Max-Min Problems for Functions of Two Variables

E. L. Lady

(September 2, 2002)

Consider a quadratic function $f(x,y)$ of two variables. In terms of extrema, there are three possibilities, which we will illustrate with three examples.

First Example. Let

$$f(x,y) = 3x^2 - 7xy + 2y^2 .$$

We have

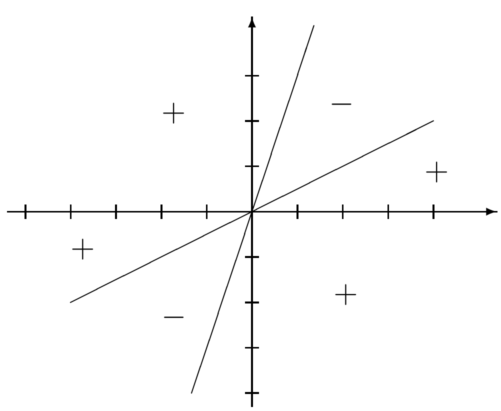
$$\frac{\partial f}{\partial x} = 6x - 7y$$
$$\frac{\partial f}{\partial y} = -7x + 4y$$

and these are only both zero at $(0,0)$. Thus the origin is the only critical point.

We can factor the function as

$$f(x) = 3x^2 - 7xy + 2y^2 = (3x-y)(x-2y) .$$

This function is zero along the two lines $y = 3x$ and $y = x/2$; these lines are thus level curves for the function. The function is positive when both $y > 3x$ and $y > x/2$ and also when both $y < 3x$ and $y < x/2$. It is negative at other places. Thus the xy-plane is divided into four regions as follows.



Then $f(x,y)$ is positive when the point (x,y) belongs to one of the regions marked $+$, and negative when (x,y) belongs to a region marked $-$, and zero when (x,y) is on the lines indicated. Since $f(0,0) = 0$ and every neighborhood of $(0,0)$ contains both points where $f(x,y) > 0$ and also where $f(x,y) < 0$, we see that $f(x,y)$ does not have either a maximum or a minimum at the origin.

The factorization $3x^2 - 7xy + 2y^2 = (3x - y)(x - 2y)$ was fairly easy to find by inspection. But it could also be determined in a systematic way. Namely, if we set $f(x,y) = 3x^2 - 7xy + 2y^2$ equal to zero and then solve for x in terms of y, from the quadratic formula (for instance) one gets the two solutions $x = \frac{1}{6}(7y \pm \sqrt{49y^2 - 24y^2}) = \frac{1}{6}(7y \pm 5y)$, i.e.

$$\begin{aligned} x &= \frac{1}{6}(7y + 5y) = 2y \\ x &= \frac{1}{6}(7y - 5y) = y/3\,, \end{aligned}$$

and these are the equations of the two lines along which $f(x,y) = 0$. One can then see (because $f(x,y)$ is a quadratic) that $f(x,y) = c(x - 2y)(x - \frac{y}{3})$, where c is a constant. In fact, here $c = 3$.

Second Example. Now consider the function

$$f(x,y) = x^2 + xy + y^2\,.$$

It is not obvious how to factor this function, so we attempt to solve for x in terms of y so that $f(x,y) = 0$. The quadratic formula yields

$$x = \frac{1}{2}(-y \pm \sqrt{-3y^2})\,.$$

Since y^2 is automatically positive, the quantity under the square root sign is always negative except when $y = 0$, in which case we get $x = 0$.

Thus the quadratic $f(x,y) = x^2 + xy + y^2$ is never zero except at the origin. (It follows that this quadratic cannot factor.)

Key Observation. If $f(x,y)$ is any continuous function, it cannot change sign except when (x,y) moves through a point where $f(x,y) = 0$. In other words, if $f(x,y)$ is positive in a certain region and negative in another, then it has to be 0 on the boundary between the two regions.

Since the boundary between the two regions would have to be more than a single point, we therefore conclude that if a function $f(x,y)$ is zero at the origin but not zero anywhere else, then the function must be either always positive everywhere, except at the origin, or always negative except at the origin. In the first case, $f(x,y)$ takes its smallest possible value at $(0,0)$, so that $f(x,y)$ has a a minimum at $(0,0)$, and in the second case the function has a maximum at $(0,0)$.

More generally, if c is any real number and $f(x,y)$ is greater than c at some points and less than c at others, then we must have $f(x,y)=c$ on the boundary between the regions with $f(x,y)>c$ and $f(x,y)<c$. Consequently, if $f(x_0,y_0)=c$ but $f(x,y)\neq c$ at all other points, then $f(x,y)$ cannot take on values both greater than c and less than c. It follows that $f(x,y)$ has either a maximum or a minimum at (x_0,y_0). (We are assuming in this reasoning that the function $f(x,y)$ is defined in the whole plane, or at least that the domain of $f(x,y)$ is a connected set. This is almost always true for those functions that arise in applications.)

Returning to the second example, since $f(1,0)=1$, we see that $f(x,y)$ must be strictly positive everywhere except at the origin, so the function has a minimum at the origin.

The logic of these two examples shows that a function

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

will have either a maximum or a minimum at $(0,0)$ if $B^2-AC<0$, since in that case the quadratic formula shows that the function cannot equal 0 except at the origin, and therefore must either be positive in the entire remainder of the plane, or must be negative in the rest of the plain. Furthermore, in this case it will have a minimum at $(0,0)$ if $A>0$ (since in the case the function is positive along the y-axis) and a maximum at $(0,0)$ if $A<0$. (Notice also that if B^2-AC is strictly negative, then AC must be positive, so A and C must necessarily have the same sign.)

On the other hand, if $B^2-AC>0$ then the function will be zero along two lines which intersect at the origin and will change sign whenever (x,y) crosses either of these lines, and thus will take both positive and negative values in every neighborhood of the origin. Since $f(0,0)=0$, this shows that the origin is neither a maximum nor a minimum. Since $(0,0)$ is the only critical point for the function, we see that the origin is a saddle point and that the function has no maxima or minima.

The critical points for a function

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

occur when

$$\frac{\partial f}{\partial x} = 2Ax + 2By = 0$$
$$\frac{\partial f}{\partial y} = 2Bx + 2Cy = 0\,.$$

Clear the origin $(0,0)$ is a critical point. In most cases, we would expect this system of two equations in two unknowns to have only this one solution. However if $B^2 = AC$, then multiplying the first equation by C and the second by B (assuming $B \neq 0$, and therefore also $C \neq 0$) yields

$$C\frac{\partial f}{\partial x} = 2ACx + 2BCy = 0$$
$$B\frac{\partial f}{\partial y} = 2B^2x + 2BCy = 2ACx + 2BCy = 0\,,$$

showing that the two equations are equivalent, so that all solutions to the first equation, i. e. all points along the line $Ax + By = 0$, are critical points. And in the case that $B^2 = AC$ and also $B = 0$, then either $A = 0$ or $C = 0$, so $f(x,y) = Ax^2$ or $f(x,y) = Cy^2$, and so the line $x = 0$ or $y = 0$ consists of critical points.

So we see that if $B^2 = AC$, then there is a whole line of critical points rather than just one.

Third Example. To see further what happens when $B^2 - AC = 0$, consider the quadratic function

$$f(x,y) = 4x^2 - 12xy + 9y^2\,.$$

(Here $A = 4,\ B = 6$, and $C = 9$, so $B^2 - AC = 0$.) If we use the quadratic formula to find the lines where $f(x,y) = 0$, we get

$$x = \frac{1}{8}\,(12y \pm \sqrt{0})$$

so that the function vanishes only along the single line $x = 3y/2$. This corresponds to the factorization

$$f(x,y) = 4x^2 - 12xy + 9y^2 = (2x - 3y)^2\,.$$

We now see that $f(x,y)$ is never negative and so, as before, it has a minimum at $(0,0)$. However this is not an *isolated* minimum, since $f(x,y)$ takes the same minimum value along the entire line $x = 3y/2$. It would seem to follow, then, that all the points along the line $x = 3y/2$ must be critical points for the function. In fact,

$$\frac{\partial f}{\partial x} = 8x - 12y$$
$$\frac{\partial f}{\partial y} = -12x + 18y\,.$$

so that both partial derivatives vanish on the line $x = 3y/2$.

The General Quadratic Function. Consider a quadratic function of two variables: $f(x,y) = Ax^2 + 2Bxy + Cy^2 + Dx + Ey + G$. Note that $A = \frac{\partial^2 f}{\partial x^2}$, $B = \frac{\partial^2 f}{\partial x\,\partial y}$, and $C = \frac{\partial^2 f}{\partial y^2}$. There will usually be one critical point (x_0, y_0), which can be determined by solving the system of two equations

$$\frac{\partial f}{\partial x} = 2Ax + 2By + D = 0$$
$$\frac{\partial f}{\partial y} = 2Bx + 2Cy + E = 0\,.$$

By completing the square, we can rewrite the function as.

$$f(x,y) = A(x-x_0)^2 + 2B(x-x_0)(y-y_0) + C(y-y_0)^2 + f(x_0,y_0)\,.$$

PROOF: If one for the moment writes $g(x,y) = A(x-x_0)^2 + 2B(x-x_0)(y-y_0) + C(y-y_0)^2 + f(x_0,y_0)$, then one sees immediately that f and g have the same quadratic terms, so that $f(x,y) - g(x,y)$ is a polynomial of degree one:

$$f(x,y) - g(x,y) = Lx + My + N\,.$$

But simple differentiation shows that $\frac{\partial g}{\partial x}(x_0,y_0) = 0 = \frac{\partial f}{\partial x}(x_0,y_0)$ and $\frac{\partial g}{\partial y}(x_0,y_0) = 0 = \frac{\partial f}{\partial y}(x_0,y_0)$, so that $L = M = 0$. Furthermore, $N = 0$ since clearly $g(x_0,y_0) = f(x_0,y_0)$. Thus $g(x,y) = f(x,y)$. ☑

The point is that from the equation

$$f(x,y) = A(x-x_0)^2 + 2B(x-x_0)(y-y_0) + C(y-y_0)^2 + f(x_0,y_0)\,.$$

we can now see, as previously, that if $B^2 - AC < 0$ then either $f(x,y)$ takes on only values greater than or equal to $f(x_0,y_0)$ (in case $A > 0$) and consequently has a minimum at (x_0,y_0), or else takes on only values less than or equal to $f(x_0,y_0)$ (in case $A < 0$), in which case $f(x,y)$ has a maximum at (x_0,y_0). (Note also that if $B^2 - AC < 0$, then AC must be positive, so necessarily A and C have the same sign.)

But if $B^2 - AC > 0$ then $A(x-x_0)^2 + 2B(x-x_0)(y-y_0) + C(y-y_0)^2$ factors and consequently $f(x,y)$ takes on values both greater than and less than $f(x_0,y_0)$, so that (x_0,y_0) is a saddle point.

Finally, notice that if $B^2 - AC = 0$, then the system of equations

$$\frac{\partial f}{\partial x} = 2Ax + 2By + D = 0$$
$$\frac{\partial f}{\partial y} = 2Bx + 2Cy + E = 0$$

is either inconsistent, so that there is no critical point, or else the two equations are multiples of each other, so that there is a whole line of critical points.

For instance, if $f(x,y) = x^2 - 6xy + 9y^2 + 4x + 5y + 10$, then there is no critical point since the system

$$\frac{\partial f}{\partial x} = 2x - 6y + 4 = 0$$
$$\frac{\partial f}{\partial y} = -6x + 18y + 5 = 0$$

has no solution. But for the function $f(x,y) = x^2 - 6xy + 9y^2 + 4x - 12y + 10$, we get critical points along the line $x = 3y - 2$ since all points on this line satisfy the equations

$$\frac{\partial f}{\partial x} = 2x - 6y + 4 = 0$$
$$\frac{\partial f}{\partial y} = -6x + 18y - 12 = 0\,.$$

Furthermore, in the case that $B^2 - AC = 0$ then $A(x-x_0)^2 + 2B(x-x_0)(y-y_0) + C(y-y_0)^2$ is (except possibly for sign) a perfect square, so that if critical points for $f(x,y)$ exist, then they are all maxima (in case

$A < 0$) or minima (in case $A > 0$). Saddle points are not possible for a quadratic function in the case $B^2 - AC = 0$.

Functions Which Are Not Quadratics. Consider a critical point (x_0, y_0) for any differentiaable function of two variables $f(x, y)$. We will write

$$A = \frac{\partial^2 f}{\partial x^2}(x_0, y_0)\,, \qquad B = \frac{\partial^2 f}{\partial x\,\partial y}(x_0, y_0)\,, \qquad C = \frac{\partial^2 f}{\partial y^2}(x_0, y_0)\,.$$

It turns out that if $B^2 - AC \neq 0$, then the behavior of $f(x, y)$ near the critical point (x_0, y_0) will be the same as that of the quadratic function

$$q(x, y) = \tfrac{1}{2}A(x - x_0)^2 + B(x - x_0)(y - y_0) + \tfrac{1}{2}C(y - y_0)^2 + f(x_0, y_0)\,.$$

Example 4. In accord with the general principle that simple examples are usually more enlightening than complicated ones, consider the function

$$f(x, y) = x^2 + y^2 - y^3$$

and compare it to the quadratic function

$$q(x, y) = x^2 + y^2\,,$$

which has a minimum at $(0, 0)$. If $y < 1$ then $y^3 < y^2$, so

$$f(x, y) = x^2 + (y^2 - y^3) > 0\,,$$

so that for (x, y) close to $(0, 0)$, $f(x, y) > 0 = f(0, 0)$, showing that f has a local minimum at $(0, 0)$, just as the function $q(x, y) = x^2 + y^2$ does. (This is the case $B^2 - AC < 0$.)

Example 5. On the other hand, consider the function

$$f(x, y) = y^2 - x^2 - 10y^3\,,$$

in comparison to the quadratic function

$$q(x, y) = y^2 - x^2\,,$$

which has a saddle point at $(0, 0)$. If $0 < |y| < 1/10$ then $10y^3 < y^2$, so that for points on the y-axis sufficiently close to $(0, 0)$ (but with $y \neq 0$),

$$f(x, y) = f(0, y) = y^2 - 10y^3 > 0\,.$$

On the other hand, along the x-axis (for $x \neq 0$),

$$f(x, y) = f(x, 0) = -x^2 < 0\,.$$

Thus every neighborhood of the critical point $(0,0)$ contains both points (x,y) with $f(x,y) > f(0,0)$ and points with $f(x,y) < f(0,0)$. Therefore $f(x,y)$ has a saddle point at $(0,0)$. (This is the case $B^2 - AC > 0$.)

The point of these two examples is that when one one is looking for a *local* extremum of a polynomial, one is concerned only with points which are close to the critical point, and for points sufficiently close to the critical point, the terms with degree higher than two are so small that they don't really have any effect. (Except in the case $B^2 - AC = 0$, which is very delicate.)

Of course two examples are not a substitute for a proof, and these two examples are certainly simplistic. But it can be show that this logic is valid not only for polynomials but for any differentiable function $f(x,y)$ at a critical point (x_0, y_0) where $B^2 - AC \neq 0$. Such a function will have a maximum or minimum at (x_0, y_0) if $B^2 - AC < 0$ and a saddle point if $B^2 - AC > 0$.

However if $B^2 - AC = 0$, then at some points the quadratic approximation to the differentiable function will not will not be able to overpower the high order terms, and the second-derivative test fails to yield a clear conclusion.

Example 6. Choosing again the very simplest sort of example, compare the function

$$f(x,y) = x^2 - y^3$$

to the quadratic function

$$q(x,y) = x^2\,,$$

which has maxima along the entire line $x = 0$. (Here $A = 2,\ B = C = 0$ at $(0,0)$.) Since $q(x,y) = 0$ along the line $x = 0$, the value of $f(x,y) = q(x,y) - y^3$ along this line will be determined by on the cubic term, so that $B^2 - AC$ cannot tell us what happens. Every neighborhood of $(0,0)$ contains points $(x,0)$ with $f(x,y) = x^2 > 0$ and also contains points $(0,y)$ with $y > 0$ and $f(x,y) = -y^3 < 0$. Thus $(0,0)$ is a saddle point for $f(x,y)$, even though it is a minimum for $q(x,y)$.

To see why the sign of $B^2 - AC$ is decisive when $B^2 - AC \neq 0$ but is inconclusive when $B^2 - AC = 0$, compare this example with example 5,

$$f(x,y) = y^2 - x^2 - 10y^3\,.$$

As before, we write

$$f(x,y) = q(x,y) - 10y^3\,,$$

where

$$q(x, y) = y^2 - x^2 \,.$$

Now $q(x, y)$ is zero along the two lines $y = x$ and $y = -x$, and this fact was sufficient to conclude that $(0, 0)$ is a saddle point for the quadratic $q(x, y)$. However for $f(x, y)$ we need a stronger argument. The function $f(x, y) = y^2 - x^2 - 10y^3$ is in fact not zero along the lines $y = \pm x$, and it is not quite that easy to display those curves along which $f(x, y) = 0$. However what we can say, based on the knowledge that $(0, 0)$ is a saddle point for $q(x, y)$, is that every neighborhood of $(0, 0)$ contains points where $q(x, y) > q(0, 0) = 0$ and also points where $q(x, y) < 0$. Furthermore, if $q(x, y) > 0$, then by making y sufficiently small (for instance $y < .1$ when $x = 0$) we can make $10y^3 < q(x, y) = y^2 - x^2$ and make $f(x, y) = q(x, y) - 10y^3 > 0$. (This is easy to see for this particular example, but the reasoning can be made to work in general.)

And likewise, if $q(x, y) < 0$ we can make $|y|$ sufficiently small and find points where $q(x, y) < 0$ and $-y^3 < -q(x, y) = |q(x, y)|$ and therefore $f(x, y) = q(x, y) - y^3 < 0$. We conclude that $(0, 0)$ is also a saddle point for $f(x, y)$.

It would be tedious to spell out all the reasoning in detail, but the general idea is that if one has a polynomial

$$f(x, y) = q(x, y) + \text{higher order terms,}$$

where $q(x, y)$ is a quadratic with a critical point at $(0, 0)$, then if there are points where $q(x, y) > f(0, 0)$, then one can find such points which are so close to the origin that the effect of the higher order terms won't be large enough to disturb the inequality $f(x, y) > f(0, 0)$. Likewise if there exist points where $f(x, y) < f(0, 0)$. We then conclude that if $q(x, y)$ has a saddle point at the origin, then so does $f(x, y)$.

In the case where $q(x, y)$ has an *isolated* maximum or minimum at $(0, 0)$, the reasoning, as illustrated in Example 4, is similar although a bit more delicate. In this case, for the case of a minimum one must show that if $q(x, y) > f(0, 0) = q(0, 0)$ for all points except the origin, then for x and y sufficiently small, the higher order terms will not be large enough to effect this inequality. (Not that this only shows that $f(x, y)$ has a *local* minimum at $(0, 0)$. It does not prove that it has a *global* minimum there, and in fact that may not be true.)

But the reasoning breaks down when $B^2 - AC = 0$. In this case, there is a whole line consisting of critical points for $q(x, y)$, and these critical points are either all minima or all maxima. Assuming again, to simplify notation, that $(0, 0)$ is one of these critical points, say a minimum, and that $q(0,0) = 0$, we then have a line $Px + Qy = 0$ along which $q(x,y)$ takes the value 0 and with $q(x,y) > 0$ for all points not on this line. Now we would like to say that for when x and y are small enough, then the higher order terms will not be significant enough to affect this situation. But in fact, along the line where $q(x,y)$ is 0, the higher order terms, even when extremely small, could still be enough to make $f(x,y)$ negative. At points not close to this line, though, $f(x,y)$ would certainly have to be positive. Thus, as shown by Example 6, it is possible that a polynomal $f(x,y)$ with degree larger than 2 could still have a saddle point when $B^2 - AC = 0$, even though this is something that cannot happen for quadratics.