Lecture 7 : Monday April 13th

jacques@ucsd.edu

KEY CONCEPTS : $C^k(\mathbb{R}^n)$, $C^\infty(\mathbb{R}^n)$, POLYNOMIALS, MIXED PARTIAL DERIVATIVES
KNOW HOW TO FIND HIGHER ORDER PARTIAL DERIVATIVES

7.1 Higher partial derivatives

The partial derivatives we have studied so far have meaning in terms of the slope of surfaces in various directions. If these partial derivatives as functions themselves are differentiable, we can obtain second order partial derivatives. These have very important meaning when it comes to continuous optimization, as we shall see later. For now we concentrate on the mechanics of higher order partial derivatives.

A function $f : \mathbb{R}^n \to \mathbb{R}$ is in class C^1 if $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous function. In other words, all the partial derivatives of f exist and are continuous. If all the partial derivatives of f themselves have continuous derivatives, then we say f is in class C^2. For example, if $f : \mathbb{R}^2 \to \mathbb{R}$ then $\nabla f = (f_x, f_y)$ if it exists, and if f_x and f_y have derivatives with respect to x and y, then these derivatives are written $f_{xx}, f_{xy}, f_{yx}, f_{yy}$. Alternatively, we write these derivatives as

$$f_{xx} = \frac{\partial^2 f}{\partial x^2} \quad f_{yx} = \frac{\partial^2 f}{\partial x \partial y} \quad f_{xy} = \frac{\partial^2 f}{\partial y \partial x} \quad f_{yy} = \frac{\partial^2 f}{\partial y^2}.$$

It is very important to note that in general $f_{xy} \neq f_{yx}$. For practical purposes,

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$

so derivatives are worked out from right to left in this notation. In the notation f_{xy}, we first find f_x and then take the derivative with respect to y of f_x to get f_{xy}. So derivatives in this notation are worked out from the inside out. We refer to derivatives with respect to all different variables as mixed partial derivatives. For example f_{xx} and f_{xzx} are not mixed, but f_{xy} and f_{xyz} are.

More generally, we say a function $f : \mathbb{R}^n \to \mathbb{R}$ is in class $C^k(\mathbb{R}^n)$ if all kth order derivatives of f exist and are continuous. We say that f is k times continuously differentiable. For a function $f : \mathbb{R}^2 \to \mathbb{R}$ to be in C^3, this means that all the derivatives

$$f_{xxx} \quad f_{xxy} \quad f_{xyx} \quad f_{yxx} \quad f_{xyy} \quad f_{yxy} \quad f_{yyx} \quad f_{yyy}$$

must exist and be continuous. Finally, we say that $f \in C^\infty$ if all partial derivatives up to any order exist and are continuous.

7.2 Polynomials

An especially nice function in C^∞ which shall appear again and again is the polynomial. A multivariate monomial or monomial is any function $f : \mathbb{R}^n \to \mathbb{R}$ of the form $x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n}$ where $d_1, d_2, \dots, d_n$ are non-negative integers. The degree of this monomial is defined by $d_1 + d_2 + \cdots + d_n$. A polynomial is a finite sum of monomials. The degree of a polynomial is the largest of the degrees of the monomials in that polynomial.

Examples. The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2 + x + 1$ is a polynomial of degree two (which we refer to as quadratic), whose monomials are x^2, x and 1. The function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x, y) = x^2 - y^2 - xy + 2x + 3y - 1$ is also of degree two: it is a polynomial consisting of six monomials and three of the monomials have degree two. We refer to this too as a quadratic polynomial. Finally, $f : \mathbb{R}^n \to \mathbb{R}$ defined by $f(x_1, x_2, \dots, x_n) = x_1 x_2 \dots x_n + 1$ is a polynomial of degree n with two monomials.

Proposition.

> If $f : \mathbb{R}^n \to \mathbb{R}$ is a polynomial, then $f \in C^\infty$.

This is easy to prove by induction on the degree of the polynomial. Given a polynomial $f : \mathbb{R}^n \to \mathbb{R}$ of degree d, the derivatives $\partial f/\partial x_i$ are all multivariable polynomials too. Furthermore, the degree of $\partial f/\partial x_i$ is less than the degree of f, because all monomials containing a non-zero power of x_i have degree one less in $\partial f/\partial x_i$, and all monomials containing no x_i have degree zero in $\partial f/\partial x_i$. By induction, this means that $\partial f/\partial x_i \in C^\infty$ for $i = 1, 2, \dots, n$. But if the derivatives of f are in C^∞, then so is f.

Examples. We find all derivatives of the function $f(x) = f(x_1, x_2, \dots, x_n) = x_1 x_2 \dots x_n$. Suppose we take derivatives with respect to variables $x_{i_1}, x_{i_2}, \dots, x_{i_k}$ in that order. If any two of these variables are the same, we get zero. For example, $f_{x_1 x_2 x_1 x_n} = 0$. So the only non-zero derivatives are those with respect to different variables. In particular we have $k \le n$. In this case,

$$f_{x_{i_1} x_{i_2} \dots x_{i_k}} = \frac{f(x)}{x_{i_1} x_{i_2} \dots x_{i_k}}$$

since the derivative is exactly the product of all x_is other than $x_{i_1}, x_{i_2}, \dots, x_{i_k}$. Let's consider another example: $f(x) = xy^2z^3$. Then the derivatives are

$$f_x = y^2z^3 \quad f_y = 2xyz^3 \quad f_z = 3xy^2z^2 \quad f_{xy} = f_{yx} = 2yz^3 \quad f_{xz} = f_{zx} = 3y^2z^2 \quad f_{yz} = f_{zy} = 6xyz^2.$$

Notice that the mixed partial derivatives are pairwise equal.

7.3 Equality of mixed partial derivatives

In general, the number of different derivatives of order k for a function $f : \mathbb{R}^n \to \mathbb{R}$ is n^k, which is very large if k is large. It is the case, however, that many of these derivatives are equal if $f \in C^k(\mathbb{R}^2)$:

Proposition.

If $f \in C^2$ then

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Using this proposition we can show that if $f \in C^k(\mathbb{R}^n)$, then the order of partial derivatives taken up to kth order derivatives does not matter. So for example if $f \in C^3(\mathbb{R}^2)$ then $f_{xxy} = f_{xyx} = f_{yxx}$ and if $f \in C^3(\mathbb{R}^3)$ then $f_{xyz} = f_{yzx} = f_{zxy} = f_{xzy} = f_{yxz} = f_{zyx}$. This can be proved as an exercise by induction on k.

Example. Given a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ we may define the Hessian Matrix

$$H(f) = \nabla(\nabla f).$$

So the rows of $H(f)$ are given by ∇f_{x_i} if $f = f(x_1, x_2, \ldots, x_n)$. If $f \in C^2(\mathbb{R}^n)$, then $L(f)$ is a symmetric matrix, meaning that $H(f)_{ij} = H(f)_{ji}$ since

$$H(f)_{ij} = f_{x_i x_j}.$$

and the order of differentiation does not matter. The Hessian Matrix encodes all second order partial derivatives of f and is extremely useful in optimization. For example, consider the function $f(x_1, x_2, \ldots, x_n) = x_1^2 + x^2 + \cdots + x_n^2$. The Hessian Matrix is

$$\begin{pmatrix} 2x_1 & 0 & \cdots & 0 & 0 \\ 0 & 2x_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 2x_n \end{pmatrix}.$$

We will see how to tell from the Hessian Matrix whether a function is convex or concave at a point or neither.