Optimization

Ryan C. Daileda

Trinity University

Calculus III

Local Extrema

Definition

- $f(x,y)$ has a *local maximum* at (a,b) if $f(x,y) \leq f(a,b)$ for all (x,y) near (a,b).
- $f(x,y)$ has a *local minimum* at (a,b) if $f(x,y) \geq f(a,b)$ for all (x,y) near (a,b).

Example. The function

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

has two local maxima and two local minima, all situated on the line $y = x$. See Maple.

The First Derivative Test in Two Variables

Question. How can we identify local extrema of $f(x, y)$?

Theorem 1

If $f(x, y)$ is differentiable at (a, b) and has a local extremum there, then $\nabla f(a, b) = \mathbf{0}$.

Idea of Proof. If $f(x, y)$ has a local maximum, say, at (a, b) then f has a local maximum as we move in *any* fixed direction $\mathbf{v}$.

From Calc. I we know that this means

$$D_{\mathbf{v}} f(a, b) = \frac{\nabla f(a, b) \cdot \mathbf{v}}{|\mathbf{v}|} = 0 \;\Rightarrow\; \nabla f(a, b) \cdot \mathbf{v} = 0.$$

This means $\nabla f(a, b)$ is orthogonal to *every* vector $\mathbf{v}$. Only $\nabla f(a, b) = \mathbf{0}$ has this property. □

Critical Points

Definition

We say that (a, b) is a *critical point* of $f(x, y)$ provided $\nabla f(a, b) = \mathbf{0}$.

Moral. If $f(x, y)$ is differentiable, then its local extrema must occur among its critical points.

To find the critical points of $f(x, y)$, we must solve the *vector equation*

$$\langle f_x(x, y), f_y(x, y)\rangle = \nabla f = \mathbf{0} = \langle 0, 0\rangle.$$

This is equivalent to the *simultaneous system* of equations

$$f_x(x, y) = 0 \quad \text{and} \quad f_y(x, y) = 0.$$

Example 1

Find and classify the critical points of $f(x,y) = 3x - x^3 - 2y^2 + y^4$.

Solution. The critical points are given by

$$f_x = 3 - 3x^2 = 0 \;\Leftrightarrow\; x^2 - 1 = 0 \;\Leftrightarrow\; (x-1)(x+1) = 0,$$
$$f_y = -4y + 4y^3 = 0 \;\Leftrightarrow\; y^3 - y = 0 \;\Leftrightarrow\; y(y-1)(y+1) = 0.$$

So we have $x = \pm 1$ and $y = 0, \pm 1$, with no correlation between the two.

So there are *six* critical points:

$$(\pm 1, 0), (\pm 1, 1), (\pm 1, -1).$$

Based on the graph of f we find that it has:

$$\begin{aligned} &\text{a local maximum at } (1,0), \\ &\text{local minima at } (-1,\pm 1), \\ &\textit{saddle points at } (1,\pm 1), (-1,0). \end{aligned}$$

□

Question. Is there a way to identify critical points *without* using the graph of $f(x,y)$?

Recall. In Calc. I we had the Second Derivative Test, which identified critical points of $f(x)$ based on its concavity.

The Second Derivative Test in Two Variables

By (essentially) considering the concavity of $f(x, y)$ in *every* direction one arrives at the following result.

Theorem 2 (Second Derivative Test)

Suppose that $f(x, y)$ has continuous second order partial derivatives at the point (a, b) and that $\nabla f(a, b) = \mathbf{0}$. Let

$$D = D(a, b) = f_{xx}(a, b) f_{yy}(a, b) - f_{xy}(a, b)^2 = \begin{vmatrix} f_{xx}(a, b) & f_{xy}(a, b) \\ f_{xy}(a, b) & f_{yy}(a, b) \end{vmatrix}.$$

Then:

1. *If $D > 0$ and $f_{xx}(a, b) > 0$, then f has a local min. at (a, b).*
2. *If $D > 0$ and $f_{xx}(a, b) < 0$, then f has a local max. at (a, b).*
3. *If $D < 0$, then f has a saddle point at (a, b).*
4. *If $D = 0$, the test fails.*

Remarks

The determinant $D = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix}$ is called the *Hessian* of f.

The Second Derivative Test succeeds in classifying a critical point (a, b) precisely when $D \neq 0$.

If $D > 0$, we can look at *either* of f_{xx} or f_{yy} to determine the concavity of the graph of f.

If $D = 0$, then we need to do something else entirely to classify (a, b).

Examples

Example 2

Use the Second Derivative Test to classify the critical points $(\pm 1, 0)$, $(\pm 1, 1)$ and $(\pm 1, -1)$ of $f(x, y) = 3x - x^3 - 2y^2 + y^4$.

Solution. We have

$$\begin{aligned} f_x = 3 - 3x^2 \ &\Rightarrow \ f_{xx} = -6x, \\ & \ f_{xy} = 0, \\ f_y = -4y + 4y^3 \ &\Rightarrow \ f_{yy} = -4 + 12y^2. \end{aligned}$$

Thus

$$D = f_{xx} f_{yy} - f_{xy}^2 = -6x(12y^2 - 4).$$

So we have the following table:

Point	D	f_{xx}	Type
$(1, \pm 1)$	< 0	NA	Saddle
$(-1, \pm 1)$	> 0	> 0	Local Min.
$(1, 0)$	> 0	< 0	Local Max.
$(-1, 0)$	< 0	NA	Saddle

which agrees with our graphical observations. □

There are two main difficulties in that arise in the classification of the critical points of $f(x, y)$:

- Finding the critical points requires us to solve a *system* of (often nonlinear) equations in two variables.
- The Second Derivative Test has lots of "moving parts."

Example 3

Find and classify the critical points of $f(x,y) = x^3 - 12xy + 8y^3$.

Solution. To find the critical points we need to solve the system

$$
\begin{aligned}
f_x &= 3x^2 - 12y = 0 \iff x^2 - 4y = 0 \iff x^2 = 4y, \\
f_y &= -12x + 24y^2 = 0 \iff -x + 2y^2 = 0 \iff 2y^2 = x.
\end{aligned}
$$

Substituting the second into the first we obtain

$$4y = (2y^2)^2 = 4y^4 \iff y^4 - y = 0 \iff y(y^3 - 1) = 0,$$

which tells us that $y = 0, 1$.

Since $x = 2y^2$ we find the corresponding values $x = 0, 2$.

So we have two critical points:

$$(0,0) \quad \text{and} \quad (2,1).$$

Now we compute the Hessian:

$$\left.\begin{aligned} f_{xx} &= 6x, \\ f_{xy} &= -12, \\ f_{yy} &= 48y \end{aligned}\right\} \;\Rightarrow\; D = f_{xx}f_{yy} - f_{xy}^2 = 288xy - 144 = 144(2xy-1).$$

Therefore:

Point	D	f_{xx}	Type
$(0,0)$	< 0	NA	Saddle
$(2,1)$	> 0	> 0	Local Min.

So what do we do when the Second Derivative Test fails? Whatever we can!

Example 4

Find and classify the critical points of $f(x, y) = x^2 + 4y^2 - 4xy + 2$.

Solution. The critical points are given by

$$f_x = 2x - 4y = 0 \iff x - 2y = 0 \iff x = 2y,$$
$$f_y = 8y - 4x = 0 \iff 2y - x = 0 \iff x = 2y.$$

That is, there are critical points *everywhere* along the line $x = 2y$.

Since

$$\left.\begin{array}{r} f_{xx} = 2, \\ f_{xy} = -4, \\ f_{yy} = 8 \end{array}\right\} \Rightarrow D = f_{xx}f_{yy} - f_{xy}^2 = 16 - (-4)^2 = 0,$$

the second derivative test *fails* at *every* critical point.

To classify the critical points we instead notice that

$$f(x,y) = x^2 + 4y^2 - 4xy + 2 = (x-2y)^2 + 2,$$

which shows that $f(x,y) \geq 2$ for all (x,y), and $f(x,y) = 2$ when $x = 2y$.

Therefore f has (absolute) minima all along the line $x = 2y$. □