Optimization and Lagrange Multipliers

We studied single variable optimization problems in Calculus 1; given a function $f(x)$, we found the extremes of f relative to some constraint. Our ability to find local extrema of a multivariate function gives us the tools to solve some relatively simple optimization problems involving multiple variables, as in the first example below. However, many optimization problems involving multiple variables will be complicated enough that we must use the special method of Lagrange Multipliers to solve them.

Example

Find the shortest distance from the point $(1, 0, -2)$ to the plane $x + 2y + z = 4$.

Since the distance between the point $(1, 0, -2)$ and a point (x, y, z) is given by

$$D = \sqrt{(x-1)^2 + y^2 + (z+2)^2},$$

we wish to minimize this function; to make life easier, we may minimize

$$D^2 = (x-1)^2 + y^2 + (z+2)^2$$

instead. However, we must take the constraint into account: x, y, and z must be chosen so that $x + 2y + z = 4$. We can simplify the problem a bit by solving the constraint for z; $x + 2y + z = 4$ is equivalent to $z = 4 - x - 2y$.

So we really wish to minimize

$$f(x, y) = (x-1)^2 + y^2 + (4 - x - 2y + 2)^2 = (x-1)^2 + y^2 + (6 - x - 2y)^2.$$

Using the methods from Section 14.7, we find the partials:

$$f_x(x, y) = 2(x-1) - 2(6 - x - 2y) = 4x + 4y - 14,$$

and

$$f_y(x, y) = 2y - 4(6 - x - 2y) = 10y + 4x - 24.$$

To make $f_x(x, y) = 0$, we must have $y = \frac{7}{2} - x$. Taking this into account for f_y, we have $10(\frac{7}{2} - x) + 4x - 24 = 0$; solving for x, we see that $x = \frac{11}{6}$. Then since $y = \frac{7}{2} - x$, we must have $y = \frac{5}{3}$.

Thus the only possible local extreme for the distance function occurs at $\left(\frac{11}{6}, \frac{5}{3}\right)$.

Now we must determine if this point actually gives us a local extreme. Since

$$f_{xx}(x, y) = 4, \; f_{yy}(x, y) = 10, \text{ and } f_{xy}(x, y) = 4,$$

we see that

$$f_{xx}\left(\frac{11}{6}, \frac{5}{3}\right) f_{yy}\left(\frac{11}{6}, \frac{5}{3}\right) - f_{xy}^2\left(\frac{11}{6}, \frac{5}{3}\right) = 40 - 16 = 24 > 0,$$

and since $f_{xx}\left(\dfrac{11}{6},\dfrac{5}{3}\right) > 0$, we conclude by the Second Derivative Test that $\left(\dfrac{11}{6},\dfrac{5}{3}\right)$ gives a local minimum for the function. Using the constraint, we see that the corresponding value for z is

$$z = 4 - \frac{11}{6} - 2(\frac{5}{3}) = -\frac{7}{6}.$$

So the minimum distance from $(1, 0, -2)$ to the plane $x + 2y + z = 4$ is

$$D = \sqrt{(x-1)^2 + y^2 + (z+2)^2} = \sqrt{\left(-\frac{5}{6}\right)^2 + \left(\frac{5}{3}\right)^2 + \left(\frac{5}{6}\right)^2}$$

$$= \sqrt{\frac{150}{36}}$$

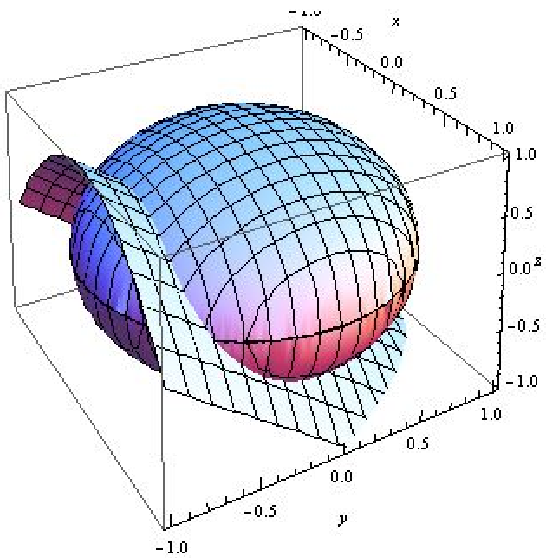
$$= \sqrt{\frac{25}{6}}$$

$$= \frac{5}{\sqrt{6}}.$$

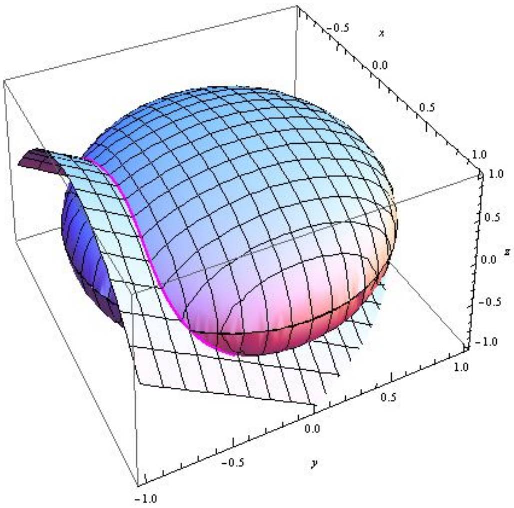
Optimization of multivariate functions is rarely as simple as the above example; in that case, we were able to solve the constraint for z, making the problem significantly easier, but that often proves difficult to do. In general, we will use the method of Lagrange Multipliers to solve multiple variable optimization problems.

The method is easier to understand if we consider the geometry behind it. We have a function f whose graph is a surface in space, and a function g that we refer to as the constraint. By this, we mean that the only values on the surface of f in which we are interested are the ones that satisfy $g(x, y) = k$ for some constant k.

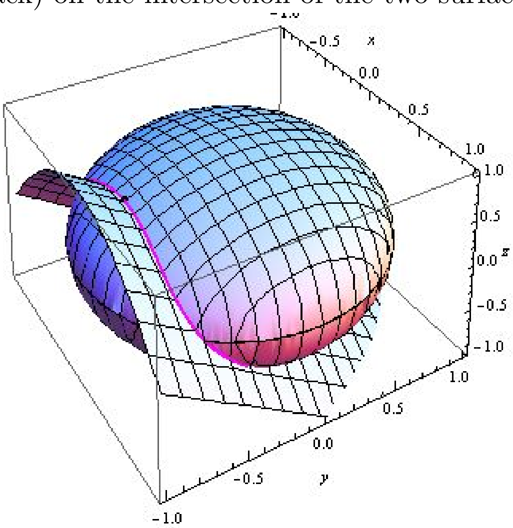
For example, in the graph below, let f be the sphere and g be the cylinder.



The intersection of the two surfaces, outlined in magenta, gives us a constraint on f:

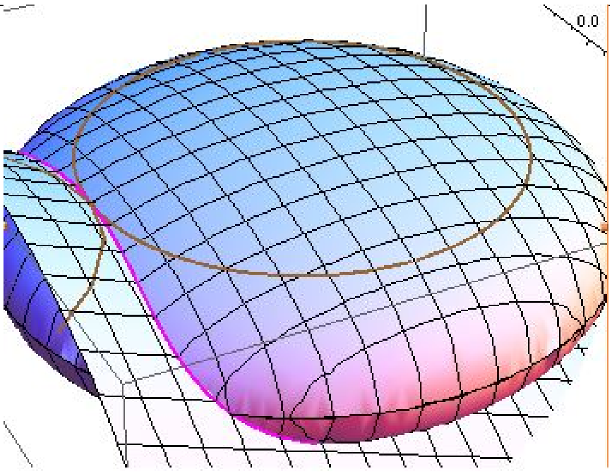


We would like to find the point (x_0, y_0) so that $g(x_0, y_0) = k$ *and* $f(x_0, y_0)$ yields the maximum value among all points that satisfy $g(x, y) = k$. Geometrically, we want to find the point (marked in black) on the intersection of the two surfaces that is at maximum height:



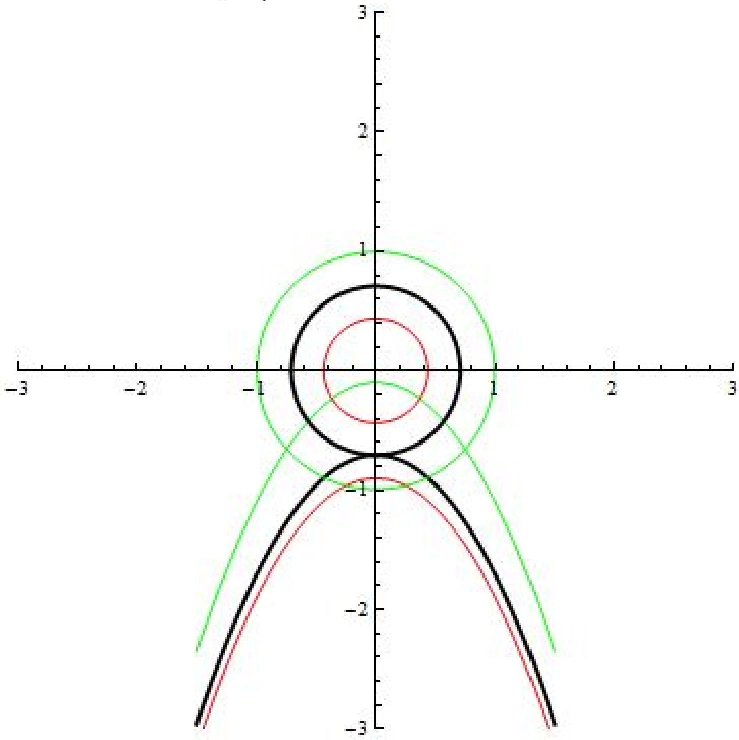
To find the coordinates of this point, let's think about the level curves of each surface.

On the 3D graph, the level curves of each surface are graphed at the height of the maximum intersection:



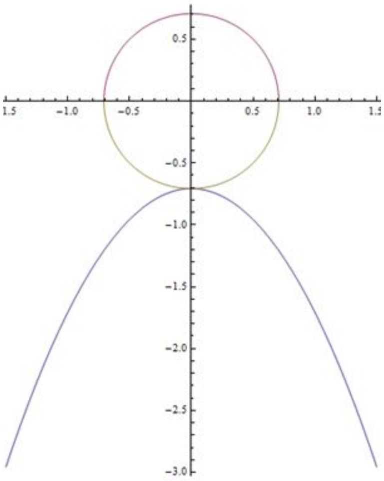
Notice that the level curves are tangent at this point; i.e. at (x_0, y_0, z_0), the two level curves $f(x, y) = z_0$ and $g(x, y) = z_0$ are tangent.

Let's look at the projection of some of the level curves into the xy plane:

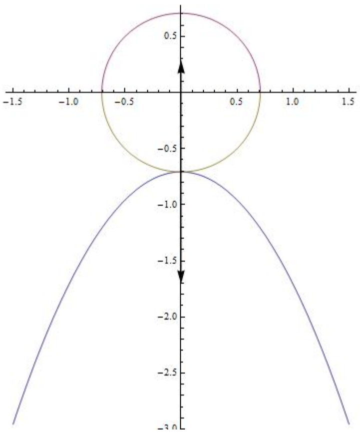


Curves of the same color indicate level curves at the same height; for example, the two green curves are both at height $z = .1$.

Notice that the only pair of curves that are tangent are the black pair, which correspond to the two level curves graphed on the surfaces above; these curves are where the maximum occurs. This will be true in general: if f has a maximum or minimum value at (x_0, y_0) with respect to the constraint $g(x, y) = k$, then the two level curves are tangent.



We can use this to our advantage; since the two surfaces have level curves that are tangent at extremes of their shared values, the corresponding tangent vectors for f and g are parallel; more importantly, so are their gradients.



In other words, if f takes on an extreme value with regards to the constraint g, then there is a number λ so that $\nabla f = \lambda \nabla g$.

Theorem 0.0.1. The Method of Lagrange Multipliers Suppose that $f(x, y, z)$ and $g(x, y, z)$ are differentiable and $\nabla g \neq 0$ when $g(x, y, z) = k$. To find the local maximum and minimum values of f subject to the constraint $g(x, y, z) = k$ (*if* such values exist), find the values of x, y, z, and λ that simultaneously satisfy the equations

$$\nabla f = \lambda \nabla g \text{ and } g(x, y, z) = k.$$

If f and g are instead functions of two variables, the condition is the same, except that we drop the variable z.

Examples:

Find the points on the sphere $x^2 + y^2 + z^2 = 4$ that are closest to and farthest from the point $(3, 1, -1)$.

We want to minimize the distance from the point $(3, 1, -1)$ subject to the constraint $x^2 + y^2 + z^2 - 4 = 0$. The distance from the point $(3, 1, -1)$ to the point (x, y, z) is

$$d = \sqrt{(x-3)^2 + (y-1)^2 + (z+1)^2};$$

however minimizing or maximizing this function is equivalent to minimizing or maximizing

$$d^2 = (x-3)^2 + (y-1)^2 + (z+1)^2,$$

which is simpler to work with. So we want to minimize/maximize

$$f(x, y, z) = (x-3)^2 + (y-1)^2 + (z+1)^2$$

subject to the constraint

$$g(x, y) = x^2 + y^2 + z^2 = 4.$$

We would like to find x, y, and z so that $x^2 + y^2 + z^2 - 4 = 0$, as well as a number λ so that $\nabla f = \lambda \nabla g$ for these same choices of x and y. Let's start by calculating the gradients:

$$\nabla f = (2x-6)\vec{i} + (2y-2)\vec{j} + (2z+2)\vec{k}, \text{ and}$$

$$\nabla g = 2x\vec{i} + 2y\vec{j} + 2z\vec{k}.$$

If $\nabla f = \lambda \nabla g$, then by equating components we must have

$$2x - 6 = 2x\lambda, 2y - 2 = 2y\lambda, \text{ and } 2z + 2 = 2z\lambda.$$

Fortunately, it is simple to solve for λ in each of the cases above; $2x - 6 = 2x\lambda$ means that $2x - 2x\lambda = 6$; $2y - 2 = 2y\lambda$ means that $2y - 2y\lambda = 2$; and $2z + 2 = 2z\lambda$ means that $2z - 2z\lambda = -2$. Solving for x, y, and z respectively, we have

$$x = \frac{3}{1-\lambda}, \; y = \frac{1}{1-\lambda}, \text{ and } z = \frac{-1}{1-\lambda}. \tag{1}$$

Recall that we want to find values for x, y, z, and λ so that $x^2 + y^2 + z^2 - 4 = 0$ and the equations in (1) are satisfied. In particular, we know how each of x, y, and z must relate to λ, so we can replace these variables with their counterparts in terms of λ in $x^2 + y^2 + z^2 - 4 = 0$: we see that

$$\left(\frac{3}{1-\lambda}\right)^2 + \left(\frac{1}{1-\lambda}\right)^2 + \left(\frac{-1}{1-\lambda}\right)^2 - 4 = 0.$$

Solving for λ, we have

$$\frac{9+1+1}{(1-\lambda)^2} - \frac{4(1-\lambda)^2}{(1-\lambda)^2} = 0,$$

i.e.

$$\frac{11 - 4 + 8\lambda - 4\lambda^2}{(1-\lambda)^2} = \frac{7 + 8\lambda - 4\lambda^2}{(1-\lambda)^2} = 0.$$

In order for the equality to be true, we must have $7 + 8\lambda - 4\lambda^2 = 0$; we can find the roots of $7 + 8\lambda - 4\lambda^2$ using the quadratic equation.

$$\begin{aligned}\lambda &= \frac{-8 \pm \sqrt{64 + 112}}{-8} \\ &= \frac{-8 \pm \sqrt{176}}{-8} \\ &= \frac{-8 \pm \sqrt{16 \times 11}}{-8} \\ &= \frac{-8 \pm 4\sqrt{11}}{-8} \\ &= 1 \pm \frac{\sqrt{11}}{2}.\end{aligned}$$

Now we can determine the values for x, y, and z using the equations in (1); if $\lambda = 1 + \frac{\sqrt{11}}{2}$ we see that

$$x = \frac{6}{\sqrt{11}}, \ y = \frac{2}{\sqrt{11}}, \text{ and } z = -\frac{2}{\sqrt{11}};$$

alternatively, if $\lambda = 1 - \frac{\sqrt{11}}{2}$ then

$$x = -\frac{6}{\sqrt{11}}, \ y = -\frac{2}{\sqrt{11}}, \text{ and } z = \frac{2}{\sqrt{11}}.$$

It is clear that plugging the point

$$\left(\frac{6}{\sqrt{11}}, \frac{2}{\sqrt{11}}, -\frac{2}{\sqrt{11}}\right)$$

into the distance equation will yield a smaller output than will plugging in the point

$$\left(-\frac{6}{\sqrt{11}}, -\frac{2}{\sqrt{11}}, \frac{2}{\sqrt{11}}\right).$$

Thus the first point is the closest point to $(3, 1, -1)$ lying on the sphere, and the second is the farthest from $(3, 1, -1)$ lying on the sphere.

Find the dimensions of a right circular cylinder with volume $V = 432\pi$ and minimum surface area.

If we center the base of the cylinder at the origin in the xy plane, we see that the radius of the cylinder is x; the cylinder's height is z. We need to find formulas for the cylinder's volume and surface area; the volume is $V = \pi x^2 z$, and the surface area is the sum of the surface areas of the top and bottom (each has area πx^2) and the surface are of the side, which is $2\pi xz$. So we wish to minimize

$$S(x, z) = 2\pi x^2 + 2\pi xz$$

subject to the constraint

$$V(x, z) = \pi x^2 z = 432\pi.$$

Let's find the gradients:

$$\nabla S = (4\pi x + 2\pi z)\vec{i} + 2\pi x\vec{k} \text{ and } \nabla V = 2\pi xz\vec{i} + \pi x^2\vec{k}.$$

Then we need to find values for x, z, and λ that satisfy the equations

$$4\pi x + 2\pi z = 2\pi xz\lambda \text{ and } 2\pi x = \pi x^2\lambda.$$

Let's work with the second equation; we can rewrite it as $2\pi x - \pi x^2\lambda = 0$, or $2x - x^2\lambda = 0$. Since the solution $x = 0$ would not make sense (the radius of the cylinder would be 0), we may assume that $x \neq 0$ and divide through by x; we end up with $2 - x\lambda = 0$. This equation has no solution if $\lambda = 0$, so we may assume $\lambda \neq 0$, and rewrite it as

$$x = \frac{2}{\lambda}.$$

The first equation above may now be rewritten using this equality; dividing through by 2π, $4\pi x + 2\pi z = 2\pi xz\lambda$ becomes $2x + z = xz\lambda$, which may be rewritten as

$$\frac{4}{\lambda} + z = 2z,$$

which means that

$$z = \frac{4}{\lambda}.$$

Now these choices for x and z must satisfy the constraint $\pi x^2 z - 432\pi = 0$; rewriting this constraint with $x = \frac{2}{\lambda}$ and $z = \frac{4}{\lambda}$ yields the expression

$$\pi\frac{4}{\lambda^2} \cdot \frac{4}{\lambda} - 432\pi = 0.$$

Solving for λ, we have

$$\lambda = \sqrt[3]{\frac{16}{432}} = \sqrt[3]{\frac{1}{27}} = \frac{1}{3}.$$

Then $x = 6$ and $z = 12$. Thus choosing the radius of the cylinder to be 3 and its height to be 12 will yield the cylinder of minimum surface area and volume 432π.

One quick note: we could have solved this problem by solving for z in the constraint, then rewriting the equation for surface area in terms of x (i.e. without using Lagrange multipliers). This would have actually been quicker, but using Lagrange multipliers in this example does illustrate some of the things we may encounter when we *are* forced to use this method.

Lagrange Multipliers with Two Constraints

If wish to maximize or minimize $f(x, y, z)$ under two constraints, as opposed to just one, the process is nearly the same. If the constraints are given by $g(x, y, z) = k$ and $h(x, y, z) = c$, and g

and h are both differentiable with non-parallel gradients, then the extreme values of f subject to these constraints may be found by determining the x, y, z, λ, and μ so that

$$\nabla f = \lambda \nabla g + \mu \nabla h, \ g(x, y, z) = k, \text{ and } h(x, y, z) = c.$$

The temperature at a point (x, y, z) on the surface of the sphere $x^2 + y^2 + z^2 = 11$ is given by $T(x, y, z) = 20 + 2x + 2y + z^2$. Find the extreme temperatures on the curve formed by the intersection of the plane $x + y + z = 3$ and the sphere.

We need to find the extremes on the surface $T(x, y, z)$ under the constraints

$$g(x, y, z) = x^2 + y^2 + z^2 = 11 \text{ and } h(x, y, z) = x + y + z = 3.$$

The gradients are

$$\nabla T = 2\vec{i} + 2\vec{j} + 2z\vec{k}, \ \nabla g = 2x\vec{i} + 2y\vec{j} + 2z\vec{k}, \text{ and } \nabla h = \vec{i} + \vec{j} + \vec{k}.$$

Using the formula $\nabla T = \lambda \nabla g_1 + \mu \nabla g_2$, we have

$$2 = 2x\lambda + \mu, \ 2 = 2y\lambda + \mu, \text{ and } 2z = 2z\lambda + \mu.$$

We need to solve this system of equations, together with the equations

$$x + y + z = 3 \text{ and } x^2 + y^2 + z^2 = 11.$$

Notice that the equations $2 = 2x\lambda + \mu$ and $2 = 2y\lambda + \mu$ can be rewritten as

$$2 - 2x\lambda = \mu, \text{ and } 2 - 2y\lambda = \mu;$$

we see that $2 - 2x\lambda = 2 - 2y\lambda$. Thus $2y\lambda - 2x\lambda = 0$, i.e.

$$2\lambda(y - x) = 0.$$

This tells us that either $y - x = 0$ (i.e. $y = x$) or $\lambda = 0$. Both possibilities could cause $T(x, y, z)$ to take on extremes values, so we must check both options.

(1) If $\lambda = 0$, then since $2 - 2x\lambda = \mu$, we must have $\mu = 2$. Then since $2z = 2z\lambda + \mu$, we have $z = 1$. Using the equation $x + y + z = 3$, we see that $x + y = 2$, so that $y = 2 - x$. We may now solve for x in the other constraint $x^2 + y^2 + z^2 = 11$; it may be rewritten as $x^2 + (2 - x)^2 + 1 = 11$. So $x^2 + 4 - 4x + x^2 + 1 - 11 = 0$; simplifying, we have $2x^2 - 4x - 6 = 0$. Using the quadratic equation to find the roots of this polynomial, we have

$$x = \frac{4 \pm \sqrt{16 + 48}}{4} = \frac{4 \pm \sqrt{64}}{4} = 1 \pm 2.$$

So the points $(3, -1, 1)$ and $(-1, 3, 1)$ could give us extrema.

(2) If, on the other hand, $y = x$, then we can account for this in the constraints: $x + y + z = 3$ becomes $z = 3 - 2x$, so that $x^2 + y^2 + z^2 = 11$ becomes $2x^2 + (3 - 2x)^2 = 11$. Solving this equality,

we have $2x^2 + 9 - 12x + 4x^2 = 11$, so that $6x^2 - 12x - 2 = 0$. Again using the quadratic equation, we have

$$x = \frac{12 \pm \sqrt{144 + 48}}{12} = \frac{12 \pm \sqrt{192}}{12} = 1 \pm \frac{8\sqrt{3}}{12} = 1 \pm \frac{2\sqrt{3}}{3}.$$

So the points $(1 + \frac{2\sqrt{3}}{3}, 1 + \frac{2\sqrt{3}}{3}, 1 - \frac{4\sqrt{3}}{3})$ and $(1 - \frac{2\sqrt{3}}{3}, 1 - \frac{2\sqrt{3}}{3}, 1 + \frac{4\sqrt{3}}{3})$ could also give extrema.

Finally, we need to plug each of these points into $T(x, y, z)$ to determine which ones give us extremes:

$$T(3, -1, 1) = 25, \ T(-3, 1, 1) = 25,$$

$$T\left(1 + \frac{2\sqrt{3}}{3}, 1 + \frac{2\sqrt{3}}{3}, 1 - \frac{4\sqrt{3}}{3}\right) = 30.\bar{3}, \text{ and } T\left(1 - \frac{2\sqrt{3}}{3}, 1 - \frac{2\sqrt{3}}{3}, 1 + \frac{4\sqrt{3}}{3}\right) = 30.\bar{3}.$$

Thus T has maximum temperature at the last two points, and minimum temperature at the first two.