Analysis II

Trong

December 1, 2018—January 17, 2019

Contents

1 Differentiation

Definition 1 (Directional derivative). *Let $A \subset \mathbf{R}^m, f : A \longrightarrow \mathbf{R}^n$. Suppose A contains a neighbourhood of a. Given $u \in \mathbf{R}^m$ with $u \neq 0$, define the directional derivative of f at a in the direction of u to be*

$$f'(a; u) = \lim_{t \to 0} \frac{f(a + u) - f(a)}{t}$$

if it exists.

Definition 2 (Derivative). *Let $A \subset \mathbf{R}^m, f : A \longrightarrow \mathbf{R}^n$. Suppose A contains a neighbourhood of a. We say that f is differentiable at a if there is an $n \times m$ matrix B s.t.*

$$\frac{f(a + h) - f(a) - B \cdot h}{|h|} \longrightarrow 0$$

as $h \longrightarrow 0$. The matrix B is unique and is denoted $Df(a)$. Sometimes people also call the derivative the gradient, and write $Df(a) = \nabla f(a)$.

Theorem 3 (Relating directional derivatives to the derivative of f). *Let $A \subset \mathbf{R}^m, f : A \longrightarrow \mathbf{R}^n$. If f is differentiable at a, then all the directional derivatives of f at a exists, and*

$$f'(a; u) = Df(a) \cdot u.$$

Definition 4 (Partial derivative). *Let $A \subset \mathbf{R}^m, f : A \longrightarrow \mathbf{R}$. Define the j-th partial derivative of f at a to be directional derivative of f at a with respect to the vector e_j, provided it exists; and we denote it by $D_j f(a)$:*

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

IOW, partial derivatives are directional derivatives along coordinate axes. Note that if we define $\phi(t) = f(a_1, \ldots, a_{j-1}, t, a_{j+1}, \ldots, a_m)$, then

$$D_j f(a) = \phi'(a_j).$$

Theorem 5 (Derivative of a real-valued function). *Let $A \subset \mathbf{R}^m, f : A \longrightarrow \mathbf{R}$. If f is differentiable at a, then the derivative of f is the row matrix*

$$Df(a) = [D_1 f(a) \quad \cdots \quad D_m f(a)].$$

Theorem 6. *Let $A \subset \mathbf{R}^m, f : A \longrightarrow \mathbf{R}^n$. Suppose A contains a neighbourhood of a. Let $f_i : A \longrightarrow R$ be the i-th component function of f, so that*

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{bmatrix}.$$

- *Then f is differentiable at a iff each component f_i is differentiable at a.*
- *If f is differentiable at a, then its derivative is the $n \times m$ matrix whose i-th row is the derivative of f_i, i.e.*

$$Df(a) = \begin{bmatrix} Df_1(a) \\ \vdots \\ Df_n(a) \end{bmatrix} = \begin{bmatrix} D_1 f_1(a) & \cdots & D_m f_1(a) \\ \vdots & & \vdots \\ D_1 f_n(a) & \cdots & D_m f_n(a) \end{bmatrix}.$$

This matrix is called the Jacobian matrix of f.

Roughly: Differentiability of $f : \mathbf{R}^m \longrightarrow \mathbf{R}^n$ is equivalent to differentiability of each component, because the components are independent of each other as far as taking limits is concerned. Note that this doesn't imply that the partial derivatives of the components must be continuous, only that they exist.

2 Continuously differentiable functions

Theorem 7 (Mean value theorem)**.** *If $\phi : [a, b] \longrightarrow \mathbf{R}$ is continuous at each point of the closed interval $[a, b]$, and differentiable at each point of the interval (a, b), then there exists a point c of (a, b) s.t.*

$$\phi(b) - \phi(a) = \phi'(c)(b - a).$$

Theorem 8 (Continuously differentiable functions)**.** *Let A be open in $\mathbf{R}^m$. Suppose that the partial derivatives $D_j f_i(x)$ of the component functions of f exist at each point $x \in A$ and are continuous on A. Then f is differentiable at each point of A.*

This theorem guarantees differentiability of f if its partial derivatives exist and are continuous. Such a function is called continuously differentiable, or C^1 on A.

Theorem 9. *Let A be open in $\mathbf{R}^m$, $f : A \longrightarrow \mathbf{R}$ be a function of class C^2. Then for each $a \in A$, the mixed second order partial derivatives are equal:*

$$D_k D_j f(a) = D_j D_k f(a).$$

3 Homogeneous functions and the Chain Rule

Proposition 10. *A function $f : \mathbf{R}^n \longrightarrow \mathbf{R}$ is called homogeneous of degree m if $f(tx) = t^m f(x)$ for all x and t. If f is also differentiable, then*

$$\sum_{i=1}^{n} x_i D_i f(x) = m f(x). \tag{1}$$

Proof. Let $g(t) = f(tx) = t^m f(x)$. By the Chain Rule,

$$\begin{aligned} g'(t) &= \sum_i D_i f(tx) \frac{d}{dt}(tx_i) \\ &= \sum_i x_i D_i f(tx) \\ &= m t^{m-1} f(x). \end{aligned}$$

Plugging in $t = 1$ yields the desired result. ■

Example 11. Intuitively a homogeneous function transforms scaling from its arguments to its value. Examples are linear operators: $T(ax) = aT(x)$, which are homogeneous of order 1.

Another example is the function

$$f(x_1, x_2) = x_1^2 + x_2^2,$$

which is homogeneous of order 2, since

$$f(tx_1, tx_2) = t^2 x_1^2 + t^2 x_2^2 = t^2 f(x_1, x_2).$$

Let's verify (1):

$$\sum_{i=1}^{n} x_i D_i f(x) = 2x_1 x_1 + 2x_2 x_2 = 2f(x).$$

Proposition 12. *If $f : \mathbf{R}^n \longrightarrow \mathbf{R}$ is differentiable and $f(0) = 0$, then there exist $g_i : \mathbf{R}^n \longrightarrow \mathbf{R}$ s.t.*

$$f(x) = \sum_{i=1}^{n} x_i g_i(x). \tag{2}$$

Specifically,

$$f(x) = \sum_{i=1}^{n} x_i \int_0^1 D_i f(tx) dt. \tag{3}$$

Proof. Let $h(t) = f(tx)$. Then

$$\int_0^1 h'(t)dt = h(1) - h(0) = f(x) - f(0) = f(x).$$

On the other hand, by the Chain Rule,

$$h'(t) = \sum_{i=1}^{n} x_i D_i f(tx),$$

hence

$$\begin{aligned} \int_0^1 h'(t)dt &= \int_0^1 \sum_{i=1}^{n} x_i D_i f(tx)dt \\ &= \sum_{i=1}^{n} x_i \int_0^1 D_i f(tx)dt. \end{aligned}$$

Thus

$$g_i(x) = \int_0^1 D_i f(tx)dt$$

are the functions we're looking for. ■

Example 13. It's always good to verify an abstract result on a simple example:

$$f(x_1, x_2) = x_1^2 + x_2^2.$$

Formula (2) says

$$f(x) = x_1 g_1(x) + x_2 g_2(x),$$

which would suggest that

$$g_1(x) = x_1 \quad \text{and} \quad g_2(x) = x_2.$$

Let's check:

$$D_1 f(tx) = 2x_1|_{tx} = 2tx_1,$$

which we plug into g_1:

$$\begin{aligned} g_1(x) &= \int_0^1 D_1 f(tx)dt \\ &= \int_0^1 2tx_1 dt \\ &= t^2 x_1|_0^1 \\ &= x_1. \end{aligned}$$

Similarly we find that $g_2(x) = x_2$, as required.

Note 14. If $f : \mathbf{R}^n \longrightarrow \mathbf{R}$ is differentiable and homogeneous of degree m, then $f(0) = 0$ since $f(0) = f(t \cdot 0) = t^m f(0)$. Therefore (3) holds:

$$f(x) = \sum_{i=1}^{n} x_i \int_0^1 D_i f(tx) dt.$$

Combining it with (1)—

$$\sum_{i=1}^{n} x_i D_i f(x) = m f(x),$$

we're led to believe that

$$\frac{D_i f(x)}{m} = \int_0^1 D_i f(tx) dt.$$

This we can easily check for $f(x_1, x_2) = x_1^2 + x_2^2$ by reusing some of our previous calculations, e.g.

$$\frac{D_1 f(x)}{m} = \frac{2x_1}{2} = x_1 = \int_0^1 D_1 f(tx) dt.$$