Local extrema and saddle points

Ever since Calculus II you have known that a sufficiently smooth ordinary function of a single variable, say f, can be well approximated in a sufficiently small neighborhood of any fixed value by either a linear Taylor polynomial L or a quadratic Taylor polynomial Q. More precisely, for $x \approx a$,

$$L(x) \;=\; f(a) + (x-a)f'(a) \tag{1}$$

satisfies $f(x) \approx L(x)$ with a relatively small error and

$$Q(x) \;=\; f(a) + (x-a)f'(a) + \tfrac{1}{2}(x-a)^2 f''(a) \tag{2}$$

satisfies $f(x) \approx Q(x)$ with an even smaller error; and the closer x is to a, the better is either approximation. What changes when f is a function of two variables? Not much! If f depends on y as well as x but we are holding y constant, then any derivative with respect to x is unaffected, except to the extent that we use a different notation to acknowledge that f now depends on y as well as on x. So, in place of $f(x) \approx L(x)$ and $f(x) \approx Q(x)$, we have

$$f(x,y) \approx f(a,y) + (x-a)f_x(a,y) \tag{3}$$

with a relatively small error; and, with an even smaller error,

$$f(x,y) \approx f(a,y) + (x-a)f_x(a,y) + \tfrac{1}{2}(x-a)^2 f_{xx}(a,y). \tag{4}$$

The closer x is to a, the better is either approximation.

When considering an ordinary function of a single variable, why did I choose x to denote the independent variable, a to denote its fixed value and f to denote the function? No especially good reason: just force of habit. So why not stir things up a bit by choosing y to denote the independent variable, b to denote its fixed value and g to denote the function? In place of $f(x) \approx L(x)$ or $f(x) \approx Q(x)$ for $x \approx a$ we find that, for $y \approx b$,

$$g(y) \approx g(b) + (y-b)g'(b) \tag{5}$$

with a relatively small error and

$$g(y) \approx g(b) + (y-b)g'(b) + \tfrac{1}{2}(y-b)^2 g''(b) \tag{6}$$

with an even smaller error; and the closer y is to b, the better is either approximation. What changes when g is a function of two variables? Same answer as before: not much! If g depends on x as well as y but we are holding x constant, then any derivative with respect to y is unaffected, except to the extent that we use a different notation to acknowledge that g now depends on x as well as on y. So, in place of (5)-(6) we obtain

$$g(x,y) \approx g(x,b) + (y-b)g_y(x,b) \tag{7}$$

$$g(x,y) \approx g(x,b) + (y-b)g_y(x,b) + \tfrac{1}{2}(y-b)^2 g_{yy}(x,b). \tag{8}$$

But g is totally arbitrary; and so, for that matter, is x. So let's set $g = f$ and $x = a$ in (8). We obtain

$$f(a,y) \approx f(a,b) + (y-b)f_y(a,b) + \tfrac{1}{2}(y-b)^2 f_{yy}(a,b). \tag{9}$$

But g is still totally arbitrary. So let's set $g = f_x$ and $x = a$ in (7) to obtain

$$f_x(a, y) \approx f_x(a, b) + (y - b)f_{xy}(a, b), \tag{10}$$

and then set $g = f_{xx}$ and $x = a$ in (7) to obtain

$$f_{xx}(a, y) \approx f_{xx}(a, b) + (y - b)f_{xxy}(a, b). \tag{11}$$

Now substitute from (9)-(11) into (4), i.e., $f(x, y) \approx f(a, y) + (x - a)f_x(a, y) + \frac{1}{2}(x - a)^2 f_{xx}(a, y)$. We obtain

$$\begin{aligned} f(x, y) &\approx f(a, b) + (y - b)f_y(a, b) + \tfrac{1}{2}(y - b)^2 f_{yy}(a, b) \\ &\qquad + (x - a)\left\{f_x(a, b) + (y - b)f_{xy}(a, b)\right\} \\ &\qquad + \tfrac{1}{2}(x - a)^2 \left\{f_{xx}(a, b) + (y - b)f_{xxy}(a, b)\right\} \\ &\approx f(a, b) + (x - a)f_x(a, b) + (y - b)f_y(a, b) + \tfrac{1}{2}(x - a)^2 f_{xx}(a, b) \\ &\qquad + (x - a)(y - b)f_{xy}(a, b) + \tfrac{1}{2}(y - b)^2 f_{yy}(a, b) \end{aligned} \tag{12}$$

plus a term that we ignore because $(x - a)^2(y - b)$ is cubic in small quantities (whereas $(x - a)^2$, $(x - a)(y - b)$ and $(y - b)^2$ are all quadratic in small quantities). The right-hand side of the above equation is the quadratic approximation to f in the vicinity of (a, b), and the first three terms of it are the corresponding linear approximation.[1]

In practice, the principal value of this quadratic approximation, at least as far as we are concerned, is to help us characterize the shape of a surface graph in the vicinity of a critical point, i.e., a point (a, b) in the domain of f at which *all* directional derivatives are zero. But if $\mathbf{u} = \cos(\theta)\mathbf{i} + \sin(\theta)\mathbf{j}$ then

$$\left.\frac{\partial f}{\partial u}\right|_{\substack{x=a\\y=b}} = \hat{\mathbf{u}} \cdot \nabla f(a, b) = \cos(\theta)f_x(a, b) + \sin(\theta)f_y(a, b). \tag{13}$$

Thus, in practice, a critical point is defined to be where both partial derivatives are simultaneously zero:[2]

$$f_x(a, b) = f_y(a, b) = 0. \tag{14}$$

In the vicinity of such a point, (12) simplifies to

$$\begin{aligned} f(x, y) &\approx f(a, b) + \tfrac{1}{2}(x - a)^2 f_{xx}(a, b) \\ &\qquad + (x - a)(y - b)f_{xy}(a, b) + \tfrac{1}{2}(y - b)^2 f_{yy}(a, b). \end{aligned} \tag{15}$$

---

[1]In other words, the first Taylor polynomial L and second Taylor polynomial Q centered on (a, b) for a function of two variables are defined by

$$\begin{aligned} L(x, y) &= f(a, b) + (x - a)f_x(a, b) + (y - b)f_y(a, b) \\ Q(x, y) &= L(x, y) + \tfrac{1}{2}(x - a)^2 f_{xx}(a, b) + (x - a)(y - b)f_{xy}(a, b) + \tfrac{1}{2}(y - b)^2 f_{yy}(a, b). \end{aligned}$$

Of course, L corresponds to a tangent-plane approximation.

[2]Because if both partial derivatives are zero then certainly every directional derivative is zero; and conversely, if every directional derivative is zero, then the directional derivative must be zero for both $\theta = 0$ and $\theta = \frac{1}{2}\pi$, implying (14). A point in the domain is also called critical if the partial derivatives do not both exist there (e.g., $(0, 0)$ for $f(x, y) = \frac{1}{x^2+y^2}$), but we have assumed throughout that f is smooth.

If we are interested only in points in space near $(a, b, f(a, b))$ on the graph $z = f(x, y)$ of the function f, however, then we may as well move the origin there by defining new coordinates X, Y, Z:

$$X = x - a, \qquad Y = y - b, \qquad Z = z - f(a, b). \tag{16}$$

It also helps to define

$$A = \tfrac{1}{2} f_{xx}(a, b), \qquad B = f_{xy}(a, b), \qquad C = \tfrac{1}{2} f_{yy}(a, b) \tag{17}$$

Then, by substituting from (16)-(17) into (15), we find that the surface $z = f(x, y)$ in the vicinity of $(a, b, f(a, b))$ becomes the surface $Z = F(X, Y)$ in the vicinity of $(0, 0, 0)$, where

$$F(X, Y) = A X^2 + B XY + C Y^2 = \frac{(2AX + BY)^2 + DY^2}{4A} \tag{18}$$

and the "discriminant" D is defined by

$$D = 4AC - B^2 = f_{xx}(a, b) f_{yy}(a, b) - \{f_{xy}(a, b)\}^2. \tag{19}$$

It now follows more or less immediately by inspection that $D > 0$ makes $F(X, Y)$ positive or negative in the vicinity of our new origin (i.e., the critical point) according to whether A is positive or negative, but that $D < 0$ enables $F(X, Y)$ to be both positive and negative in the same vicinity.[3] Thus $D > 0, A > 0$ corresponds to a local minimum, $D > 0, A < 0$ to a local maximum, and $D < 0$ to a saddle point; but if $D = 0$ the quadratic terms do not suffice to determine the shape of the function near the critical point. These points will be illustrated by examples.

Finally, it is important to note that $D > 0$ forces A and C to have the same sign. Why? Because $-B^2$ cannot be positive. Thus $D > 0$ implies $AC > 0$ and hence that A and C are either both negative or both positive. So although it is traditional to say that $D > 0, A > 0$ corresponds to a local minimum and $D > 0, A < 0$ to a local maximum, we could equally well say that $D > 0, C > 0$ corresponds to a local minimum and $D > 0, C < 0$ to a local maximum: either is equivalent to the other. Why calculus texts seem always to prefer the first and never even mention the second is just . . . well, another of those mysteries of life.

[3] For $D < 0$, if $A < 0$, then choosing $Y = 0$ makes Z negative, whereas choosing $2AX + BY = 0$ makes Z positive; and if $A > 0$, then choosing $2AX + BY = 0$ makes Z negative, whereas choosing $Y = 0$ makes Z positive.