Critical points, and local and global extrema

Sean Fitzpatrick

June 13, 2018

This is a handout with a few additional remarks on classifying critical points and finding maximum and minimum values for functions of two variables. Recall that the extreme value theorem guarantees that a continuous function will attain an absolute maximum and minimum on a set D that is **closed** and **bounded**. Of course, this theorem is false without both assumptions: the set $D = \{(x, y)|y \geq 0\}$ (that is, the upper half-plane) is closed, but not bounded. I'm sure you can easily think of examples of functions that do not have a maximum or a minimum on this set. Similarly, the set given by $0 < \sqrt{x^2 + y^2} \leq 1$ is bounded, but not closed, since the origin is a boundary point but not contained in the set. Again, it's not too hard to come up with an example of a function that does not have both a maximum and a minimum on this set. Notice that if $D \subseteq \mathbb{R}^2$ is closed and bounded, and $f : D \to \mathbb{R}$ is continuously differentiable, then finding the absolute maximum and minimum of f is reduced to two steps:

1. Find the critical values of f. That is, for each $(a, b) \in D$ such that $\nabla f(a, b) = \langle 0, 0 \rangle$, compute $f(a, b)$.

2. Find the boundary extrema of f. We'll explain this step below:

Given our closed and bounded D, let C be its boundary (that is, the set of boundary points), and let g be the *restriction* of f to C. (That is, g is the same function as f, but with the smaller domain C.) Notice that since C is the boundary of the two-dimensional region D, it's a one-dimensional object: in all the examples you'll consider, C will be a smooth, or at least piecewise-smooth, curve.

So finding the maximum and minimum of g is roughly the same as finding the maximum and minimum of a continuous function on a closed interval. In fact, if we can parameterize C, then that's exactly what we have: if $(x(t), y(t))$, $t \in [a, b]$ is a parameterization of C, we can think of g as the one-variable function $g(t) = f(x(t), y(t))$. Also, C will itself be closed and bounded, so we're guaranteed to be able to find a maximum and minimum value for the restriction of f to C, and then we just have to compare those values to the critical values of f.

By the way, suppose you wanted to extend this problem, and find the maximum and minimum values of a function of three variables on a closed, bounded subset of $\mathbb{R}^3$. Then the procedure is exactly the same. First you look for critical values, and then you look for boundary extrema. This time the boundary will be a surface S. If we're lucky, we can write S as a level set $g(x, y, z) = c$ and apply the method of Lagrange multipliers. (We could also parameterize S, but we haven't looked at parametric surfaces yet.) There are usually two cases:

1. S is a smooth surface $g(x, y, z) = c$. In this case the method of Lagrange multipliers works perfectly: S is closed and bounded, so we're guaranteed a minimum and a maximum, and we just have to compare these to the critical values. By the way, we'll learn later that since S

is a boundary, it does not have a boundary of its own. Think of the sphere as the boundary of a ball for example: a sphere has no edges. You might want to keep this in mind for case 2.

2. S is a piecewise smooth surface, such as the boundary of a rectangular box (which is made up of pieces of six planes, which are smooth surfaces with boundaries, joined together along those boundaries). Then we once again need to compare the critical values of f to the boundary extrema of f (that is, the maximum and minimum of the restriction of f to S). Unfortunately, this time the problem involves quite a bit more work: not only do we have to check each of the smooth pieces separately, but each of these pieces will have a boundary of its own. (Like the rectangles that form the boundaries to the sides of a box.) So for each piece of the surface, you need to find any local maximum and minimum values, and then restrict *further* to the edges of each surface. (Don't worry, there aren't (m)any problems like this in your homework.)

Anyway, the general story is that if you want to optimize a function of n variables, you'd look for critical points, and then look for boundary extrema, which is an $(n-1)$-dimensional problem. If the boundary consists of more than one piece, then those pieces might have boundaries of their own, and restricting to those gets you to an $(n-2)$-dimensional problem, and so on. As long as everything is at least piecewise-smooth, you can expect that at some point either Lagrange multipliers will finish off the problem for you, or you go down one more dimension. Sooner or later you'll reduce the problem to a set of one-dimensional problems (unless you solve it completely at an earlier step) and you can apply your results from single-variable calculus.

One warning to keep in mind: given a general Lagrange multiplier problem, it might be the case that the constraint equation(s) do not define a closed, bounded region. In this case, you can go about trying to solve the Lagrange multiplier equations, but there's no guarantee that you'll find both an absolute maximum and an absolute minimum. You might still be able to find solutions to the Lagrange multiplier equations, but these are perhaps only *local* constrained maxima or minima (or possibly a critical point that is neither a maximum nor a minimum) - you will still have to do a bit of reasoning to decide whether a given solution is a maximum or a minimum. (Actually, there is a second derivative test for constrained extrema, but it's pretty complicated, and beyond the scope of this course.)

Let's conclude with a few remarks about the second derivative test. It seems a bit weird at first, but the idea is actually fairly simple, if you're willing to accept Taylor's Theorem without proof for functions of more than one variable. We already know that if $f(x, y)$ is C^1 (continuously differentiable), then we get the linear approximation

$$f(x, y) \approx f(a, b) + \nabla f(a, b) \cdot \langle x - a, y - b \rangle$$

near a point (a, b) in the domain of f. Taylor's theorem tells us that if f is C^2 (continuous second-order derivatives), then we get the *quadratic* approximation

$$f(x, y) \approx f(a, b) + \nabla f(a, b) \cdot \langle x - a, y - b \rangle + \frac{1}{2} A(x - a)^2 + B(x - a)(y - b) + \frac{1}{2} C(y - b)^2,$$

where $A = \dfrac{\partial^2 f}{\partial x^2}(a, b)$, $B = \dfrac{\partial^2 f}{\partial x \partial y}(a, b)$, and $C = \dfrac{\partial^2 f}{\partial y^2}(a, b)$. (Compare this to the single-variable version: $f(x) \approx f(a) + f'(a)(x - a) + \frac{1}{2} f''(a)(x - a)^2$.) Now, if (a, b) is a critical point, then

$\nabla f(a, b) = \mathbf{0}$, and we get the approximation

$$f(x, y) \approx k + \frac{1}{2}\left(AX^2 + 2BXY + CY^2\right),$$

where $k = f(a, b), X = x - a, Y = y - b$. So it's enough to understand the critical points of the function

$$g(x, y) = Ax^2 + 2Bxy + Cy^2,$$

since f locally looks just like g. (We've basically just done a shift of the graph, and stretched by a factor of 2 to get rid of the 1/2.) Now, we can re-write g as follows, assuming $A \neq 0$:

$$\begin{aligned}
g(x, y) &= Ax^2 + 2Bxy + Cy^2 \\
&= A(x^2 + 2\frac{B}{A}xy) + Cy^2 \\
&= A(x + \frac{B}{A}y)^2 - \frac{B^2}{A}y^2 - Cy^2 \\
&= A(x + \frac{B}{A}y)^2 + \frac{1}{A}(AC - B^2)y^2.
\end{aligned}$$

Now we can see that this is basically just a paraboloid, as long as $AC - B^2 \neq 0$. (Otherwise, we end up with a parabolic cylinder.) If $AC - B^2 > 0$, then both the coefficient for both terms has the same sign; if $A > 0$ we get an elliptic paraboloid opening upwards (local minimum), and if $A < 0$ we get an elliptic paraboloid opening downwards (local maximum). If $AC - B^2 < 0$, then the two terms have coefficients with opposite signs, and that gives us a hyperbolic paraboloid (saddle point).

And what if $A = 0$? Well, in that case $AC - B^2 = -B^2 \leq 0$, so there are two cases: if $B \neq 0$, the second derivative test tells us to expect a saddle point, and indeed this is what we get. Either $C = 0$ as well, and $g(x, y) = 2Bxy$, which is just a hyperbolic paraboloid rotated by $\pi/4$ (its contour curves are the hyperbolas $xy = c$), or $C \neq 0$, in which case you can complete the square in y, and check that the result is once again a hyperbolic paraboloid (exercise). The other case is if $B = 0$, in which case $\Delta = 0$, so we can't make any conclusions from the second derivative test (although we'll have $g(x, y) = Cy^2$, which is again a parabolic cylinder).