CLARK UNIVERSITY 1887

Maxima and minima of scalar fields
Math 131 Multivariate Calculus
D Joyce, Spring 2014

Extrema. We'll discuss maxima and minima of scalar fields $f : \mathbf{R}^n \to \mathbf{R}$. They occur at critical points, that is, where the first partial derivatives are all 0. To determine which critical points are maxima, which are minima, and which are something else, we'll look at the Hessian as the basis of a second derivative test. We'll briefly survey compact sets and the Extreme Value Theorem (EVT).

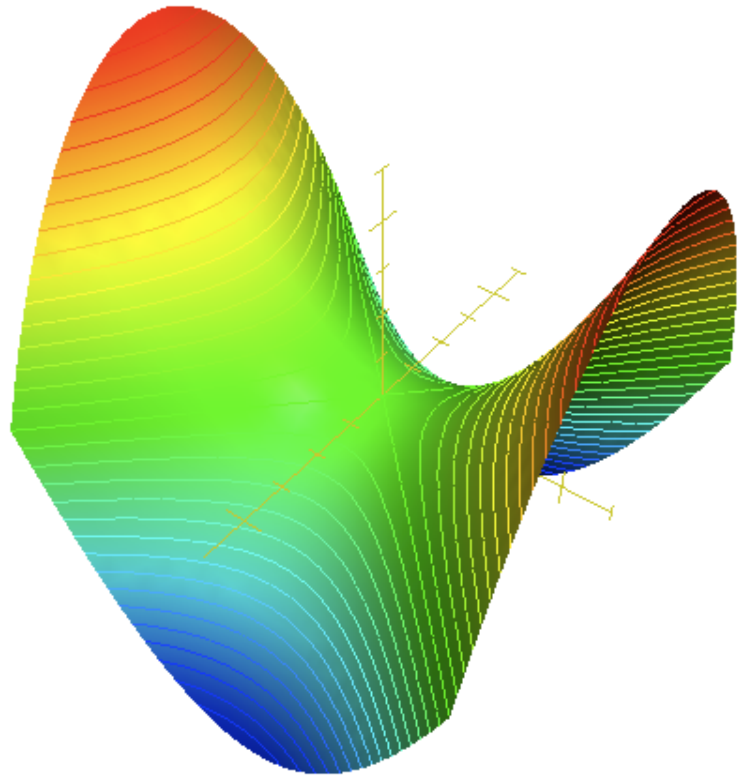
Comparison to the single variable case. In the one-dimensional case, we singled out values of a such that $f'(a) = 0$ and called such values *critical points* of f. That's because if f has a max or a min at a, then there has to be a horizontal tangent line at $(a, f(a))$ on the graph of f (assuming, of course, that f is differentiable and a occurs inside the domain of f). Recall that not every critical point of f was a max or a min; for instance, $x = 0$ is the critical point of $f(x) = x^3$, but f has neither a max nor a min at $x = 0$.

The analogous thing happens in the multidimensional case. If f has a max or min at $\mathbf{a}$, then there has to be a horizontal tangent hyperplane at $(\mathbf{a}, f(\mathbf{a})$ on the graph of f, and there's a horizontal tangent hyperplane if and only if the derivative $Df(\mathbf{a}) = 0$. Equivalently, all the partial derivatives of f at $\mathbf{a}$ are 0. And, like in the one-dimensional case, not every critical points will be a max or a min.

In the one-dimensional case, we had a second-derivative test to help us determine whether a critical point was a max or a min. It said:

> for a critical point a of f, if $f''(a) > 0$, then f has a min at $x = a$, but if $f''(a) < 0$, then f has a max at $x = a$.

If the second derivative $f''(a)$ was 0, then you had to try some other way to determine whether it was a max or min or neither. We need some sort of second-derivative test for the multidimensional case.



Saddle points. The multidimensional case is complicated by a phenomenon that doesn't occur in the one-dimensional case, and that's saddle points. Consider the function $f(x, y) = x^2 - y^2$.

This function f has a critical point at $(0, 0)$ since $\frac{\partial f}{\partial x}(0, 0) = 0$ and $\frac{\partial f}{\partial y}(0, 0) = 0$. But it's clearly not a max since the intersection of the surface $z = x^2 - y^2$ with the plane $y = 0$ gives the parabola $z = x^2$ which opens upward, and it's clearly not a min since the intersection of the surface $z = x^2 - y^2$ with the plane $x = 0$ gives the parabola $z = -y^2$ which opens downward. The surface looks like a saddle. The main problem in the multivariable case is to classify critical points as as maxima, minima, or saddle points.

The second derivative test, the Hessian criterion. Suppose $\mathbf{a}$ is a critical point of f. We can approximate the function f by the second Taylor polynomial p_2 at $\mathbf{x} = \mathbf{a}$.

$$p_2(\mathbf{x}) = f(\mathbf{a}) + Df(\mathbf{a})\,\mathbf{h} + \tfrac{1}{2}\mathbf{h}^T\,Hf(\mathbf{a})\,\mathbf{h}$$

where H is the Hessian, and $\mathbf{h} = \mathbf{x} - \mathbf{a}$. Of course, p_2 also has a critical point at $\mathbf{a}$ since it has the same horizontal tangent hyperplane there that f has. In other words, the linear term $Df(\mathbf{x})$ of p_2 is 0. Furthermore, if p_2 has a min at $\mathbf{a}$, then so does f; or if p_2 has a max at $\mathbf{a}$, then so does f; or if p_2 has a saddle point at $\mathbf{a}$, then so does f. That means we need to examine the quadratic term of p_2 more carefully to see what's going on with f.

The quadratic term, which is also called a quadratic form, $Q(\mathbf{h}) = \mathbf{h}^T\,Hf(\mathbf{a})\,\mathbf{h}$, is

$$\begin{bmatrix} h_1 & \cdots & h_n \end{bmatrix} \begin{bmatrix} f_{x_1x_1} & \cdots & f_{x_1x_n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1} & \cdots & f_{x_nx_n} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix}$$

Note that the second partial derivatives $f_{x_ix_j}$ and $f_{x_jx_i}$ are equal. (We'll assume they're continuous which implies that.) That means the matrix H is a symmetric matrix, that is, $H^T = H$.

Now, $Q(\mathbf{h})$ has a minimum at $\mathbf{h} = \mathbf{0}$ if and only if $Q(\mathbf{h}) > 0$ for all $\mathbf{h} \neq \mathbf{0}$. When that happens, we say the quadratic form Q is *positive definite*. And, in that case, p_2 has a minimum at $\mathbf{x} = \mathbf{a}$, and that implies f has a minimum at $\mathbf{x} = \mathbf{a}$.

Likewise, $Q(\mathbf{h})$ has a maximum at $\mathbf{h} = \mathbf{0}$ if and only if $Q(\mathbf{h}) < 0$ for all $\mathbf{h} \neq \mathbf{0}$. When that happens, we say the quadratic form Q is *negative definite*. And, in that case, p_2 has a maximum at $\mathbf{x} = \mathbf{a}$, and that implies f has a maximum at $\mathbf{x} = \mathbf{a}$.

But, when $Q(\mathbf{h})$ takes on both negative and positive values, there's a saddle point at $\mathbf{h} = \mathbf{0}$. And, in that case, f, too, has a saddle point at $\mathbf{x} = \mathbf{a}$.

We can summarize this as a theorem.

Theorem 1. If $f : \mathbf{R}^n \to \mathbf{R}$ is of class C^2 and has a critical point at $\mathbf{a}$, then

(1) if the Hessian $Hf(\mathbf{a})$ is positive definite, then f has a local min at $\mathbf{a}$,

(2) if $Hf(\mathbf{a})$ is negative definite, then f has a local max at $\mathbf{a}$, and

(3) if the determinant $|Hf(\mathbf{a})|$ is nonzero, but $Hf(\mathbf{a})$ is neither positive definite nor negatve definite, then f has a saddle point at $\mathbf{a}$.

We still need to be able to distinguish these three cases to get a good second-derivative test. There's a special bit of linear algebra that applies here. It uses the fact that the Hessian is a symmetric matrix. The second derivative test involves the sequence of principal minors $d_1, d_2, \ldots, d_n$ of the Hessian $Hf(\mathbf{a})$. The i^{th} principal minor d_i is defined as the determinant of the upper left $i \times i$ submatrix of $Hf(\mathbf{a})$, so the n^{th} principle minor d_n is the determinant of the entire Hessian, $d_n = |Hf(\mathbf{a})|$.

So, here's the test.

Theorem 2 (Second derivative test for local extrema)**.** Calculate the sequence of principal minors $d_1, d_2, \ldots, d_n$ of the Hessian $Hf(\mathbf{a})$. If d_n is not 0, then

(1) If all n principal minors $d_1, d_2, \ldots, d_n$ are positive, then f has a local minimum at $\mathbf{a}$;

(2) If all the odd principal minors $d_1, d_3, \ldots$ are negative while all the even principal minors $d_2, d_4, \ldots$ are positive, then f has a local maximum at $\mathbf{a}$;

(3) In all other cases, f has a saddle point at $\mathbf{a}$.

We'll look at a couple of examples in class.

Special cases occur when the determinant, $d_n = |Hf(\mathbf{a})|$ is 0. That can occur, for instance, when f is constant on a line through a. For example, $f(x, y) = x^2$ takes its minimum on the whole line $x = 0$.

Global extrema on compact regions. Recall from calculus of one variable that the extreme values of a function f defined on a closed interval $[a, b]$ could occur not only at the critical points in the interior (a, b), but also at the endpoints $x = a$ or $x = b$. The extreme value theorem, EVT, said that any continuous function f defined on a closed interval takes on a maximum value and a minimum value. Is there anything like that for functions of several variables?

Yes, there is an EVT for several variables, but the statement of it is more complicated because the subsets of $\mathbf{R}^n$ that correspond to closed intervals in $\mathbf{R}$ are more complicated to describe. We need to look at the topological concepts of 'open subset,' 'closed subset,' and 'compact subset' of $\mathbf{R}^n$. These were mentioned back in section 2.2.

Some topology. We'll look at some examples to illustrate the following concepts.

First, we define 'open' and 'closed' balls in $\mathbf{R}^n$. These directly correspond to open and closed intervals in $\mathbf{R}$. The *closed ball* of radius r centered at the point $\mathbf{a} \in \mathbf{R}^n$ is the set

$$\{\mathbf{x} \in \mathbf{R}^n \mid \|\mathbf{x} - \mathbf{a}\| \leq r\},$$

in other words, it includes the interior and the boundary of the sphere of radius r. The *open ball* of radius r centered at the point $\mathbf{a} \in \mathbf{R}^n$ is the set

$$\{\mathbf{x} \in \mathbf{R}^n \mid \|\mathbf{x} - \mathbf{a}\| < r\},$$

in other words, it includes the interior of the sphere of radius r but not the boundary.

Next, we define open and closed subsets of $\mathbf{R}^n$ in general. A subset $S \subseteq \mathbf{R}^n$ is said to be *open* if every point $\mathbf{a} \in S$ is the center of some open ball that is entirely contained in S, that is, there is some positive number r such that

$$\|\mathbf{x} - \mathbf{a}\| < r \text{ implies } \mathbf{x} \in S.$$

A *neighborhood* of a point $\mathbf{x}$ is any open set that contains $\mathbf{x}$. A point $\mathbf{x} \in \mathbf{R}^n$ is said to be in the *boundary* of a set $S \subseteq \mathbf{R}^n$ if every open ball centered at $\mathbf{x}$ intersects S (i.e., the intersection is nonempty) and also intersects the complement of S. In other words, for each positive number r there is one point in S within r of $\mathbf{x}$ and another point not in S within r of $\mathbf{x}$. A subset $S \subseteq \mathbf{R}^n$ is said to be *closed* if it contains all of its boundary points. It can be proved that the complement of an open set is closed, and vice versa.

Compact subsets of $\mathbf{R}^n$ aren't introduced until section 4.2. A subset $S \subseteq \mathbf{R}^n$ is said to be *compact* if it is both closed and bounded.

The Extreme Value Theorem, EVT**,** for $\mathbf{R}^n$. A proof of this theorem is beyond the scope of this course.

Theorem 3 (EVT)**.** If f is a continuous real-valued function defined on a compact subset $X \subseteq \mathbf{R}^n$, then f takes on a maximum and a minimum value. That is, there are points $\mathbf{a}_{\text{min}}$ and $\mathbf{a}_{\text{max}}$ in X such that for all $\mathbf{x} \in X$,

$$f(\mathbf{a}_{\text{min}}) \leq f(\mathbf{x}) \leq f(\mathbf{a}_{\text{max}}).$$

When we apply this theorem to a differentiable function f, it means we will find the extrema if we check (1) the critical points on the interior of the domain, and (2) all the points on the boundary. Note that a nondifferentiable function can also have extrema (3) where the function is not differentiable.

Math 131 Home Page at
`http://math.clarku.edu/~djoyce/ma131/`