Multi-dimensional Calculus

ZhenTing Liu

Disclaimer: this note assumes sufficient understanding of linear algebra

Contents

Chapter 1

Partial derivative and directional derivative

1.0 Partial derivative

we define partial derivative $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ as follows

Definition 1.1 (partial derivative).

$$\frac{\partial f}{\partial x}\begin{pmatrix} a \\ b \end{pmatrix} = \lim_{h \to 0} \frac{f\begin{pmatrix} a+h \\ b \end{pmatrix} - f\begin{pmatrix} a \\ b \end{pmatrix}}{h}$$

$$\frac{\partial f}{\partial y}\begin{pmatrix} a \\ b \end{pmatrix} = \lim_{k \to 0} \frac{f\begin{pmatrix} a \\ b+k \end{pmatrix} - f\begin{pmatrix} a \\ b \end{pmatrix}}{k}$$

Example. Let $f(x, y) = x^3y^5 + e^{xy}\sin(2x + 3y)$ then

$$\frac{\partial f}{\partial x} = 3x^2y^5 + e^{xy}(2\cos(2x + 3y)) + ye^{xy}(\sin(2x + 3y))$$

◇

The partial derivative of f measures the rate of change of f in the direction of coordinate axis, i.e., in the direction of the standard basis vector $e_1, \dots e_n$. Given any non-zero vector v, it is natural to consider the rate of change of f in the direction of v.

Definition 1.2. Let $U \subset \mathcal{R}^n$ be open and $a \in U$. We define the directional derivative of $f : U \mapsto \mathcal{R}^m$ at a in the direction v to be

$$D_v f(a) = \lim_{t \to 0} \frac{f(a + tv) - f(a)}{t}$$

provided this limit exists.

Example. Let $f : \mathcal{R}^2 \mapsto \mathcal{R}$ by

$$f(x, y) = \frac{|x|y}{\sqrt{x^2 + y^2}}, \quad x \neq 0 \wedge f(0) = 0$$

Then the directional derivative of f at 0 in the direction of unit vector $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$

$$D_v f(0) = \lim_{t \to 0} \frac{f(tv) - f(a)}{t}$$

then it is trivial that this expression yields $|v_1|v_2$ $\diamond$

Example. Let $f = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 y + e^{3x+y-z}$, and let $a = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ and $v = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$.

What is the directional derivative $D_v f(a)$?

We define $\phi : \mathcal{R} \mapsto \mathcal{R}$ by $\phi(t) = f(a + tv)$. Note that

$$\phi'(0) = \lim_{t \to 0} \frac{\phi(t) - \phi(0)}{t} = \lim_{t \to 0} \frac{f(a + tv) - f(a)}{t} = D_v f(a)$$

so we just calculate ϕ and compute its derivative at 0

$$\phi(t) = f \begin{pmatrix} 1 + 2t \\ -1 + 3t \\ 2 - t \end{pmatrix} = (1 + 2t)^2(-1 + 3t) + e^{10t}$$

$$\phi'(t) = 4(1 + 2t)(-1 + 3t) + 3(1 + 2t)^2 + 10e^{10t}$$

from which we conclude $D_v(a) = \phi'(0) = 9$ $\diamond$

1.1 Differentiability

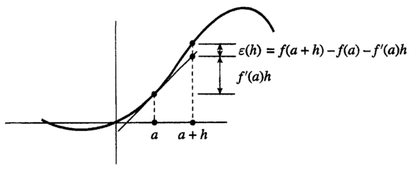
Recall that the derivative is defined to be

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

alternatively, if it exists, it is the unique number m with the property that

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - mh}{h} = 0.$$

That is, the tangent line—the line passing through $\begin{bmatrix} a \\ f(a) \end{bmatrix}$ with slope $m = f'(a)$—is the best (affine) linear approximation to the graph of f at a, in the sense that the error goes to 0 faster than h as $h \to 0$. (See Figure 2.1.) Generalizing the latter notion, we make this



Definition 1.3. Let $U \subset \mathcal{R}^n$ be open and $a \in U$. A function $f : U \mapsto \mathcal{R}^m$ is **differentiable** at a if there is a linear map $Df(a) : \mathcal{R}^n \mapsto R^m$ such that

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - Df(a)h}{h} = 0$$

This says that $Df(a)$ is the best linear approximation to the function $f - f(a)$ at a, in the sense that $f(a+h) - f(a) - Df(a)h$ converges faster than h.

Remark. The derivative $Df(a)$, if exists, must be unique, the proof is skipped as it is trivial.

Proposition 1.4. if $f : \mathcal{R}^n \mapsto \mathcal{R}^m$ is **differentiable** at a, then the partial derivatives $\frac{\partial f_i}{\partial x_j}(a)$ exist and

$$\Big(Df(a)\Big) = \Big(\tfrac{\partial f_i}{\partial x_j}(a)\Big)$$

The latter matrix is often called the Jacobian matrix of f at a.

Proof. since we assume f is differentiable at a, we know that there is a linear map $Df(a)$ with property that

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

, for any $j = 1, ..., n$, we consider $h = te_j$ and let $t \mapsto 0$. Then we have

$$0 = \lim_{t\to 0} \frac{f(a+te_j) - f(a) - Df(a)(te_j)}{|t|} = 0$$

Considering separately the cases $t > 0$ and $t < 0$, we find that

$$\lim_{t\to 0^+} \frac{f(a+te_j) - f(a) - Df(a)(te_j)}{|t|} = \lim_{t\to 0^+} \frac{f(a+te_j) - f(a)}{t} - Df(a)(e_j)$$

$$\lim_{t\to 0^-} \frac{f(a+te_j) - f(a) - Df(a)(te_j)}{|t|} = (-1)\lim_{t\to 0^-} \frac{f(a+te_j) - f(a)}{t} - Df(a)(e_j)$$

so $Df(a)(e_j) = \lim\limits_{t\to 0} \frac{f(a+te_j)-f(a)}{t} = \frac{\partial f}{\partial x_j}(a)$ □

Proposition 1.5. if $f : \mathcal{R}^n \mapsto \mathcal{R}^m$ is differentiable at a, then f is continuous at a

Proof. Trivial, assume f is differentiable at a and deduce $\lim\limits_{x\to a} f(x) = f(a)$ □

s

1.2 Differentiation rules

Theorem 1.6 (Chain rule)**.** Suppose $g : \mathcal{R}^n \mapsto \mathcal{R}^m$ and $f : \mathcal{R}^m \mapsto \mathcal{R}^l$, g is differentiable at a and f is differentiable at $g(a)$. Then $f \circ g$ is differentiable at a and

$$D(f \circ g)(a) = Df(g(a))Dg(a)$$

Example. Let

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}, \quad g\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u\cos v \\ u\sin v \end{pmatrix}$$

Since

$$Df\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}$$

and

$$Dg\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos v & -u\sin v \\ \sin v & u\cos v \end{pmatrix}$$

we have

$$D(f\circ g)\begin{pmatrix} u \\ v \end{pmatrix} = Df\left(g\begin{pmatrix} u \\ v \end{pmatrix}\right) Dg\begin{pmatrix} u \\ v \end{pmatrix} \tag{1.1}$$

$$= \begin{pmatrix} 2u\cos v & -2u\sin v \\ 2u\sin v & 2u\cos v \end{pmatrix} \begin{pmatrix} \cos v & -u\sin v \\ \sin v & u\cos v \end{pmatrix} = 2\begin{pmatrix} u\cos 2v & -u^2\sin 2v \\ u\sin 2v & u^2 \end{pmatrix} \tag{1.2}$$

$\diamond$

1.3 Gradient

Definition 1.7. Let $f : \mathcal{R}^n \mapsto \mathcal{R}$ be differentiable at a. We define the gradient of f at a to be the vector

$$\nabla f(a) = (Df(a))^T = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \frac{\partial f}{\partial x_2}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{pmatrix}$$

Now we can interpret the directional derivative of a differentiable function as a dot product

$$D_v f(a) = Df(a)v = \nabla f(a) \cdot v$$

If we consider the directional derivative in the direction of various unit vector v, we infer from the Cauchy-Schwarz inequality that

$$D_v f(a) \leq ||\nabla f(a)||$$

With the equality holding iff $\nabla f(a)$ is a positive scalar multiple of v. As consequence, we have

Proposition 1.8. Suppose f is differentiable at a, then $\nabla f(a)$ points in the direction in which f increases at the greatest rate, and $||\nabla f(a)||$ is the greatest possible rate of change

Definition 1.9 (Subspace topology)**.** If $Y \subset X$ then (Y, σ_Y) is a topological space, where

$$\sigma_Y = \{U \cap Y \mid U \in \sigma\}.$$

We call σ_Y the subspace topology.

Example. Endowing $\mathbb{R}^2$ with the Euclidean topology, the subspace topology on $\mathbb{R} \times \{0\} \subset \mathbb{R}^2$ is also the Euclidean topology. $\diamond$

Definition 1.10 (Quotient topology)**.** Let $\sim$ be an equivalence relation on X. Consider $\pi : X \to X/\sim$. Then $X/\sim$ is a topological space, where the open sets are by definition sets such that $\pi^{-1}(U)$ is open in X.

Definition 1.11 (Continuous functions). A function $f : X_1 \to X_2$ is called continuous iff $\forall U \in \sigma_2 : f^{-1}(U) \in \sigma_1$.

Definition 1.12. A topological space is called Hausdorff iff $\forall x, y \in X$, there exist neighbourhoods U of x, V of y such that $U \cap V = \emptyset$.

Example. Endow $\mathbb{R}^2 \setminus \{0\}$ with the equivalence relation given by the thick lines and the two half lines in the following figure. That is:

$$(x, y) \sim (x', y') \Leftrightarrow \begin{cases} x = x' & \text{if } x \neq 0 \\ yy' > 0 & \text{if } x = 0. \end{cases}$$

Then the quotient topology on $(\mathbb{R}^2 \setminus \{0\})/\sim$ is not Hausdorff. ◇

Definition 1.13 (Basis for a topology). A basis for a topology is $S \subset \sigma$ such that every open set of X is a union of elements of S.

Definition 1.14 (C2). A space (X, σ) is second countable if there exists a countable basis.

Example. $\mathbb{R}^n$ is second countable. Indeed $\{B_{\frac{1}{m}}(x) \mid x \in \mathbb{Q}^n, m \in \mathbb{N}\}$ is a countable basis for the topology. Here $B_r(x)$ is the open ball with radius r around x. ◇

1.4 Differentiable manifolds

Definition 1.15 (Topological manifold). A topological manifold M of dimension m is a second countable, Hausdorff topological space which is locally homeomorphic to $\mathbb{R}^m$.

Remark. 'Locally homeomorphic to $\mathbb{R}^m$' means that $\forall p \in M$, there exists a neighborhood U of p and a homeomorphism $\phi : U \to V \underset{\text{open}}{\subset} \mathbb{R}^m$. Recall that homeomorphism means: bijective map that is continuous in *both* directions.

Definition 1.16 (Chart). The pair (U, ϕ) is called a chart.

Remark.

- Any subset of a Hausdorff space is Hausdorff

- Any subset of a C2 space is C2.

Definition 1.17 (Compatibility)**.** Two charts (U_1, ϕ_1) and (U_2, ϕ_2) are called smoothly compatible if

$$\phi_2 \circ (\phi_1)^{-1}|_{\phi_1(U_1 \cap U_2)} : \phi_1(U_1 \cap U_2) \to \phi_2(U_1 \cap U_2)$$

is a diffeomorphism (i.e. bijective, differentiable and inverse differentiable, where "differentiable" means that all partial derivatives exist).

Definition 1.18 (Smooth atlas)**.** A smooth atlas for M is called a collection of charts $\{(U_\alpha, \phi_\alpha)\}$ such that $\bigcup_\alpha U_\alpha = M$ and any two charts are smoothly compatible.

Definition 1.19 (Maximal smooth atlas)**.** A smooth atlas $\mathcal{A}$ is maximal if: whenever $\mathcal{B}$ is a smooth atlas and $\mathcal{A} \subset \mathcal{B}$ then $\mathcal{B} = \mathcal{A}$.

Definition 1.20 (Differentiable manifold)**.** A differentiable manifold (also called a smooth manifold) is a topological manifold M together with a maximal smooth atlas.

Remark. Given a smooth atlas $\mathcal{A}$ on a topological manifold M, there exists a unique maximal smooth atlas containing it, namely

$$\{(V, \psi) \mid (V, \psi) \text{ is smoothly compatible with all charts of } \mathcal{A}\}.$$

Example. Let $U \subset \mathbb{R}^n$. It's a smooth manifold: an atlas is $\{(U, \mathrm{Id})\}$. Take the maximal smooth atlas containing it. ◇

Example. Let $S^n := \{\mathbf{x} \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\| = 1\}$. The sphere S^n with the subspace topology is Hausdorff and C2, simply because $\mathbb{R}^{n+1}$ is. Two charts are given by the stereographic projections from the Northpole N and Southpole S:

$$\phi_N : S^n \setminus \{N\} \to \mathbb{R}^n : (x_1, \ldots, x_{n+1}) \mapsto \frac{(x_1, \ldots, x_n)}{1 - x_{n+1}}$$
$$\phi_S : S^n \setminus \{S\} \to \mathbb{R}^n : (x_1, \ldots, x_{n+1}) \mapsto \frac{(x_1, \ldots, x_n)}{1 + x_{n+1}}.$$

Now, ϕ_N and ϕ_S are homomorphisms. Furthermore, $\|\phi_N(p)\| \cdot \|\phi_S(p)\| = 1$, which allows us to calculate the inverse of ϕ_N. Hence

$$(\phi_S \circ \phi_N^{-1})|_{\phi_N(S^n \setminus \{N,S\})} : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n \setminus \{0\} : y \mapsto \frac{y}{\|y\|^2},$$

so ϕ_N and ϕ_S are smoothly compatible. Take the maximal smooth atlas containing ϕ_N and ϕ_S. $\diamond$

Remark. We could have started with other points $P, Q \in S^n$ instead of N, S. The smooth atlases $\{\phi_P, \phi_Q\}$ and $\{\phi_N, \phi_S\}$ would be different, but they define the same maximal smooth atlas.

1.5 Differentiable map

Let M be a smooth manifold.

Definition 1.21 (Smooth function). A function $f : M \to \mathbb{R}$ is differentiable (or smooth) at $p \in M$ iff $\exists$ a chart (U, ϕ) around p such that $f \circ \phi^{-1}$ is differentiable in $\phi(p)$.

Remark. If $f \circ \phi^{-1}$ is differentiable at p for a chart (U, ϕ), then $f \circ \psi^{-1}$ is also differentiable at p, for any other chart (V, ψ) (in the maximal atlas of M).

Proof. We want to argue that $f \circ \psi^{-1}$ is smooth.

$$f \circ \psi^{-1} = \underbrace{(f \circ \phi^{-1})}_{C^\infty} \circ \underbrace{(\phi \circ \psi^{-1})}_{C^\infty}.$$

$\square$

Notation. We write $C^\infty(M)$ to denote all smooth functions $M \to \mathbb{R}$.

Definition 1.22 (Smooth map). $f : M \to N$ is differentiable at $p \in M$ iff

- it is continuous
- there exists charts (U_M, ϕ_M) around p and (U_N, ϕ_N) around $f(p)$ such that $\phi_N \circ f \circ \phi_M^{-1}$ is differentiable at $\phi_M(p)$

Remark. The map $\phi_N \circ f \circ \phi_M^{-1}$ is defined on $\phi_M(U_M \cap f^{-1}(U_N))$. The continuity of f ensures that this is an open neighborhood of $\phi_M(p)$ in $\mathbb{R}^m$, hence it makes sense to talk about the differentiability of the above map at $\phi_M(p)$.

Remark. A map f being a differentiable and a homeomorphism does not imply that f is a diffeomorphism (which also include the differentiability of the inverse). For instance, $f \colon \mathbb{R} \to \mathbb{R}, x \mapsto x^3$ is not a diffeomorphism, because the inverse $x \mapsto \sqrt[3]{x}$ is not differentiable at zero.