Analysis, Part IV: Several Variable Differential Calculus

Jay Havaldar

4.1 Derivatives in Several Variables

Recall that the definition of the derivative in a single variable is:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

However, this definition does not generalize to higher dimensions. In particular, the quantity on the right hand side would be a quotient of vectors. Instead we note that we can rewrite the above as:

$$0 = \lim_{h \to 0} \frac{|f(x+h) - f(x) - f'(x)h|}{|h|}$$

We generalize the definition as follows.

Definition: Let E be a subset of $\mathbb{R}^n$ and let $f : E \to \mathbb{R}^m$ be a function. Then we say that f is **differentiable** at $x_0 \in E$ if there exists a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^n$ so that:

$$0 = \lim_{h \to 0} \frac{|f(x+h) - f(x) - Lh|}{|h|}$$

Because limits are unique, (with some work), we can show that this definition makes sense, i.e. L is unique at each point if it exists.

4.1.1 Partial and Directional Derivatives

Definition: Let $E \subset \mathbb{R}^n$ and let f be a function $f : E \to \mathbb{R}^{\gtrdot}$, with $x_0 \in \text{int}(E)$. Let v be a vector in $\mathbb{R}^n$. Then if the following limit exists, it is called the directional derivative of f in the direction of v, or $D_v f(x_0)$:

$$\lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t}$$

Note that this object is a vector in the codomain $\mathbb{R}^{\gtrdot}$, not a linear transformation as the derivative is.

One special type of directional derivative is a partial derivative.

Definition: The partial derivative $\frac{\partial f}{\partial x_j}$ or sometimes denoted f_{x_j} is the directional derivative of f in the direction of the jth basis vector in the standard basis for $\mathbb{R}^n$.

Proposition

Let $E \subset \mathbb{R}^n$ and let f be a function $f : E \to \mathbb{R}^{\gtrdot}$, with $x_0 \in \text{int}(E)$. Let v be a vector in $\mathbb{R}^n$. Then if f is differentiable at x_0, then f has a directional derivative $D_v f(x_0)$, and in particular:

$$D_v f(x_0) = f'(x_0)v$$

Thus, in particular, if f is differentiable then we can recover its rows. Note that from above we have:

$$f_{x_j}(x_0) = f'(x_0)(e_j)$$

The quantity on the right is exactly the jth column of $f'(x_0)$. Thus, if a function is differentiable, then its partial derivatives all exist and they completely describe the derivative. However, the converse is not true; there are many examples of functions which have partial derivatives but which are not differentiable. The stronger condition that guarantees the converse is as follows.

At times, $f'(x_0)$ is used to denote the linear transformation of the derivative, while $Df(x_0)$ represents the matrix of the derivative with respect to some basis; in practice, it is harmless to use the two interchangeably.

Theorem

Let $E \subset \mathbb{R}^n$ and let f be a function $f : E \to \mathbb{R}^{>}$, with $x_0 \in \text{int}(E)$. Then if the partial derivatives of f exist and are continuous in a neighborhood of x_0 contained in E, then f is differentiable at x_0, and the jth column of the derivative is exactly the jth partial derivative of f.

Alternatively, we could describe the rows of the derivative instead of the columns. Let $f = (f_1, \dots, f_m)$. then we can write:

$$f'(x_0) = \begin{bmatrix} \nabla f_1(x_0) \\ \vdots \\ \nabla f_m(x_0) \end{bmatrix}$$

4.1.2 The Chain Rule

There is also a suitable analogue for the chain rule.

Theorem (Chain Rule)

Let $E \subset \mathbb{R}^n$ and let f be a function $f : E \to F \subset \mathbb{R}^{>}$, with $x_0 \in \text{int}(E)$. Let $g : F \to \mathbb{R}^p$. Supose that f is differentiable at x_0 and g is differentiable at $f(x_0)$. Then $g \circ f : E \to \mathbb{R}^p$ is also differentiable at x_0, and:

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0)$$

4.1.3 Clairaut's Theorem

We similarly define what it means for a function to be twice differentiable.

Definition: let $E \subset \mathbb{R}^n$ be an open set, and $f : E \to \mathbb{R}^m$. Then f is **continuously differentiable** (also written C^1) if the partial derivatives all exist and are continuous. f is twice continuously differentiable if it is continuously differentiable, and the partial derivatives (also being functions from $\mathbb{R}^n$ to $\mathbb{R}^m$) are also continuously differentiable. We have $C^1 \subset C^2$.

We can indeed generalize this definition as follows. A function f is C^k if its partial derivatives exist up to order k and are continuous. The following useful fact is immensely important to calculus:

Theorem (Clairaut's Theorem)

Let $E \subset \mathbb{R}^n$ open, and let $f : E \to \mathbb{R}^m$ be a twice continuously differentiable function on E. Then for each $1 \leq i, j \leq n$

$$(f_{x_i}(x_0))_{x_j} = (f_{x_j}(x_0))_{x_i}$$

In other words, we can swap the order of differentiation.

4.2 The Inverse & Implicit Function Theorems

Recall that if a function $f : \mathbb{R} \to \mathbb{R}$ is invertible, differentiable, and $f'(x_0)$ is nonzero, then f^{-1} is differentiable at $f(x_0)$, and:

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}$$

In particular, note that we really only need that f is continuously differentiable. If $f'(x_0)$ is nonzero, then f' is either strictly positive or strictly negative. In a small enough neighborhood, by the continuity of f', f is strictly positive or strictly negative; in either case, it is invertible if we pick a small enough neighborhood around x_0 and $f(x_0)$, respectively.

The analogue for this theorem is as follows:

Theorem (Inverse Function Theorem)

Let E be an open subset of $\mathbb{R}^n$, and $f : E \to \mathbb{R}^n$ be a function which is continuously differentiable on E (i.e. its partial derivatives exist and are continuous). Suppose $x_0 \in E$ and $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^n$ is invertible.

Then, there exists a neighborhood of x_0 in E, and a neighborhood $V \in \mathbb{R}^n$ of $f(x_0)$ such that f is a bijection from U to V (i.e. there exists an inverse $f^{-1} : V \to U$).

Finally, f^{-1} is differentiable at $f(x_0)$, and $(f^{-1})'(f(x_0)) = (f'(x_0))^{-1}$.

There is also an analogue for implicit differentiation which follows from the inverse function theorem.

Theorem (Implicit Function Theorem)

Let $E \subset \mathbb{R}^n$ be an open set, and $f : E \to \mathbb{R}$ be continuously differentiable. Let $y = (y_1, \ldots, y_n)$ be a point in E such that $f(y) = 0$ and $f_{x_n}(y) \neq 0$.

Then, there exists an open subset U of $\mathbb{R}^{n-1}$ containing $(y_1, \ldots, y_{n-1})$, an open subset V of E containing y, and a function $g : U \to \mathbb{R}$, such that:

$$g(y_1, \ldots, y_{n-1}) = y_n$$

Furthermore, the zero set of $f(x)$ is a graph of a function over U. Finally, g is differentiable with derivative:

$$g_{x_j}(y_1, \ldots, y_{n-1}) = -\frac{f_{x_j}(y)}{f_{x_n}(y)}$$