Lagrange Multipliers

(Notes)

by Ashwin Nanjappa

Web: http://www.comp.nus.edu.sg/~ashwinna/

November 1, 2006

1 Concepts

Lagrange is pronounced as *Lagrounge*.

Lagrange Multipliers are a mathematical tool for constrained optimization of differentiable functions.

An *optimization problem* aims to maximize or minimize a given function.

A *constrained optimization problem* is a kind of optimization problem in which the solution has to satisfy the *constraints* imposed on the problem to be acceptable.

Extreme points of a function f are points where the gradient $\nabla f = 0$ or equivalently, each of the partial derivatives is 0.

Feasible region consists of the points of the function which satisfies the constraints that have been imposed upon the function.

Equality constraints restrict the feasible region to points lying on some surface inside R^n where $f(x_1, ..., x_n)\colon R^n \rightarrow R$.

2 The Lagrangian

Given a function,

$$f(x_1, ..., x_n)\colon R^n \rightarrow R \tag{1}$$

A *constraint* upon the function $f(x_1, ..., x_n)$ is defined by another function $g(x_1, ..., x_n)$ and we are interested in the points x where:

$$g(x) = 0 \tag{2}$$

In the graphical representation of a function f, *contours* or *level curves* represent the points which have the same value of f.

***

At an extreme point p, the gradient of f will be parallel to the gradient of g, i.e., the 2 gradient vectors are either equal or multiples of each other:

$$\nabla f(p) = \lambda \nabla g(p) \tag{3}$$

This multiplier denoted by λ is called the *Lagrange Multiplier*.

(1), (2) and (3) can be combined and compactly represented as:

$$\Lambda(x, \lambda) = f(x) - \lambda g(x) \tag{4}$$

and

$$\nabla\Lambda(x,\lambda)=0 \tag{5}$$

The technique of Lagrange Multipliers involves representing the constrained optimization problem in the form of (4) and (5) and solving it.

3 An Example

Given,

$$f(x,y)=2-x^2-2y^2 \tag{6}$$

$$g(x,y)=x^2+y^2-1=0 \tag{7}$$

Consider the function $f(x,y)$ (which describes a paraboloid) and the constraint $g(x,y)$ (an unit circle). Find the maximum and minimum of $f(x,y)$ under the constraint $g(x,y)$.

3.1 Using Substitution

Lets first solve the problem using the old method of substitution.

Solving (7) we get,

$$x^2=1-y^2$$

Substituting this in (6) we get,

$$f(x,y)=1-y^2$$

From the above equation, we can deduce that $f(x,y)$ has maximum at $y=0$ which results in $f(x,y)=1$ and $x=\pm1$.

Similarly, we can deduce that $f(x,y)$ has minimum at $y=\pm1$ which results in $f(x,y)=0$ and $x=0$.

3.2 Using Lagrange Multipliers

Writing down (6) and (7) in the form of (4) and (5), we get,

$$\Lambda(x,y,\lambda)=2-x^2-2y^2-\lambda(x^2+y^2-1)$$

$$\nabla\Lambda(x,y,\lambda)=0 \tag{8}$$

Solving (8) means,

$$\frac{\partial}{\partial x}\Lambda(x,y,\lambda)=-2x-2\lambda x=0 \tag{9}$$

$$\frac{\partial}{\partial y}\Lambda(x,y,\lambda)=-4y-2\lambda y=0 \tag{10}$$

$$\frac{\partial}{\partial\lambda}\Lambda(x,y,\lambda)=-x^2-y^2+1=0 \tag{11}$$

We now have 3 equations and 3 unknowns.

Solving (9), we get $\lambda = -1$. Using this in (10), we get $y = 0$. Using that result in (11), we get $x = \pm 1$. Using these results in (6), we get $f(x, y) = 1$. We've got the maximum.

Solving (10), we get $\lambda = -2$. Using this in (9), we get $x = 0$. Using that result in (11), we get $y = \pm 1$. Using these results in (6), we get $f(x, y) = 0$. We've got the minimum.

4 Multiple Constraints

For n number of constraints, (3) and (4) are expressed as:

$$\nabla f(x) = \sum_i \lambda_i \nabla g_i(x) \tag{12}$$

and

$$\Lambda(x, \lambda) = f(x) - \sum_i \lambda_i g_i(x) \tag{13}$$

***

The *Karush-Kuhn-Tucker conditions* are a generalization of the method of Lagrange Multipliers.

5 References

1. *Lagrange Multipliers Without Permanent Scarring* by Dan Klein

   http://www.cs.berkeley.edu/~klein/papers/lagrange-multipliers.pdf

   I primarily used this document. Though it has typos all over the place, Dan Klein has given an excellent lucid and graphical introduction to this concept.

2. *Constraints And Lagrange Multipliers* (§13.7) from the book *Calculus* by Gilbert Strang

   http://ocw.mit.edu/ans7870/resources/Strang/strangtext.htm

□