Notes on Min/Max

Kevin Yeh

Boston College, Autumn 2020

Contents

1 Some Basic Topology

Here I list the precise definitions of certain properties of subsets of $\mathbb{R}^n$. They are a little more technical, and are not really part of what you are required to know, so feel free to only skim them or skip them. However, you should still have some intuition for what they mean, and be able to identify them on sets which we often encounter. You could practice this by looking at familiar subsets of $\mathbb{R}^n$, such as a circle, disc, square, rectangle, lines, intervals, planes, etc. and try to identify their relevant properties listed below.

Definition 1.1 (closed, limit definition)**.** A subset $A \subset \mathbb{R}^n$ of $\mathbb{R}^n$ is *closed* if for any convergent sequence of points

$$\{x_n\}_{n=1}^{\infty} \subset A$$

in A, the limit of the sequence is in A:

$$\lim_{n\to\infty} x_n \in A.$$

In other words, the subset A is *closed* if for any sequence

$$x_1, x_2, x_3, \ldots$$

of points in A (so $x_i \in A$ for all i) that converges, i.e. the limit of the sequence exists, the limit point lies in A.

Yet another way to say this is: A subset $A \subset \mathbb{R}^n$ is closed if it contains all of its limit points.

This definition is kind of abstract, so here's another way to think about a closed set: a set is closed if it contains all of its boundary. A boundary of a set, intuitively, is like the "outer border" or "edge" of a set. Often times it is easy to see where the boundary is, especially for bounded sets. But be careful, the concept of a hard edge of set may not apply in every situation, for instance , the entire space $\mathbb{R}^n$ itself is closed because any convergent sequence in $\mathbb{R}^n$ must also converge to something in $\mathbb{R}^n$. But $\mathbb{R}^n$ has empty boundary, i.e. its boundary is the empty set $\emptyset$. This makes sense as the entire $\mathbb{R}^n$ "goes to infinity" in every direction and it doesn't have a "border". Of course any set contains the empty set as a subset so $\mathbb{R}^n$ contains all of its boundary. Here's the precise definition of the boundary:

Definition 1.2 (boundary)**.** The *boundary* of A is a subset $B \subset \mathbb{R}^n$ defined by the following: A point a is in the boundary if and only if for every ball $B_\varepsilon(a)$ centered at a of arbitrary radius ε, $B_\varepsilon(a)$ contains at least one point in A, an at least one point outside A.

Another terminology that you may sometimes see: The *interior* of a subset A

So restating what we mentioned:

Definition 1.3 (closed, boundary definition)**.** A subset $A \subset \mathbb{R}^n$ is *closed* if $B \subset A$ where B is the boundary of A.

Definition 1.4. A subset $A \subset \mathbb{R}^n$ is *bounded* if there exists some constant $C \geq 0$, such that for every $x \in A$,

$$|x| \leq C.$$

Here's how you think about a bounded set: a set is bounded if all the points in the set are within a certain finite size. Here by size we mean the absolute value of a point, or the magnitude of the position vector of the point in $\mathbb{R}^n$. So a set is bounded if it doesn't "go to infinity in any direction".

Exercise 1.5. Think of a set in $\mathbb{R}^n$ for each of the following:

1. Closed and bounded.
2. Closed but not bounded.
3. Bounded but not closed.
4. Neither closed nor bounded.

Here's a definition we will need later:

Definition 1.6. Fix a point $x \in \mathbb{R}^n$. The *(open) ball of radius* $\varepsilon > 0$ *centered at* x is the set

$$B_\varepsilon(x)\{y \in \mathbb{R}^n : |x - y| < \varepsilon\}.$$

So these are the points that are ε distance away from x.

2 What Exactly are Minimum and Maximum Values?

For a function $f\colon \mathbb{R}^n \to \mathbb{R}$, we denote by $D \subset \mathbb{R}^n$ the domain of the function.

Definition 2.1. The function $f\colon \mathbb{R}^n \to \mathbb{R}$ has a *global/absolute maximum*, or sometimes just *maximum* at the point $(x_1, \dots, x_n) \in D \subset \mathbb{R}^n$ if for all $(y_1, \dots, y_n) \in D \subset \mathbb{R}^n$,

$$f(y_1, \dots, y_n) \leq f(x_1, \dots, x_n).$$

We say that f has *global/absolute maximal value* $f(x_1, \dots, x_n)$.

Definition 2.2. The function $f\colon \mathbb{R}^n \to \mathbb{R}$ has a *global/absolute minimum*, or sometimes just *minimum* at the point $(x_1, \dots, x_n) \in D \subset \mathbb{R}^n$ if for all $(y_1, \dots, y_n) \in D \subset \mathbb{R}^n$,

$$f(y_1, \dots, y_n) \geq f(x_1, \dots, x_n).$$

We say that f has *global/absolute minimum value*

There are some immediate observations we can make just from these definitions. Firstly, there can be multiple points at which f has a maximum or minimum, so a maximum or minimum point is not unique. For instance, just take f to be a constant function, then since f takes the same value at every point, f has a global maximum at every point in $\mathbb{R}^n$; and f also has a global minimum at every point in $\mathbb{R}^n$ as well. On the other hand, the maximal value, and minimal value, are unique, but they can possibly coincide, as in the case of a constant function. Thus a point can be both a global maximal point and a global minimal point at the same time. That is why we say *the* maximal/minimal value, whereas we say f has *a* maximum/minimum at the point $(x_1, \dots, x_n)$. Try to think of examples for yourself to make yourself comfortable with these observations.

Definition 2.3. The function $f\colon \mathbb{R}^n \to \mathbb{R}$ has a *local maximum* at the point $x = (x_1, \dots, x_n) \in D \subset \mathbb{R}^n$ if there exists some open ball $B_\varepsilon(x) \subset D$ centered at x, such that for all $(y_1, \dots, y_n) \in B_\epsilon(x)$,

$$f(y_1, \dots, y_n) \leq f(x_1, \dots, x_n).$$

Definition 2.4. The function $f\colon \mathbb{R}^n \to \mathbb{R}$ has a *local minimum* at the point $x = (x_1, \dots, x_n) \in D \subset \mathbb{R}^n$ if there exists some open ball $B_\varepsilon(x) \subset D$ centered at x, such that for all $(y_1, \dots, y_n) \in B_\epsilon(x)$,

$$f(y_1, \dots, y_n) \geq f(x_1, \dots, x_n).$$

So one can think of a local maximal/minimal point of f as being a global maximal/minimal point of f but only within some neighborhood of that point. We have the following observations: again there can be multiple points in the domain that are local maximum/minimum. This time however, the values that they take are not unique, so there can be multiple local maximal values, and multiple minimal values that f can take.

Here are some useful terminology: instead of saying a point is either a global/local maximal or minimal point, we can just say a point is a *global/local extremum.* The plural of extremum is extrema.

And instead of saying a value is either a global/local maximal or minimal value , we can just say *global/local extremal value.*

How to Find Them?

In general, global extrema of a function are not easy to find. In fact there is no method that can produce the global extrema for any function $f\colon \mathbb{R}^n \to \mathbb{R}$. Although in some special cases it might be obvious what where they should be, for instance the constant function, or a monotonically increasing or decreasing function.

We do have a some hope for local extrema though, and this is given by the first derivative test.

3 First Derivative Test

First, a definition,

Definition 3.1. Let $f\colon \mathbb{R}^n \to \mathbb{R}$, with variables $(x_1, \ldots, x_n)$. A point $a \in D \subset \mathbb{R}^n$ is a *critical point of* f if

$$\frac{\partial f(a)}{\partial x_i} \text{ exists for all } 1 \leq i \leq n$$

and moreover,

$$\frac{\partial f(a)}{\partial x_i} = 0 \quad \text{for all } 1 \leq i \leq n.$$

Theorem 3.2 (Fermat's Theorem/First Derivative Test Theorem). *If $x \in D \subset \mathbb{R}^n$ is a local extrema of f, and*

$$\frac{\partial f(a)}{\partial x_i} \textit{ exists for all } 1 \leq i \leq n,$$

then x is a critical point, i.e.

$$\frac{\partial f(a)}{\partial x_i} = 0 \quad \textit{for all } 1 \leq i \leq n.$$

Pay extra attention to the definition of critial point, and Fermat's Theorem, that these are statements about points a in the domain of f where all the first partials exist. If this condition is not met, then even if it is indeed an extremum, we cannot say anything about it.

An observation then is the following: a local extremum that has all the first partials must be a critial point, but the converse is not always true, for instance we will see later that a saddle point is also a critial point, but a saddle point is not an extremum.

So the First Derivative Test allows us to detect local extrema, but only among the points of the domain that are at least once-differentiable (meaning that all the first partials exist). This will be important later when we want to try to find the global extrema of a function defined on a compact set because in that case the First Derivative Test will give us the local extrema in the interior of the set that are candidates for global extrema, but it won't tell us anything about the points that lie on the boundary of the domain, where the function is not differentiable. This is where we make use of Lagrange Multiplier to find points on the boundary to add into our list of candidates for global extrema. More on this later.

4 Second Derivative Test

The First Derivative Test tells us which points on the domain of f are critical points, at least among the points that are once-differntiable. By Fermat's Theorem, all the local minimum and maximum points of f must be among these critical points. Now we need a way to tell which of these critial points are local minimum points, which are local maximum points, and which ones are neither. This is the goal of the Second Derivative Test, although it is not 100 percent guaranteed to give the full classification, as there are situations where it does not yield any conclusion for a critical point. However, it is the best we can do. For instance, for this to work, the second partial derivatives of f at the critial point not only must exist, but the second partial derivative must also be continuous. Although most functions you will encounter in this course will all be nice enough that this is the case, it is still important to be aware of.

Now, here's something extremely important: **the Second Derivative Test only works for functions of two variables**. There exists a general version that works for any number of variables, but that is beyond the scope of this course.

Theorem 4.1 (Second Derivative Test). *Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables. Let x be a critical point of f. Suppose also the second partial derivatives of f exists and are continuous. Consider the Hessian maxtrix defined as*

$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

which has determinant

$$D = f_{xx}f_{yy} - f_{yx}f_{xy}.$$

Of course if the second partials are continuous, we can use the Symmetry of Second Partials to write this as

$$D = f_{xx}f_{yy} - (f_{xy})^2.$$

We can evaluate the determinant at some point (a, b)*:*

$$D(a, b) = f_{xx}(a, b)f_{yy}(a, b) - [f_{xy}(a.b)]^2.$$

And in particular if we evaluate D *at a critical point* (x_0, y_0)*, then:*

1. *If* $D(x_0, y_0) > 0$*, and* $f_{xx}(x_0, y_0) > 0$*, then* $f(x_0, y_0)$ *is a local minimum.*
2. *If* $D(x_0, y_0) > 0$*, and* $f_{xx}(x_0, y_0) < 0$*, then* $f(x_0, y_0)$ *is a local maximum.*
3. *If* $D(x_0, y_0) < 0$*, then* $f(x_0, y_0)$ *is neither a local maximum nor a local minimum.*
4. *If* $D(x_0, y_0) = 0$*, then the test is inconclusive, we cannot say anything about the type of critical point* $f(x_0, y_0)$ *is.*

5 The Extreme Value Theorem

What about global/absolute extrema of a function? In the nice situation of a two-variable function with , we might think that using the Second Derivative Test,

Theorem 5.1 (Extreme Value Theorem (EVT))**.** *Let* $f: \mathbb{R}^n \to \mathbb{R}$ *be a continuous function defined on a closed and bounded set* $D \subset \mathbb{R}^n$*. Then* f *attains an absolute maximum value* $f(x_1)$*, and an absolute minimum value* $f(x_2)$*, where* $x_1, x_2 \in D \subset \mathbb{R}^n$*.*

6 The Method to Find Global Extrema

To find the absolute maximum and minimum values of a continuous function f that is at least once-differentiable defined on a closed and bounded subset $D \subset \mathbb{R}^n$. We do the following:

1. Find the critical points of f using First Derivative Test.
2. Find the extreme values on the boundary of f.
3. The largest value of the points found in the two previous steps is the absolute maximal value, while the smallest value of the points found in the two previous steps is the absolute minimal value.

In Step 1, because the function is continuous and once-differntiable, the First Derivative Test will cover all the interior points of the domain because the only points that are non-differentiable will be on the boundary. So we have a list of critical points, and the absolutely extrema must be among them, **if they are in the interior**. Combining this with Step 2, which looks at the boundary, i.e. the non-differentiable points,

Why do we need to check the boundary?

The First Derivative Test which gives us critical points by the result of Fermat's Theorem won't be able to detect possible extremal points on the boundary. This is a technical detail that is not important for you to know unless you intend to study calculus rigorously. Basically it is because of how we define a local extremum, which involves open balls around a local extremum that are contained in the domain, and points on the boundary do not have open balls around them that are contained in the domain.

Here's a better, more intuitive way of thinking about this : you can think of the boundary as where we "cut off" the function, beyond which we do not define the function. And so, it is possible that where we cut off this function, the function attains a max or min, but this may not be detected by looking at whether the partial derivatives vanish because the max or min is attained not from the local inflective behavior of the function, but only because we restricted the function. What I mean is, the reason that a local min or max has banishing first partials is representing the fact that at that point the function has a "peak" (for max), or a "trough" (for min), at which the function "flattens out". But for a possible extrema on the boundary, this may not be the case.

7 Lagrange Multiplier

This is a method that we may use to help us with Step 2. in the above process. In the case when we can express the boundary of the domain as a number of constraints, i.e equations, then we can employ Lagrange Multiplier to obtain the extrema points on the boundary.

The extreme values of $f\colon \mathbb{R}^n \to \mathbb{R}$ subject to the constraint $g(x_1, \ldots, x_n) = k$, can be found by solving the values $x_1, \ldots, x_n, \lambda$, such that

$$\nabla f(x_1, \ldots, x_n) = \lambda \nabla g(x_1, \ldots, x_n).$$

Then f attains its extremal values at those points $(x_1, \ldots, x_n)$ that satisfy the above equation. Here it is necessary to assume that ∇g actually exists and

$$\nabla g \neq 0.$$

For more than one constraints, for instance with two constraints

$$g(x_1, \ldots, x_n) \text{ and } h(x_1, \ldots, x_n),$$

we have to solve

$$\nabla f = \lambda \nabla g + \mu \nabla h.$$

And the process is the same.

One thing to note: **Lagrange multiplier gives you the absolute extrema of a function, but only on a paritucular subset of the domain specified by the constraint**. This is why we have to combine using critical points via First Derivative Test to find candidate extrema points on the interior of the domain, then use Lagrange Multiplier to find candidate extrema points on the boundary. We have to combine the list of candidate points we found with both methods to cover all the points on the domain.

8 Things to Consider

The method we have outlined to find global extrema does not work in every case, in particular if any of the assumptions regarding the domain of the function, continuity of the function, or non-existence of first partials, or that of the constraint, we cannot employ the method. When some of these conditions are not met, we often have to resolve to analyzing the function itself in an ad hoc way to attempt to find the extrema.