Lagrange Multipliers and Duality

Michael

School of Mathematics and Statistics, UCD

School of Economics, University of Nottingham

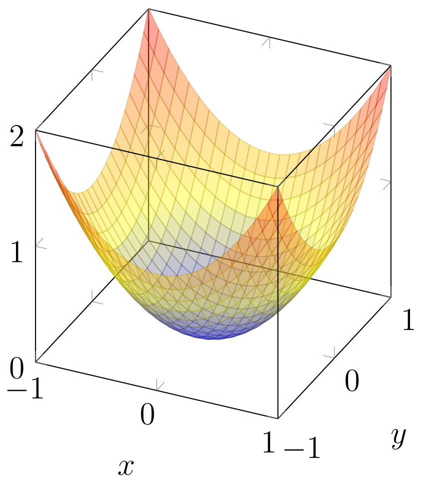
1 A Simple Example

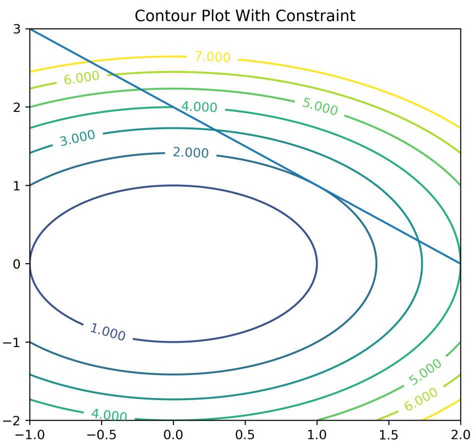
Suppose we want to minimize the function:

$$f(x, y) = x^2 + y^2$$

subject to the constraint

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





We set the Lagrangian L:

$$L = x^2 + y^2 + p(x + y - 2)$$

To solve this problem, we take the first derivative

$$\begin{aligned} 2x &= -p \\ 2y &= -p \\ x + y &= 2 \end{aligned}$$

Then the unique solution for this problem is $x = 1, y = 1, p = -2$.

The Lagrangian method could also be used to solve the maximisation or minimisation problem with multiple constraints. For example, if we want to minimize:

$$x^2 + y^2 + z^2$$

subject to

$$\begin{aligned} x + y - 2 &= 0 \\ x + z - 2 &= 0 \end{aligned}$$

Again, we set this Lagrangian L:

$$L(x, y, z, p, q) = x^2 + y^2 + z^2 + p(x + y - 2) + q(x + z - 2)$$

Solve this problem, we can get the following unique solution:

$$p = q = -\frac{4}{3}, x = \frac{4}{3}, y = \frac{2}{3}, z = \frac{2}{3}$$

There is a very important theory related to the Lagrange multiplier method. We will only state it without giving proof: *THe solution, if it exists, is always at a saddle point of the Lagrangian: no change in the original variables can decrease the Lagrangian, while no change in the multipliers can increase it.*

2 Lagrangians as Games

Because the constrained optimum always occurs at a saddle point of the Lagrangian, we can view a constrained optimization problem as a game between two players: one player controls the original variables and tries to minimize the Lagrangian, while the other controls the multipliers and tries to maximize the Lagrangian[1].

If the constrained optimization problem is well-posed (that is, has a finite and achievable minimum), the resulting game has a finite value (which is equal to the value of the Lagrangian at its saddle point).

On the other hand, if the constraints are unsatisfiable, the player who controls the Lagrange multipliers can win (i.e., force the value to $+\infty$), while if the objective function has no finite lower bound within the constraint region, the player who controls the original variables can win (i.e., force the value to $-\infty$).

We will not consider the case where the problem has a finite but unachievable infimum (e.g., minimize $\exp(x)$ over the real line).

3 Inequality Constraints

What if we want to minimize $x^2 + y^2$ subject to $x + y - 2 \geq 0$? We can use the same Lagrangian as before:

$$L = x^2 + y^2 + p(x + y - 2)$$

[1] The intuition is quite straightforward: if both equation move at the same direction, then you would never have the minimum or maximum value.

but with the additional restriction that $p \leq 0$.

Now, as long as $x + y - 2 \geq 0$, the player who controls p can't do anything: making p more negative is disadvantageous, since it decreases the Lagrangian, while making p more positive is not allowed. Unfortunately, when we add inequality constraints, the simple condition $\nabla L = 0$ is neither necessary nor sufficient to guarantee a solution to a constrained optimization problem.

4 The Lagrange Multiplier Theorem

We can now state the Lagrange multiplier theorem in its most general form, which tells how to minimize a function over an arbitrary convex polytope $Nx + c \leq 0$.

Theorem 4.1. (Lagrange Multiplier Theorem): The problem of finding x to

$$\begin{aligned} &\min f(x) \\ \text{subject to } \; &Nx + c \leq 0 \end{aligned}$$

is equivalent to the problem of finding

$$\arg\min_x \max_{\lambda \in \mathbb{R}^+} L(x, \lambda)$$

where the Lagrangian L is defined as

$$L(x, p) = f(x) + \lambda^T (Nx + c)$$

for a vector of Lagrange multipliers λ.

The Lagrange multiplier theorem uses properties of convex cones and duality to transform our original problem to a problem which mentions only the very simple cone $\mathbb{R}^+$ (it is simple because it is axis-parallel, and it is self-dual).

Comment: If you have been studied the Minimax theorem by John von Neumann, then it would be much easier to understand this problem. Again, think of the Lagrange method as a game.

Reference

This set of notes are based on the lecture notes by *Geoff Gordon*: `http://www.cs.cmu.edu/~ggordon/lp.pdf`