Real Analysis III
(MAT312β)

Department of Mathematics
University of Ruhuna

A.W.L. Pubudu Thilan

A Chain Rule for Derivatives of Scalar Fields

A function of a function

- Consider the expression $\sin t^2$.
- It is clear that this is different from the straightforward sine function, $\sin t$.
- We are finding the sine of t^2, not simply the sine of t.
- We call such an expression a "function of a function" or a "composite function".

- Suppose, in general, that we have two functions, $f(t)$ and $r(t)$.
- Then $g(t) = f[r(t)]$ is a function of a function.
- In our case, the function f is the sine function and the function r is the square function.
- We could identify them more mathematically by saying that $f(t) = \sin t$ and $r(t) = t^2$, so that $f[r(t)] = f(t^2) = \sin t^2$.

The chain rule in one-dimensional space

In one-dimensional derivative theory, the chain rule enables us to compute the derivative of a function of a function $g(t) = f[r(t)]$ by the formula

$$g'(t) = f'[r(t)].r'(t).$$

(i) $y = \sin x^2$

(ii) $y = (2x - 3)^{12}$

(iii) $y = e^{x^3}$

(iv) $y = e^{1+x^2}$

(v) $y = \sin(x + e^x)$

The chain rule for derivatives of scalar fields

This Chapter provides an extension of the formula when f is replaced by a scalar field defined on a set in n-space and r is replaced by a vector-valued function of a real variable with values in the domain of f.

The chain rule for derivatives of scalar fields
Cont...

- It is easy to conceive of examples in which the composition of a scalar field and a vector field might arise.
- For instance, suppose $f(\mathbf{x})$ measures the temperature at a point $\mathbf{x}$ of a solid in 3-space, and suppose we wish to know how the temperature changes as the point $\mathbf{x}$ varies along a curve C lying in the solid.
- If the curve is described by a vector-valued function $\mathbf{r}$ defined on an interval $[a, b]$, we can introduce a new function g by means of the formula

$$g(t) = f[\mathbf{r}(t)] \quad \text{if } a \leq t \leq b.$$

- This composite function g expresses the temperature as a function of the parameter t, and its derivative $g'(t)$ measures the rate of chage of the temperature along the curve.
- The following extension of the chain rule enables us to compute the derivative $g'(t)$ without determining $g(t)$ explicitly.

Theorem 8.1
Chain rule

Let f be a scalar field defined on an open set $\mathbb{S}$ in $\mathbb{R}^n$, and let $\mathbf{r}$ be a vector-valued function which maps an interval $\mathbb{J}$ from $\mathbb{R}^1$ into $\mathbb{S}$. Define the composite function $g = f \circ \mathbf{r}$ on $\mathbb{J}$ by the equation

$$g(t) = f[\mathbf{r}(t)] \quad \text{if } t \in \mathbb{J}.$$

Let t be a point in $\mathbb{J}$ at which $\mathbf{r}'(t)$ exists and assume that f is differentiable at $\mathbf{r}(t)$. Then $g'(t)$ exists and is equal to the dot product

$$g'(t) = \nabla f(\mathbf{a}).\mathbf{r}'(t), \text{ where } \mathbf{a} = \mathbf{r}(t). \tag{1}$$

Let $\mathbf{a} = \mathbf{r}(t)$, where t is a point in $\mathbb{J}$ at which $\mathbf{r}'(t)$ exists.

Since $\mathbb{S}$ is open there is an n-ball $\mathbf{B}(\mathbf{a})$ lying in $\mathbb{S}$.

We take $h \neq 0$ but small enough so that $\mathbf{r}(t+h)$ lies in $\mathbf{B}(\mathbf{a})$, and we let $\mathbf{y} = \mathbf{r}(t+h) - \mathbf{r}(t)$.

Note that $\mathbf{y} \to \mathbf{0}$ as $h \to 0$.

Now we have

$$\begin{aligned} g(t+h) - g(t) &= f[\mathbf{r}(t+h)] - f[\mathbf{r}(t)] \\ &= f(\mathbf{a}+\mathbf{y}) - f(\mathbf{a}). \end{aligned} \tag{2}$$

Applying the first-order Taylor formula for f we have

$$f(\mathbf{a}+\mathbf{y})-f(\mathbf{a})=\nabla f(\mathbf{a}).\mathbf{y}+\|\mathbf{y}\|E(\mathbf{a},\mathbf{y}), \tag{3}$$

where $E(\mathbf{a},\mathbf{y})\to 0$ as $\|\mathbf{y}\|\to 0$.

From (2) and (3) we have

$$g(t+h)-g(t)=\nabla f(\mathbf{a}).\mathbf{y}+\|\mathbf{y}\|E(\mathbf{a},\mathbf{y}).$$

Since $\mathbf{y}=\mathbf{r}(t+h)-\mathbf{r}(t)$ this gives us

$$\begin{aligned}\frac{g(t+h)-g(t)}{h} &= \nabla f(\mathbf{a}).\frac{\mathbf{r}(t+h)-\mathbf{r}(t)}{h}\\ &\quad+\frac{\|\mathbf{r}(t+h)-\mathbf{r}(t)\|}{h}E(\mathbf{a},\mathbf{y})\end{aligned}$$

By letting $h \to 0$ we obtain:

$$
\begin{aligned}
\lim_{h\to 0} \frac{g(t+h)-g(t)}{h} &= \lim_{h\to 0} \nabla f(\mathbf{a}).\frac{\mathbf{r}(t+h)-\mathbf{r}(t)}{h} \\
&\quad + \lim_{h\to 0} \frac{\|\mathbf{r}(t+h)-\mathbf{r}(t)\|}{h} E(\mathbf{a},\mathbf{y}) \\
\lim_{h\to 0} \frac{g(t+h)-g(t)}{h} &= \nabla f(\mathbf{a}). \lim_{h\to 0} \frac{\mathbf{r}(t+h)-\mathbf{r}(t)}{h} + 0 \\
g'(t) &= \nabla f(\mathbf{a}).\mathbf{r}'(t).
\end{aligned}
$$

Example 1
Directional derivative along a curve

- When the function $\mathbf{r}$ describes a curve C, the derivative $\mathbf{r}'$ is the velocity vector (tangent to the curve) and derivative g' in Equation (1) is the derivative of f with respect to the velocity vector, assuming that $\mathbf{r}' \neq \mathbf{0}$.

- If $\mathbf{T}(t)$ is a unit vector in the direction of $\mathbf{r}'(t)$ ($\mathbf{T}$ is the unit tangent vector), the dot product $\nabla f[\mathbf{r}(t)].\mathbf{T}(t)$ is called the directional derivative of f along the curve C or in the direction of C.

- For a plane curve we can write

$$\mathbf{T}(t) = \cos\alpha(t)\mathbf{i} + \cos\beta(t)\mathbf{j},$$

where $\alpha(t)$ and $\beta(t)$ are the angles made by the vector $\mathbf{T}(t)$ and the positive x- and y-axes; the directional derivative of f along C becomes

$$\nabla f[\mathbf{r}(t)].\mathbf{T}(t) = D_1 f[\mathbf{r}(t)]\cos\alpha(t) + D_2 f[\mathbf{r}(t)]\cos\beta(t).$$

- This formula is often written more briefly as

$$\nabla f.\mathbf{T} = \frac{\partial f}{\partial x}\cos\alpha + \frac{\partial f}{\partial y}\cos\beta.$$

- Since the directional derivative along C is defined in terms of **T**, its value depends on the parametric representation chosen for C.

- A change of the representation could reverse the direction of **T**; this in turn, would reverse the sign of the directional derivative.

Example 2

Find the directional derivative of the scalar field $f(x, y) = x^2 - 3xy$ along the parabola $y = x^2 - x + 2$ at the point (1,2).

Example 2
Cont...

At an arbitary point (x, y) the gradient vector is

$$\begin{aligned}\nabla f(x,y) &= \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} \\ &= (2x-3y)\mathbf{i} - 3x\mathbf{j}.\end{aligned}$$

At the point (1,2) we have $\nabla f(1,2) = -4\mathbf{i} - 3\mathbf{j}$.

The parabola can be represented parametrically by the vector equation $\mathbf{r}(t) = t\mathbf{i} + (t^2 - t + 2)\mathbf{j}$.

Therefore $r(1) = \mathbf{i} + 2\mathbf{j}$, $\mathbf{r}'(t) = \mathbf{i} + (2t-1)\mathbf{j}$, and $\mathbf{r}'(1) = \mathbf{i} + \mathbf{j}$.

For this representation of C the unit tangent vector $\mathbf{T}(1)$ is $(\mathbf{i}+\mathbf{j})/\sqrt{2}$ and the required directional derivative is $\nabla f(1,2).\mathbf{T}(1) = -7/\sqrt{2}$.

Example 3

Let f be nonconstant scalar field, differentiable everywhere in the plane, and let c be a constant. Assume the Cartesian equation $f(x, y) = c$ describes a curve C having a tangent at each of its points. Prove that f has the following properties at each point of C:

(a) The gradient vector ∇f is normal to C.

(b) The directional derivative of f is zero along C.

(c) The directional derivative of f has its largest value in a direction normal to C.

If $\mathbf{T}$ is a unit tangent vector to C, the directional derivative of f along C is the dot product $\nabla f.\mathbf{T}$.

This product is zero if ∇f is perpendicular to $\mathbf{T}$, and it has its largest value if ∇f is parallel to $\mathbf{T}$.

Therefore both statements (b) and (c) are consequences of (a).

To prove (a), consider any plane curve Γ with a vector equation of the form $\mathbf{r}(t) = X(t)\mathbf{i} + Y(t)\mathbf{j}$ and introduce the function $g(t) = f[\mathbf{r}(t)]$.

Example 3
Cont...

By the chain rule we have $g'(t) = \nabla f[\mathbf{r}(t)].\mathbf{r}'(t)$.

When $\Gamma = C$, the function g has the constant value c so $g'(t) = 0$ if $\mathbf{r}(t) \in C$.

Since $g' = \nabla f.\mathbf{r}'$, this shows that ∇f is perpendicular to $\mathbf{r}'$ on C; hence ∇f is normal to C.

Level sets

Let f be a scalar field defined on a set $\mathbb{S}$ in $\mathbb{R}^n$ and consider those points $\mathbf{x}$ in $\mathbb{S}$ for which $f(\mathbf{x})$ has a constant value, say $f(\mathbf{x}) = c$. Denote this set by $L(c)$, so that

$$L(c) = \{\mathbf{x} | \mathbf{x} \in \mathbb{S} \text{ and } f(\mathbf{x}) = c\}.$$

The set $L(c)$ is called a level set of f. In $\mathbb{R}^2$, $L(c)$ is called a level curve; in $\mathbb{R}^3$, it is called a level surface.

- A level curve of a function $f(x, y)$ is the curve of points (x, y) where $f(x, y)$ is some constant value.
- A level curve is simply a cross section of the graph of $z = f(x, y)$ taken at a constant value, say $z = c$.
- A function has many level curves, as one obtains a different level curve for each value of c in the range of $f(x, y)$.
- We can plot the level curves for a bunch of different constants c together in a level curve plot, which is sometimes called a contour plot.

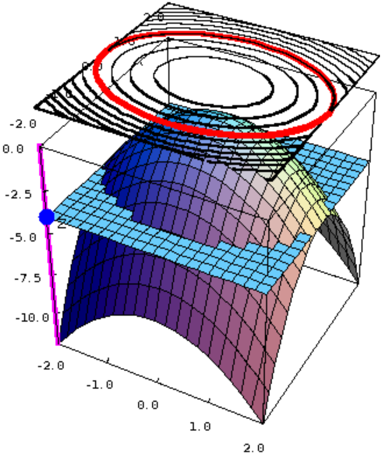
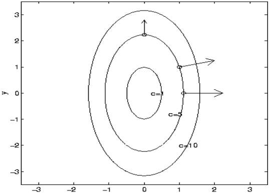


Figure: The graph of the function $f(x,y) = -x^2 - 2y^2$ is shown along with a level curve plot.

- Consider $z = f(x, y) = 4x^2 + y^2$.
- The figure below shows the level curves, defined by $f(x, y) = c$, of the surface.
- The level curves are the ellipses $4x^2 + y^2 = c$.
- As the plot shows, the gradient vector at (x, y) is normal to the level curve through (x, y).



- Now consider a scalar field f differentiable on an open set $\mathbb{S}$ in $\mathbb{R}^3$, and examine one of its level surfaces, $L(c)$.
- Let $\mathbf{a}$ be a point on this surface, and consider a curve Γ which lies on the surface and passes through $\mathbf{a}$.
- We shall prove that the gradient vector $\nabla f(\mathbf{a})$ is normal to this curve at $\mathbf{a}$.

- That is, we shall prove that $\nabla f(\mathbf{a})$ is perpendicular to the tangent vector of Γ at $\mathbf{a}$.
- For this purpose we assume that Γ is described parametrically by a differentiable vector-valued function $\mathbf{r}$ defined on some interval $\mathbb{j}$ in $\mathbb{R}^1$.
- Since Γ lies on the level surface $L(c)$, the function $\mathbf{r}$ satisfies the equation

$$f[\mathbf{r}(t)] = c \text{ for all } t \text{ in } \mathbb{j}.$$

- If $g(t) = f[\mathbf{r}(t)]$ for t in $\mathbb{j}$, the chain rule states that

$$g'(t) = \nabla f[\mathbf{r}(t)].\mathbf{r}'(t).$$

- Since g is a constant on $\mathbb{j}$, we have $g'(t) = 0$ on $\mathbb{j}$. In particular, choosing t_1 so that $\mathbf{r}(t_1) = \mathbf{a}$, we find that

$$\nabla f(\mathbf{a}).\mathbf{r}'(t_1) = 0.$$

- In other words, the gradient of f at $\mathbf{a}$ is perpendicular to the tangent vector $\mathbf{r}'(t_1)$, as asserted.

- Now we take family of curves on the level surface $L(c)$, all passing through the point $\mathbf{a}$.

- According to the foregoing discussion, the tangent vectors of all these curves are perpendicular to the gradient vector $\nabla f(\mathbf{a})$.

- If $\nabla f(\mathbf{a})$ is not the zero vector, these tangent vectors determine a plane, and the gradient $\nabla f(\mathbf{a})$ is normal to this plane.

- This particular plane is called as the tangent plane of the surface $L(c)$ at $\mathbf{a}$.

- We know that a plane through $\mathbf{a}$ with normal vector N consists of all points $\mathbf{x} \in \mathbb{R}^3$ satisfying $N.(\mathbf{x} - \mathbf{a}) = 0$.
- Therefore the tangent plane to the level surface $L(c)$ at $\mathbf{a}$ consists of all $\mathbf{x}$ in $\mathbb{R}^3$ satisfying

$$\nabla f(\mathbf{a}).(\mathbf{x} - \mathbf{a}) = 0.$$

- To obtain a Cartesian equation for this plane we express $\mathbf{x}$, $\mathbf{a}$, and $\nabla f(\mathbf{a})$ in terms of thier components.

- Writing $\mathbf{x} = (x, y, z)$, $\mathbf{a} = (x_1, y_1, z_1)$ and

$$\nabla f(\mathbf{a}) = D_1 f(\mathbf{a})\mathbf{i} + D_2 f(\mathbf{a})\mathbf{j} + D_3 f(\mathbf{a})\mathbf{k},$$

we obtain the Cartesian equation

$$D_1 f(\mathbf{a})(x - x_1) + D_2 f(\mathbf{a})(y - y_1) + D_3 f(\mathbf{a})(z - z_1) = 0.$$

- A similar discussion applies to a scalar fields defined in $\mathbb{R}^2$.
- In Example 3 we proved that the gradient vector $\nabla f(\mathbf{a})$ at a point $\mathbf{a}$ of a level curve is perpendicular to the tangent vector of the curve at $\mathbf{a}$.
- Therefore the tangent line of the level curve $L(c)$ at a point $\mathbf{a} = (x_1, y_1)$ has the Cartesian equation

$$D_1 f(\mathbf{a})(x - x_1) + D_2 f(\mathbf{a})(y - y_1) = 0.$$

The equation of the tangent plane

Consider the surface $z = f(x, y)$. If $Z = f(X, Y)$, then $(X, Y, Z)^T$ is a point on the surface $z = f(x, y)$. If the surface admits a non vertical tangent plane at $(X, Y, Z)^T$, then we say that f is differentiable at $(X, Y)^T$.

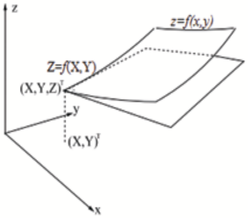


Figure: The tangent plane

The equation of the tangent plane
Cont...

If f is differentiable at $(X, Y)^T$ its tangent plane must have equation

$$z - Z = f_x(X, Y)(x - X) + f_y(X, Y)(y - Y).$$

We usually write this in the less precise form

$$z - Z = \frac{\partial f}{\partial x}(X, Y)(x - X) + \frac{\partial f}{\partial y}(X, Y)(y - Y).$$

N.B Partial derivatives are to be evaluated at the point $(X, Y)^T$.

Example

Let $f(x, y) = \dfrac{x - y}{x + y}$.

(a) Compute $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$.

(b) Find the equation of the tangent plane to the surface $z = f(x, y)$ where $x = 1$ and $y = 1$.

(a)

$$
\begin{aligned}
f(x,y) &= \frac{x-y}{x+y} \\
\frac{\partial f}{\partial x} &= \frac{(x+y).1-(x-y).1}{(x+y)^2} \\
&= \frac{2y}{(x+y)^2}. \\
\frac{\partial f}{\partial y} &= \frac{(x+y).(-1)-(x-y).1}{(x+y)^2} \\
&= \frac{-2x}{(x+y)^2}.
\end{aligned}
$$

(b) The tangent plane must have equation

$$z - Z = f_x(X, Y)(x - X) + f_y(X, Y)(y - Y).$$

The equation of the tangent plane to the surface $z = f(x, y)$, where $X = 1$ and $Y = 1$ is

$$z - Z = f_x(1, 1)(x - 1) + f_y(1, 1)(y - 1),$$

where $Z = f(1, 1)$. The required equation is

$$z = \frac{1}{2}(x - 1) - \frac{1}{2}(y - 1).$$

Thank you!