alpha\cdot\nabla)^2 f(c) > 0$ *for all unit vectors* α*. Then* f *has a local minimum value at* c *if* $(x-c)\cdot\nabla f(x) > 0\ \forall\ x\in D\setminus\{c\}$*. That is to say,* (ii) *of Theorem 10 holds, which leads to* f *having a local minimum value at* c*.*

Proof. Let

$$m := min\left\{(\alpha\cdot\nabla)^2 f(c)/\alpha \text{ is a unit vector}\right\}$$

Since $(\alpha\cdot\nabla)^2 f(c)$ is a continuous function of α on the closed, bounded set $\{\alpha : ||\alpha|| = 1\}$ in $\mathbb{R}^n$, by the Weierstrass Theorem [7], m exists.

Now, by hypothesis, $(\alpha\cdot\nabla)^2 f(c) > 0$, therefore, $m > 0$.

Next, it will be shown that the following expression is true:

$$\left|(\alpha\cdot\nabla)^2 f(x) - (\alpha\cdot\nabla)^2 f(c)\right| \leq \sum_{i,j}\left|\frac{\partial^2 f(x)}{\partial x_j\partial x_i} - \frac{\partial^2 f(c)}{\partial x_j\partial x_i}\right|$$

First, by definition:

$$\left|(\alpha\cdot\nabla)^2 f(x) - (\alpha\cdot\nabla)^2 f(c)\right| = \left|\sum_{i,j}\alpha_i\alpha_j\frac{\partial^2 f(x)}{\partial x_j\partial x_i} - \sum_{i,j}\alpha_i\alpha_j\frac{\partial^2 f(c)}{\partial x_j\partial x_i}\right|$$

Then:

$$\left|\sum_{i,j}\alpha_i\alpha_j\frac{\partial^2 f(x)}{\partial x_j\partial x_i} - \sum_{i,j}\alpha_i\alpha_j\frac{\partial^2 f(c)}{\partial x_j\partial x_i}\right| = \left|\sum_{i,j}\alpha_i\alpha_j\Big(\frac{\partial^2 f(x)}{\partial x_j\partial x_i} - \frac{\partial^2 f(c)}{\partial x_j\partial x_i}\Big)\right|$$

By applying the triangular inequality and the fact that α is a unit vector:

$$\left|\sum_{i,j}\alpha_i\alpha_j\Big(\frac{\partial^2 f(x)}{\partial x_j\partial x_i} - \frac{\partial^2 f(c)}{\partial x_j\partial x_i}\Big)\right| \leq \sum_{i,j}\left|\alpha_i\alpha_j\right|\left|\frac{\partial^2 f(x)}{\partial x_j\partial x_i} - \frac{\partial^2 f(c)}{\partial x_j\partial x_i}\right| \leq \sum_{i,j}\left|\frac{\partial^2 f(x)}{\partial x_j\partial x_i} - \frac{\partial^2 f(c)}{\partial x_j\partial x_i}\right|$$

In addition with the fact that all second order partial derivatives of f are continuous at c, exists $\delta > 0$ such that

$$\left|(\alpha\cdot\nabla)^2 f(x) - (\alpha\cdot\nabla)^2 f(c)\right| < \frac{m}{2}$$

if $||x-c|| < \delta$ and α is any vector. Using the absolute value property:

$$-\frac{m}{2} < (\alpha \cdot \nabla)^2 f(x) - (\alpha \cdot \nabla)^2 f(c) < \frac{m}{2}$$

The last expression implies that:

$$(\alpha \cdot \nabla)^2 f(x) > (\alpha \cdot \nabla)^2 f(c) - \frac{m}{2} > 0$$

But by definition of m, $(\alpha \cdot \nabla)^2 f(c) \geq m$, so the above expression complies that:

$$(\alpha \cdot \nabla)^2 f(x) > (\alpha \cdot \nabla)^2 f(c) - \frac{m}{2} \geq m - \frac{m}{2} = \frac{m}{2} > 0$$

if $||x-c|| < \delta$ and α is a unit vector.

Therefore $(\alpha \cdot \nabla)^2 f(c) < 0$ if $||x-c|| < \delta$ and α is a unit vector.

Now let x be any point such that $0 < ||x-c|| < \delta$. Let, as it was done in the proof of Theorem 10, $F(t) = f(c + t(x-c))$ for $0 \leq t \leq 1$. Obviously notice that F can be defined on some open interval containing $[0,1]$. With its derivative being:

$$F'(t) = (x-c) \cdot \nabla f(c + t(x-c))$$

and its second derivative:

$$F''(t) = ((x-c) \cdot \nabla)^2 f(c + t(x-c))$$

for $0 \leq t \leq 1$. Letting $\alpha = \frac{x-c}{||x-c||}$, then it holds that $x - c = \alpha \cdot ||x-c||$, which means that:

$$F''(t) = ||x-c||^2 (\alpha \cdot \nabla)^2 f(c + t(x-c)) > 0$$

Due to the fact that the second derivative is positive in an interval, $F'(t)$ is increasing, and since $F'(0) = (x-c) \cdot \nabla f(c + 0(x-c) = (x-c) \cdot \nabla f(c)$, being c a critical point, then $F'(0) = (x-c) \cdot 0 = 0$, which means $F'(t) > 0$ for $t \in [0,1]$. Also, since $F'(1) = (x-c) \cdot \nabla f(c + 1(x-c)) = (x-c) \cdot \nabla f(x)$, $(x-c) \cdot \nabla f(x) > 0$ for $0 < ||x-c|| < \delta$. Which means that our hypothesis is verified and therefore f has a local minimum value at c. □

Thus, it has been shown that if Theorem 9 can be used to conclude that a critical point is a local minimum point, then Theorem 10 can also be used to draw this conclusion. Same thing obviously happens for local maximum points.

On the other hand, it is very important to know that the criteria above are not equivalent, meaning that is possible to give examples of functions for which Theorem 9 fails and yet Theorem 10 yields a conclusion. This can be seen within the next examples:

Example 10.2. $g(x,y) = 1 - x^{2/3} - y^{4/5}$

We look for the critical points:

$$\begin{cases} \dfrac{\partial g(x,y)}{\partial x} = -\dfrac{2}{3\sqrt[3]{x}} \\ \dfrac{\partial g(x,y)}{\partial y} = -\dfrac{4}{5\sqrt[5]{y}} \end{cases}$$

Note that $g_x(x,y) \neq 0$ and $g_y(x,y) \neq 0$ for all points of the domain of g $D = \mathbb{R}^2$, but we can see that both partial derivatives do not exist at $(x,y) = (0,0)$, thus, following the definition 6, this is the only critical point of the function $g(x,y)$.

We are going to see that Theorem 10 can not be applied in this case. We calculate the second derivatives in order to apply Theorem 9:

$$\frac{\partial^2 g(x,y)}{\partial x^2} = \frac{2}{9\sqrt[3]{x^4}}$$

$$\frac{\partial^2 g(x,y)}{\partial y^2} = \frac{4}{25\sqrt[5]{y^6}}$$

$$\frac{\partial^2 g(x,y)}{\partial x \partial y} = \frac{\partial^2 g(x,y)}{\partial y \partial x} = 0$$

Now, we can see that $g_{xx}(x,y)$ and $g_{yy}(x,y)$ do not exist at $(0,0)$, which means we can not use Theorem 9.

Let us check now how the second criterion we just proved provides a solution.

Being $(0,0)$ *the critical point:*

$$(x-0,y-0)\cdot\nabla g(x,y)=(x,y)\cdot\left(-\frac{2}{3\sqrt[3]{x}},-\frac{4}{5\sqrt[5]{y}}\right)=-\frac{2x}{3\sqrt[3]{x}}-\frac{4y}{5\sqrt[5]{y}}=$$
$$=-\frac{2\sqrt[3]{x^2}}{3}-\frac{4\sqrt[5]{y^4}}{5}<0\ \forall\ (x,y)\in A\subset\mathbb{R}\setminus(0,0)$$

Being A *a disk around* $(0,0)$.

And therefore, by Theorem 10, $g(x,y)$ *has a local maximum value at* $(0,0)$.

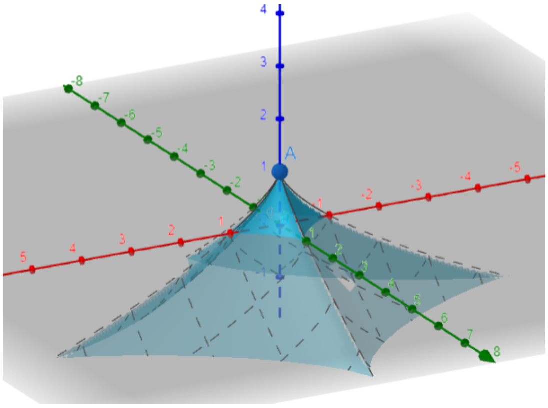

Figure 3.2: $g(x,y)=1-x^{2/3}-y^{4/5}$

Example 10.3. $h(x,y)=25+(x-y)^4+(x-1)^4$

$$\begin{cases}\dfrac{\partial h(x,y)}{\partial x}=4(x-y)^3+4(x-1)^3\\ \dfrac{\partial h(x,y)}{\partial y}=-4(x-y)^3\end{cases}$$

$$\begin{cases}4(x-y)^3+4(x-1)^3=0\\ -4(x-y)^3=0\end{cases}$$

Solving the above system we obtain that $x=y$, *therefore the only critical point is* $(1,1)$.

Calculating second derivatives:

$$\frac{\partial^2 h(x,y)}{\partial x^2} = 12(x-y)^2 + 12(x-1)^2$$

$$\frac{\partial^2 h(x,y)}{\partial y^2} = 12(x-y)^2$$

$$\frac{\partial^2 h(x,y)}{\partial x \partial y} = \frac{\partial^2 h(x,y)}{\partial y \partial x} = -12(x-y)$$

When evaluating them at $(1,1)$*:*

$$\frac{\partial^2 h(1,1)}{\partial x^2} = 0$$
$$\frac{\partial^2 h(1,1)}{\partial y^2} = 0$$
$$\frac{\partial^2 h(x,y)}{\partial x \partial y} = 0$$

Thus:

$$h_{xx}((1,1))h_{yy}((1,1)) - h_{xy}((1,1))^2 = 0$$

Providing us with no information about the critical point of the function, according to Theorem 9. Let us try to apply Theorem 10 once again:

$$(x-1, y-1) \cdot \nabla h(x,y) = (x-1, y-1) \cdot (4(x-y)^3 + 4(x-1)^3, -4(x-y)^3) =$$
$$= 4(x-y)^3(x-1) + 4(x-1)^4 - 4(x-y)^3(y-1) = 4(x-y)^3(x-y) + 4(x-1)^4 =$$
$$= 4(x-y)^4 + 4(x-1)^4 > 0 \;\forall\; (x,y) \in D \setminus (1,1).$$

And therefore, by Theorem 10, $h(x,y)$ *has a local minimum value at* $(1,1)$.

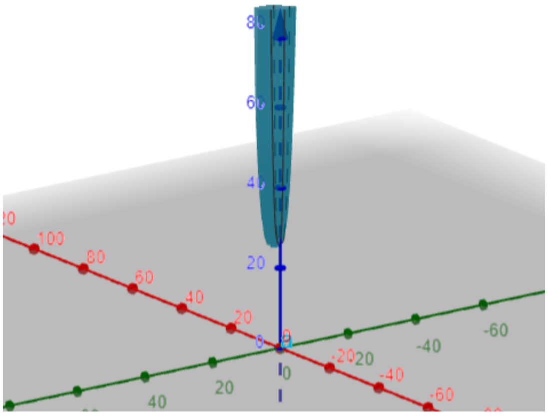

Figure 3.3: $h(x, y) = 25 + (x - y)^4 + (x - 1)^4$

In practice, when Theorem 9 is applicable, it may often be easier to use than Theorem 10. The latter theorem should be of interest because it extends the First Derivative Test (Theorem 1) from single variable functions to functions of several variables.

Chapter 4

Absoluteness and uniqueness of critical points

4.1 Non extreme unique critical points

Suppose we have a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ with a unique critical point at c that is a local maximum or local minimum.

It may be intuitive to assume that since c is a unique critical point, then it must be an absolute one, but we will check that this is not true. We will show later that in fact we can find examples where a unique local maximum or minimum exists which is not absolute maximum or minimum [11.1],[11.2], [11.3]. Naturally this leads to the question: for which values of n is the local minimum necessarily an absolute minimum or maximum?

Let us first check that the previous statement is indeed true for $n = 1$:

Theorem 11. *[5]. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuously differentiable function. If c is the only critical point of f in $I \subset \mathbb{R}$ and c is a local minimum (or maximum), then in fact c is a global minimum (or maximum).*

Proof. Since f has a local minimum at c, let $m = f(c)$, i.e. $f(x) > m$ for all $|x| < \delta \in \mathbb{R}$, with $\delta > 0$.

If c is not an absolute minimum then we must have $f(b) < m$ for some b, but then $f(\delta/2) > m$, $f(b) < m$ and the intermediate value theorem [3] gives $f(d) = m$ for some $d \in (\delta/2, b)$.

But $f(c) = f(d) = m$, so by Rolle's Theorem [4], exists $a \in (m, d)$ with $f'(a) = 0$ and by definition of critical point f has at least one more critical point, but this contradicts our first assumption, therefore c has to be an absolute minimum (or maximum). □

The proof for when c is a local maximum is analogous.

We ask ourselves if the same rule is true for functions defined on higher dimensions. Let us consider the following examples of functions with two variables:

Example 11.1. $f(x, y) = x^2 + y^2(1-x)^3$

First we find the critical points through solving the system provided by its partial derivatives:

$$\begin{cases} \dfrac{\partial f(x,y)}{\partial x} = 2x - 3(1-x)^2y^2 \\ \dfrac{\partial f(x,y)}{\partial y} = -2(x-1)^3 \end{cases}$$

$$\begin{cases} 2x - 3(1-x)^2y^2 = 0 \\ -2(x-1)^3y = 0 \end{cases}$$

Solving the system we see that only critical point is $(0, 0)$*.*

Applying now the second derivative test:

We evaluate them at $(0, 0)$*:*

$$\frac{\partial^2 f(0,0)}{\partial x^2} = 2$$
$$\frac{\partial^2 f(0,0)}{\partial y^2} = 2$$
$$\frac{\partial^2 f(0,0)}{\partial x \partial y} = \frac{\partial^2 f(0,0)}{\partial y \partial x} = 0$$

Thus:

$$f_{xx}((0,0)) = 2 > 0$$
$$|\boldsymbol{H}_f| = f_{xx}((0,0))f_{yy}((0,0)) - f_{xy}((0,0))^2 = 4 > 0$$

Therefore, by the second derivative test, $(0, 0)$ *is a local minimum.*

We have seen that f has a unique local minimum value at $(0,0)$, where $f(0,0) = 0$, but is not an absolute minimum value since $f(4,1) = -11$.

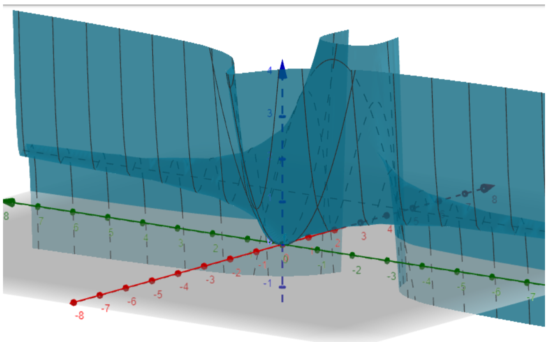

Figure 4.1: $f(x,y) = x^2 + y^2(1-x)^3$

Example 11.2. *[2] $g(x,y) = \frac{-1}{1+x^2} + (2y^2 - y^4)\left(e^x + \frac{1}{1+x^2}\right)$*

To calculate the critical points in this case, we will modify the usual method to make it easier, instead of calculating both partial derivatives directly.

Consider first the sections of the function when $x =$ constant, using the variable change $\frac{1}{1+x^2} = a$ and $e^x = b$, which results in:

$$g = g(y) = -a + (2y^2 - y^4)(b + a)$$

Calculating its critical points:

$$g'(y) = (b + a)(4y - 4y^3)$$

$$(b + a)(4y - 4y^3) = 0$$

With the solutions $y = 0, 1, -1$. Now we will consider $g(x,y)$ with the variable $y =$ constant but only with the three values we have obtained:

$$g(x,-1) = g(x,1) = e^x$$

Which is always positive, and:

$$g(x,0) = \frac{-1}{1+x^2}$$

Where we can find its critical points:

$$\frac{\partial g}{\partial x}(x,y) = \frac{2x}{(1+x^2)^2}$$

$$\frac{2x}{(1+x^2)^2} = 0$$

The only solution of the equation is $x = 0$.

Therefore, g has a unique critical point, $(0,0)$. Let us try to classify it.

Notice that $f(0,0) = -1$ and that we have that:

$$g(x,y) = \frac{-1}{1+x^2} + (2y^2 - y^4)\Big(e^x + \frac{1}{1+x^2}\Big) = -1 + \frac{x^{-2}}{1+x^2} + (2y^2 - y^4)\Big(e^x + \frac{1}{1+x^2}\Big) =$$

$$= f(0,0) + \left\{ \frac{x^2}{1+x^2} + y^2(2-y^2)\Big(e^x + \frac{1}{1+x^2}\Big) \right\}$$

Now, if (x,y) is in the unit disc around the critical point $(0,0)$, then the value of the curly brackets is always positive unless $(x,y) = (0,0)$, which means that $(0,0)$ is a local minimum.

We know therefore know that g is a differentiable function that has a unique local minimum at $(0,0)$. Since $g(0,0) = -1 > -17 = g(0,2)$, it remains that $(0,0)$ is indeed a local minimum but not an absolute minimum.

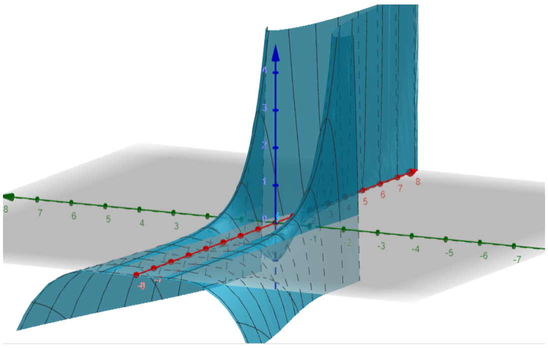

Figure 4.2: $g(x,y) = \frac{-1}{1+x^2} + (2y^2 - y^4)\left(e^x + \frac{1}{1+x^2}\right)$

Example 11.3. *[5] A more general counterexample can be found considering a polynomial of higher order. Let, for $n > 1$, the following fifth degree polynomial in $x = (x_1, \dots, x_n)$:*

$$f(x) = x_n^2 + \sum_{i=1}^{n-1} x_i^2 (1+x_n)^3 = \sum_{i=2}^{n} x_i^2 + H$$

where $H = H(x_1, \dots, x_n)$ consists of terms of higher order.

When calculating the partial derivatives of f we have:

$$\frac{\partial f(x)}{\partial x_i} = 2x_i(1+x_n)^3, \ 1 \leq i \leq n-1$$

$$\frac{\partial f(x)}{\partial x_n} = 2x_n + 3\sum_{i=1}^{n-1} x_i^2 (1+x_n)^2$$

We proceed to solve the equation system provided by the partial derivatives:

$$\begin{cases} 2x_i(1+x_n)^3 = 0, \ 1 \leq i \leq n-1 \\ 2x_n + 3\displaystyle\sum_{i=1}^{n-1} x_i^2(1+x_n)^2 = 0 \end{cases}$$

Notice in the first equality of the system that the possible roots for each variable are the following:

$$x_i = 0$$

$$x_n = 0 \text{ or } x_n = -1$$

When substituting the values in the second expression we see that the only valid one is $x_n = 0$, meaning it provides the following solution:

$$x_i = 0$$

$$x_n = 0$$

And therefore the function possesses a single critical point for any given $n \geq 0$: $0 = (0, \ldots, 0)$

Applying again the first derivative test [10]:

$$(x_1-0, x_2-0, \ldots, x_n-0)\cdot\nabla f(x_1, x_2, \ldots, x_n) = 2x_1(1+x_n)^3+2x_2(1+x_n)^3+\cdots+2x_n+3\sum_{i=1}^{n-1} x_i^2(1+x_n)^2$$

Which for all values around a disk $D \setminus (0, \ldots, 0)$ it holds that:

$$(x_1 - 0, x_2 - 0, \ldots, x_n - 0) \cdot \nabla f(x_1, x_2, \ldots, x_n) > 0$$

so by the first derivative test, $0 = (0, \ldots, 0)$ is local minimum value.

But at the same time we see that because of f contains values of a cubic polynomial, its range is all real numbers and therefore no absolute minimum exists.

We can conclude that this criterion does not hold for multi variable functions, i.e., only with one variable functions we can assure that a unique critical that is a local maximum or minimum is an absolute maximum or minimum. Very common functions like example [11.1] have this characteristic.

4.2 Two variable functions with non unique critical points

In the previous section we have studied the behaviour of functions that contain a single critical point. Knowing how they behave in that regard we can ask ourselves the next question:

Suppose that $f : \mathbb{R}^2 \to \mathbb{R}$ is continuously differentiable and suppose f has a local but non-global minimum at c. Are there any conditions that guarantee the existence of additional critical points?

Consider the following theorem:

Theorem 12. *[2] Let $f : \mathbb{R}^2 \to \mathbb{R}$ be continuously differentiable and has a local, non-global minimum. If, further, f is proper [8] then f must have at least one additional critical point.*

Instead of proving this theorem directly we will use a different approach, a good reason for this is that, as we have seen, the theorem requires the function to be proper [8] in order to state that it has more than one critical point, which can be a high requirement to meet. In fact, there are very common functions which are not proper, such as the following polynomial:

Example 12.1. *[3] $g(x, y) = x^3 - 3xy + y^3$*

Consider first that the following can be proved:

Proposition 12.1. *If $f : X \to Y$ is a proper map, then, for each closed subset $C \subset X$, the image $f(C) \subset Y$ is also closed.*

Therefore, in order to proof that the function g is not proper, it will be sufficient to find a closed subset $C \subset \mathbb{R}^2$ for which $g(C) \subset \mathbb{R}$ is not closed.

ENCONTRAR CONJUNTO

Now that we have checked that theorem 12 can not be applied, let us find the critical points of the function:

$$\begin{cases} \dfrac{\partial g(x,y)}{\partial x} 3x^2 - 3y \\ \dfrac{\partial g(x,y)}{\partial y} 3y^2 - 3x \end{cases}$$

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

The system yields to a solution with two critical points: $(0, 0)$ *and* $(1, 1)$.

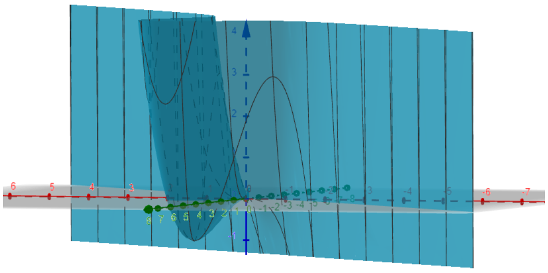

Figure 4.3: $g(x, y) = x^3 - 3xy + y^3$

The result of this last example [12.1] could not have been obtained using the theorem above [12], an analysis had to be made, which can be sometimes not so straight forward, as we previously saw in example 11.2.

Consider first the two following definitions:

Definition 13. *(p. 278, [3]) Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function. A sequence of points x_n such that*

(i) *$f(x_n)$ is bounded and*

(ii) $\lim_{n\to\infty} ||\nabla f(x_n)|| = 0$.

is called a Palais-Smale sequence for f.

Moreover:

Definition 14. *If $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable, we say that f satisfies the Palais–Smale condition if every Palais-Smale sequence of f has a convergent subsequence.*

As said before example 12.1, we will use a different approach to prove theorem 12. We can check that the function in [12.1] satisfies the Palais-Smale condition PROBAR!! [13], but is not proper, i.e., the Palais-Smale condition is weaker than the condition of the function being proper.

This leads to the main result of the section:

Lema 12.1. ***Mountain Pass Lemma*** *(Ambrosetti and Rabinowitz). Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable, and satisfies the Palais-Smale condition [14]. If f has a local but non-global minimum, then f has at least one other critical point.*

Proof. DEMOSTRACIÓN DEL TEOREMA POR HACER □

Therefore, if we can prove the fact that if f is proper, then f will satisfy the Palais-Smale condition, then the lema 12.1 can be applied not only to functions that comply with the Palais-Smale condition [13], but also to those who are proper, as we firstly intended with theorem 12.

Theorem 13. *If $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable proper function, then, it satisfies the Palais-Smale condition.*

Proof. □

Notice however, as we have checked with example 12.1 that the converse is not true.

We can conclude that functions which comply with the Palais-Smale condition can garantee us the existence of more critical points, without the necessity of them being proper.

Chapter 5

Conclusions

Having finished my Final Degree Project I can say that I am satisfied with the result. Through all these months I have been able to realize the huge amount of theory behind critical points, and their importance in the world of mathematics and their applications. Also, it has been a great opportunity to improve my ability in using LaTeXand the redaction of mathematical documents for the future. Although I had experience with it before, I had never written a document of this complexity and I am glad with how it turned out.

Regarding the contents of the project, for me it has been an interesting topic of study due to the fact that during the whole degree critical points are shown in a very introductory level. Furthermore, I have been able to use several concepts, knowledge and techniques that I saw in previous courses during the degree, such as the subjects Foundations of Differential Calculus and Topology, which I found gratifying.

As a future line of study, we could consider different methods of finding and classifying critical points or investigating under what conditions we can affirm that local extrema points behave as absolute maximums or minimums in functions of multiple variables.

Finally, I consider in high regard that the document can be of great interest since it can be approached by even undergraduate students without any previous knowledge further than basic mathematical concepts and can give them some new methods and criteria that could be useful in the future.

Bibliography

[1] A.V. Arkhangelskii and L.S. Pontryagin. *General Topology I.* Springer-Verlag, 1990.

[2] Ash, J. Marshall, and Harlan Sexton. A surface with one local minimum. *Mathematics Magazine*, 58(23):147–49, 1985.

[3] James Bisgard. Mountain passes and saddle points. *SIAM Review*, 57(3):275–292, 2015.

[4] Michael W. Botsko. A first derivative test for functions of several variables. *The American Mathematical Monthly*, 93(7):558–561, 1986.

[5] J. Arias de Reyna and Ole Jorsboe. Non extreme unique critical points. *The American Mathematical Monthly*, 93(4):307, 1986.

[6] Watson Fulks. *Advanced Calculus.* Wiley, 1978.

[7] Jorge Galindo Pastor. *Compactness.* Jaume I University, 2020. Unit 7.

[8] Jorge Galindo Pastor. *The topology of a metric space.* Jaume I University, 2020. Unit 2.

[9] James Stewart. *Multivariable Calculus, Early Transcendentals.* Cengage Learning, 2007. Chapter 14.

[10] James Stewart. *Single Variable Essential Calculus.* Cengage Learning, 2012. Chapter 3.