Second Order Derivatives for a Function of Two Variables

Let $f(x,y)$ be a function from a subset of $\mathbb{R}^2$ into the set of real numbers. We defined its derivative at a point (x,y) to be the linear map $Df\,(x,y)$ that approximates $f(x+h,y+k)-f(x,y)$ in such a way that the error term $f(x+h,y+k)-f(x,y)-Df\,(x,y)\,(h,k)$ shrinks to 0 much faster than $||(h,k)||$. If f has continuous partial derivatives then a formula for $Df\,(x,y)\,(h,k)$ is known; it is

$$Df\,(x,y)\,(h,k)=f_x(x,y)h+f_y\,(x,y)\,k$$

Example 1 *Let $f(x,y)=x^3+4xy-y^2$. The domain of f is the set of all the points in $\mathbb{R}^2$. The derivative of f at an arbitrary point (x,y) is the linear approximator with formula*

$$D\,(x,y)\,(h,k)=\big(3x^2+4y)h+(4x-2y\big)\,k \tag{1}$$

For instance, if we choose $(x,y)=(2,-1)$ then (1) states that the derivative of f at $(2,-1)$ has formula

$$D\,(2,-1)\,(h,k)=8h+10k$$

If you choose a different point, you may get a different derivative. For example, the derivative of f at $(1,3)$ has formula

$$D\,(1,3)\,(h,k)=15h-2k$$

This is expected because the shape of the graph of f changes from point to point. To define the second derivative of f at a point (x,y) we ask the question: What linear expression, (linear in (r,s)), approximates

$$D\,(x+r,y+s)\,(h,k)-D\,(x,y)\,(h,k)$$

in such a way that the error in the approximation shrinks to 0 faster than $||(r,s)||$? To answer the question, we simply expand $D\,(x+r,y+s)\,(h,k)-D\,(x,y)\,(h,k)$ and simplify. The result is

$$D\,(x+r,y+s)\,(h,k)-D\,(x,y)\,(h,k)=\big(6xr+4s+3r^2\big)\,h+(4r-2s)\,k \tag{2}$$

It we write (2) as

$$D\,(x+r,y+s)\,(h,k)-D\,(x,y)\,(h,k)=(6xr+4s)\,h+(4r-2s)\,k+3r^2h$$

then the linear term in (r,s) is clear; it is $(6xr+4s)\,h+(4r-2s)\,k$. The error term is also clear; it is $3r^2h$ and it shrinks to 0 faster than $||(r,s)||$. Note that the expression $(6xr+4s)\,h+(4r-2s)\,k$ is a function of (r,s) and (h,k). The function which maps $((r,s),(h,k))$ into $(6xr+4s)\,h+(4r-2s)\,k$ is called the second derivative of f at (x,y) and it is denoted by $D^2f\,(x,y)$. More precisely,

$$D^2f\,(x,y)\,((r,s)\,,(h,k))=(6xr+4s)\,h+(4r-2s)\,k$$

Exercise 2 *For each given function f, do the following:*

(a) Evaluate the expression $Df(x+r,y+s)\,(h,k)-Df(x,y)\,(h,k)$.

(b) Determine the linear function that may be used to approximate $Df(x+r,y+s)\,(h,k)-Df(x,y)\,(h,k)$.

(c) Verify that the error in the approximation shrinks to 0 faster than $||(r,s)||$ and write down a formula for $D^2f\,(x,y)\,((r,s)\,,(h,k))$.

1. $f(x,y)=x^2y^2$
2. $f(x,y)=2x^3+4xy^3-y$
3. $f(x,y)=4xy-7x^3+3y^2$

Definition 3 *Let $f(x,y)$ be a function of two variables with a derivatives $Df(x,y)$ at every point (x,y) in its domain. It has a second derivative at a point (x,y) in its domain if there is a function that is linear in (r,s) and approximates $Df(x+r,y+s)(h,k) - Df(x,y)(h,k)$ in such a way that the error in the approximation shrinks to 0 faster than $||(r,s)||$. The linear approximator is then denoted by $D^2(x,y)$ and is called the second derivative of f at (x,y).*

It can be shown that if f has continuous first and second order partial derivatives then a formula for its second derivative is

$$D^2(x,y)((r,s),(h,k)) = \begin{pmatrix} r & s \end{pmatrix} \begin{pmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{xy}(x,y) & f_{yy}(x,y) \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}$$

In expanded form, $D^2(x,y)((r,s),(h,k)) = f_{xx}(x,y)rh + (rk+sh)f_{xy}(x,y) + skf_{yy}(x,y)$. To verify it, we apply it to the function $f(x,y)$ in Example 1. Since $f_{xx} = 6x$, $f_{xy} = 4$ and $f_{yy} = -2$, the result, (as expected) is

$$\begin{aligned} D^2(x,y)((r,s),(h,k)) &= \begin{pmatrix} r & s \end{pmatrix} \begin{pmatrix} 6x & 4 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} \\ &= \begin{pmatrix} 6xr+4s & 4r-2s \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} = (6xr+4s)h + (4r-2s)k \end{aligned}$$

We can now state, (without proof), a version of Taylor's theorem for a function of two variables and derive the test for the nature of a critical point which we promised earlier.

Theorem 4 *(Taylor's Theorem). Let f have continuous first and second order partial derivatives at all points in its domain. Let (x,y) and $(x+h,y+k)$ be in the domain of f. Also assume that the line segment joining (x,y) to $(x+h,y+k)$ is in the domain of f. Then there is a point (α,β) on the line segment such that*

$$f(x+h,y+k) = f(x,y) + Df(x,y)(h,k) + D^2f(\alpha,\beta)((h,k),(h,k))$$

We now justify the test for the nature of a critical point which we introduced on page **??**. Thus assume that (c,d) is a critical point of f. Then $f_x(c,d)$ is zero and so is $f_y(c,d)$. Take any point $(c+h,d+k)$ near (c,d). By Taylor's Theorem, there is point (α,β) on the line segment joining (c,d) and $(c+h,d+k)$ such that

$$f(c+h,d+k) = f(c,d) + Df(c,d)(h,k) + D^2f(\alpha,\beta)((h,k),(h,k)) \tag{3}$$

In matrix form,

$$\begin{aligned} f(c+h,d+k) &= f(c,d) + (0)h + (0)k + \begin{pmatrix} h & k \end{pmatrix} \begin{pmatrix} f_{xx}(\alpha,\beta) & f_{xy}(\alpha,\beta) \\ f_{xy}(\alpha,\beta) & f_{yy}(\alpha,\beta) \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} \\ &= f(c,d) + h^2 f_{xx}(\alpha,\beta) + 2hkf_{xy}(\alpha,\beta) + k^2 f_{yy}(\alpha,\beta) \end{aligned}$$

Since the partial derivatives of f are continuous, $f_{xx}(\alpha,\beta) \simeq f_{xx}(c,d)$, $f_{xy}(\alpha,\beta) \simeq f_{xy}(c,d)$ and $f_{yy}(\alpha,\beta) \simeq f_{yy}(c,d)$. Denote $f_{xx}(c,d)$ by A, $f_{xy}(c,d)$ by B, $f_{yy}(c,d)$ by C, $AC - B^2$ by H and move $f(c,d)$ to the left hand side. The result is

$$f(c+h,d+k) - f(c,d) \simeq Ah^2 + 2Bhk + Ck^2 \tag{4}$$

The right hand side of (4) is an example of what is called a quadratic form in two variables, (in this case the variables are h and k).

Suppose $H > 0$, (i.e. suppose $AC - B^2 > 0$).

Then A and C are non-zero and they have the same sign. The main idea is to write $Ah^2 + 2Bhk + Ck^2$ as a linear combination of two squared terms. To this end, add and subtract $\dfrac{B^2k^2}{A}$ to the right hand side of (4). The result is

$$\begin{aligned} f(c+h, d+k) - f(c,d) &\simeq Ah^2 + 2Bhk + \frac{B^2k^2}{A} + Ck^2 - \frac{B^2k^2}{A} \\ &= A\left(h^2 + \frac{2Bhk}{A} + \frac{B^2k^2}{A^2}\right) + \left(AC - B^2\right)\frac{k^2}{A} \\ &= A\left(h + \frac{Bk}{A}\right)^2 + H\frac{k^2}{A} \end{aligned}$$

If A is positive then as long as h and k are close to 0 (so that the approximation (4) is valid),

$$f(c+h, d+k) - f(c,d) \simeq A\left(h + \tfrac{Bk}{A}\right)^2 + H\tfrac{k^2}{A} \geq 0$$

therefore $f(x,y) \geq f(c,d)$ for all (x,y) in some neighborhood of (c,d). This implies that (c,d) is a point of relative minimum. However, if H is positive and A is negative then

$$f(c+h, d+k) - f(c,d) \simeq A\left(h + \frac{Bk}{A}\right)^2 + H\frac{k^2}{A} \leq 0$$

for all sufficiently small values of h and k. This implies that $f(c,d) \geq f(x,y)$ for all (x,y) in some neighborhood of (c,d), hence (c,d) is a point of relative maximum.

Suppose H is negative.
If A is also negative then $\frac{H}{A}$ is positive. If we choose $h = -\frac{Bk}{A}$ then

$$f(c+h, d+k) - f(c,d) \simeq A\left(h + \frac{Bk}{A}\right)^2 + H\frac{k^2}{A} = H\frac{k^2}{A} \geq 0$$

On the other hand, the choice $k = 0$ and $h \neq 0$ gives

$$f(c+h, d+k) - f(c,d) \simeq A\left(h + \frac{Bk}{A}\right)^2 + H\frac{k^2}{A} = Ah^2 \leq 0$$

This shows that every neighborhood of (c,d) contains points (x,y) such that $f(x,y) > f(c,d)$ and others (u,v) such that $f(u,v) < f(c,d)$. Therefore (c,d) is neither a point of relative maximum nor a point of relative minimum. A similar conclusion is arrived at if A is positive. If $A = 0$ we arrive at the same, but conclusion along a different route. More precisely, when $A = 0$ then (4) implies that

$$f(c+h, d+k) - f(c,d) \simeq 2Bhk + Ck^2 \tag{5}$$

with $B \neq 0$, (else H would not be negative). By fixing k, we can find values of h that make $2Bhk + Ck^2$ positive and others that make it negative. Therefore every neighborhood of (c,d) contains points (x,y) such that $f(x,y) > f(c,d)$ and others (u,v) such that $f(u,v) < f(c,d)$, which implies that (c,d) is neither a point of relative minimum nor a point of relative maximum.

To prove the last part it suffices to give examples. Take

$$f(x,y) = x^2 + y^3, \;\; g(x,y) = x^4 + y^4, \;\; h(x,y) = -x^4 - y^4 \text{ and the point } (0,0).$$

It turns out that $(0,0)$ is a critical point for all the three functions, (verify). Furthermore, $H = 0$ for all three. Note that $g(x,y) = x^4 + y^4 \geq 0$ for all $(x,y) \in \mathbb{R}^2$ and $g(0,0) = 0$, therefore $(0,0)$ is a point of relative minimum for g. A similar argument shows that $(0,0)$ is a point of relative maximum for h. In the case of f, we observe that if $y > 0$ then $f(0,y) > 0 = f(0,0)$ and if $y < 0$ then $f(0,y) < f(0,0)$. This shows that every neighborhood of $(0,0)$ contains points (x,y) such that $f(x,y) > f(0,0)$ and others (u,v) such that $f(u,v) < f(0,0)$, which implies that $(0,0)$ is neither a point of relative minimum nor a point of relative maximum.