Constrained Optimization and Lagrange Multiplier Method

BENNETT UNIVERSITY
TIMES OF INDIA GROUP

Constrained optimization

- Constrained optimization is the process of optimizing an objective function with respect to some variables in the presence of constraints on those variables.
- The objective function is either a **cost function** or **energy function**, which is to be minimized, or a **reward function** or **utility function**, which is to be maximized.
-

General form of a constrained optimization problem with equality constraints only

minimize or **max**imize $f(x, y, z, ...)$

subject to $\quad g_i(x, y, z, ...) = c_i,$

for $i = 1, 2, \ldots, m$.

Example 1: Maximize the function $f(x, y) = x^2 y$ subject to the constraints $x^2 + y^2 = 1$.

Example 2: Find the shortest distance from origin to the plane $z = 2x + y - 5$, i.e. minimize the function $f(x, y, z) = x^2 + y^2 + z^2$ subject to the constraint $2x + y + z - 5 = 0$.

How to solve constraint optimization problems

Problem 1: $\max f(x, y) = x^2y$ subject to $x^2 + y^2 = 1$.

Solution: Substituting the constraint in the function, we get

$$h(y) = (1 - y^2)y = y - y^3.$$

Now, $h'(y) = 0 \implies y = \pm\frac{1}{\sqrt{3}}$ and from the constraint equation, we have $x = \pm\sqrt{\frac{2}{3}}$.

Therefore, the critical points are $\left(\frac{\sqrt{2}}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(\frac{\sqrt{2}}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right), \left(-\frac{\sqrt{2}}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ and $\left(-\frac{\sqrt{2}}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)$. By evaluating the function at these critical points, we get

$\left(\pm\frac{\sqrt{2}}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ as the points of maxima and
the final answer to the problem is $\frac{2}{3\sqrt{3}}$.

How to solve constraint optimization problems

Problem 2: $\min f(x, y, z) = x^2 + y^2 + z^2$ subject to $2x + y + z - 5 = 0$.

Solution: Substituting the constraint in the function, we get

$$h(x, y) = f(x, y, 2x + y - 5) = x^2 + y^2 + (2x + y - 5)^2.$$

The critical points of this function are

$$h_x = 2x + 2(2x + y - 5)(2) = 0, \; h_y = 2y + 2(2x + y - 5) = 0.$$

This leads to $x = 5/3, y = 5/6$. Then $z = 2x + y - 5$ implies $z = -5/6$. Observe that $D > 0$ and $h_{xx} > 0$. So the point $(5/3, 5/6, -5/6)$ is a point of minimum.

Does the substitution method always work?

Example: The shortest distance from origin to $x^2 - z^2 = 1$, i.e. minimizing $f(x, y, z) = x^2 + y^2 + z^2$ subject to the constraint $x^2 - z^2 = 1$.

Ans: Substituting $z^2 = x^2 - 1$ in f, we get

$$h(x, y) = f(x, y, \sqrt{x^2 - 1}) = 2x^2 + y^2 - 1.$$

Now, $h_x = 4x = 0, h_y = 2y = 0$. Implies, $x = 0, y = 0, z^2 = -1$. Then z is imaginary. To overcome this difficulty, we can substitute $x^2 = z^2 + 1$ in f and find that $z = y = 0$ and $x = \pm 1$. These points are on the hyperbolic cylinder $(x^2 - z^2 = 1)$ and we can check that $D > 0, h_{xx} > 0$. This implies the points are of local minimum nature.

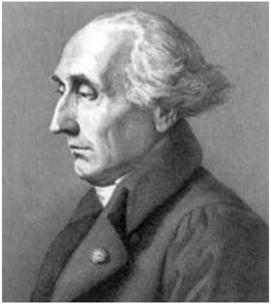
Caution: In the substitution method, once we substitute the constraint in the minimizing function, then the the domain of the function will be the domain of the minimizing function.

Then the critical points can belong to this domain which may not be the domain of constraints.

How do we overcome this problem?

Lagrange multipliers method

- The Lagrange multipliers method is a strategy for finding the local maxima and minima of a function subject to equality constraints (i.e., subject to the condition that one or more equations have to be satisfied exactly by the chosen values of the variables).



Joseph-Louis Lagrange

(1736–1813)

Lagrange multiplier method for a function of two variables with single constraint

Consider the optimization problem:

$$\begin{aligned} &\text{maximize} \quad f(x,y), \\ &\text{subject to: } g(x,y) = 0. \end{aligned}$$

Note: Assume both f and g have continuous first partial derivatives.

Now, introduce a new variable λ called a **Lagrange multiplier** and study the **Lagrange function** defined by

$$\boxed{\mathcal{L}(x,y,\lambda) = f(x,y) + \lambda g(x,y).}$$

- If $f(x_0, y_0)$ is a maximum of $f(x, y)$ for the original constrained problem and $\nabla g(x_0, y_0) \neq 0$, then there exists λ_0 such that (x_0, y_0, λ_0) is a **stationary point**† for the Lagrange function.

  († Stationary points are those points where the first partial derivatives of $\mathcal{L}$ are zero.)

- The assumption $\nabla g \neq 0$ is called **constraint qualification**.

Algorithm for the general case

1. Number of constraints m should be less than the number of independent variables n say $g_1 = 0, g_2 = 0, ...g_m = 0$.
2. Write the Lagrange multiplier equation: $\nabla f = \sum_{i=1}^{m} \lambda_i \nabla g_i$.
3. Solve the set of $m + n$ equations to find the extremal points

$$\nabla f = \sum_{i=1}^{m} \lambda_i \nabla g_i, \quad g_i = 0,\ i = 1, 2, ...m$$

4. Once we have extremum points, compare the values of f at these points to determine the maxima and minima.

Examples

Example 1: Suppose we wish to maximize $f(x, y) = x + y$ subject to the constraint $x^2 + y^2 = 1$.

$$\begin{aligned}\mathcal{L}(x, y, \lambda) &= f(x, y) + \lambda \cdot g(x, y) \\ &= x + y + \lambda(x^2 + y^2 - 1).\end{aligned}$$

So gradient is given by

$$\begin{aligned}\nabla_{x,y,\lambda}\mathcal{L}(x, y, \lambda) &= \left(\frac{\partial \mathcal{L}}{\partial x}, \frac{\partial \mathcal{L}}{\partial y}, \frac{\partial \mathcal{L}}{\partial \lambda}\right) \\ &= \left(1 + 2\lambda x, 1 + 2\lambda y, x^2 + y^2 - 1\right)\end{aligned}$$

continue...

$$\therefore \nabla_{x,y,\lambda}\mathcal{L}(x,y,\lambda) = 0 \Leftrightarrow \begin{cases} 1 + 2\lambda x = 0 \\ 1 + 2\lambda y = 0 \\ x^2 + y^2 - 1 = 0 \end{cases}$$

The first two equations yield

$$x = y = -\frac{1}{2\lambda}, \qquad \lambda \neq 0.$$

By substituting into the last equation we have

$$\frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} - 1 = 0 \implies \lambda = \pm\frac{1}{\sqrt{2}}.$$

continue...

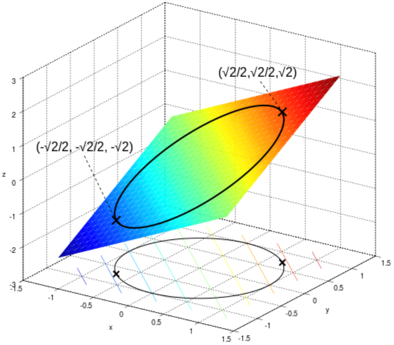
Therefore, the stationary points of $\mathcal{L}$ are

$$\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, -\frac{1}{\sqrt{2}}\right), \quad \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, \frac{1}{\sqrt{2}}\right).$$

Evaluating the objective function f at these points yields

$$f\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = \sqrt{2}, \;\; f\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = -\sqrt{2}.$$

Thus the constrained maximum is $\sqrt{2}$ and the constrained minimum is $-\sqrt{2}$.



https://en.wikipedia.org/wiki/File:Lagrange_very_simple.svg

Example 2: Find the maximum values of $f(x, y) = x^2 y$ subject to the constraint $g(x, y) = x^2 + y^2 - 3 = 0$.

$$\begin{aligned}\mathcal{L}(x, y, \lambda) &= f(x, y) + \lambda \cdot g(x, y) \\ &= x^2 y + \lambda(x^2 + y^2 - 3).\end{aligned}$$

$$\begin{aligned}\implies \nabla_{x,y,\lambda}\mathcal{L}(x, y, \lambda) &= \left(\frac{\partial \mathcal{L}}{\partial x}, \frac{\partial \mathcal{L}}{\partial y}, \frac{\partial \mathcal{L}}{\partial \lambda}\right) \\ &= \left(2xy + 2\lambda x, x^2 + 2\lambda y, x^2 + y^2 - 3\right).\end{aligned}$$

$$\therefore \nabla_{x,y,\lambda}\mathcal{L}(x, y, \lambda) = 0 \iff \begin{cases} 2xy + 2\lambda x = 0 \\ x^2 + 2\lambda y = 0 \\ x^2 + y^2 - 3 = 0 \end{cases} \iff \begin{cases} x(y + \lambda) = 0 & \text{(i)} \\ x^2 = -2\lambda y & \text{(ii)} \\ x^2 + y^2 = 3 & \text{(iii)} \end{cases}$$

continue...

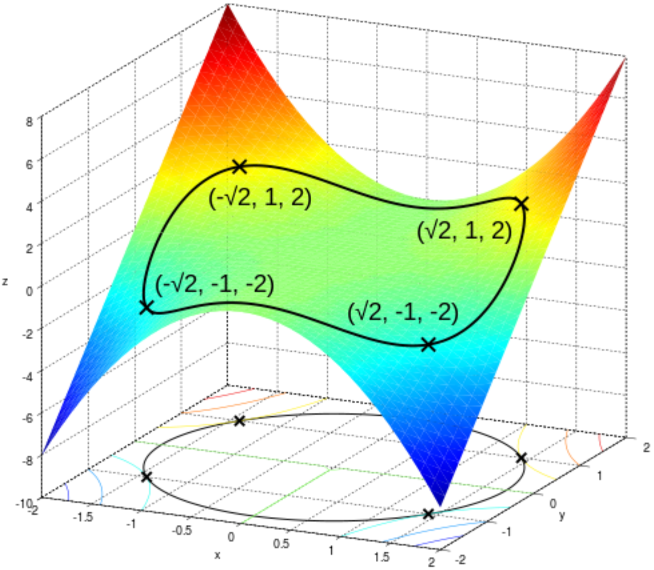
(i) implies $x = 0$ $\lambda = -y$. If $x = 0$ then $y = \pm\sqrt{3}$ by (iii) and consequently $\lambda = 0$ from (ii). If $\lambda = -y$, substituting this into (ii) we get $x^2 = 2y^2$. Now substituting this into (iii) and solving for y gives $y = \pm 1$. Thus there are six critical points of $\mathcal{L}$:

$$(\sqrt{2}, 1, -1); (-\sqrt{2}, 1, -1); (\sqrt{2}, -1, 1); (-\sqrt{2}, -1, 1); (0, \sqrt{3}, 0); (0, -\sqrt{3}, 0).$$

Evaluating the objective at these points, we find that

$$f(\pm\sqrt{2}, 1) = 2; f(\pm\sqrt{2}, -1) = -2; f(0, \pm\sqrt{3}) = 0.$$

Therefore, the objective function attains the global maximum (subject to the constraints) at $(\pm\sqrt{2}, 1)$ and the global minimum at $(\pm\sqrt{2}, -1)$. The point $(0, \sqrt{3})$ is a local minimum of f and $(0, -\sqrt{3})$ is a local maximum of f.



maximize $f(x, y) = x^2 y$ subject to the constraint $g(x, y) = x^2 + y^2 - 3 = 0.$

https://en.wikipedia.org/wiki/File:Lagrange_simple.svg

Thank You.