Optimization

Scott Surgent

Optimization is the process of determining the highest (maximum) and lowest (minimum) points on a graph.

Maximum and minimum points are collectively called **extreme points**, or **extrema**.

Let $z = f(x, y)$ be a function in R^3, and assume that f exists and is continuous over the entire *xy*-plane. That is, its domain is R^2, there being no restrictions on variables *x* and *y*.

A **critical point** (x_c, y_c, z_c), where $z_c = f(x_c, y_c)$, is a point where $f_x(x_c, y_c) = 0$ or does not exist, *and* where $f_y(x_c, y_c) = 0$ or does not exist.

All minimum and maximum points are **local** (or **relative**), meaning that the point is the lowest or highest point within some open region in R^2 that includes the point.

If it is the lowest or highest point over the entire domain, then the point is an **absolute** minimum or maximum.

The **second derivative test** for R^2 is one way to determine if a critical point (x_c, y_c, z_c) is a minimum, a maximum, or neither. The formula is

$$D = f_{xx}(x_c, y_c) f_{yy}(x_c, y_c) - \left(f_{xy}(x_c, y_c)\right)^2 .$$

• If $D > 0$ and if $f_{xx}(x_c, y_c) > 0$, then the graph of f is concave upward, and (x_c, y_c, z_c) is a relative minimum.
• If $D > 0$ and if $f_{xx}(x_c, y_c) < 0$, then the graph of f is concave downward, and (x_c, y_c, z_c) is a relative maximum.
• If $D < 0$, then (x_c, y_c, z_c) is not a minimum nor a maximum. It is a saddle point.
• If $D = 0$, then no conclusion about (x_c, y_c, z_c) can be inferred. Other methods need to be used to classify the critical point.

When $D > 0$, the signs of $f_{xx}(x_c, y_c)$ and $f_{yy}(x_c, y_c)$ will be the same. Thus, when $D > 0$, it is sufficient to note the sign of one since the sign of the other will be identical.

When there are no restrictions on the domain, this process is called **unconstrained optimization**.

Example 1: Let $z = f(x, y) = x^2 + y^2 + 6x - 4y + 2$. Find its critical points and classify these points as minima, maxima or saddle.

Solution: Find the first partial derivatives:

$$f_x(x, y) = 2x + 6$$
$$f_y(x, y) = 2y - 4.$$

Note that the derivatives (as well as the function itself) are defined for all x and all y in R^2. Thus, there are no possible locations where the derivatives "do not exist". Then set the partial derivatives to 0, and solve:

$$\begin{matrix} 2x + 6 = 0 \\ 2y - 4 = 0 \end{matrix} \quad \text{which gives} \quad \begin{matrix} x = -3 \\ y = 2. \end{matrix}$$

Thus, we have one critical point, $\big(-3,2, f(-3,2)\big)$, where $f(-3,2) = -11$. To classify this critical point, we use the second derivative test. The second derivatives are found first (recall that $f_{xy}(x, y) = f_{yx}(x, y)$):

$$f_{xx}(x, y) = 2, \qquad f_{yy}(x, y) = 2 \quad \text{and} \quad f_{xy}(x, y) = 0.$$

By the second derivative test, we have

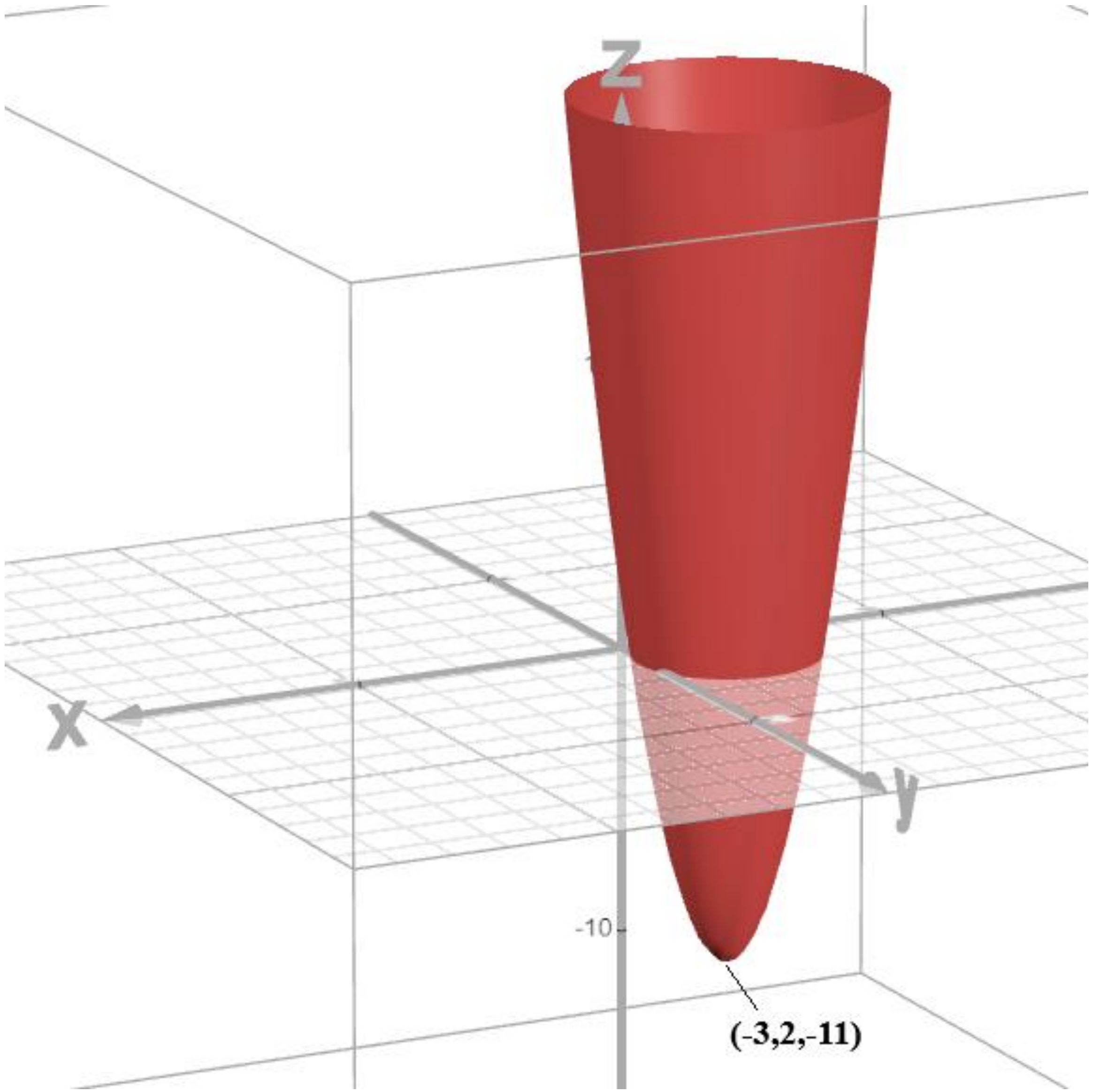
$$\begin{aligned} D &= f_{xx}(-3,2)f_{yy}(-3,2) - \left(f_{xy}(-3,2)\right)^2 \\ &= (2)(2) - 0^2 \\ &= 4. \end{aligned}$$

Note that $D > 0$ and that $f_{xx} > 0$. Therefore, $(-3,2,-11)$ is a local minimum point.

The graph of $z = f(x,y) = x^2 + y^2 + 6x - 4y + 2$ is a paraboloid that opens upward (in the direction of positive *z*).

Its vertex is $(-3,2,-11)$.

This point is also the absolute minimum point over the entire domain.



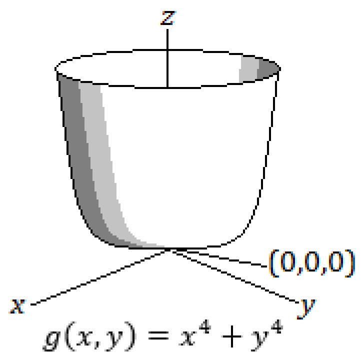
Example 2: Find the critical points of $z = g(x, y) = x^4 + y^4$, and classify these points as minima, maxima or saddle.

Solution: The first partial derivatives are $g_x(x, y) = 4x^3$ and $g_y(x, y) = 4y^3$.

Setting each to 0, we get $x = 0$ and $y = 0$. Note that $z = g(0,0) = 0$, so that (0,0,0) is the lone critical point.

The second derivatives are $g_{xx}(x, y) = 12x^2$, $g_{yy}(x, y) = 12y^2$ and $g_{xy}(x, y) = 0$. Using the second derivative test:

$$\begin{aligned} D &= g_{xx}(0,0)g_{yy}(0,0) - \left(g_{xy}(0,0)\right)^2 \\ &= [12(0)^2][12(0)^2] - 0 \\ &= 0. \end{aligned}$$



The second derivative test yields no useful information. However, note that the cross sections of this surface are $z = x^4$ (when $y = 0$) and $z = y^4$ (when $x = 0$). In each case, the point (0,0) is a minimum, so we can infer that (0,0,0) is a local minimum point on the surface of $z = x^4 + y^4$. The surface is bowl-shaped, with a flattened bottom, where (0,0,0) is its vertex. Viewing its graph suggests that the point is the absolute minimum.

Example 3: Let $z = f(x, y) = x^3 + y^3 - 3x - 27y + 7$. Find its critical points and classify these points as minima, maxima or neither.

Solution: We find the partial derivatives:

$$f_x(x, y) = 3x^2 - 3$$
$$f_y(x, y) = 3y^2 - 27.$$

These are set equal to 0 and solved for the variable:

$$\begin{matrix} 3x^2 - 3 = 0 \\ 3y^2 - 27 = 0 \end{matrix}, \qquad \text{which simplifies as} \qquad \begin{matrix} 3(x^2 - 1) = 0 \\ 3(y^2 - 9) = 0. \end{matrix}$$

From the first equation, $x^2 - 1 = 0$, we get $x = 1$ and $x = -1$. From the second equation, $y^2 - 9 = 0$, we get $y = 3$ and $y = -3$. We combine these solutions in all possible ways, and we have four critical points:

$$\big(1,3, f(1,3)\big), \qquad \big(1, -3, f(1, -3)\big), \qquad \big(-1,3, f(-1,3)\big), \qquad \big(-1, -3, f(-1, -3)\big).$$

The z values are $f(1,3) = -49$, $f(1, -3) = 59$, $f(-1,3) = -45$ and $f(-1, -3) = 63$.

To classify these critical points, use the second derivative test. The second derivatives are

$$f_{xx}(x,y) = 6x, \qquad f_{yy}(x,y) = 6y \quad \text{and} \quad f_{xy}(x,y) = 0.$$

Thus, using the formula, we have

$$D = f_{xx}(x,y)f_{yy}(x,y) - \left(f_{xy}(x,y)\right)^2 = (6x)(6y).$$

• When $x = 1$ and $y = 3$, we have $D = (6)(18)$, which is a positive number. Note that $f_{xx}(1,3)$ is also positive. Thus, the critical point $(1, 3, -49)$ is a local minimum.

• When $x = 1$ and $y = -3$, we have $D = (6)(-18)$, which is a negative number. Thus, the critical point $(1, -3, 59)$ is a saddle point.

• When $x = -1$ and $y = 3$, we have $D = (-6)(18)$, which is a negative number. Thus, the critical point $(-1, 3, -45)$ is a saddle point.

• When $x = -1$ and $y = -3$, we have $D = (-6)(-18)$, which is a positive number. Note that $f_{xx}(-1,-3)$ is negative. Thus, the critical point $(-1, -3, 63)$ is a local maximum.

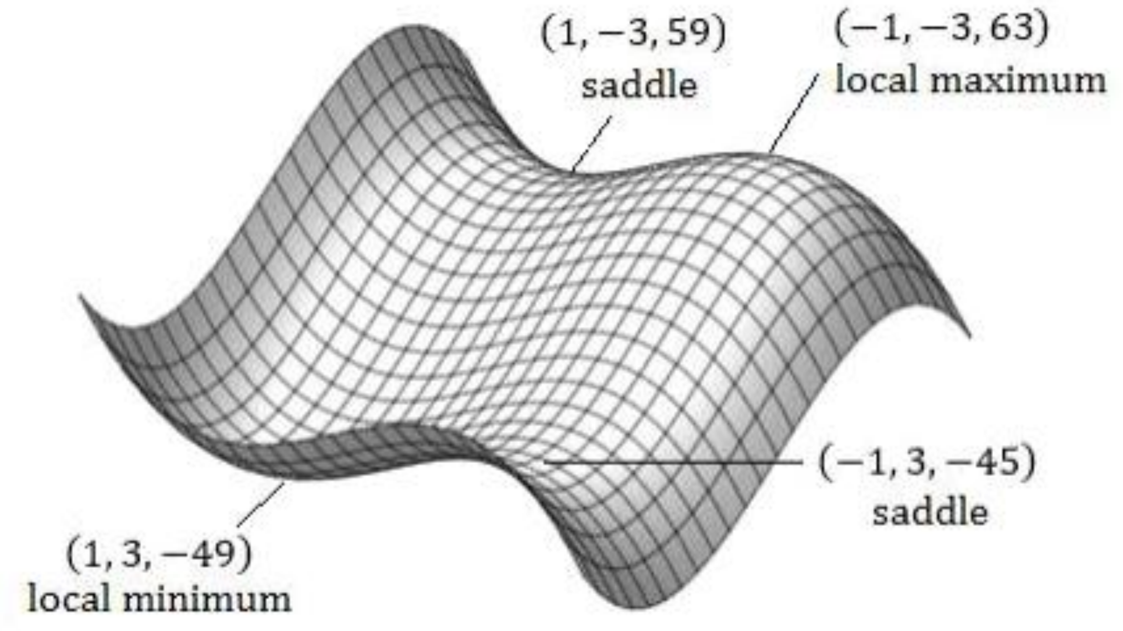
When finding *D*, it's not important to determine its actual value.

It's more important to determine its sign.

Thus, calculating (6)(18) is not as important as observing that the product of two positive values is positive.

Furthermore, by leaving the expression as (6)(18) rather than simplifying it, we can also quickly see that the value 6, representing $f_{xx}(1,3)$, is positive.

The graph of $z = f(x, y) = x^3 + y^3 - 3x - 27y + 7$ is:



Example 4: Let $z = f(x, y) = x^3 - y^3 - 2x^2 + xy + 3y$. Find its critical points and classify these points as minima, maxima or saddle.

Solution: The partial derivatives are

$$f_x(x, y) = 3x^2 - 4x + y$$
$$f_y(x, y) = -3y^2 + x + 3.$$

Setting these to zero, develop a non-linear system:

$$3x^2 - 4x + y = 0$$
$$-3y^2 + x + 3 = 0.$$

We cannot use the elimination method. Instead, we use substitution. In the first equation, solve for *y*:

$$y = 4x - 3x^2.$$

This is substituted into the second equation, then simplified:

$$-3(4x - 3x^2)^2 + x + 3 = 0$$

$$-3(16x^2 - 24x^3 + 9x^4) + x + 3 = 0$$

$$-27x^4 + 72x^3 - 48x^2 + x + 3 = 0.$$

Using a graphing calculator, we find four roots to this quartic equation. They are

$$x \approx -0.21, \qquad x \approx 0.36, \qquad x \approx 0.92 \quad \text{and} \quad x \approx 1.59.$$

For each x value above, use the equation $y = 4x - 3x^2$ to find the corresponding y value. The z-values are then found by evaluating f at each x and y value. There are four critical points:

$$(-0.21, -0.97, -1.89), \qquad (0.364, 1.06, 2.19), \qquad (0.92, 1.14, 2.073) \quad \text{and} \quad (1.59, -1.24, -4.82).$$

The second derivatives are

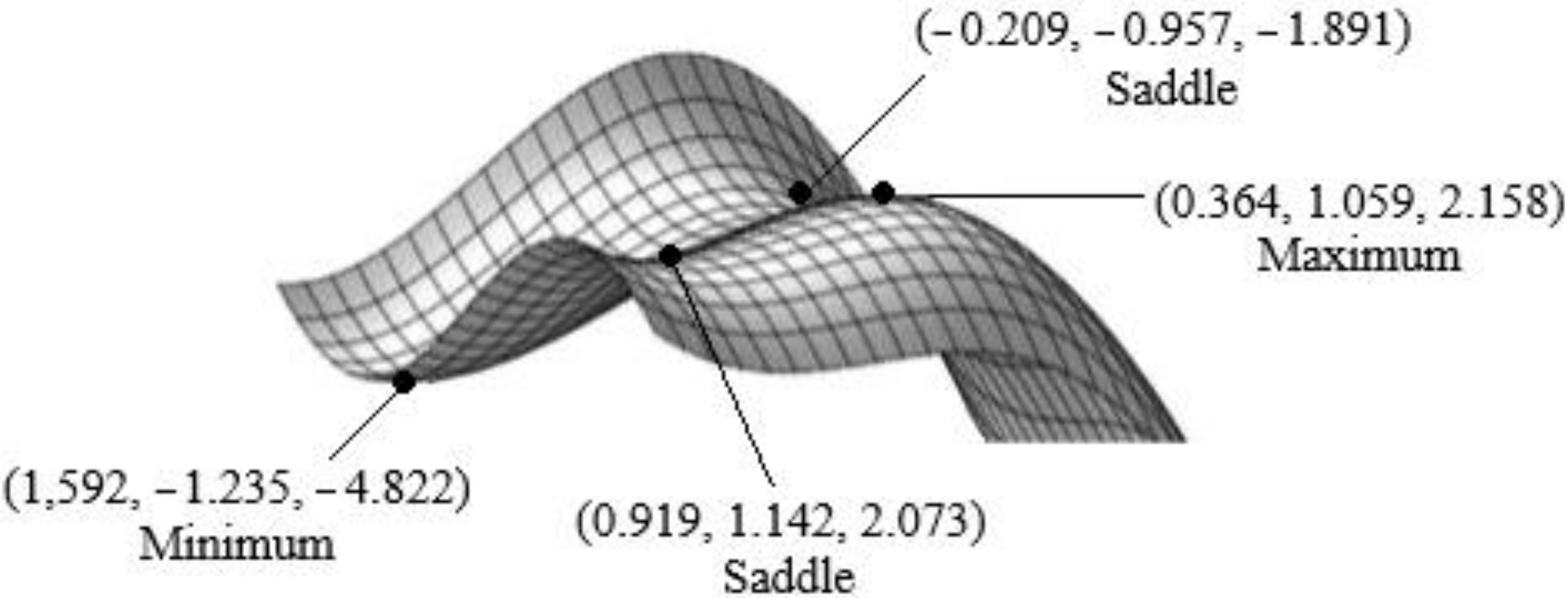
$$f_{xx}(x,y) = 6x - 4, \qquad f_{yy}(x,y) = -6y, \qquad f_{xy}(x,y) = 1.$$

Thus, we have

$$D = f_{xx}(x_c, y_c) f_{yy}(x_c, y_c) - \left(f_{xy}(x_c, y_x)\right)^2 = (6x_c - 4)(-6y_c) - (1)^2,$$

where x_c and y_c are the input values of a critical point. Each critical point is evaluated into the second derivative test formula. Only *x* and *y* are used, *z* is not:

- At $(-0.21, -0.97, -1.89)$, we get $D = (6(-0.21) - 4)\big(-6(-0.97)\big) - 1 = -31.48$. Since *D* is negative, the point $(-0.21, -0.97, -1.89)$ is a saddle point.
- At $(0.36, 1.06, 2.19)$, we get $D = (6(0.36) - 4)\big(-6(1.06)\big) - 1 = 10.54$. Since *D* is positive and since $f_{xx}(0.36, 1.06)$ is negative (as is $f_{yy}(0.36, 1.06)$), the point $(0.36, 1.06, 2.19)$ is a local maximum.
- At $(0.92, 1.14, 2.07)$, we get $D = (6(0.92) - 4)\big(-6(1.14)\big) - 1 = -11.37$. Since *D* is negative, the point $(0.92, 1.14, 2.07)$ is a saddle point.
- At $(1.59, -1.24, -4.82)$, we get $D = (6(1.59) - 4)\big(-6(-1.24)\big) - 1 = 40.14$. Since *D* is positive and since $f_{xx}(1.59, -1.24)$ is positive (as is $f_{yy}(1.59, -1.24)$), the point $(1.59, -1.24, -4.82)$ is a local minimum.

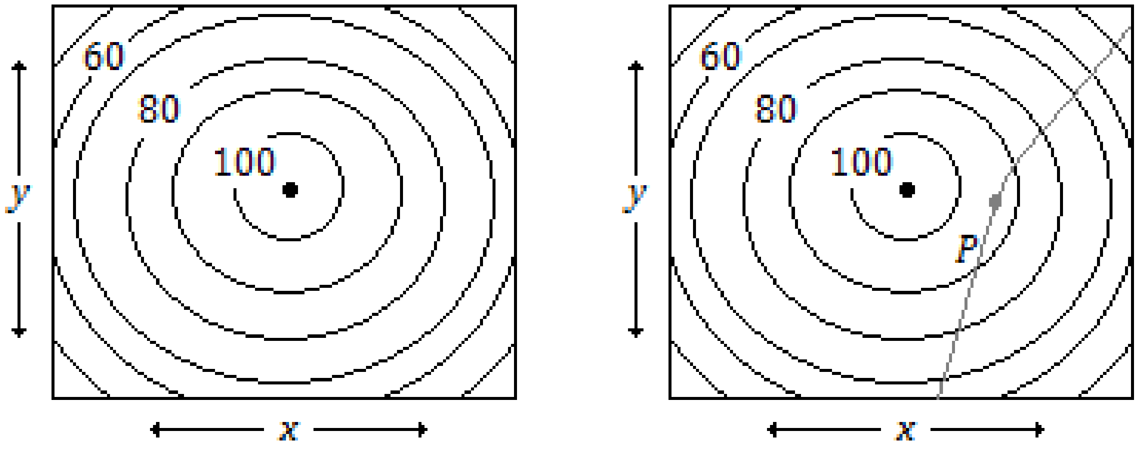

(−0.209, −0.957, −1.891)
Saddle
(0.364, 1.059, 2.158)
Maximum
(1,592, −1.235, −4.822)
Minimum
(0.919, 1.142, 2.073)
Saddle

Normally, x and y are chosen independently of one another so that one may "roam" over the entire surface of f (within any domain restrictions on x and y). Determining minimum or maximum points on f under this circumstance is called unconstrained optimization.

If x and y are related to one another by an equation, then only one of the variables can be independent. In such a case, we may determine a minimum or maximum point on the surface of f subject to the constraint placed on x and y.

This is called **constrained optimization**. Such constraints are usually written where x and y are combined implicitly, $g(x, y) = c$.

Suppose you are hiking on a hill. If there is no restriction on where you may walk, then you are "unconstrained" and you may seek the hill's maximum point. However, if you are constrained to a hiking path, then it is possible to determine a maximum point on the hill, but only that part along the hiking path.

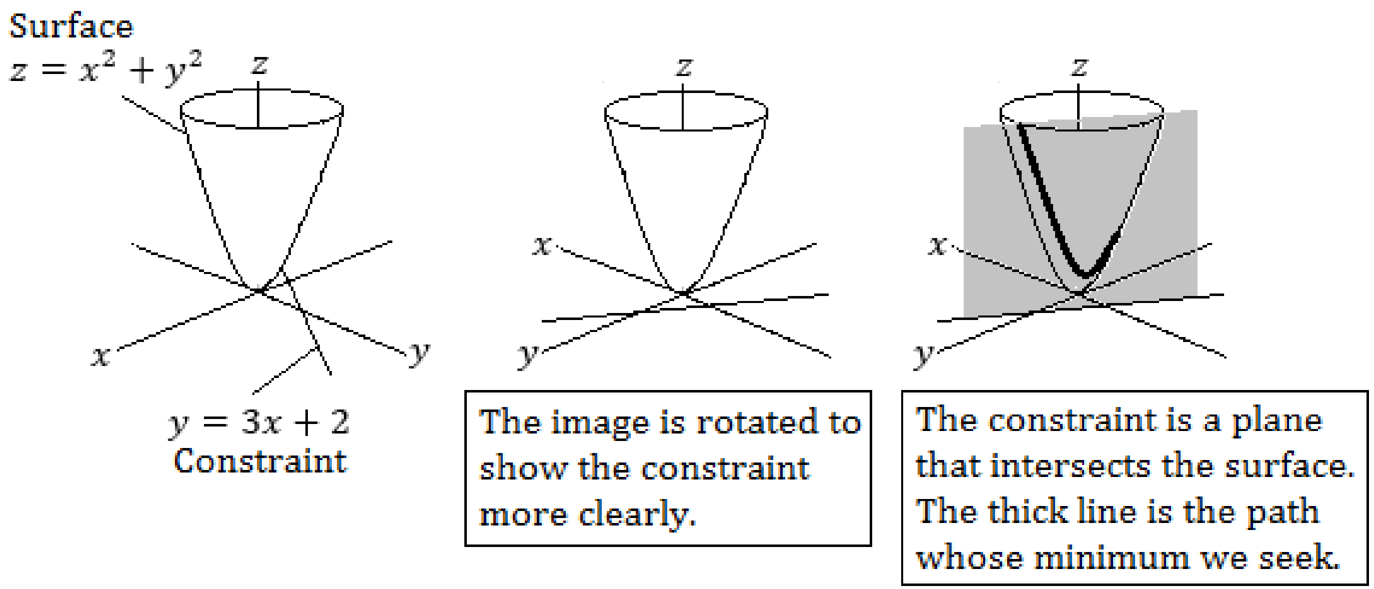


(Left) Unconstrained optimization: The maximum point of this hill is marked by a black dot, and is roughly $z = 105$.
(Right) Constrained optimization: The highest point on the hill, subject to the constraint of staying on path P, is marked by a gray dot, and is roughly $z = 93$.

Example 5: Find the minimum or maximum point on the surface of $z = f(x, y) = x^2 + y^2$ subject to the constraint $-3x + y = 2$.

Solution: The surface of f is a paraboloid with its vertex (0,0,0) at the origin, opening in the positive z direction (or "up"). Its unconstrained minimum point is (0,0,0). There is no maximum point on this surface.

Now, note that with the constraint $y = 3x + 2$ in place, x and y are no longer independent variables. Once a value for x is chosen, then y is determined. We are now restricted to this "path" on the surface of f.



To find the minimum or maximum point on this paraboloid subject to the constraint $y = 3x + 2$, substitute the constraint into the function f and simplify:

$$f(x, 3x + 2) = x^2 + (3x + 2)^2 = 10x^2 + 12x + 4.$$

Differentiating, we have $f'(x) = 20x + 12$, and to find the critical value of x, we set $f'(x) = 0$:

$$20x + 12 = 0 \qquad \text{which gives} \qquad x = -\frac{3}{5}\,.$$

Observe that $f(x) = 10x^2 + 12x + 4$ is a parabola in R^2 that opens upward. Thus, the critical value for x will correspond to the minimum point on this parabola. Find y by substitution into the constraint:

$$y = 3\left(-\frac{3}{5}\right) + 2 = \frac{1}{5}\,.$$

The minimum *point* on the surface of f is $\left(-\frac{3}{5}, \frac{1}{5}, \frac{2}{5}\right)$. The minimum *value* of z on the surface of f is $y = 3x + 2$ is $= \frac{2}{5}$.

Lastly, we find z:

$$z = f\left(-\frac{3}{5}, \frac{1}{5}\right) = \left(-\frac{3}{5}\right)^2 + \left(\frac{1}{5}\right)^2 = \frac{10}{25} = \frac{2}{5}\,.$$

Example 6: Consider the portion of the plane $2x + 4y + 5z = 20$ in the first octant. Find the point on the plane closest to the origin.

Solution: The point on the plane closest to the origin will lie on a line orthogonal to the plane. Let (x, y, z) be a point on the plane, so the distance d between this point and the origin (0,0,0) is

$$d(x, y, z) = \sqrt{(x-0)^2 + (y-0)^2 + (z-0)^2} = \sqrt{x^2 + y^2 + z^2}.$$

However, note that not all variables are independent—they are constrained to one another by the plane's equation. We can isolate one of the variables in the plane. For example, $z = 4 - \frac{2}{5}x - \frac{4}{5}y$. Thus, d can be written as a function of *x* and *y* only, and the radicand is expanded:

$$d(x, y) = \sqrt{x^2 + y^2 + \left(4 - \frac{2}{5}x - \frac{4}{5}y\right)^2} = \sqrt{\frac{29}{25}x^2 + \frac{41}{25}y^2 + \frac{16}{25}xy - \frac{16}{5}x - \frac{32}{5}y + 16}.$$

Variables x and y also obey another constraint: both must be non-negative. This will ensure that z is also non-negative.

Taking partial derivatives and simplifying, we have

$$d_x = \frac{\frac{29}{25}x + \frac{8}{25}y - \frac{8}{5}}{\sqrt{\frac{29}{25}x^2 + \frac{41}{25}y^2 + \frac{16}{25}xy - \frac{16}{5}x - \frac{32}{5}y + 16}}, \qquad d_y = \frac{\frac{8}{25}x + \frac{41}{25}y - \frac{16}{5}}{\sqrt{\frac{29}{25}x^2 + \frac{41}{25}y^2 + \frac{16}{25}xy - \frac{16}{5}x - \frac{32}{5}y + 16}}.$$

When set to 0, the denominators can be ignored. Thus, only the numerators are considered, and we have

$$\frac{29}{25}x + \frac{8}{25}y - \frac{8}{5} = 0 \quad \text{and} \quad \frac{8}{25}x + \frac{41}{25}y - \frac{16}{5} = 0.$$

Placing the constants to the right of the equality and multiplying by 25 to clear fractions, we then have a simplified system in two variables:

$$29x + 8y = 40$$
$$8x + 41y = 80.$$

Using any method (such as elimination) to solve this system, we find that $x = \frac{8}{9}$, $y = \frac{16}{9}$, and substituting these into the plane's equation, we have $z = \frac{20}{9}$.

Thus, the point $\left(\frac{8}{9}. \frac{16}{9}, \frac{20}{9}\right)$ is the point on the plane closest to the origin.