2.4.3 Constraints via Lagrange multipliers

In this section we will see a particular method to solve so-called problems of constrained extrema. There are two kinds of typical problems:

Finding the shortest distance from a point to a plane: Given a plane

$$Ax + By + Cz + D = 0, \tag{2.191}$$

obtain the shortest distance from a point (x_0, y_0, z_0) to this plane. In this case we have to minimize the function

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2, \tag{2.192}$$

with the constraint that the point (x, y, z) is in the plane (2.191).

Shortest distance from a point to a generic surface: This is a more general problem where the equation of a three dimensional surface is given,

$$\phi(x, y, z) = 0, \tag{2.193}$$

and we are asked to obtain the shortest distance from a point (x_0, y_0, z_0) to this surface. Again we need to minimize the function distance including the constraint that the points we want to consider all belong to the surface (2.193).

The method of Lagrange multipliers provides an easy way to solve this kind of problems. Assume that $f(x, y, z)$ is the function we want to minimize (in the examples it would be the distance). Then if the function has a minimum at a point (x_0, y_0, z_0) its first order differential vanishes at that point:

$$df = f_x dx + f_y dy + f_z dz = 0. \tag{2.194}$$

Suppose that

$$\phi(x, y, z) = 0, \tag{2.195}$$

is the constraint (in the previous examples the equation of a certain surface). Then it follows trivially that

$$d\phi = \phi_x dx + \phi_y dy + \phi_z dz = 0, \tag{2.196}$$

and therefore we can also write that

$$d(f + \lambda\phi) = df + \lambda d\phi = 0, \tag{2.197}$$

where λ is an arbitrary constant which we call **Lagrange's multiplier**. The previous equation is equivalent to

$$\left(\frac{\partial f}{\partial x} + \lambda\frac{\partial \phi}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y} + \lambda\frac{\partial \phi}{\partial y}\right) dy + \left(\frac{\partial f}{\partial z} + \lambda\frac{\partial \phi}{\partial z}\right) dz = 0. \tag{2.198}$$

Due to the constraint (2.195), z is an implicit function of the independent variables x, y. Since λ is arbitrary we can choose it to have a particular value. For example, let us choose it so that

$$\frac{\partial f}{\partial z} + \lambda\frac{\partial \phi}{\partial z} = 0. \tag{2.199}$$

In that case equation (2.198) reduces to

$$\left(\frac{\partial f}{\partial x} + \lambda\frac{\partial \phi}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y} + \lambda\frac{\partial \phi}{\partial y}\right) dy = 0, \tag{2.200}$$

and since x and y are independent variables, there should not be a relationship between their differentials, so the only way to solve the equation above is to take

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} = 0, \qquad \frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} = 0. \tag{2.201}$$

Therefore, we have now a set of 4 equations, namely

$$\begin{aligned}
\phi(x, y, z) &= 0, && (2.202)\\
\frac{\partial f}{\partial z} + \lambda \frac{\partial \phi}{\partial z} &= 0, && (2.203)\\
\frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} &= 0, && (2.204)\\
\frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} &= 0, && (2.205)
\end{aligned}$$

which determine completely the 4 unknowns of the problem, namely the coordinates of the point (x, y, z) which is closest to the point (x_0, y_0, z_0) and the value of λ.

Remark: Notice that we can use the same method for problems involving only two variables. In that case we will have three equations, instead of four.

Example 1: Using the method of Lagrange multipliers find the shortest distance from the point $(0, 0, 1)$ to the surface $yx + yz + xz = 0$.
Following the general scheme given above we can identify the constraint

$$\phi(x, y, z) = yx + yz + xz = 0, \tag{2.206}$$

and the function we want to minimize is

$$f(x, y, z) = x^2 + y^2 + (z-1)^2. \tag{2.207}$$

Therefore we can directly apply the method we have just seen and reduce the problem to the solution of the 4 equations (2.202)-(2.205). For that we need to evaluate the partial derivatives

$$\begin{aligned}
f_x &= 2x, \qquad f_y = 2y, \qquad f_z = 2(z-1), && (2.208)\\
\phi_x &= z + y, \qquad \phi_y = x + z, \qquad \phi_z = x + y. && (2.209)
\end{aligned}$$

And so we have to solve the equations

$$\begin{aligned}
yx + yz + xz &= 0, && (2.210)\\
2(z-1) + \lambda(x+y) &= 0, && (2.211)\\
2x + \lambda(z+y) &= 0, && (2.212)\\
2y + \lambda(x+z) &= 0. && (2.213)
\end{aligned}$$

From the last 3 equations we obtain:

$$\lambda = -\frac{2(z-1)}{x+y} = -\frac{2x}{z+y} = -\frac{2y}{x+z}, \tag{2.214}$$

and we can rewrite the last equality as,

$$\frac{x}{z+y} = \frac{y}{x+z} \quad \Leftrightarrow \quad x(x+z) = y(y+z), \tag{2.215}$$

which is equivalent to

$$z(y-x) = x^2 - y^2 = (x-y)(x+y). \tag{2.216}$$

The last equation admits two solutions: $x = y$ and $z = -(x+y)$ with $x \neq y$.

Let us take the solution $x = y$ for now. If we put this into (2.212) we get that

$$z = -x\frac{2+\lambda}{\lambda}.$$

If we put this and into equation (2.210) we get:

$$x^2\left(1 - 2\frac{2+\lambda}{\lambda}\right) = 0, \tag{2.217}$$

which is solved by

$$\lambda = -4, \quad \text{or} \quad x = 0. \tag{2.218}$$

If $\lambda = -4$ then $z = -x/2$. Substituting in (2.211) with $x = y$ we obtain

$$x = -2/9 = y, \tag{2.219}$$

and $x = 1/9$. If we take the other solution to (2.217), namely $x = 0 = y$, then, substituting in (2.211) we obtain $z = 1$ and substituting this in (2.212) we obtain $\lambda = 0$.

Therefore we have the points $(x, y, z) = (0, 0, 1)$ and $(x, y, z) = (-2/9, -2/9, 1/9)$. The fist solution is the same point whose distance to we wanted to minimize. Therefore we have discovered that the point $(0, 0, 1)$ is itself on the surface (2.135) and therefore the solution to our problem is the same point. The second solution we obtain must be therefore the point on the surface whose distance to $(0, 0, 1)$ is maximal. In fact the distance is

$$r = \sqrt{(2/9)^2 + (2/9)^2 + (8/9)^2} = \sqrt{72/81} = \frac{2\sqrt{2}}{3}. \tag{2.220}$$

Finally, let us go back to the line after (2.216). There we had obtained a second possible solution besides $x = y$. It was $z = -x - y$ with $x \neq y$. We must see if this gives any further solutions. If we substitute this in (2.211) we obtain $(x+y)(-2+\lambda) = 2$, that is $x + y = 2/(-2+\lambda)$. If we now take equations (2.212) and (2.213) and add them up, we obtain

$$2(x+y) + \lambda(y + 2z + x) = 0 \quad \Leftrightarrow \quad (x+y)(2-\lambda) = 0, \tag{2.221}$$

this gives an equation for λ,

$$\frac{(2-\lambda)}{-2+\lambda} = 0, \tag{2.222}$$

which does not make sense. Therefore the solution $z = -x - y$ is not valid.

Example 2: Using the method of Lagrange multipliers, determine the maximum of the function

$$f(x, y, z) = xyz, \tag{2.223}$$

subject to the condition

$$x^3 + y^3 + z^3 = 1, \tag{2.224}$$

with $x \geq 0, y \geq 0, z \geq 0$.

In this case our constraint is

$$\phi(x, y, z) = x^3 + y^3 + z^3 - 1 = 0, \tag{2.225}$$

and the corresponding partial derivatives of f and ϕ are

$$f_x = zy, \qquad f_y = xz, \qquad f_z = xy, \tag{2.226}$$
$$\phi_x = 3x^2, \qquad \phi_y = 3y^2, \qquad \phi_z = 3z^2. \tag{2.227}$$

Therefore we need to solve the following system of equations

$$x^3 + y^3 + z^3 - 1 = 0, \tag{2.228}$$
$$zy + \lambda 3x^2 = 0, \tag{2.229}$$
$$zx + \lambda 3y^2 = 0, \tag{2.230}$$
$$xy + \lambda 3z^2 = 0. \tag{2.231}$$

It is useful to try to see if there are any obvious solutions that we could get without too much work. For example, if we set any two variables to zero, then the remaining one will automatically need to be 1, from equation (2.228). This immediately gives us three solutions: the points $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$. These solutions correspond to $\lambda = 0$. Another solution which is easy to see is to take $x = y = z$. Substituting this into (2.228) will give us

$$x = y = z = \sqrt[3]{\frac{1}{3}}, \tag{2.232}$$

and any of the other equations will tell us that $\lambda = -1/3$. These are in fact all the solutions to our problem. If we now compute

$$f(0, 0, 1) = f(1, 0, 0) = f(0, 1, 0) = 0, \tag{2.233}$$

and

$$f(\sqrt[3]{\frac{1}{3}}, \sqrt[3]{\frac{1}{3}}, \sqrt[3]{\frac{1}{3}}) = 1/3, \tag{2.234}$$

we find that f has minimum value at the points $(0, 0, 1), (1, 0, 0), (0, 1, 0)$ and is maximal at $(\sqrt[3]{\frac{1}{3}}, \sqrt[3]{\frac{1}{3}}, \sqrt[3]{\frac{1}{3}})$.