Multivariable Analysis - MATH0019

Based on lectures by Prof Yiannis Petridis

Notes taken by Imran Radzi

Notes based on the Autumn 2021 Multivariable Analysis lectures by Prof Yiannis Petridis.

Contents

1 Review of Euclidean space and some linear algebra **1**

1.1 Euclidean space 1

1.2 Some linear algebra 2

2 Functions and continuity **3**

2.1 Limits 4

3 Derivatives **7**

1 Review of Euclidean space and some linear algebra

1.1 Euclidean space

Recall the *Euclidean n-space* $\mathbb{R}^n$

$$\mathbb{R}^n = \{(x^1, x^2, \ldots, x^n) \,:\, x^i \in \mathbb{R}\}$$

Note that use of superscripts instead of subscripts. We also have the *Euclidean norm* given by

$$|x| = \Big((x^1)^2 + \ldots + (x^n)^2\Big)^{\frac{1}{2}}$$

and the *inner product*

$$x \cdot y = \sum_{i=1}^{n} x^i \cdot y^i$$

where

$$\begin{aligned} x &= (x^1, \ldots, x^n) \\ y &= (y^1, \ldots, y^n) \end{aligned}$$

Recall the *standard basis*

$$\{e_1, e_2, \ldots, e_n\}$$

where $e_i \in \mathbb{R}^n$ where the only non-zero component is the i^{th} component whose value is 1. Hence we can represent $x \in \mathbb{R}^n$ as

$$x = \sum_{i=1}^{n} x^i e_i$$

Proposition.

1. $|x| \geq 0$, $|x| = 0 \iff x = 0$
2. $|x \cdot y| \leq |x||y|$ (Cauchy-Schwarz inequality)

3. $|x+y| \leq |x| + |y|$ (Triangle inequality)

4. $|a \cdot x| = |a||x|$ (for $a \in \mathbb{R}$)

5. $x \cdot y = y \cdot x$

We may also write $x \cdot y = \langle x, y \rangle$ as the inner product is a bilinear form. Recall the properties of a bilinear form which are

1. $\langle x + 1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle$

2. $\langle a \cdot x, y \rangle = a \langle x, y \rangle$

Also note that

$$\langle x, x \rangle = |x|^2$$

1.2 Some linear algebra

Recall that a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be *linear* if and only if, for $x, y \in \mathbb{R}^n$, $a \in \mathbb{R}$,

$$T(x+y) = T(x) + T(y)$$

$$T(a \cdot x) = a \cdot T(x)$$

Note that in both equations, the operations taking place on the left hand side is done in $\mathbb{R}^n$ and $\mathbb{R}^m$ on the right hand side.
Recall that we can recover M, the matrix representation of a linear mapping T by applying T to each of the standard basis. So if

$$T(e_j) = a_{ij} e_1 + a_{2j} e_2 + \ldots + a_{mj} e_m$$

then

$$\begin{aligned} M &= (a_{ij}) \\ &= [T] \end{aligned}$$

where M (or as we shall denote $[T]$) is the $m \times n$ matrix representing the linear transformation T.

$$\begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{pmatrix} = M \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}$$

If we have two mappings $T : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to \mathbb{R}^p$ where $[S]_{p \times m}$ represents the $p \times m$ matrix representing the linear transformation S, then $S \circ T$ is a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^p$.

$$S \circ T : \mathbb{R}^n \to \mathbb{R}^p$$

and

$$[S \circ T] = [S][T]$$

2 Functions and continuity

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called a *vector field* if $m > 1$, and a *scalar field* if $m = 1$.

If $f : \mathbb{R}^n \to \mathbb{R}^m$ then we have

$$f = (f^1, f^2, \ldots, f^m)$$

where

$$f^i : \mathbb{R}^n \to \mathbb{R}$$

so we can write

$$\begin{aligned} f(x) &= f^1(x)e_1 + \ldots + f^m(x)e_m \\ &= \Big(f^1(x), \ldots, f^m(x)\Big) \end{aligned}$$

and we call each of these f^i's the components of f.

Now we define a function $\pi^i : \mathbb{R}^m \to \mathbb{R}$ given by $\pi^i(y) = y^i$, so that

$$\pi^i(y^1, y^2, \ldots, y^m) = y^i$$

and we call this the *projection in the i^{th} direction* (*or the projection function*). This function is a linear transformation.

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "f"] \arrow[dr, "f^i=\pi^i \circ f"'] & \mathbb{R}^m \arrow[d, "\pi^i"] \\
& \mathbb{R}
\end{tikzcd}$$

Sometimes instead of $\mathbb{R}^n$ we may define f on a subset of $A \subset \mathbb{R}^n$, usually where A is open, i.e., sometimes we have $f : A \to \mathbb{R}^m$.

2.1 Limits

Definition. Let $a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$. We write $\lim_{x \to a} f(x) = b$ to mean

$$\forall \epsilon > 0, \exists \delta > 0 \text{ such that } 0 < |x - a| < \delta \implies |f(x) - b| < \epsilon$$

We say f is continuous on a set A us f is continuous at a, for all $a \in A$.

Theorem (Combination theorem)**.** Suppose that

$$\lim_{x \to a} f(x) = b \text{ and } \lim_{x \to a} g(x) = c$$

then,

1. $\lim_{x \to a} \Big(f(x) + g(x) \Big) = b + c$
2. If $\lambda \in \mathbb{R}$, then $\lim_{x \to a} \Big(\lambda \cdot f(x) \Big) = \lambda \cdot b$
3. $\lim_{x \to a} f(x) \cdot g(x) = b \cdot c$

4. $\lim_{x \to a} |f(x)| = |b|$

Proof.

3.

$$\begin{aligned} f(x) \cdot g(x) - b \cdot c &= f(x) \cdot g(x) - b \cdot g(x) + b \cdot g(x) - b \cdot c \\ &= (f(x) - b) \cdot g(x) + b \cdot (g(x) - c) \end{aligned}$$

$$\begin{aligned} |f(x) \cdot g(x) - b \cdot c| &\leq |(f(x) - b) \cdot g(x)| + |b \cdot (g(x) - c)| \\ &\leq |f(x) - b| \cdot |g(x)| + |b| \cdot |g(x) - c| \end{aligned}$$

and $|f(x) - b|, |g(x) - c|$ both tend to 0 as $x \to a$, while $|g(x)|$ is bounded close to a, so the right hand side tends to 0.

4. By the reverse triangle inequality,

$$||f(x)| - |b|| \leq |f(x) - b|$$

$\square$

Now note that a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is continuous everywhere.

Lemma. Given T linear $T : \mathbb{R}^n \to \mathbb{R}^m$, there exists $M > 0$ such that

$$\forall x \in \mathbb{R}^n,\ |T(x)| \leq M|x|$$

Proof. Given that $x = (x^1, x^2, \ldots, x^n)$, we can rewrite

$$x = \sum_{i=1}^{n} x^i e_i$$

and so as T is linear,

$$T(x) = \sum_{i=1}^{n} x^i T(e_i)$$

and then,

$$\begin{aligned}
|T(x)| &= |\sum_{i=1}^{n} x^i T(e_i)| \\
&\leq \sum_{i=1}^{n} |x^i||T(e_i)| \leq (\sum_{i=1}^{n} |x^i|^2)^{\frac{1}{2}} (\sum_{i=1}^{n} |T(e_i)|^2)^{\frac{1}{2}} \\
&= |x|(\sum_{i=1}^{n} |T(e_i)|^2)^{\frac{1}{2}}
\end{aligned}$$

and so set

$$M = |x|(\sum_{i=1}^{n} |T(e_i)|^2)^{\frac{1}{2}}$$

$\square$

Using this lemma we can prove T to be continuous at $y \in \mathbb{R}^n$ as follows,

$$T(x) - T(y)| = |T(x-y)| \leq M|x-y|$$

Given $\epsilon > 0$, we can simply take $\delta = \frac{\epsilon}{M} > 0$.

Remark.

1. Let $f : \mathbb{R}^n \to \mathbb{R}^m$. Then,

   $$f \text{ is continuous } \iff f^i \text{ is continuous for } i = 1, \ldots, m$$

   Sufficiency follows noting that $f^i = \pi^i \circ f$ and the fact that π^i is continuous. Necessity is a simple exercise to prove.

2. Polynomials in n-variables are continuous, for example, the function

   $$1024(x^4)^5(x^2)^3(x^5)^{29}$$

   is continuous. The same holds for rational functions

   $$R(x^1, \ldots, x^n) = \frac{P(x^1, \ldots, x^n)}{Q(x^1, \ldots, x^n)}$$

   where P, Q are polynomials, continuous, where denominator is not equal to 0.

3 Derivatives

Definition. We define

$$D_i f(a) = \lim_{h \to 0} \frac{f(a^1, a^2, \ldots, a^{i-1}, a^i + h, a^{i+1}, \ldots, a^n) - f(a)}{h}$$

Example. Let $f : \mathbb{R}^2 \to \mathbb{R}$, so $f(x, y)$ is a function of two variables x, y. Then,

$$D_1 f(a, b) = \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h} = \frac{\partial f}{\partial x}(a, b)$$

$$D_2 f(a, b) = \frac{\partial f}{\partial y}(a, b)$$

and similarly for $\mathbb{R}^3 \to \mathbb{R}$.

Definition. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ (or replace $\mathbb{R}^n$ with A open in $\mathbb{R}^n$), and let $a \in \mathbb{R}$. We say f is *differentiable* at a if we can find a linear transformation $\lambda : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h \to 0} \frac{|f(a + h) - f(a) - \lambda(h)|}{|h|} = 0 \tag{1}$$

We call $\lambda : \mathbb{R}^n \to \mathbb{R}^m$ the *derivative* of f at a, and we denote it by $Df(a)$.

Theorem. λ as defined above is unique.

Proof. If f is differentiable at a and $\lambda, \mu : \mathbb{R}^n \to \mathbb{R}^m$ both linear transformations that satisfy (1), then

$$\mu = \lambda$$

□

Proof. We need to prove that $\forall x \in \mathbb{R}^n$, that $\mu(x) = \lambda(x)$. We can assume $x \neq 0$ since as μ, λ are linear (so they map 0 to 0).

$$\begin{aligned}
\frac{|\lambda(h) - \mu(h)|}{|h|} &= \frac{|\lambda(h) + f(a) - f(a+h) + f(a+h) - f(a) - \mu(h)|}{|h|} \\
&= \frac{|\lambda(h) + f(a) - f(a+h)|}{|h|} + \frac{|f(a+h) - f(a) - \mu(h)|}{|h|}
\end{aligned}$$

Now if we take limits as $h \to 0$, then

$$\lim_{h \to 0} \frac{|\lambda(h) - \mu(h)|}{|h|} = 0$$

If we take t small, and let $tx = h$, then since $h \to 0 \iff t \to 0$ we have

$$\begin{aligned}
\lim_{t \to 0} \frac{|\lambda(tx) - \mu(tx)|}{|tx|} &= \lim_{t \to 0} \frac{|t\lambda(x) - t\mu(x)|}{|t||x|} \\
&= \lim_{t \to 0} \frac{|t||\lambda(x) - \mu(x)|}{|t||x|} \\
&= \lim_{t \to 0} \frac{|\lambda(x) - \mu(x)|}{|x|} \\
&= \frac{|\lambda(x) - \mu(x)|}{|x|}
\end{aligned}$$

which is equal to 0 as

$$\lim_{t \to 0} \frac{|\lambda(tx) - \mu(tx)|}{|tx|} = 0$$

hence

$$|\lambda(x) - \mu(x) = 0$$

i.e.,

$$\lambda(x) = \mu(x)$$

□

Note that for $h \in \mathbb{R}^m$, that

$$\lambda(h) = Df(a)(h)$$

so $Df(a)$ is a linear transformation, $Df(a) : \mathbb{R}^n \to \mathbb{R}^m$.

We call the matrix representation of $Df(a) : \mathbb{R}^n \to \mathbb{R}^m$ the *Jacobian* of f at a (here we use the standard basis) and we denote it by $f'(a)$. So,

$$Df(a)(h) = f'(a) \begin{pmatrix} h^1 \\ h^2 \\ \vdots \\ h^n \end{pmatrix}$$

Note that the product of this operation is column m-vector ($m \times 1$ matrix), an element of $\mathbb{R}^m$.

Definition. Let $f : \mathbb{R}^n \to \mathbb{R}$ (or replace $\mathbb{R}^n$ with any A open in $\mathbb{R}^n$), and let $u \in \mathbb{R}^n$, where $u \neq 0$. We define the *directional derivative* of f at a in the u-direction by

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

(if it exists), where $h \in \mathbb{R}$. If $u = e_i$, then $D_u f(a) = D_i f(a)$.

Theorem. If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$ hen it is continuous at a.

Proof. We want to show that

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

Hence we proceed

$$\begin{aligned}
\lim_{h \to 0} |f(a + h) - f(a)| &= \lim_{h \to 0} |f(a + h) - f(a) - \lambda(h)| \\
&\leq \lim_{h \to 0} |f(a + h) - f(a) - \lambda(h)| + \lim_{h \to 0} |\lambda(h)| \\
&\leq \lim_{h \to 0} \frac{|f(a + h) - f(a) - \lambda(h)|}{|h|} |h| + \lim_{h \to 0} |\lambda(h)| \\
&= 0
\end{aligned}$$

so

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

$\square$

Remark. If

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

then letting $h = x - a$, we can write

$$\lim_{x\to a} \frac{|f(x) - f(a) - \lambda(x-a)|}{|x-a|} = 0$$

If we let $\phi(x) = f(x) - f(a) - \lambda(x-a)$, then we have that $\phi : \mathbb{R}^n \to \mathbb{R}^m$ and we can rewrite

$$f(x) = f(a) + \lambda(x-a) + \phi(x)$$

or, recalling that $\lambda = Df(a)$,

$$f(x) = f(a) + Df(a)(x-a) + \phi(x)$$

By definition, we have

$$\lim_{x\to a} \frac{\phi(x)}{|x-a|} = 0$$