Notes on Topology and Analysis

Topology

A **topology** on a set X is a collection τ of subsets of X having the following properties:

1. $\varnothing$ and X are in τ.
2. The union of the elements of any subcollection of τ is in τ.

$$\{U_i : i \in I\} \subseteq \tau \implies \bigcup_{i \in I} U_i \in \tau$$

3. The intersection of the elements of any finite subcollection of τ is in τ.

$$U_1, \cdots, U_n \in \tau \implies U_1 \cap \cdots \cap U_n \in \tau$$

A set X for which a topology τ has been specified is called a **topological space**.

If (X, τ) is a topological space, a subset U of X is an **open set** of X if $U \in \tau$.

If X is a set, a **basis** for a topology on X is a collection $\mathfrak{B}$ of subsets of X (called **basis elements**) such that

1. For each $x \in X$, there is at least one basis element B containing x.
2. If x belongs to the intersection of two basis elements B_1 and B_2, then there is a basis element B_3 containing x such that $B_3 \subset B_1 \cap B_2$.

If $\mathfrak{B}$ is a basis for a topology on X, the **topology τ generated by** $\mathfrak{B}$ is described as follows: A subset U of X is said to be open in X (that is, $U \in \tau$) if for each $x \in U$, there is a basis element $B \in \mathfrak{B}$ such that $x \in B$ and $B \subset U$.

A subset A of a topological space (X, τ) is said to be **closed** if $X - A$ is open. It can be shown that:

1. $\varnothing$ and X are closed.
2. Arbitrary intersection of closed sets are closed.
3. Finite unions of closed sets are closed.

If (X, τ) is a topological space and p is a point in X, a **neighborhood** of p is a subset V of X that includes an open set U containing p, $p \in U \subseteq V \subseteq X$.

Metric Spaces

A **metric space** is an ordered pair (M, d) where M is a set and d is a **metric** defined on M, that is, $d : M \times M \to \mathbb{R}$ is a function such that for any $x, y, z \in M$, the following holds:

1. $d(x, y) = 0 \iff x = y$
2. $d(x, y) = d(y, x)$
3. $d(x, z) \leqslant d(x, y) + d(y, z)$

It can also be shown that any metric must also satisfy $d(x, y) \geqslant 0$.

Some examples of metric spaces are given below:

1. $\mathbb{R}^n$ along with the Euclidean distance $d(x, y) = \sqrt{(x-y)^T(x-y)}$ is a *complete* metric space.
2. Any **normed vector space** is a metric space by defining $d(x, y) = \|y - x\|$

About any point x in a metric space M we define the **open ball** of radius $r > 0$ (where r is a real number) about x as the set $B(x; r) = \{y \in M : d(x, y) < r\}$

The collection of these open balls form a basis for a topology on M, making every metric space a topological space as well in a natural manner.

A subset U of M is called **open** if for every $x \in U$, there exists a real number $r > 0$ such that $B(x; r)$ is contained in U.

A **neighborhood** of a point x is any subset of M that contains an open ball about x as a subset.

A sequence $\{x_n\}$ in a metric space M is said to *converge* to the limit $x \in M$ if and only if for every $\varepsilon > 0$ there exists a natural number N such that $d(x_n, x) < \varepsilon$ for all $n > N$.

If U is an open set in $\mathbb{R}^n$ and we *remove a finite number of points* from it, the resulting set is still open.

Multivariable Calculus

Derivative: Suppose E is an open set in $\mathbb{R}^n$, $\mathbf{f}$ maps E into $\mathbb{R}^m$, and $\mathbf{x} \in E$. If there exists a *linear transformation* A of $\mathbb{R}^n$ into $\mathbb{R}^m$ such that

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

then we say that $\mathbf{f}$ is **differentiable** at $\mathbf{x}$, and we write $\mathbf{f}'(\mathbf{x}) = A$. If $\mathbf{f}$ is differentiable at every $\mathbf{x} \in E$, we say that $\mathbf{f}$ is *differentiable in* E.

Here are a few remarks on the definition presented above:

1. Note that since E is open, if $|\mathbf{h}|$ is small enough, then $\mathbf{x}+\mathbf{h} \in E$. Hence, $\mathbf{f}(\mathbf{x}+\mathbf{h}) \in \mathbb{R}^m$ is defined.
2. It can be shown that if $\mathbf{f}$ is differentiable at $\mathbf{x}$, then $\mathbf{f}'(\mathbf{x}) = A$ is unique.
3. The aforementioned relation can be rewritten as

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\mathbf{h} + \mathbf{r}(\mathbf{h}) \tag{1}$$

   where the remainder $\mathbf{r}(\mathbf{h})$ satisfies

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{|\mathbf{r}(\mathbf{h})|}{|\mathbf{h}|} = 0.$$

   This may be interpreted by saying that *for fixed* $\mathbf{x}$ *and small* $\mathbf{h}$*, the left side of (1) is approximately equal to* $\mathbf{f}'(x)\mathbf{h}$, that is, to the value of a linear transformation applied to $\mathbf{h}$.
4. If $\mathbf{f}$ is differentiable in E, then for every $\mathbf{x} \in E$, $\mathbf{f}'(\mathbf{x})$ is a function, namely, a linear transformation of $\mathbb{R}^n$ into $\mathbb{R}^m$. But $\mathbf{f}'$ *itself is also a function*: it maps E into the space of linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, denoted by $L(\mathbb{R}^n, \mathbb{R}^m)$.
5. $\mathbf{f}$ is continuous at any point at which it is differentiable.
6. The derivative that is defined here is often called the **differential** or the **total derivative** of $\mathbf{f}$ at $\mathbf{x}$, to distinguish it from partial derivatives that will come later.

Chain rule: Suppose E is an open set in $\mathbb{R}^n$, $\mathbf{f}$ maps E into $\mathbb{R}^m$ and is differentiable at $\mathbf{x}_0 \in E$, $\mathbf{g}$ maps an open set containing $\mathbf{f}(E)$ into $\mathbb{R}^k$ and is differentiable at $\mathbf{f}(\mathbf{x}_0)$. Then the mapping $\mathbf{F}$ of E into $\mathbb{R}^k$ defined by $\mathbf{F}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$ is differentiable at $\mathbf{x}_0$, and

$$\mathbf{F}'(\mathbf{x}_0) = \mathbf{g}'(\mathbf{f}(\mathbf{x}_0))\mathbf{f}'(\mathbf{x}_0),$$

where on the right side, we have the product of two linear transformations.

Partial Derivatives: Consider a function $\mathbf{f} : E \to \mathbb{R}^m$, where $E \subset \mathbb{R}^n$ is an open set. Let $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ and $\{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ be the standard basis of $\mathbb{R}^n$ and $\mathbb{R}^m$. The *components of* $\mathbf{f}$ are the real functions $f_1, \cdots f_m$ defined by

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^{m} f_i(\mathbf{x})\mathbf{u}_i \qquad (\mathbf{x} \in E),$$

or, equivalently, by $f_i(\mathbf{x}) = \mathbf{f}(\mathbf{x}) \cdot \mathbf{u}_i,\ 1 \leqslant i \leqslant m.$

For $\mathbf{x} \in E, 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$, we define

$$(D_j f_i)(\mathbf{x}) = \lim_{t \to 0} \frac{f_i(\mathbf{x} + t\mathbf{e}_j) - f_i(\mathbf{x})}{t},$$

provided the limit exists. It can be seen that $D_j f_i$ is the derivative of f_i with respect to x_j , keeping other variables fixed. The notation

$$\frac{\partial f_i}{\partial x_j}$$

is therefore often used in place of $D_j f_i$, which is called a *partial derivative*.

It must be noted that even for continuous functions, the existence of all partial derivatives <u>does not</u> imply differentiability. However, if $\mathbf{f}$ is known to be differentiable at a point $\mathbf{x}$, then its partial derivatives exist at $\mathbf{x}$, and they completely determine the linear transformation $\mathbf{f}'(\mathbf{x})$. See the following theorem for more details.

Suppose $\mathbf{f}$ maps an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$, and $\mathbf{f}$ is differentiable at a point $\mathbf{x} \in E$. Then the partial derivatives $\frac{\partial f_i}{\partial x_j}$ exist, and

$$\mathbf{f}'(\mathbf{x})\mathbf{e}_j = \sum_{i=1}^{m} (\frac{\partial f_i}{\partial x_j})(\mathbf{x})\mathbf{u}_i \quad (1 \leqslant j \leqslant n),$$

where $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ and $\{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ are the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively.

As a consequence of the above theorem, if $[\mathbf{f}'(\mathbf{x})]$ is the matrix that represents $\mathbf{f}'(\mathbf{x})$ with respect to our standard bases, then $\mathbf{f}'(\mathbf{x})\mathbf{e}_j$ is the jth column of that matrix. Therefore, $\frac{\partial f_i}{\partial x_j}$ is the number at ith row and jth column of $[\mathbf{f}'(\mathbf{x})]$:

$$[\mathbf{f}'(\mathbf{x})] = \begin{bmatrix} (D_1 f_1)(\mathbf{x}) & \cdots & (D_n f_1)(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ (D_1 f_m)(\mathbf{x}) & \cdots & (D_n f_m)(\mathbf{x}) \end{bmatrix}$$

This matrix is often regarded as the **Jacobian** matrix.

Gradient: Let f be a real valued differentiable function with its domain $E \subseteq \mathbb{R}^n$ being an open set. Associate with each $\mathbf{x} \in E$ a vector, the so-called *gradient* of f at $\mathbf{x}$, defined by

$$(\boldsymbol{\nabla} f)(\mathbf{x}) = \sum_{i=1}^{n} (D_i f)(\mathbf{x}) \mathbf{e}_i$$

One can represent this vector (with respect to the standard basis) as $[\frac{\partial f}{\partial x_1}(\mathbf{x}), \cdots, \frac{\partial f}{\partial x_n}(\mathbf{x})]^T$.

Some points regarding the gradient:

1. The gradient defines a vector field, over the domain of a real valued function, that is, if $f : \mathbb{R}^n \to \mathbb{R}$ then the gradient $\boldsymbol{\nabla} f : \mathbb{R}^n \to \mathbb{R}^n$ assigns a vector to each point in $\mathbb{R}^n$.
2. The gradient vector can be interpreted as *the direction and rate of fastest increase*. If the gradient of a function is non-zero at a point $\mathbf{x}$, its direction is the one moving along which and away from $\mathbf{x}$ increases the function most quickly. The magnitude of gradient is the *rate of change* in that direction.
3. The gradient is the **dual to the total derivative** Df: The value of the gradient at any point is a **tangent vector**, while the value of the derivaitve at any point is a **cotangent vector** [1], that is, a linear function on vectors.
4. If $f : \mathbb{R}^n \to \mathbb{R}$ [2] is differentiable at a point $\mathbf{x}$ and $\mathbf{v}$ is a vector, then the dot product of the gradient of f at $\mathbf{x}$ with $\mathbf{v}$ is equal to the directional derivative of f at $\mathbf{x}$ along $\mathbf{v}$. Thus,

$$(\boldsymbol{\nabla} f)(\mathbf{x}) \cdot \mathbf{v} = \frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}) = D_{\mathbf{v}} f(\mathbf{x})$$

5. TODO: Gradient of a vector field is the transpose of Jacobian, Einstien notaion

TODO: Multi variable Taylor expansion

TODO: Manifolds

References

Topology, a first course by Munkres

Principles of mathematical analysis by Rudin

Wikipedia

(By: Saeed Hedayatian)

1. For any smooth manifold $\mathcal{M}$ and a point x on it, the *cotangent space* is a vector space denoted by $T_x^*\mathcal{M}$ and defined to be the dual space of the tangent space at x, $T_x\mathcal{M}$. The elements of this space are called *cotangent vectors* or *tangent covectors*. ↩

2. More generally, the domain of f can be any smooth manifold. In this case, the vector $\mathbf{v}$ must be a tangent vector at point $\mathbf{x}$. ↩