Multivariable Differentiation – Lagrange Multipliers

You may have noticed in the previous lesson on optimization the applications based examples involved finding the extreme values of a function subject to a constraint. The terminology generally adopted for these types of problems is as follows:

Objective Function $f(x_1, \dots, x_n)$	Expresses the quantity we would like to optimize in terms of n independent variables.
Constraint Function $g(x_1, \dots, x_n) = 0$	Expresses a relationship between the independent that must be satisfied within the context of optimizing the objective function.

As an example, suppose we want to build a rectangular garden to maximize the area given we have purchased P feet of fencing. In this case the objective function is

$$A = f(x, y) = xy$$

Where, x and y are the length and width of the garden respectively.

The size of the garden is, of course, *constrained* by the amount of fencing we purchased. The amount of fencing corresponds to the perimeter of the garden, $P = 2x + 2y$. With this we can define our constraint function as

$$g(x, y) = 2x + 2y - P = 0$$

In the previous lesson we solved this problem by first solving the constraint equation for one of the variables and substituting into the objective function, as shown below.

$$f(x) = x\left(\frac{1}{2}P - x\right) = \frac{P}{2}x - x^2$$

Which is a single variable function that can be optimized with minimal effort as shown below.

$$f'(x) = \frac{P}{2} - 2x = 0 \rightarrow \ x = \frac{P}{4}, \therefore \ y = \frac{P}{4} \ and \ A_{max} = \frac{P^2}{16}$$

This method for solving optimization problems can be referred to as the 'substitution' method. An alternate method that can be used is called the method of Lagrange multipliers. It's a much more general procedure that can be more easily applied to optimization problems with many variables and even more than one constraint function.

The method of Lagrange multipliers relies on the following theorem.

Lagrange Multiplier Theorem
Assume $f(x, y)$ and $g(x, y)$ are differentiable functions. If $f(x, y)$ has a local extremum on the constraint curve, $g(x, y) = 0$, at $P = (a, b)$ and if $\nabla g_P \neq 0$, then there is a scalar, λ, such that $$\nabla f_P = \lambda \nabla g_P$$

Rather than providing a formal proof we will instead illustrate the above relationship through a relatively simple example. Doing this will allow us to gain a much more intuitive feel for the relationship and how it is used to solve constrained optimization problems.

Example 1: Find the maximum value for the function, $f(x, y) = 4 - (x^2 + y^2)$, while maintaining the following relationship between the two independent variables.

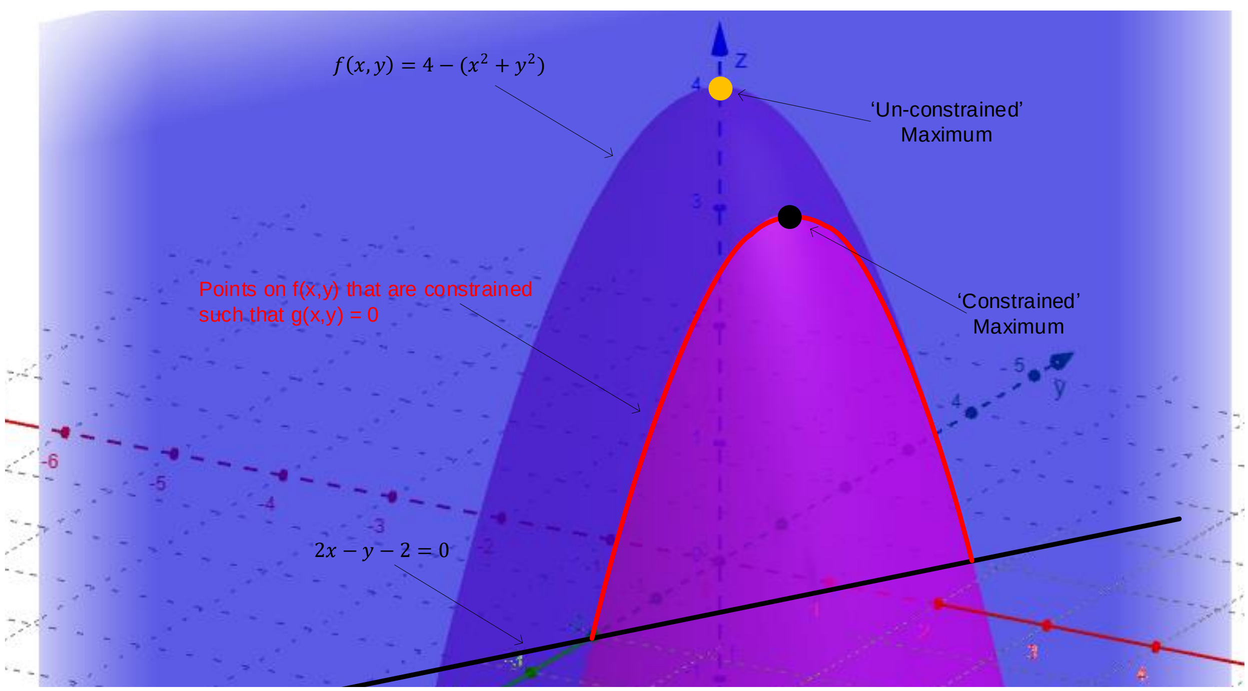
$$y = 2x - 2$$

Solution: Based on the information given in the problem we can easily identify the objective function as $f(x, y) = 4 - (x^2 + y^2)$. The problem then states that the coordinates of the maximum value of $f(x, y)$ must also satisfy the relationship, $y = 2x - 2$. This relationship is what we refer to as the constraint. However, to use the method of Lagrange multipliers we first recast this relationship into a two variable function, $g(x, y) = 2x - y - 2$. Although this equation defines an entire surface, we are only concerned with its level curve, $g(x, y) = 0$.

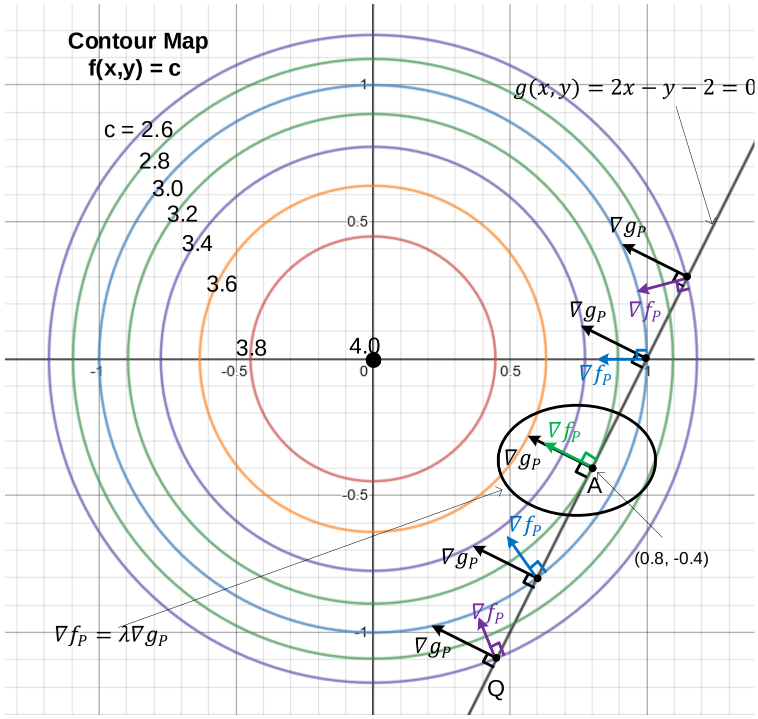
Objective Function	***Constraint Function***
$f(x, y) = 4 - (x^2 + y^2)$	$g(x, y) = 2x - y - 2 = 0$

The figure below illustrates the scenario given. Some important observations are listed below.

- $f(x, y)$ is a paraboloid which contains an *unconstrained* maximum indicated in yellow.
- The black line, $g(x, y) = 0$, is shown in the x-y plane.
  - Extending this line in the z direction creates a vertical plane shown in the blue.
- The red curve shows the intersection of this plane and $f(x, y)$.
  - This curve represents the subset of $f(x, y)$ that also satisfies $g(x, y) = 0$.
- The black point represents the maximum value from this subset of values on $f(x, y)$.



To help understand the Lagrange multiplier technique, we graph the level curves of the objective function, $f(x, y) = c$, along with a single level curve that corresponds to the constraint function, $g(x, y) = 0$



The contour map shows $f(x,y)$ increasing as we move towards the point $(0,0)$. If we were not constrained by $g(x,y) = 0$ then we would have a maximum value of $f(x,y) = 4$ occurring at $(0,0)$. However, in determining the maximum value we are *only* allowed to access the x-y coordinates that also satisfy $g(x,y) = 0$. In other words, we need to find the largest value of $f(x,y)$, while remaining on the line $g(x,y) = 0$.

Before proceeding we review two of the most important properties of the gradient vector.

1. The gradient vector, ∇f_P, point in the maximum direction of increase of $f(x,y)$ at P.
2. The gradient vector, ∇f_P, is orthogonal to the level curve or $f(x,y)$ at P.

The second bullet point is used to justify the gradient vectors, ∇f_P and ∇g_P, shown in the figure. Recall $g(x,y) = 0$ is the level curve representing the constraint function, $y = 2x - 2$.

Now, imagine starting at the point Q on the level curve, $g(x,y) = 0$. To increase the value of f we would like to move in the direction of the gradient, ∇f_Q. However, we must remain on the constraint curve. Moving along the curve to the left decreases the value of f while moving to the right increases the value of f. Therefore, we move a small distance to the right. Before choosing the direction of our next movement we again check the direction that increases the value of f. This process would continue until we arrive at the point A, where moving in either direction would result in a decrease in the value of f. When this happens, we have found a local maximum of f that is constrained by the curve g!

The key to the Lagrange Multiplier technique is to notice that the point A corresponds to the point where the gradient vectors, ∇f_P and ∇g_P, are parallel. At all other points along our journey these vectors were ***not*** parallel. Two parallel vectors are related by a constant, in this case, we refer to this constant as the Lagrange multiplier, λ, and we write.

$$\nabla f_P = \lambda \nabla g_P$$

Let's now see how we use this relationship to help us solve the optimization problem.

$$\nabla f_P = \lambda \nabla g_P$$
$$\langle -2x, -2y \rangle = \lambda \langle 2, -1 \rangle$$

This gives us two equations, one for each vector component, for which we can then solve for λ.

$$\lambda = -x \qquad\qquad \lambda = 2y$$

Setting these equations equal to each other we find $y = -(1/2)x$, which we can then substitute into the constraint equation to solve for x.

$$\begin{aligned} 2x - y - 2 &= 0 \\ 2x - 4\left(-\frac{1}{2}x\right) &= 2 \\ x &= \frac{4}{5} \end{aligned} \quad \rightarrow \quad \begin{aligned} y &= 2x - 2 \\ y &= 2\left(\frac{4}{5}\right) - 2 \\ y &= -\frac{2}{5} \end{aligned}$$

Finally, we substitute these values into the objective function to find the maximum value.

$$f\left(\frac{4}{5}, -\frac{2}{5}\right) = 4 - \left(\left(\frac{4}{5}\right)^2 + \left(-\frac{2}{5}\right)^2\right) \rightarrow f_{max} = 3.2$$

For completion, let's use the substitution method, with $y = 2x - 2$, to verify we get the same result.

$$\begin{aligned} f(x) &= 4 - (x^2 + (2x - 2)^2) \\ &= 4 - (x^2 + 4x^2 - 8x - 4) \\ &= -5x^2 + 8x \end{aligned}$$

Next, we differentiate and set to zero.

$$f'(x) = -10x + 8 = 0 \rightarrow x = \frac{4}{5}$$

Again, substituting we find $y = -2/5$, which results in the same maximum value, $f_{max} = 3.2$.

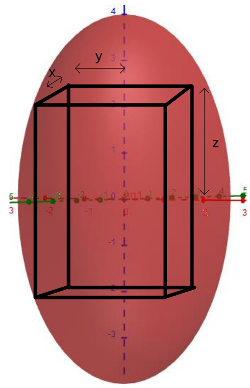
This example illustrates the Lagrange multiplier technique for a two variable objective function. The method is also valid for any number of variables, but of course becomes more difficult to visualize as the number of variables increases. The method also extends to problems with more than one constraint function. For example, if the problem is to minimize $f(x, y, z)$ subject to constraints $g(x, y, z)$ and $h(x, y, z)$, the Lagrange condition becomes.

$$\boxed{\nabla f_P = \lambda_1 \nabla g_P + \lambda_2 \nabla h_P}$$

Next, let's revisit a problem from the previous lesson to illustrate the benefits of the Lagrange method.

Example 2 (Example 6 from previous lesson): Find the dimensions of the largest rectangular box that can be inscribed in a general ellipsoid.

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1$$



Solution: The objective and constraint functions are shown below.

Objective Function (Volume of Rectangle)	***Constraint Function (General Ellipsoid)***
$f(x, y, z) = 2x \cdot 2y \cdot 2x = 8xyz$	$g(x, y, z) = \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 - 1 = 0$

Directly using the Lagrange Multiplier technique, we have

$$\nabla f_P = \lambda \nabla g_P$$

$$\langle 8yz, 8xz, 8xy \rangle = \lambda \left\langle \frac{2x}{a^2}, \frac{2y}{b^2}, \frac{2z}{c^2} \right\rangle$$

Which gives us the following three equations

$$8yz = \lambda \frac{2x}{a^2} \qquad \lambda = \frac{4a^2yz}{x}$$

$$8xz = \lambda \frac{2y}{b^2} \qquad \lambda = \frac{4b^2xz}{y}$$

$$8xy = \lambda \frac{2z}{c^2} \qquad \lambda = \frac{4c^2xy}{z}$$

Therefore,

$$\frac{a^2yz}{x} = \frac{b^2xz}{y} = \frac{c^2xy}{z}$$

Next, we can solve for x and z in terms of y as shown below.

$$\frac{b^2x\cancel{z}}{y} = \frac{a^2y\cancel{z}}{x} \qquad x^2 = \frac{a^2}{b^2}y^2$$

$$\frac{b^2\cancel{x}z}{y} = \frac{c^2\cancel{x}y}{z} \qquad z^2 = \frac{c^2}{b^2}y^2$$

Substituting these results into the constraint equation we can solve for y as follows

$$\frac{\frac{a^2}{b^2}y^2}{a^2}+\frac{y^2}{b^2}+\frac{\frac{c^2}{b^2}y^2}{c^2}=1$$

$$\frac{y^2}{b^2}+\frac{y^2}{b^2}+\frac{y^2}{b^2}=1$$

$$y=\frac{b}{\sqrt{3}}$$

We can now go back and solve for x and y in terms of z this time.

$$x^2=\frac{a^2}{c^2}z^2 \quad \text{and} \quad y^2=\frac{b^2}{c^2}z^2$$

Substituting as we did before we can now solve for z.

$$\frac{\frac{a^2}{c^2}z^2}{a^2}+\frac{\frac{b^2}{c^2}z^2}{b^2}+\frac{z^2}{c^2}=1$$

$$\frac{z^2}{c^2}+\frac{z^2}{c^2}+\frac{z^2}{c^2}=1$$

$$z=\frac{c}{\sqrt{3}}$$

Repeating this procedure for a third time results in

$$x=\frac{a}{\sqrt{3}}$$

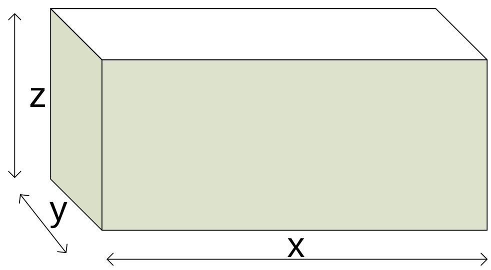
Therefore, we arrive at the same result from the previous lesson with the maximum volume box having dimensions of

$$\left(\frac{a}{\sqrt{3}},\frac{b}{\sqrt{3}},\frac{c}{\sqrt{3}}\right)$$

with a volume of

$$V=8xyz=8\left(\frac{a}{\sqrt{3}}\cdot\frac{b}{\sqrt{3}}\cdot\frac{c}{\sqrt{3}}\right)=\frac{8abc}{3\sqrt{3}}$$

Example 3: (Example 5 from previous lesson): An open top rectangular bin is designed to hold V cubic feet of garbage. The four walls are fabricated from material costing \$2 per square foot while the bottom is made from material costing \$4 per square foot. Find the dimensions of the bin to minimize the cost.



Solution: The objective and constraint functions are shown below.

Objective Function (Cost of Material)	***Constraint Function (Volume of Bin)***
$C = f(x,y,z) = 4(xy + yz + xz)$	$g(x,y,z) = xyz - V = 0$

Using the Lagrange Multiplier technique there is no need to solve the constraint function for one of the variables and substitute. Instead we proceed as follows

$$\nabla f_P = \lambda \nabla g_P$$
$$\langle 4(y+z), 4(x+z), 4(y+x)\rangle = \lambda\langle yz, xz, xy\rangle$$

Which gives us the following three equations

$$4(y+z) = \lambda yz \qquad 4(x+z) = \lambda xz \qquad 4(y+x) = \lambda xy$$
$$\lambda = \frac{4(y+z)}{yz} \qquad \lambda = \frac{4(x+z)}{xz} \qquad \lambda = \frac{4(y+x)}{xy}$$

Setting the first two equations equal we find

$$\frac{4(y+z)}{y\cancel{z}} = \frac{4(x+z)}{x\cancel{z}}$$
$$x(y+z) = y(x+z)$$
$$\cancel{xy} + x\cancel{z} = \cancel{yx} + y\cancel{z}$$
$$x = y$$

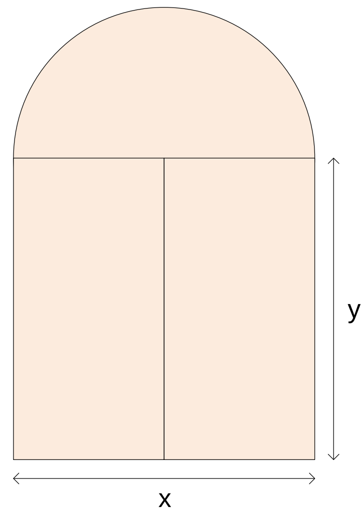
And using the first and third we find

$$\frac{4(y+z)}{\cancel{y}z} = \frac{4(y+x)}{x\cancel{y}}$$
$$x\cancel{y} + \cancel{xz} = z\cancel{y} + \cancel{zx}$$
$$x = z$$

Substituting into the constraint equation we get $x = y = z = \sqrt[3]{V}$ resulting in a minimum cost

$$C = 4\left(\sqrt[3]{V}\sqrt[3]{V} + \sqrt[3]{V}\sqrt[3]{V} + \sqrt[3]{V}\sqrt[3]{V}\right) = 12\sqrt[3]{V^2}$$

Example 5: Assume we have 30 meters of framing material to build the window shown below. Find the dimensions of the window frame that will maximize the area of the glass. (Ignore the thickness of the framing material)



Solution: The area of the window can be computed as

$$A = xy + \frac{\pi r^2}{2} = xy + \frac{\pi\left(\frac{x}{2}\right)^2}{2} = xy + \frac{\pi x^2}{8}$$

The amount of framing material, $P = 30$, can be expressed as

$$P = 2x + 3y + \frac{\cancel{2}\pi r}{\cancel{2}} = 2x + 3y + \pi\left(\frac{x}{2}\right) = x\left(\frac{4+\pi}{2}\right) + 3y$$

These functions are mapped to the objective and constraint functions as shown below.

Objective Function (Area of Window)	***Constraint Function (Length of Framing Material)***
$A = f(x,y) = xy + \frac{\pi x^2}{8}$	$g(x,y) = x\left(\frac{4+\pi}{2}\right) + 3y - P = 0$

Using the Lagrange multiplier technique, we have

$$\nabla f_P = \lambda \nabla g_P$$

$$\langle y + \frac{2\pi x}{8}, x\rangle = \lambda \langle \frac{4+\pi}{2}, 3\rangle \quad \rightarrow \quad \begin{aligned} y + \frac{\pi x}{4} &= \lambda\left(\frac{4+\pi}{2}\right) \\ x &= \lambda 3 \end{aligned}$$

Substituting $\lambda = x/3$ from the second equation into the first we can write y as a function of x

$$y = \frac{x}{3}\left(\frac{4+\pi}{2}\right) - \frac{\pi x}{4}$$

$$y = x\left(\frac{4+\pi}{6} - \frac{\pi}{4}\right)$$

Next, we substitute into the constraint equation and solve for x.

$$x\left(\frac{4+\pi}{2}\right)+3y=P$$

$$x\left(\frac{4+\pi}{2}\right)+3\left(x\left(\frac{4+\pi}{6}-\frac{\pi}{4}\right)\right)=P$$

$$x\left(\frac{4+\pi}{2}+\frac{4+\pi}{2}-\frac{3\pi}{4}\right)=P$$

$$x\left(4+\pi-\frac{3\pi}{4}\right)=P$$

$$x\left(\frac{16+\pi}{4}\right)=P$$

$$x=P\left(\frac{4}{16+\pi}\right)$$

Substituting once more to solve for y.

$$3y=P-x\left(\frac{4+\pi}{2}\right)$$

$$3y=P-P\left(\frac{4}{16+\pi}\right)\left(\frac{4+\pi}{2}\right)$$

$$3y=P-P\left(\frac{8+2\pi}{16+\pi}\right)$$

$$3y=P\left(\frac{16+\pi-8-2\pi}{16+\pi}\right)$$

$$y=P\left(\frac{8-\pi}{48+3\pi}\right)$$

Finally, with $P=30$ we find the dimension and the maximum area below.

$$x=30\left(\frac{4}{16+\pi}\right)$$

$$x\cong 6.27$$

$$y=30\left(\frac{8-\pi}{48+3\pi}\right)$$

$$y\cong 2.54$$

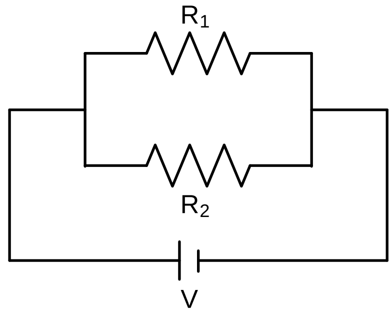
$$A_{max}=xy+\frac{\pi x^2}{8}$$

$$A_{max}\cong 6.27\cdot 2.54+\frac{\pi(6.27)^2}{8}$$

$$A_{max}\cong 31.36$$

Example 6: A constant voltage is applied to the circuit shown below. Assume the sum of the two resistors is a constant, L. What values of R_1 and R_2 will result in the smallest power output. The power can be computed as, $P = V^2/R_T$, where R_T is the equivalent resistance of the circuit, which in this case is given as

$$R_T = \frac{R_1 R_2}{R_2 + R_2}$$



Solution: The objective and constraint functions are shown below.

Objective Function (Circuit Power)	***Constraint Function (Resistor Sum)***
$P = f(R_1, R_2,) = \frac{V^2}{R_T}$ $= V^2 \left(\frac{R_1 + R_2}{R_1 R_2}\right)$	$g(R_1, R_2) = R_1 + R_2 - L = 0$

Using the Lagrange multiplier technique, we have

$$\nabla f_P = \lambda \nabla g_P$$

$$V^2 \langle \frac{R_1 R_2 - (R_1 + R_2) R_2}{(R_1 R_2)^2}, \frac{R_1 R_2 - (R_1 + R_2) R_1}{(R_1 R_2)^2} \rangle = \lambda \langle 1, 1 \rangle$$

$$V^2 \langle \frac{-R_2 R_2}{(R_1 R_2)^2}, \frac{R_1 R_1}{(R_1 R_2)^2} \rangle = \lambda \langle 1, 1 \rangle$$

$$\langle \frac{-V^2}{{R_1}^2}, \frac{-V^2}{{R_2}^2} \rangle = \lambda \langle 1, 1 \rangle$$

Which implies

$$\frac{V^2}{{R_1}^2} = \frac{V^2}{{R_2}^2}$$

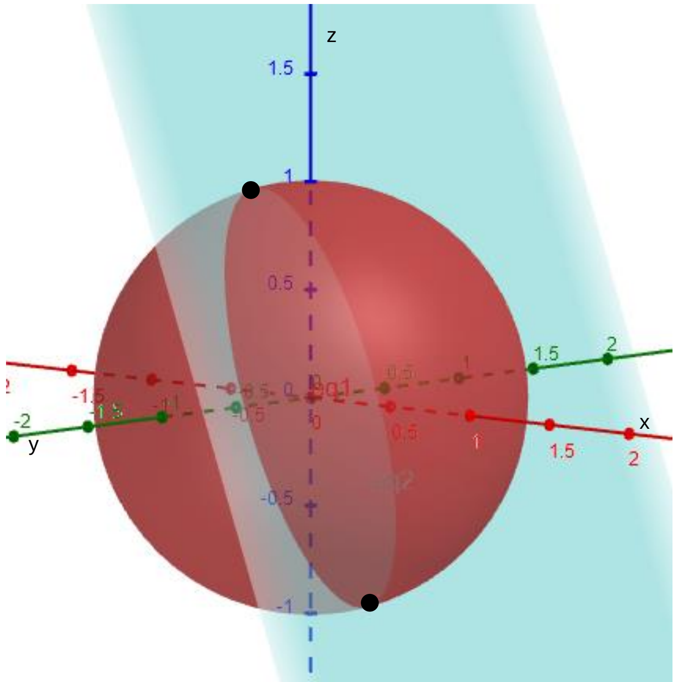
Therefore, to minimize the power the two resistors should be equal.

$$R_1 = R_2 = \frac{L}{2}$$

Let's do one more example using two constraint functions.

Example 7:

The intersection of the plane $x + \frac{1}{2}y + \frac{1}{3}z = 0$ with the unit sphere, $x^2 + y^2 + z^2 = 1$ is called a great circle. Find the point on the great circle with the largest x-coordinate.



Solution: In this case we simply want to find the maximum value of x. Therefore, the objective function is $f(x, y, z) = x$. However, this value is constrained by the fact that it must satisfy the following *two* equations.

$$g(x, y, z) = x + \frac{1}{2}y + \frac{1}{3}z = 0 \qquad\qquad h(x, y, z) = x^2 + y^2 + z^2 - 1 = 0$$

The Lagrange condition with two constraints is as follows

$$\nabla f_P = \lambda_1 \nabla g_P + \lambda_2 \nabla h_P$$
$$\langle 1, 0, 0\rangle = \lambda_1 \langle 1, \frac{1}{2}, \frac{1}{3}\rangle + \lambda_2\langle 2x, 2y, 2z\rangle$$

$$1 = \lambda_1 + \lambda_2 2x \qquad\qquad 0 = \lambda_1 \frac{1}{2} + \lambda_2 2y \qquad\qquad 0 = \lambda_1 \frac{1}{3} + \lambda_2 2z$$

Solving the second two equation for λ_1 we can write y as a function of z.

$$-\lambda_2 4y = -\lambda_2 6z$$
$$y = \frac{3}{2}z$$

Using this in the first constraint equation we can also find x as a function of z.

$$x + \frac{1}{2}\left(\frac{3}{2}z\right) + \frac{1}{3}z = 0$$

$$x = -\frac{13}{12}z$$

Now, using the second constraint equation, we can solve for z.

$$\left(-\frac{13}{12}\right)^2 + \left(\frac{3}{2}z\right)^2 + z^2 = 1$$

$$z = \pm\sqrt{\frac{144}{637}} = \pm\frac{12}{7\sqrt{13}}$$

Therefore, in this case we have two critical points

$$C_1 = \left(-\frac{13}{7\sqrt{13}}, \frac{18}{7\sqrt{13}}, \frac{12}{7\sqrt{13}}\right) \qquad C_2 = \left(\frac{13}{7\sqrt{13}}, -\frac{18}{7\sqrt{13}}, -\frac{12}{7\sqrt{13}}\right)$$

The point, C_2, has the largest x value, therefore

$$x_{max} = \frac{13}{7\sqrt{13}} \cong 0.515$$

As you can see, even with the Lagrange multiplier technique, optimization problems can still become algebraically intensive. The main advantages of using Lagrange multipliers are

1. Isolating variables from the constraint equation(s) to substitute into the objective function is not required. Recall, for implicitly defined functions isolating variables is not generally possible.
2. The technique is a general procedure that can be used for any number of variables or constraint equations.

Final Summary for Multivariable Differentiation – Lagrange Multipliers

Optimizing with Constraints

Optimizing with constraints involves finding the minimum or maximum value of a function, e.g. $f(x_1, \ldots, x_n)$ subject to the fact that the independent variables are related in some fashion, e.g. $g(x_1, \ldots, x_n) = 0$. The terminology used is as follows:

Objective Function $f(x_1, \ldots, x_n)$	Expresses the quantity we would like to optimize in terms of n independent variables.
Constraint Function $g(x_1, \ldots, x_n) = 0$	Expresses a relationship between the independent that must be satisfied within the context of optimizing the objective function.

Lagrange Multiplier Theorem

Assume $f(x, y)$ and $g(x, y)$ are differentiable functions. If $f(x, y)$ has a local extremum on the constraint curve, $g(x, y) = 0$, at $P = (a, b)$ and if $\nabla g_P \neq 0$, then there is a scalar, λ, such that

$$\nabla f_P = \lambda \nabla g_P$$

Lagrange Multipliers Technique Applied to Optimization with Constraints

The above Lagrange Multiplier Theorem can be applied to optimization problems with constraints. The theorem can be generalized to any number of variables and any number of constraints functions as follows:

Given an n variable differentiable objective function, $f(x_1, \ldots, x_n)$, and m differentiable constraint functions, $\{g_1(x_1, \ldots, x_n) = 0, \ldots, g_m(x_1, \ldots, x_n) = 0\}$. The Lagrange condition is written as follows:

$$\nabla f_P = \sum_{i=1}^{m} \lambda_i \nabla g_{i,P}$$

Expanding this expression creates n equations that can then be used to find the extreme values of $f(x_1, \ldots, x_n)$ subject to $\{g_1(x_1, \ldots, x_n) = 0, \ldots, g_m(x_1, \ldots, x_n) = 0\}$.

By: ferrantetutoring