Lecture 12 - Monday April 27th

jacques@ucsd.edu

Key words: Several constraints, Jacobian, Implicit function theorem, Inverse function theorem
Know how to do Lagrange multipliers with several constraints, computing Jacobian determinants, statement of implicit function theorems

12.1 Several Constraints

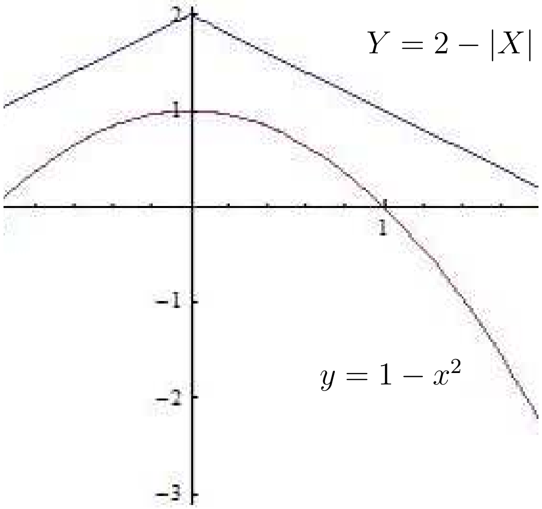
The method of Lagrange multipliers works for optimizing functions under several constraints. If we want to find the extreme points of a differentiable function $f : U \to \mathbb{R}$ where U is a closed set subject to constraints $g_1(x) = g_2(x) = \cdots = g_r(x) = 0$, we introduce multipliers $\lambda_1, \lambda_2, \dots, \lambda_r$ and find the critical points of

$$\phi(x, \lambda) = f(x) + \lambda_1 g_1(x) + \cdots + \lambda_r g_r(x)$$

where $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_r)$.

Example 1. Find the closest distance between the parabola $y = 1 - x^2$ and the parabola $y = 2 - |x|$ for $0 \leq x \leq 2$.

Solution. Let (x, y) be a point on the curve $y = 1 - x^2$ and let (X, Y) be a point on the curve $Y = 2 - |X|$. We want to minimize $(x - X)^2 + (y - Y)^2$ – the square of the distance between the two points. We can assume that all the variables are positive, by symmetry (see the figure).



The constraints are $g_1(x,y) = y - 1 + x^2 = 0$ and $g_2(X,Y) = Y - 2 + X = 0$. By the method of Lagrange multipliers, we set $\nabla\phi = 0$ where λ and μ are the Lagrange multipliers and

$$\phi(x,y,X,Y,\lambda,\mu) = (x-X)^2 + (y-Y)^2 + \lambda(y-1+x^2) + \mu(Y-2+X).$$

Then from $\nabla\phi = 0$ we get

$$\begin{aligned} 2(x-X)+2\lambda x &= 0 \\ 2(y-Y)+\lambda &= 0 \\ -2(x-X)+\mu &= 0 \\ -2(y-Y)+\mu &= 0 \\ y &= 1-x^2 \\ Y &= 2-X. \end{aligned}$$

If $\lambda = 0$ or $\mu = 0$, then we get $x = X$ and $y = Y$. The last two equations are then $Y = 1 - X^2$ and $Y = 2 - X$ and this is a contradiction – there is no point on both of these curves. Therefore neither λ nor μ is zero. Adding the second and fourth equations gives $\lambda = -\mu$. Putting this in the first and third equations, we get

$$2(x-X)+2\lambda x = 0 \quad \text{and} \quad -2(x-X)-\lambda = 0$$

which imply $x = \frac{1}{2}$. This means $y = \frac{3}{4}$ from the fifth equation. Putting these values into the second and third equations, we get

$$X = \frac{\lambda+1}{2} \quad \text{and} \quad Y = \frac{2\lambda+3}{4}.$$

In the last equation we get

$$\frac{2\lambda+3}{4} = 2 - \frac{\lambda+1}{2} \Rightarrow \lambda = \frac{3}{4} \Rightarrow X = \frac{7}{8} \text{ and } Y = \frac{9}{8}.$$

Therefore the two points which are closest together are $(\frac{1}{2}, \frac{3}{4})$ and $(\frac{7}{8}, \frac{9}{8})$. Another pair of points is $(-\frac{1}{2}, \frac{3}{4})$ and $(-\frac{7}{8}, \frac{9}{8})$. For both pairs the distance is $3/4\sqrt{2}$.

12.2 Implicit function theorems

The soundness of the method of multipliers in locating extrema of a function f is largely based on the implicit function theorem. The idea of the implicit function theorem is summarised as follows: given an equation $f(x) = 0$ for values of x in an open ball containing a point $a \in \mathbb{R}^n$, under what conditions does the equation give a functional relationship $x_n = g(x_1, x_2, \ldots, x_{n-1})$ and what can be said about the function g? For example, we know that if $f(x,y,z) = x^2 + y^2 + z^2 - 1$ for $z \geq 0$, then $f(x,y,z) = 0$

on the region $z \geq 0$ defines a hemisphere of radius one. Now a standard representation of a surface is as a function $z = g(x, y)$, and indeed here we can solve for z to get the hemisphere $z = \sqrt{1 - x^2 - y^2}$. So here $g(x, y) = \sqrt{1 - x^2 - y^2}$, and g is differentiable on the open ball $x^2 + y^2 < 1$. In general, it is impossible to give an explicit function $g(x, y)$, even in this very simple case with only three variables. Nevertheless, the implicit function theorem gives a general condition whereby g exists and is differentiable on a certain open ball:

Implicit Function Theorem

> Let $f(x, y) : \mathbb{R}^{n+m} \to \mathbb{R}^m$ where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ and let U be an open ball containing a point $a \in \mathbb{R}^n$. Suppose that $\det(\nabla f) \neq 0$ on U when f is treated only as a function of y. Then there is a differentiable function $g : \mathbb{R}^n \to \mathbb{R}^m$ such that $f(x, g(x)) = 0$ in some open ball containing a.

The matrix ∇f is often also called the Jacobian matrix of f – recall the ijth entry is $\partial f_i / \partial y_j$ for $1 \leq i, j \leq m$ if $f = (f_1, f_2, \ldots, f_m)$. A matrix whose determinant is non-zero is also called non-singular – this refers to the matrix being invertible in the linear algebraic sense. In the form of equations, the implicit function theorem says that if we are trying to solve

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) &= 0 \\ f_2(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) &= 0 \\ &\vdots \\ f_m(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) &= 0 \end{aligned}$$

for each y_i in terms of $x_1, x_2, \ldots, x_n$ near a point $a \in \mathbb{R}^n$, then all we have to do is check that the Jacobian matrix of the function $f = (f_1, f_2, \ldots, f_m)$ is non-singular near a when f is treated as a function of $(y_1, y_2, \ldots, y_m)$.

The second theorem we state is a special case of the implicit function theorem and is called the inverse function theorem. Consider a vector valued function $f : \mathbb{R}^n \to \mathbb{R}^n$. An inverse for f on a set $U \subset \mathbb{R}^n$ is a function $g : \mathbb{R}^n \to \mathbb{R}^n$ such that $f(g(x)) = x$ for all $x \in U$. The function g is called an inverse of f on U. This entails finding functions $y_1(x), y_2(x), \ldots, y_n(x)$ such that

$$f(y_1, y_2, \ldots, y_n) = x$$

or in other words, if $f = (f_1, f_2, \ldots, f_n)$, we have to solve the n simultaneous functional equations

$$f_1(y_1, y_2, \ldots, y_n) = x_1 \quad f_2(y_1, y_2, \ldots, y_n) = x_2 \quad \cdots \quad f_n(y_1, y_2, \ldots, y_n) = x_n$$

for y_i in terms of $x_1, x_2, \ldots, x_n$. This is exactly the setup of the implicit function theorem. Note that in the case that all the functions are linear functions, the theorem reduces to linear algebra: we require the coefficient matrix to have a non-zero determinant if it is to be invertible, and this is enough to give a solution to the system of equations (in fact this linear algebra connection is not at all a coincidence). The inverse function theorem may be stated as follows:

Inverse Function Theorem

> Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be defined on an open ball around a point $a \in \mathbb{R}^n$ and suppose f has continuous partial derivatives on this ball. If ∇f is non-singular on this ball, then there exists a unique function $g : \mathbb{R}^n \to \mathbb{R}^n$ such that $f(g(x)) = x$ in some open ball containing a.