

Note 1

Quadratic functions, Taylor series, critical points, second derivative test, least squares

1. Quadratic functions

These are functions of the form

$$f(x, y) = ax^2 + 2bxy + cz^2,$$

where a, b, c are constants. (The reason for the 2 will appear later.)

Examples:

$$(1.1) \quad f(x, y) = x^2 + y^2 \quad \text{paraboloid opening upward}$$

$$(1.2) \quad f(x, y) = x^2 - y^2 \quad \text{saddle}$$

$$(1.3) \quad f(x, y) = xy \quad \text{another saddle}$$

$$(1.4) \quad f(x, y) = -x^2 - y^2 \quad \text{paraboloid opening downward}$$

$$(1.5) \quad f(x, y) = x^2 + 2xy + y^2 = (x + y)^2 \quad \text{parabolic trough sitting on line } x + y = 0.$$

What conditions on a, b, c determine whether the graph of $f(x, y) = ax^2 + 2bxy + cz^2$ is a paraboloid, saddle, or trough? Answer:

$$ac - b^2 < 0 \quad \Rightarrow \text{saddle.}$$

$$ac - b^2 > 0 \quad \Rightarrow \text{paraboloid,}$$

$$ac - b^2 = 0 \quad \Rightarrow \text{trough,}$$

Here's why: It's a saddle exactly when $f(x, y)$ factors into a product of two lines. Now

$$f(x, y) = y^2[a(x/y)^2 + 2b(x/y) + c],$$

so $f(x, y)$ factors when exactly when the polynomial $at^2 + 2bt + c$ factors, which happens exactly when $4b^2 - 4ac \geq 0$, ie when

$$ac - b^2 \leq 0.$$

If $ac - b^2 = 0$ the factors are the same line, and we get a trough on this line. If $ac - b^2 < 0$ then the lines are distinct, and we get a saddle.

More examples:

$$(1.6) \quad f(x, y) = 3x^2 + 7xy + 2y^2.$$

Here $ac - b^2 = 6 - (7/2)^2 < 0$, so it's a saddle. In fact,

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

$$(1.7) \quad f(x, y) = 4x^2 - 4xy + y^2$$

Here $ac - b^2 = 4 - 2^2 = 0$, so it's a trough. In fact,

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

so the trough is on the line $2x - y = 0$.

$$(1.8) \quad f(x, y) = x^2 + xy + y^2$$

Here $ac - b^2 = 1 - (1/2)^2 > 0$ so it doesn't factor. It's a paraboloid, opening up.

In general, if $ac - b^2 > 0$, so it's a paraboloid, then a and c must have the same sign. The paraboloid opens up exactly when a and c are positive.

2. Taylor approximations

We now consider more general functions $f(x, y)$. Any decent function $z = f(x, y)$ has a Taylor expansion at a point $p = (a, b)$:

$$\begin{aligned} f(x, y) &= f(p) + [f_x(p)(x - a) + f_y(p)(y - b)] \\ &\quad + \frac{1}{2}[f_{xx}(p)(x - a)^2 + 2f_{xy}(p)(x - a)(y - b) + f_{yy}(p)(y - b)^2] \\ &\quad + \text{higher powers of } (x - a), (y - b). \end{aligned}$$

Here $f(p)$ is the constant term, $[f_x(p)(x - a) + f_y(p)(y - b)]$ is the linear term, and

$$[f_{xx}(p)(x - a)^2 + 2f_{xy}(p)(x - a)(y - b) + f_{yy}(p)(y - b)^2]$$

is the quadratic term. It is the analogue of $f''(a)$ in the one-variable Taylor expansion.

If we neglect all but the constant plus linear terms, we get

$$z = f(p) + f_x(p)(x - a) + f_y(p)(y - b),$$

which is just the equation of the tangent plane to the graph of f above p . To go beyond this linear approximation, we look at the quadratic term.

Example 1:

$$(2.1) \quad f(x, y) = (x^2 - y^2 - 1)^2 + 4x^2y^2.$$

If we multiply out, and neglect higher terms, we get

$$f(x, y) = 1 - 2(x^2 - y^2) + \dots$$

The constant term is

$$1 = f(0, 0),$$

The linear term is

$$0 = f_x(0, 0)x + f_y(0, 0)y,$$

and the quadratic term is (up to constant) $x^2 - y^2$. Near $(0, 0)$, the function looks like

$$1 - 2(x^2 - y^2),$$

which is a saddle, lifted up by 1. The level curve $f(x, y) = 1$ is a figure eight (“lemniscate”), crossing at $(0, 0)$, and this cross is approximated by the saddle cross. The level curves of $f(x, y)$ are the equi-potentials in the xy -plane caused by two long parallel wires with equal charge in the z direction passing through $(1, 0)$ and $(-1, 0)$.

Example 2:

$$(2.2) \quad f(x, y) = \cos(x) \cos(y).$$

The one-variable Taylor expansion of $\cos(x)$ is

$$\cos(x) = 1 - x^2/2 + \dots,$$

likewise for $\cos(y)$. Therefore,

$$f(x, y) = (1 - x^2/2 + \dots)(1 - y^2/2 + \dots) = 1 - (1/2)(x^2 + y^2) + \dots$$

Near $(0, 0)$, this function looks like

$$1 - (1/2)(x^2 + y^2),$$

which is a downward paraboloid with vertex $(0, 0, 1)$. This paraboloid approximates the graph of $f(x, y)$ near $(0, 0)$.

3. Critical Points

A **critical point** of a function $f(x, y)$ is a point where the gradient of f is zero. This is the same as saying the tangent plane is horizontal. Maxima and minima of f can only occur at critical points, but other things can happen there as well.

To find the critical points of $f(x, y)$, set $f_x = 0$, $f_y = 0$ and solve for x, y .

Example (2.1):

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

After some simplification, we get

$$f_x = 4x(x^2 + y^2 - 1), \quad f_y = 4y(x^2 + y^2 + 1).$$

For $f_y = 0$ we must have $y = 0$. Then $f_x = 0$ for $x = 0, \pm 1$, so there are three critical points: $(0, 0), (1, 0), (-1, 0)$. The critical point $(0, 0)$ is where the figure eight crosses itself. The other two critical points are inside the loops of the figure eight.

Example (2.2):

$$f(x, y) = \cos(x) \cos(y).$$

We have

$$f_x = -\sin(x) \cos(y), \quad f_y = -\cos(x) \sin(y).$$

The critical points are $(n\pi, m\pi)$ and $((2n+1)\frac{\pi}{2}, (2m+1)\frac{\pi}{2})$, where n and m are integers. There are infinitely many critical points, and you see they come in two flavors. The critical points $((2n+1)\frac{\pi}{2}, (2m+1)\frac{\pi}{2})$ are the intersection points in the grid where $f = 0$. At the critical points $(n\pi, m\pi)$ the function has the value $(-1)^{n+m}$, which are the maximum and minimum values of f . The level curves of f look like a checkerboard full of bloodshot eyes, and the graph of f looks like an egg-carton.

4. The second derivative test

One of four things can happen at a critical point: $f(x, y)$ will have a local maximum, a local minimum, a saddle, or a “degenerate” critical point. The behavior of $f(x, y)$ near a critical point p is determined by the quadratic term in the Taylor expansion of f at p . The second derivative test will determine which of the four possibilities occurs, by detecting the behavior of the quadratic term.

Suppose you have a critical point $p = (a, b)$ of $f(x, y)$. Then $f_x(p) = f_y(p) = 0$, so the Taylor expansion has no linear term:

$$f(x, y) = f(p) + \frac{1}{2}[f_{xx}(p)(x-a)^2 + 2f_{xy}(p)(x-a)(y-b) + f_{yy}(p)(y-b)^2] \\ + \text{higher powers of } (x-a), (y-b).$$

From part 1, we know the quadratic term is determined by the sign of

$$H_f(p) = f_{xx}(p)f_{yy}(p) - f_{xy}(p)^2.$$

(H stands for “Hessian”, which is the strange name of this quantity.) Namely, the quadratic term is

a saddle if $H_f(p) < 0$

an upward opening paraboloid if $H_f(p) > 0$ and $f_{xx}(p) < 0$ (concave up)

a downward opening paraboloid if $H_f(p) > 0$ and $f_{xx}(p) > 0$ (concave down)

a trough if $H_f(p) = 0$.

In this last case, where $H_f(p) = 0$, the quadratic term tells us nothing about f at p , so the second derivative test fails. If $H_f(p) = 0$, we say the critical point p is “degenerate”. An interesting example of a degenerate critical point is the function

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

This is called the “monkey saddle” because it is a saddle with room for a monkey’s tail. For the rest of this note, we ignore degenerate critical points.

Example (2.1) revisited:

$$f(x, y) = (x^2 - y^2 - 1)^2 + 4x^2y^2, \\ H_f(0, 0) = (-4)(4) - (0)^2 = -16 \quad (\text{saddle}) \\ H_f(\pm 1, 0) = (8)(8) - (0)^2 = 64 \quad (\text{minima}).$$

Example (2.2) revisited:

$$\begin{aligned} f(x, y) &= \cos(x) \cos(y), \\ H_f(x, y) &= (-\cos x)(-\cos y)^2 - (\sin x \sin y)^2 \\ H_f\left((2n+1)\frac{\pi}{2}, (2m+1)\frac{\pi}{2}\right) &= -1 \quad (\text{saddle}) \\ H_f(n\pi, m\pi) &= 1, \quad f_{xx}(n\pi, m\pi) = (-1)^{n+m+1} \quad (\text{max/min}), \end{aligned}$$

and it's a max exactly when $n + m$ is even. For example, $(0, 0)$ is a max, $(\pi, 0)$ is a min.

5. Least squares

A simple application of the second derivative test is finding the best line through a collection of data points. Typically, the points won't all lie on one line, so we want to find the "best" line that approximates these data points. The Least Squares way of measuring "best" is to minimize the sum of the squares of the vertical distances from the line to your data points.

To get the idea, consider three data points

$$p_i = (a_i, b_i), \quad i = 1, 2, 3.$$

If $a_1 = a_2 = a_3$, then clearly the best line is the vertical line containing all three points. Suppose not all a_i 's are equal. Then our line is

$$y = mx + k,$$

with m, k to be determined. Let

$$f(m, k) = [ma_1 + k - b_1]^2 + [ma_2 + k - b_2]^2 + [ma_3 + k - b_3]^2.$$

This is the sum of squares of vertical distances from the line to our three points. We want to minimize $f(m, k)$. We think of m, k as the variables, and find the partial derivatives:

$$\begin{aligned} f_m &= 2a_1[ma_1 + k - b_1] + 2a_2[ma_2 + k - b_2] + 2a_3[ma_3 + k - b_3] \\ &= 2(a_1^2 + a_2^2 + a_3^2)m + 2(a_1 + a_2 + a_3)k - 2(a_1b_1 + a_2b_2 + a_3b_3), \\ f_k &= 2[ma_1 + k - b_1] + 2[ma_2 + k - b_2] + 2[ma_3 + k - b_3] \\ &= 2(a_1 + a_2 + a_3)m + 6k - 2(b_1 + b_2 + b_3). \end{aligned}$$

Now set $f_m = 0$, $f_k = 0$, and solve for m, k . To see that this is a minimum, we compute

$$\begin{aligned} H_f &= f_{mm}f_{kk} - f_{km}^2 \\ &= 12(a_1^2 + a_2^2 + a_3^2) - 4(a_1 + a_2 + a_3)^2 \\ &= 4[(a_1 - a_2)^2 + (a_1 - a_3)^2 + (a_2 - a_3)^2](!) \\ &> 0, \end{aligned}$$

since not all a_i 's are equal, and $f_{mm} = 2(a_1^2 + a_2^2 + a_3^2) > 0$, so the solution for (m, k) , whatever it turns out to be, must be a minimum.