Real Analysis III
(MAT312β)

Department of Mathematics
University of Ruhuna

A.W.L. Pubudu Thilan

The Total Derivative

Introduction

- In the previous Chapter, we discussed partial derivatives, which represent the instantaneous rates of change of a function, f, with respect to a single variable, while keeping all of the other independent variables constant.
- We can think of each partial derivative as the instantaneous rate of change of f, at a point $\mathbf{a}$, as the point moves in a direction parallel to the corresponding coordinate axis.

- Another way to say this is that the partial derivative, with respect to x_i is the instantaneous rate of change of f, at a point $\mathbf{a}$, as the point moves in the direction of the corresponding standard basis vector, $\mathbf{e}_i$.

- This naturally leads us to look at the instantaneous rates of change of f, at a point $\mathbf{a}$, as the point moves in an arbitrary direction, with an arbitrary speed, i.e., as the point moves with an arbitrary velocity $\mathbf{v}$.

- Thus, we define the total derivative of f, at $\mathbf{a}$, not as a number, but rather as a function which returns a number for each specified velocity vector.

Approximating a differentiable function by a linear function

- How your calculator gives answer for $\sin x$ for any particular value of x that you request?
- It can not remember sin value for every x, because this requires more memory.
- So it uses a polynomial approximation for that.

Approximating a differentiable function by a linear function

Example

$$\begin{aligned}
f'(a) &\simeq \frac{f(x) - f(a)}{(x-a)} \\
f(x) &\simeq f(a) + f'(a)(x-a) \\
\text{For example } x &= 0.2 \Rightarrow \\
\sin(0.2) &\simeq \sin 0 + \cos 0(0.2 - 0) \\
&\simeq 0.2
\end{aligned}$$

- We can obtain a better result using higher order Taylor polynomials.

Approximating a differentiable function by a Taylor polynomial

We recall that in the one-dimensional case a function f with a derivative at a can be approximated near a by a linear Taylor polynomial. If $f'(a)$ exists we let $E(a,h)$ denote the difference

$$E(a,h) = \begin{cases} \dfrac{f(a+h)-f(a)}{h} - f'(a) & \text{if } h \neq 0, \\ 0 & \text{if } h = 0. \end{cases} \tag{1}$$

Approximating a differentiable function by a Taylor polynomial
Cont...

- From (1) we obtain the formula;

$$f(a+h) = f(a) + f'(a)h + hE(a,h),$$

  an equation which holds also for $h = 0$.

- This is the first-order Taylor formula for approximating $f(a+h) - f(a)$ by $f'(a)h$.

- The error committed is $hE(a,h)$.

- From (1) we see that $E(a,h) \to 0$ as $h \to 0$.

- Therefore the error $hE(a,h)$ is of smaller order than h for small h.

The concept of differentiability in higher-dimensional space

- This property of approximating a differentiable function by a linear function suggests a way of extending the concept of differentiability to the higher-dimensional case.
- Let $f : \mathbb{S} \to \mathbb{R}$ be a scalar field defined on a set $\mathbb{S}$ in $\mathbb{R}^n$.
- Let $\mathbf{a}$ be an interior point of $\mathbb{S}$, and let $\mathbf{B}(\mathbf{a}; r)$ be an n-ball lying in $\mathbb{S}$.
- Let $\mathbf{v}$ be a vector with $\|\mathbf{v}\| < r$, so that $\mathbf{a} + \mathbf{v} \in \mathbf{B}(\mathbf{a}; r)$.

Definition of a differentiable scalar field

We say that f is differentiable at $\mathbf{a}$ if there exists a linear transformation

$$T_{\mathbf{a}} : \mathbb{R}^n \to \mathbb{R}$$

from $\mathbb{R}^n$ to $\mathbb{R}$, and a scalar function $E(\mathbf{a}, \mathbf{v})$ such that

$$f(\mathbf{a} + \mathbf{v}) = f(\mathbf{a}) + T_{\mathbf{a}}(\mathbf{v}) + \|\mathbf{v}\| E(\mathbf{a}, \mathbf{v}), \tag{2}$$

for $\|\mathbf{v}\| < r$, where $E(\mathbf{a}, \mathbf{v}) \to 0$ as $\|\mathbf{v}\| \to 0$. The linear transformation $T_{\mathbf{a}}$ is called the total derivative of f at $\mathbf{a}$.

Definition of a differentiable scalar field
Cont...

- The total derivative $T_{\mathbf{a}}$ is a linear transformation, not a number.
- The function value $T_{\mathbf{a}}(\mathbf{v})$ is a real number; it is defined for every point $\mathbf{v}$ in $\mathbb{R}^n$.
- The total derivative was introduced by W.H. Young in 1908 and by M. Frechet in 1911 in more general context.

- The equation (2), which holds for $\|\mathbf{v}\| < r$, is called a first-order Taylor formula for $f(\mathbf{a}+\mathbf{v})$.
- It gives a linear approximation, $T_{\mathbf{a}}(\mathbf{v})$, to the difference $f(\mathbf{a}+\mathbf{v}) - f(\mathbf{a})$.
- The error in the approximation is $\|\mathbf{v}\| E(\mathbf{a},\mathbf{v})$, a term which is of smaller order than $\|\mathbf{v}\|$ as $\|\mathbf{v}\| \to 0$; that is, $E(\mathbf{a},\mathbf{v}) = O(\|\mathbf{v}\|)$ as $\|\mathbf{v}\| \to 0$.

Theorem (6.1)

Assume f is differentiable at $\mathbf{a}$ with total derivative $T_{\mathbf{a}}$. Then the derivative $f'(\mathbf{a};\mathbf{y})$ exists for every $\mathbf{y}$ in $\mathbb{R}^n$ and we have

$$T_{\mathbf{a}}(\mathbf{y}) = f'(\mathbf{a};\mathbf{y}). \tag{3}$$

Moreover, $f'(\mathbf{a};\mathbf{y})$ is a linear combination of the components of $\mathbf{y}$. In fact, if $\mathbf{y} = (y_1, ..., y_n)$, we have

$$f'(\mathbf{a};\mathbf{y}) = \sum_{k=1}^{n} D_k f(\mathbf{a}) y_k. \tag{4}$$

Theorem (6.1)
Proof

The equation (3) holds trivially if $\mathbf{y} = \mathbf{0}$ since both $T_{\mathbf{a}}(\mathbf{0}) = 0$ and $f'(\mathbf{a}; \mathbf{0}) = 0$.

Therefore we can assume that $\mathbf{y} \neq \mathbf{0}$.

Since f is differentiable at $\mathbf{a}$ we have a Taylor formula,

$$f(\mathbf{a} + \mathbf{v}) = f(\mathbf{a}) + T_{\mathbf{a}}(\mathbf{v}) + \|\mathbf{v}\| E(\mathbf{a}, \mathbf{v}), \tag{5}$$

for $\|\mathbf{v}\| < r$ for some $r > 0$, where $E(\mathbf{a}, \mathbf{v}) \to 0$ as $\|\mathbf{v}\| \to 0$.

In this formula we take $\mathbf{v} = h\mathbf{y}$, where $h \neq 0$ and $|h|\|\mathbf{y}\| < r$.

Then $\|\mathbf{v}\| < r$.

Since $T_{\mathbf{a}}$ is linear we have $T_{\mathbf{a}}(\mathbf{v}) = T_{\mathbf{a}}(h\mathbf{y}) = hT_{\mathbf{a}}(\mathbf{y})$.

Theorem (6.1)
Proof

Therefore (5) gives us

$$\begin{aligned}
f(\mathbf{a}+\mathbf{v}) &= f(\mathbf{a})+T_{\mathbf{a}}(\mathbf{v})+\|\mathbf{v}\|E(\mathbf{a},\mathbf{v}) \\
f(\mathbf{a}+h\mathbf{y}) &= f(\mathbf{a})+hT_{\mathbf{a}}(\mathbf{y})+|h|\|\mathbf{y}\|E(\mathbf{a},\mathbf{v}) \\
f(\mathbf{a}+h\mathbf{y})-f(\mathbf{a}) &= hT_{\mathbf{a}}(\mathbf{y})+|h|\|\mathbf{y}\|E(\mathbf{a},\mathbf{v}) \\
\frac{f(\mathbf{a}+h\mathbf{y})-f(\mathbf{a})}{h} &= T_{\mathbf{a}}(\mathbf{y})+\frac{|h|\|\mathbf{y}\|}{h}E(\mathbf{a},\mathbf{v}). \qquad (6)
\end{aligned}$$

Theorem (6.1)
Proof

- Since $\|\mathbf{v}\| \to 0$ as $h \to 0$ and since $|h|/h = \pm 1$, the right hand member of (6) tends to the limit $T_{\mathbf{a}}(\mathbf{y})$ as $h \to 0$.
- Therefore the left-hand member tends to the same limit.
- This proves (3).

Theorem (6.1)
Proof

Now we use the linearity of $T_{\mathbf{a}}$ to deduce (4). If $\mathbf{y} = (y_1, ..., y_n)$ we have $\mathbf{y} = \sum_{k=1}^{n} y_k \mathbf{e}_k$, hence

$$\begin{aligned} T_{\mathbf{a}}(\mathbf{y}) &= T_{\mathbf{a}}\left(\sum_{k=1}^{n} y_k \mathbf{e}_k\right) \\ &= \sum_{k=1}^{n} y_k T_{\mathbf{a}}(\mathbf{e}_k) \\ &= \sum_{k=1}^{n} y_k f'(\mathbf{a}; \mathbf{e}_k) \\ &= \sum_{k=1}^{n} y_k D_k f(\mathbf{a}). \end{aligned}$$

Thank you!