Real Analysis III
(MAT312β)

Department of Mathematics
University of Ruhuna

A.W.L. Pubudu Thilan

Derivatives of Vector Fields

What is a vector field?

- Derivative theory for vector fields is a straightforward extension of that for scalar fields.
- Let $\mathbf{f} : \mathbb{S} \to \mathbb{R}^m$ be a vector field defined on a subset $\mathbb{S}$ of $\mathbb{R}^n$.
- Then $\mathbf{f}$ consists of m scalar fields of n variables. That is,

$$\mathbf{f}(\mathbf{a}) = (f_1(\mathbf{a}), ..., f_m(\mathbf{a})).$$

- In here each $f_i : \mathbb{S} \to \mathbb{R}$ is a scalar field, where $i = 1, ..., m$.

Definition
The derivative of a vector field

Let $\mathbf{f} : \mathbb{S} \rightarrow \mathbb{R}^m$ be a vector field defined on a subset $\mathbb{S}$ of $\mathbb{R}^n$. If $\mathbf{a}$ is an interior point of $\mathbb{S}$ and if $\mathbf{y}$ is any vector in $\mathbb{R}^n$ we define the derivative $\mathbf{f}'(\mathbf{a}; \mathbf{y})$ by the formula

$$\mathbf{f}'(\mathbf{a}; \mathbf{y}) = \lim_{h \rightarrow 0} \frac{\mathbf{f}(\mathbf{a} + h\mathbf{y}) - \mathbf{f}(\mathbf{a})}{h},$$

whenever the limit exists. The derivative $\mathbf{f}'(\mathbf{a}; \mathbf{y})$ is a vector in $\mathbb{R}^m$.

Differentiability in component wise

Let f_k denote the k^{th} component of $\mathbf{f}$. We note that the derivative $f'(\mathbf{a};\mathbf{y})$ exists if and only if $f_k'(\mathbf{a};\mathbf{y})$ exists for each $k = 1, 2, ..., m$ in which case we have

$$\mathbf{f}'(\mathbf{a};\mathbf{y}) = (f_1'(\mathbf{a};\mathbf{y}), ..., f_m'(\mathbf{a};\mathbf{y})) = \sum_{k=1}^{m} f_k'(\mathbf{a};\mathbf{y})\mathbf{e}_k, \to \mathrm{(A)}$$

where $\mathbf{e}_k$ is the k^{th} unit coordinate vector.

The total derivative of a vector field

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

$$\mathbf{T_a} : \mathbb{R}^n \to \mathbb{R}^m$$

such that

$$\mathbf{f}(\mathbf{a}+\mathbf{v}) = \mathbf{f}(\mathbf{a}) + \mathbf{T_a}(\mathbf{v}) + \|\mathbf{v}\|\mathbf{E}(\mathbf{a},\mathbf{v}), \qquad (1)$$

where $\mathbf{E}(\mathbf{a},\mathbf{v}) \to \mathbf{0}$ as $\mathbf{v} \to \mathbf{0}$. The first order Taylor formula (1) is to hold for all $\mathbf{v}$ with $\|\mathbf{v}\| < r$ for some $r > 0$. The term $\mathbf{E}(\mathbf{a},\mathbf{v})$ is a vector in $\mathbb{R}^m$. The linear transformation $\mathbf{T_a}$ is called the total derivative of $\mathbf{f}$ at $\mathbf{a}$.

The total derivative of a vector field
Cont...

- For scalar fields we proved that $T_{\mathbf{a}}(\mathbf{y})$ is the dot product of the gradient vector $\nabla f(\mathbf{a})$ with $\mathbf{y}$.
- For vector fields we will prove that $\mathbf{T}_{\mathbf{a}}(\mathbf{y})$ is a vector whose k^{th} component is the dot product $\nabla f_k(\mathbf{a}).\mathbf{y}$.

Theorem (9.1)

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

$$\mathbf{T_a}(\mathbf{y}) = \mathbf{f}'(\mathbf{a};\mathbf{y}). \tag{2}$$

Moreover, if $\mathbf{f} = (f_1, ..., f_m)$ and if $\mathbf{y} = (y_1, ..., y_n)$, we have

$$\mathbf{T_a}(\mathbf{y}) = \sum_{k=1}^{m} \nabla f_k(\mathbf{a}).\mathbf{y}\mathbf{e}_k = (\nabla f_1(\mathbf{a}).\mathbf{y}, ..., \nabla f_m(\mathbf{a}).\mathbf{y}). \tag{3}$$

Theorem (9.1)
Proof

We argue as in the scalar case. If $\mathbf{y} = \mathbf{0}$, then $\mathbf{f}'(\mathbf{a};\mathbf{y}) = \mathbf{0}$ and $\mathbf{T}_\mathbf{a}(\mathbf{0}) = \mathbf{0}$. Therefore we can assume that $\mathbf{y} \neq \mathbf{0}$. Taking $\mathbf{v} = h\mathbf{y}$ in the Taylor formula (1) we have

$$\begin{aligned}
\mathbf{f}(\mathbf{a}+\mathbf{v}) &= \mathbf{f}(\mathbf{a}) + \mathbf{T}_\mathbf{a}(\mathbf{v}) + \|\mathbf{v}\|\mathbf{E}(\mathbf{a},\mathbf{v}) \\
\mathbf{f}(\mathbf{a}+h\mathbf{y}) &= \mathbf{f}(\mathbf{a}) + \mathbf{T}_\mathbf{a}(h\mathbf{y}) + \|h\mathbf{y}\|\mathbf{E}(\mathbf{a},\mathbf{v}) \\
\mathbf{f}(\mathbf{a}+h\mathbf{y}) - \mathbf{f}(\mathbf{a}) &= h\mathbf{T}_\mathbf{a}(\mathbf{y}) + |h|\|\mathbf{y}\|\mathbf{E}(\mathbf{a},\mathbf{v}) \\
\lim_{h\to 0}\frac{\mathbf{f}(\mathbf{a}+h\mathbf{y}) - \mathbf{f}(\mathbf{a})}{h} &= \lim_{h\to 0}\frac{h\mathbf{T}_\mathbf{a}(\mathbf{y})}{h} + \lim_{h\to 0}\frac{|h|\|\mathbf{y}\|\mathbf{E}(\mathbf{a},\mathbf{v})}{h} \\
\mathbf{f}'(\mathbf{a};\mathbf{y}) &= \mathbf{T}_\mathbf{a}(\mathbf{y})
\end{aligned}$$

Theorem (9.1)

Proof⇒Cont...

To prove (3) we simply note that

$$\begin{aligned}
\mathbf{f}'(\mathbf{a};\mathbf{y}) &= \sum_{k=1}^{m} f_k'(\mathbf{a};\mathbf{y})\mathbf{e}_k \quad \text{(From (A))}\\
&= \sum_{k=1}^{m} \nabla f_k(\mathbf{a}).\mathbf{y}\mathbf{e}_k = (\nabla f_1(\mathbf{a}).\mathbf{y}, ..., \nabla f_m(\mathbf{a}).\mathbf{y})\\
\mathbf{T}_{\mathbf{a}}(\mathbf{y}) &= \mathbf{f}'(\mathbf{a};\mathbf{y}) \quad \text{(From (2))}\\
\mathbf{T}_{\mathbf{a}}(\mathbf{y}) &= \sum_{k=1}^{m} \nabla f_k(\mathbf{a}).\mathbf{y}\mathbf{e}_k = (\nabla f_1(\mathbf{a}).\mathbf{y}, ..., \nabla f_m(\mathbf{a}).\mathbf{y}).
\end{aligned}$$

Hence the result.

The Jacobian matrix of $\mathbf{f}$ at $\mathbf{a}$

Equation (3) can also be written more simply as a matrix product,

$$\mathbf{T}_{\mathbf{a}}(\mathbf{y}) = D\mathbf{f}(\mathbf{a})\mathbf{y},$$

where $D\mathbf{f}(\mathbf{a})$ is the $m \times n$ matrix whose k^{th} row is $\nabla f_k(\mathbf{a})$, and where $\mathbf{y}$ is regarded as an $n \times 1$ column matrix. The matrix $D\mathbf{f}(\mathbf{a})$ is called the **Jacobian matrix** of $\mathbf{f}$ at $\mathbf{a}$. Its kj entry is the partial derivative $D_j f_k(\mathbf{a})$. Thus, we have

$$D\mathbf{f}(\mathbf{a}) = \begin{bmatrix} D_1 f_1(\mathbf{a}) & D_2 f_1(\mathbf{a}) & \dots & D_n f_1(\mathbf{a}) \\ D_1 f_2(\mathbf{a}) & D_2 f_2(\mathbf{a}) & \dots & D_n f_2(\mathbf{a}) \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ D_1 f_m(\mathbf{a}) & D_2 f_m(\mathbf{a}) & \dots & D_n f_m(\mathbf{a}) \end{bmatrix}.$$

The Jacobian matrix of $\mathbf{f}$ at $\mathbf{a}$
Cont...

- The Jacobian matrix $D\mathbf{f}(\mathbf{a})$ is defined at each point where the mn partial derivatives $D_j f_k(\mathbf{a})$ exists.
- The total derivative $\mathbf{T_a}$ is also written as $\mathbf{f}'(\mathbf{a})$.
- The derivative $\mathbf{f}'(\mathbf{a})$ is a linear transformation; the Jacobian $D\mathbf{f}(\mathbf{a})$ is a matrix representation for this transformation.

The Jacobian matrix of $\mathbf{f}$ at $\mathbf{a}$
Cont...

The first-order Taylor formula takes the form

$$\mathbf{f}(\mathbf{a}+\mathbf{v}) = \mathbf{f}(\mathbf{a}) + \mathbf{f}'(\mathbf{a})(\mathbf{v}) + \|\mathbf{v}\|\mathbf{E}(\mathbf{a},\mathbf{v}), \tag{4}$$

where $\mathbf{E}(\mathbf{a},\mathbf{v}) \to \mathbf{0}$ as $\mathbf{v} \to \mathbf{0}$. This resembles the one-dimensional Taylor formula. To compute the components of the vector $\mathbf{f}'(\mathbf{a})(\mathbf{v})$ we can use the matrix product $D\mathbf{f}(\mathbf{a})\mathbf{v}$ or formula (3) of Theorem (9.1).

Theorem (9.2)
Differentiability implies continuity

If a vector field $\mathbf{f}$ is differentiable at $\mathbf{a}$, then $\mathbf{f}$ is continuous at $\mathbf{a}$.

As in the scalar case, we use the Taylor formula to prove this theorem.

If we let $\mathbf{v} \to \mathbf{0}$ in the first-order Taylor formula,

$$\lim_{\mathbf{v}\to\mathbf{0}} \mathbf{f}(\mathbf{a}+\mathbf{v}) = \lim_{\mathbf{v}\to\mathbf{0}} \mathbf{f}(\mathbf{a}) + \lim_{\mathbf{v}\to\mathbf{0}} \mathbf{f}'(\mathbf{a})(\mathbf{v}) + \lim_{\mathbf{v}\to\mathbf{0}} \|\mathbf{v}\|\mathbf{E}(\mathbf{a},\mathbf{v}).$$

The error term $\|\mathbf{v}\|\mathbf{E}(\mathbf{a},\mathbf{v}) \to \mathbf{0}$.

The linear part $\mathbf{f}'(\mathbf{a})(\mathbf{v})$ also trends to $\mathbf{0}$ because linear transformations are continuous at $\mathbf{0}$.

This completes the proof.

Theorem (9.3)
Chain rule

Let $\mathbf{f}$ and $\mathbf{g}$ be vector fields such that the composition $\mathbf{h} = \mathbf{f} \circ \mathbf{g}$ is defined in a neighborhood of a point $\mathbf{a}$. Assume that $\mathbf{g}$ is differentiable at $\mathbf{a}$, with total derivative $\mathbf{g}'(\mathbf{a})$. Let $\mathbf{b} = \mathbf{g}(\mathbf{a})$ and assume that $\mathbf{f}$ is differentiable at $\mathbf{b}$, with total derivative $\mathbf{f}'(\mathbf{b})$. Then $\mathbf{h}$ is differentiable at $\mathbf{a}$, and the total derivative $\mathbf{h}'(\mathbf{a})$ is given by

$$\mathbf{h}'(\mathbf{a}) = \mathbf{f}'(\mathbf{b}) \circ \mathbf{g}'(\mathbf{a}),$$

the composition of the linear transformations $\mathbf{f}'(\mathbf{b})$ and $\mathbf{g}'(\mathbf{a})$.

Matrix form of the chain rule

Let $\mathbf{h} = \mathbf{f} \circ \mathbf{g}$, where $\mathbf{g}$ is differentiable at $\mathbf{a}$ and $\mathbf{f}$ if differentiable at $\mathbf{b} = \mathbf{g}(\mathbf{a})$. The chain rule states that

$$\mathbf{h}'(\mathbf{a}) = \mathbf{f}'(\mathbf{b}) \circ \mathbf{g}'(\mathbf{a}).$$

We can express the chain rule in terms of the Jacobian matrices $D\mathbf{h}(\mathbf{a})$, $D\mathbf{f}(\mathbf{b})$, and $D\mathbf{g}(\mathbf{a})$ which represent the linear transformations $\mathbf{h}'(\mathbf{a})$, $\mathbf{f}'(\mathbf{b})$, and $\mathbf{g}'(\mathbf{a})$, respectively.

Matrix form of the chain rule
Cont...

Since composition of linear transformations corresponds to multiplication of their matrices, we obtain

$$D\mathbf{h}(\mathbf{a}) = D\mathbf{f}(\mathbf{b})D\mathbf{g}(\mathbf{a}), \quad \text{where } \mathbf{b} = \mathbf{g}(\mathbf{a}). \tag{5}$$

This is called the **matrix form of the chain rule**.

Matrix form of the chain rule
Cont...

It can also be written as a set of scalar equations by expressing each matrix in terms of its entries.

Suppose that $\mathbf{a} \in \mathbb{R}^m$, $\mathbf{b} = \mathbf{g}(\mathbf{a}) \in \mathbb{R}^n$, and $\mathbf{f}(\mathbf{b}) \in \mathbb{R}^m$.

Then $\mathbf{h}(\mathbf{a}) \in \mathbb{R}^m$ and we can write

$$\mathbf{g} = (g_1, ..., g_n),\ \mathbf{f} = (f_1, ..., f_m),\ \mathbf{h} = (h_1, ..., h_m).$$

Matrix form of the chain rule
Cont...

Then $D\mathbf{h}(\mathbf{a})$ is an $m \times p$ matrix, $D\mathbf{f}(\mathbf{b})$ is an $m \times n$ matrix, and $D\mathbf{g}(\mathbf{a})$ is an $n \times p$ matrix, given by

$$D\mathbf{h}(\mathbf{a}) = [D_j h_i(\mathbf{a})]_{i,j=1}^{m,p},$$
$$D\mathbf{f}(\mathbf{b}) = [D_k f_i(\mathbf{b})]_{i,k=1}^{m,n},$$
$$D\mathbf{g}(\mathbf{a}) = [D_j g_k(\mathbf{a})]_{k,j=1}^{n,p}.$$

Matrix form of the chain rule
Cont...

The matrix equation (5) is equivalent to mp scalar equations,

$$D_j h_i(\mathbf{a}) = \sum_{k=1}^{n} D_k f_i(\mathbf{b}) D_j g_k(\mathbf{a}), \text{ for } i = 1, 2, ..., m \text{ and } j = 1, 2, ..., p.$$

These equations express the partial derivatives of the components of **h** in terms of the partial derivatives of the components of **f** and **g**.

Example 1

Extended chain rule for scalar fields

Suppose f is a scalar field ($m = 1$). Then h is also a scalar field and there are p equations in the chain rule, one for each of the partial derivatives of h:

$$D_j h(\mathbf{a}) = \sum_{k=1}^{n} D_k f(\mathbf{b}) D_j g_k(\mathbf{a}), \text{ for } j = 1, 2, ..., p.$$

The special case $p = 1$ was already considered in the Chapter of "A Chain Rule for Derivatives of Scalar Fields". In this case we get only one equation,

$$h'(\mathbf{a}) = \sum_{k=1}^{n} D_k f(\mathbf{b}) g_k'(\mathbf{a}).$$

Now take $p = 2$ and $n = 2$. Write $\mathbf{a} = (s, t)$ and $\mathbf{b} = (x, y)$. Then the components x and y are related to s and t by the equations

$$x = g_1(s, t), \qquad y = g_2(s, t).$$

The chain rule gives a pair of equations for the partial derivatives of h:

$$\begin{aligned} D_1 h(s, t) &= D_1 f(x, y) D_1 g_1(s, t) + D_2 f(x, y) D_1 g_2(s, t), \\ D_2 h(s, t) &= D_1 f(x, y) D_2 g_1(s, t) + D_2 f(x, y) D_2 g_2(s, t). \end{aligned}$$

In the ∂-notation, this pair of equations is usually written as

$$\frac{\partial h}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} \tag{6}$$

$$\frac{\partial h}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}. \tag{7}$$

Example 2
Polar coordinates

The temperature of a thin plate is described by a scalar field f, the temperature at (x, y) being $f(x, y)$. Polar coordinates $x = r\cos\theta$, $y = r\sin\theta$ are introduced, and the temperature becomes a function of r and θ determined by the equation

$$\varphi(r, \theta) = f(r\cos\theta, r\sin\theta).$$

Express the partial derivatives $\partial\varphi/\partial r$ and $\partial\varphi/\partial\theta$ in terms of the partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$.

We use the chain rule as expressed in Equations (6) and (7), writing (r, θ) instead of (s, t), and φ instead of h. The equations

$$x = r\cos\theta,\ \ y = r\sin\theta$$

gives us

$$\frac{\partial x}{\partial r} = \cos\theta, \frac{\partial y}{\partial r} = \sin\theta, \frac{\partial x}{\partial \theta} = -r\sin\theta, \frac{\partial y}{\partial \theta} = r\cos\theta.$$

Substituting these formulas in (6) and (7) we obatain

$$\frac{\partial \varphi}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta, \quad (8)$$

$$\frac{\partial \varphi}{\partial \theta} = -r\frac{\partial f}{\partial x}\sin\theta + r\frac{\partial f}{\partial y}\cos\theta. \quad (9)$$

These are the required formulas for $\frac{\partial \varphi}{\partial r}$ and $\frac{\partial \varphi}{\partial \theta}$.

Example 3
Second-order partial derivatives

Refer to Example 2 and express the second-order partial derivatives $\dfrac{\partial^2 \varphi}{\partial \theta^2}$ in terms of partial derivatives of f.

We begin with the formula for $\dfrac{\partial\varphi}{\partial\theta}$ in (9) and differentiate with respect to θ, treating r as a constant. There are two terms on the right, each of which must be differentiated as a product. Thus we have

$$\begin{aligned}
\frac{\partial^2\varphi}{\partial\theta^2} &= -r\frac{\partial f}{\partial x}\frac{\partial(\sin\theta)}{\partial\theta} - r\sin\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial x}\right) + r\frac{\partial f}{\partial y}\frac{\partial(\cos\theta)}{\partial\theta} \\
&\quad + r\cos\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right) \\
\frac{\partial^2\varphi}{\partial\theta^2} &= -r\cos\theta\frac{\partial f}{\partial x} - r\sin\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial x}\right) - r\sin\theta\frac{\partial f}{\partial y} \\
&\quad + r\cos\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right). \qquad (10)
\end{aligned}$$

To compute the derivatives of $\partial f/\partial x$ and $\partial f/\partial y$ with respect to θ we must keep in mind that, as functions of r and θ, $\partial f/\partial x$ and $\partial f/\partial y$ are composite functions given by

$$\frac{\partial f}{\partial x} = D_1 f(r\cos\theta, r\sin\theta) \text{ and } \frac{\partial f}{\partial y} = D_2 f(r\cos\theta, r\sin\theta).$$

Therefore, thier derivatives with respect to θ must be determined by use of the chain rule. We again use (6) and (7) with f replaced by $D_1 f$, to obtain

$$\begin{aligned} \frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial x}\right) &= \frac{\partial(D_1 f)}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial(D_1 f)}{\partial y}\frac{\partial y}{\partial\theta} \\ &= \frac{\partial^2 f}{\partial x^2}(-r\sin\theta) + \frac{\partial^2 f}{\partial y\partial x}(r\cos\theta). \end{aligned}$$

Example 3

Second-order partial derivatives⇒Cont...

Similarly, using (6) and (7) with f replaced by $D_2 f$, we find

$$\begin{aligned} \frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right) &= \frac{\partial(D_2 f)}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial(D_2 f)}{\partial y}\frac{\partial y}{\partial\theta} \\ &= \frac{\partial^2 f}{\partial x \partial y}(-r\sin\theta) + \frac{\partial^2 f}{\partial y^2}(r\cos\theta). \end{aligned}$$

When these formulas are used in (10) we obtain

$$\begin{aligned} \frac{\partial^2\varphi}{\partial\theta^2} &= -r\cos\theta\frac{\partial f}{\partial x} + r^2\sin^2\theta\frac{\partial^2 f}{\partial x^2} - r^2\sin\theta\cos\theta\frac{\partial^2 f}{\partial y\partial x} \\ &\quad -r\sin\theta\frac{\partial f}{\partial y} - r^2\sin\theta\cos\theta\frac{\partial^2 f}{\partial x\partial y} + r^2\cos^2\theta\frac{\partial^2 f}{\partial y^2}. \end{aligned}$$

This is the required formula for $\dfrac{\partial^2\varphi}{\partial\theta^2}$.

Thank you!