Lecture 13 - Wednesday April 29th

jacques@ucsd.edu

Key words: Systems of equations, Implicit differentiation
Know how to do implicit differentiation, how to use implicit and inverse function theorems

13.1 Examples

Recall the two main theorems of the last section:

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.

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.

We now show how to use these theorems. In the first example, we can explicitly find the inverse of the function, without using the implicit function theorem.

Example 1. The function $f(x, y) = (x + y, x - y)$ has an inverse on $U = \mathbb{R}$: we have to solve $u + v = x$ and $u - v = y$. Clearly this means $u = \frac{x+y}{2}$ and $v = \frac{x-y}{2}$. In other words, the function $g(x, y) = ((x + y)/2, (x - y)/2)$ is an inverse of $f(x, y)$. What this means is that the equations $u + v = x$ and $u - v = y$ are solvable for x and y in terms of u and v.

However for most functions $f(x, y)$ there is no hope of computing an inverse. The inverse function theorem nevertheless tells us when an inverse exists, even though it may be impossible to find explicitly. Let's use the inverse function theorem to show that the function f in the last example has an inverse.

Example 2. In the last example the Jacobian matrix is

$$\nabla f(x,y) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

The determinant is -2, so this function f is invertible (i.e. the equations $u+v=x$ and $u-v=y$ are solvable for x and y in terms of u and v. Next we consider $f(x,y) = (x^2+y^2, xy)$. Then

$$\nabla f(x,y) = \begin{pmatrix} 2x & 2y \\ y & x \end{pmatrix}.$$

This is non-singular everywhere except at $(0,0)$. Therefore f is invertible on $\mathbb{R}^2\backslash\{(0,0)\}$. We won't actually compute the inverse here.

The last two examples were examples of the inverse function theorem. Now we do an example of the implicit function theorem.

Example 3. A surface in $\mathbb{R}^4$ can be parametrized by two equations $u = u(s,t)$ and $v = v(s,t)$. Suppose that $u^2+v^2+s^2+t^2 = 1$ and $uv+st = 1$. Determine for what values of s and t these two equations define a surface given in the form $u = u(s,t)$ and $v = v(s,t)$, and determine if the surface is differentiable on $\mathbb{R}^2$ – this means u and v are both differentiable functions of s and t. The implicit function theorem says that for this to work we require

$$\det\begin{pmatrix} 2u & 2v \\ v & u \end{pmatrix} \neq 0.$$

The determinant is $2u^2 - 2v^2$. So as long as $|u| \neq |v|$ we get a smooth surface. To find the values of s and t for which $|u| = |v|$, we note that $u = v$ or $u = -v$ in this case, and so

$$2u^2 + s^2 + t^2 = 1 \quad \text{and } u^2 + st = 1.$$

If we subtract twice the second equation from the first, we get $s^2+t^2-2st = -1$ and so $(s-t)^2 = -1$ which never happens. In the second case, $u = -v$, we get

$$2u^2 + s^2 + t^2 = 1 \quad \text{and } -u^2 + st = 1.$$

Add twice the second equation to the first gives $s^2+t^2+2st = 3$ and so $(s+t)^2 = 3$. Therefore we are in trouble on the regions $\{(s,t) : |s+t| = \sqrt{3}\}$. So the surface is not differentiable on $\mathbb{R}^2$.

Example 4. Let (x,y) denote a point in Cartesian co-ordinates and let (r,θ) denote a point in polar co-ordinates. Then recall that $x = r\cos\theta$ and $y = r\sin\theta$. For what values of x and y can we solve these two equations for r and θ in terms of x and y? Let $x = f_1(r,\theta) = r\cos\theta$ and $y = f_2(r,\theta) = r\sin\theta$. Then the Jacobian of $f = (f_1, f_2)$ is

$$\nabla f = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

The determinant is

$$\det(\nabla f) = r(\cos^2\theta + \sin^2\theta) = r.$$

Provided $r > 0$, the determinant is non-zero. Therefore for $r > 0$, we can solve the equations. For $r = 0$, we get the origin $(x, y) = (0, 0)$. So for all $(x, y) \neq (0, 0)$, we can solve the two equations. This we knew since in fact $r = \sqrt{x^2 + y^2}$ and $\theta = \arcsin(y/\sqrt{x^2 + y^2})$ for $(x, y) \neq (0, 0)$.

Example 5. Determine where the system $y + x + uv = 0$ and $uxy + v = 0$ is solvable for u, v in terms of x, y near the point $(x, y) = (1, 1)$. Actually we can solve directly for u and v in terms of x and y to get

$$u^2xy = x + y \qquad \text{and} \qquad v = -uxy.$$

If $(x + y)/xy \geq 0$ then we get

$$u = \sqrt{(x + y)/xy} \quad \text{ and } v = \sqrt{xy(x + y)}.$$

So near $(x, y) = (1, 1)$, we can solve for u, v in terms of x and y.

Now let's do it using the inverse function theorem. In this case the Jacobian of derivatives with respect to u and v is

$$\begin{pmatrix} v & u \\ xy & 1 \end{pmatrix}$$

At $(x, y) = (1, 1)$ we get $(u, v) = (-\sqrt{2}, \sqrt{2})$. Therefore at $(1, 1)$ the determinant is $v - uxy = \sqrt{2} + \sqrt{2} = 2\sqrt{2}$ and this is non-zero. So we can solve the system near $(1, 1)$.

13.2 Differentiating Implicit Functions

The equation $f(x, y) = 0$ in general defines a curve in the plane provided f is differentiable and has a non-zero Jacobian. This curve has tangent lines at all points in this case, whose slopes are given by dy/dx. We now look more closely at how to find these slopes. In the special case $f(x, y) = 0$, we assume that we can represent the curve as $y = y(x)$ in which case the chain rule gives:

$$\frac{d}{dx} f(x, y(x)) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} = 0.$$

It follows that

$$\frac{dy}{dx} = -\frac{\partial f}{\partial x} / \frac{\partial f}{\partial y}.$$

Example 1. Find the slope of the tangent line to the circle $x^2 + y^2 = 1$ at the point $(1/\sqrt{2}, 1/\sqrt{2})$. In this case we already know the slope is -1. Let's check it using implicit differentiation:

$$\frac{dy}{dx} = -\frac{\partial f}{\partial x} / \frac{\partial f}{\partial y} = -x/y.$$

If $x = y$ then this is -1 as required.

Suppose that we are given a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, in which case the implicit function theorem gives conditions whereby we can solve for any variable as a differentiable function of the others. Then we can find the partial derivatives of this variable with respect to the others using the chain rule: suppose $f(x_1, x_2, \ldots, x_{n-1}, z) = 0$ defines a differentiable function $z = z(x_1, x_2, \ldots, x_n)$. Then for $i \leq n-1$, we have

$$\frac{\partial f}{\partial x_i} = \sum_{j=1}^{n-1} \frac{\partial f}{\partial x_j} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x_i}$$

and we can now solve for $\partial z/\partial x_i$.

Example 2. Consider the surface described by $x^2 + 2xz + z^2 - 2yz - 1 = 0$ where $z \geq 0$. If $f(x, y, z) = x^2 + 2xz + z^2 + 2yz - 1$, then $f(x, y, z) = 0$ describes a differentiable function $z = z(x, y)$. Now suppose we want $\partial z/\partial x$. We can find this by differentiating $f(x, y, z) = 0$ with respect to x: treating z as a function of x and y we have

$$f(x, y, z(x, y)) = x^2 + 2xz(x, y) + z(x, y)^2 + 2yz(x, y) - 1 = 0.$$

Differentiating this we get

$$2x + 2z + 2x\frac{\partial z}{\partial x} + 2z\frac{\partial z}{\partial x} - 2y\frac{\partial z}{\partial x} = 0.$$

It follows that

$$\frac{\partial z}{\partial x} = -\frac{x+z}{x+z+y} = -\frac{1}{1 + y/(x+z)}.$$

If we wanted the derivative $\partial z/\partial x$ at $(1, 1)$, we would observe that $f(1, 1, z) = z^2 + 4 = 0$ so $z = 0$ or $z = -4$. Since $z \geq 0$, we only consider $z = 0$. In this case

$$\frac{\partial z}{\partial x}(1, 1) = -\frac{1}{1 + 1/(1+0)} = -\frac{1}{2}.$$

Example 3. Let $f(x, z) : \mathbb{R}^{n+1} \to \mathbb{R}$ be a differentiable function where $x \in \mathbb{R}^n$ and such that $f(x, z) = 0$. Determine $\nabla z(x)$. Well if we differentiate implicitly we get for each $i \leq n$:

$$0 = \frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x_i}.$$

Therefore

$$\frac{\partial z}{\partial x_i} = -\frac{\partial f/\partial x_i}{\partial f/\partial z}.$$

It follows that

$$\nabla z = -\frac{\nabla f}{\partial f/\partial z}.$$

13.3 Justifying Lagrange Multipliers

Here we sketch a proof of why the method of Lagrange multipliers works. Suppose $f, g : \mathbb{R}^2 \to \mathbb{R}$ are differentiable functions and suppose we want to show that the extremes of f on the region defined by $g(x, y) = 0$ have $\nabla f + \lambda g = 0$ for some λ. This is to say that ∇f and ∇g are parallel vectors at the extreme points of f. Geometrically, it means that if λ is non-zero and (x_0, y_0) is an extreme point of f, then the curves represented by $f(x, y) = f(x_0, y_0)$ and $g(x, y) = 0$ are tangent at (x_0, y_0).

Let's prove that ∇f and ∇g are parallel at extreme points of f on the region $g(x, y) = 0$. By the implicit function theorem, $g(x) = 0$ describes a curve given parametrically by $x = x(t)$ and $y = y(t)$ where each of these functions is differentiable. By the chain rule and implicit differentiation,

$$\frac{\partial g}{\partial t} = \frac{\partial g}{\partial x}\frac{dx}{dt} + \frac{\partial g}{\partial y}\frac{dy}{dt} = 0.$$

Let (x_0, y_0) be an extreme of f, and suppose $x_0 = x(t_0)$ and $y = y(t_0)$. By the chain rule again,

$$\nabla f(t) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

From a preceding theorem, $\nabla f(t_0) = 0$ since t_0 represents an extreme point of $f(x(t), y(t))$. It follows that

$$\nabla f(t_0) \cdot (\frac{dx}{dt}, \frac{dy}{dt})\Big|_{t=t_0} = 0 \quad \text{and } \nabla g(t_0) \cdot (\frac{dx}{dt}, \frac{dy}{dt})\Big|_{t=t_0} = 0.$$

In particular, this means ∇f and ∇g are parallel vectors since they are both perpendicular to the vector $(\frac{dx}{dt}, \frac{dy}{dt})\big|_{t=t_0}$.