Lagrangian Multipliers

Podcast Learn & Fun *

February 3, 2025

In mathematical optimization, the method of *Lagrange multipliers* is a strategy used to find the local maxima and minima of a function subject to constraints. This method transforms a constrained optimization problem into an unconstrained one by incorporating the constraints into the objective function using additional variables, called Lagrange multipliers. It plays a central role in fields such as economics, physics, engineering, and machine learning.

1 Equality Constraints

Problem Setup

Consider a function $f(x_1, x_2, \ldots, x_n)$ that we wish to maximize or minimize, subject to one or more constraints. We are given the following optimization problem:

$$\text{Maximize or Minimize} \quad f(x_1, x_2, \ldots, x_n)$$

subject to the constraints:

$$\begin{gathered} g_1(x_1, x_2, \ldots, x_n) = 0 \\ g_2(x_1, x_2, \ldots, x_n) = 0 \\ \vdots \\ g_m(x_1, x_2, \ldots, x_n) = 0 \end{gathered}$$

Here, f is the objective function to be maximized or minimized, and each g_i represents one constraint function.

The Method of Lagrange Multipliers

To solve this constrained optimization problem, we introduce new variables, called *Lagrange multipliers*, denoted $\lambda_1, \lambda_2, \ldots, \lambda_m$, corresponding to each constraint $g_1, g_2, \ldots, g_m$. We then form the *Lagrangian* function $\mathcal{L}$, which combines the objective function and the constraints:

*YouTube Channel: https://www.youtube.com/@Podcast_Learn_Fun

$$\mathcal{L}(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m) = f(x_1, x_2, \ldots, x_n) + \sum_{i=1}^{m} \lambda_i g_i(x_1, x_2, \ldots, x_n)$$

To find the critical points, we take the partial derivatives of $\mathcal{L}$ with respect to all variables, including the original variables $x_1, x_2, \ldots, x_n$ and the Lagrange multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m$. The first-order conditions for optimization are:

$$\frac{\partial \mathcal{L}}{\partial x_j} = 0 \quad \text{for} \quad j = 1, 2, \ldots, n$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} = 0 \quad \text{for} \quad i = 1, 2, \ldots, m$$

These conditions give the system of equations:

$$\frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0 \quad \text{for each} \quad j = 1, 2, \ldots, n$$

$$g_i(x_1, x_2, \ldots, x_n) = 0 \quad \text{for each} \quad i = 1, 2, \ldots, m$$

The first equation expresses the fact that at the optimal point, the gradient of the objective function $f(x_1, x_2, \ldots, x_n)$ must be a linear combination of the gradients of the constraint functions $g_i(x_1, x_2, \ldots, x_n)$. In other words, the optimal solution occurs when the gradient of the objective function is "aligned" with the gradients of the constraints. The second equation ensures that the values of the variables $x_1, x_2, \ldots, x_n$ satisfy the constraints $g_i(x_1, x_2, \ldots, x_n) = 0$ for all i.

Geometrical Interpretation

From a geometric perspective, Lagrange multipliers offer a way of finding the points where the level sets (contours) of the objective function f are tangent to the level sets of the constraint functions g_i. At the optimal point, the objective function and the constraints "intersect" in a manner such that their gradients are parallel.

The Lagrange multiplier λ_i can be interpreted as a measure of how sensitive the optimal value of the objective function is to changes in the constraint g_i. If $\lambda_i > 0$, the constraint g_i is active (affecting the solution); if $\lambda_i = 0$, the constraint is inactive (it does not affect the solution).

Example

Consider the problem of maximizing the function $f(x, y) = xy$ subject to the constraint $x^2 + y^2 = 1$.

Solution: The Lagrangian function is:

$$\mathcal{L}(x, y, \lambda) = xy + \lambda(1 - x^2 - y^2)$$

We take the partial derivatives of $\mathcal{L}$ with respect to x, y, and λ:

$$\frac{\partial \mathcal{L}}{\partial x} = y - 2\lambda x = 0$$

$$\frac{\partial \mathcal{L}}{\partial y} = x - 2\lambda y = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 1 - x^2 - y^2 = 0$$

From the first two equations, we obtain:

$$y = 2\lambda x \quad \text{and} \quad x = 2\lambda y$$

Substituting $y = 2\lambda x$ into $x = 2\lambda y$, we get:

$$x = 2\lambda(2\lambda x) \quad \Rightarrow \quad x(1 - 4\lambda^2) = 0$$

Therefore, $x = 0$ or $\lambda = \pm 1/2$.

If $x = 0$, then $y = \pm 1$ from the constraint equation $x^2 + y^2 = 1$, yielding the points $(0, 1)$ and $(0, -1)$. If $\lambda = \pm 1/2$, substituting into the constraint equation and solving gives the points $\left(\pm 1/\sqrt{2}, \pm 1/\sqrt{2}\right)$. By evaluating $f(x, y) = xy$ at these points, we find that the maximum is $1/2$ and the minimum is $-1/2$, occurring at $\left(1/\sqrt{2}, 1/\sqrt{2}\right)$ and $\left(-1/\sqrt{2}, -1/\sqrt{2}\right)$, respectively.

2 Inequality Constraints

When dealing with optimization problems that involve inequality constraints rather than just equality constraints, we need to modify the Lagrangian multiplier method. This extension is known as the Karush-Kuhn-Tucker (KKT) conditions, which provide necessary and sufficient conditions for optimality for problems with inequality constraints.

Problem Setup

Consider an optimization problem where we want to maximize or minimize an objective function $f(x_1, x_2, \ldots, x_n)$, subject to inequality constraints. The general form of the problem is:

$$\text{Maximize or Minimize} \quad f(x_1, x_2, \ldots, x_n)$$

subject to the constraints:

$$g_i(x_1, x_2, \ldots, x_n) \leq 0 \quad \text{for} \quad i = 1, 2, \ldots, m$$

$$h_j(x_1, x_2, \ldots, x_n) = 0 \quad \text{for} \quad j = 1, 2, \ldots, p$$

Here, $f(x_1, x_2, \ldots, x_n)$ is the objective function to be optimized, $g_i(x_1, x_2, \ldots, x_n) \leq 0$ are the inequality constraints, and $h_j(x_1, x_2, \ldots, x_n) = 0$ are the equality constraints.

KKT (Karush-Kuhn-Tucker) Conditions

For a constrained optimization problem with inequality constraints, the KKT conditions provide the necessary and sufficient conditions for optimality. These conditions include:

Stationarity The first-order condition for the objective function is:

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(x^*) = 0$$

where λ_i are the Lagrange multipliers corresponding to the inequality constraints, and μ_j are the Lagrange multipliers for the equality constraints.

Primal Feasibility The constraints must be satisfied at the optimal point:

$$g_i(x^*) \leq 0, \quad i = 1, 2, \ldots, m$$

$$h_j(x^*) = 0, \quad j = 1, 2, \ldots, p$$

Dual Feasibility The Lagrange multipliers for the inequality constraints must be non-negative:

$$\lambda_i \geq 0, \quad i = 1, 2, \ldots, m$$

Complementary Slackness For each inequality constraint, the product of the Lagrange multiplier and the constraint must be zero:

$$\lambda_i g_i(x^*) = 0, \quad i = 1, 2, \ldots, m$$

This condition ensures that if a constraint is active (i.e., $g_i(x^*) = 0$), then the corresponding Lagrange multiplier λ_i can be positive; if the constraint is inactive (i.e., $g_i(x^*) < 0$), then $\lambda_i = 0$.

Dual Optimality The optimal values of the objective function and Lagrange multipliers should satisfy the above conditions simultaneously.

Example

Consider the following optimization problem:

$$\text{Maximize} \quad f(x, y) = xy$$

subject to the constraints:

$$x^2 + y^2 \leq 1 \quad \text{(inequality constraint)}$$

$$x \geq 0 \quad \text{(inequality constraint)}$$

$$y = 1 \quad \text{(equality constraint)}$$

Solution: The Lagrangian function is:

$$\mathcal{L}(x, y, \lambda_1, \lambda_2, \mu) = xy + \lambda_1(1 - x^2 - y^2) + \lambda_2(-x) + \mu(y - 1)$$

We take the partial derivatives of $\mathcal{L}$ with respect to x, y, λ_1, λ_2, and μ, and set them equal to zero:

$$\frac{\partial \mathcal{L}}{\partial x} = y - 2\lambda_1 x - \lambda_2 = 0$$
$$\frac{\partial \mathcal{L}}{\partial y} = x - 2\lambda_1 y + \mu = 0$$
$$\frac{\partial \mathcal{L}}{\partial \lambda_1} = 1 - x^2 - y^2 = 0$$
$$\frac{\partial \mathcal{L}}{\partial \lambda_2} = -x = 0$$
$$\frac{\partial \mathcal{L}}{\partial \mu} = y - 1 = 0$$

Now, we solve the system of equations derived from the first-order conditions and the KKT conditions. From $\partial\mathcal{L}/\partial\lambda_2 = -x = 0$, we get $x = 0$. From $\partial\mathcal{L}/\partial\mu = y - 1 = 0$, we get $y = 1$.

Substituting $x = 0$ and $y = 1$ into $\partial\mathcal{L}/\partial y = x - 2\lambda_1 y + \mu = 0$, we get:

$$0 - 2\lambda_1 \cdot 1 + \mu = 0 \quad \Rightarrow \quad \mu = 2\lambda_1,$$

Substituting $x = 0$ and $y = 1$ into $\partial\mathcal{L}/\partial x = y - 2\lambda_1 x - \lambda_2 = 0$, we get:

$$1 - 0 - \lambda_2 = 0 \quad \Rightarrow \quad \lambda_2 = 1.$$

Now we check the complementary slackness conditions: Since $x = 0$, the constraint $x \geq 0$ is active, so λ_2 can be positive (and we found $\lambda_2 = 1$). Since $x^2 + y^2 = 1$, the constraint $x^2 + y^2 \leq 1$ is active, so λ_1 can also be positive. Thus, the KKT conditions are satisfied, and the optimal solution is $x = 0$, $y = 1$.