MARCH 21 LECTURE

TEXTBOOK REFERENCE:
- *Vector Calculus*, Colley, 4th Edition: §2.3, 2.4

PARTIAL DERIVATIVES

LEARNING OBJECTIVES:
- Understand what it means for a function of several variables to be differentiable.
- Learn how to compute the matrix of partial derivatives.
- Understand the definition and basic properties of the derivative of a vector-valued function of several variables.
- Learn how to compute higher order partial derivatives.

KEYWORDS: differentiability, matrix of partial derivatives, the derivative, mixed partial derivatives

Differentiability
Let \( f : X \subseteq \mathbb{R}^2 \rightarrow \mathbb{R} \) be a function of two variables, \((a,b) \in X\). Suppose that the partial derivatives of \( f \) at \((a,b)\) exist. The linearisation of \( f \), \( L(x,y) \), is the function
\[
L(x,y) = f(a,b) + f_x(a,b)(x - a) + f_y(a,b)(y - b)
\]
and the tangent plane to the graph of \( f \) at \((a,b,f(a,b))\) is defined by the equation
\[
z = f(a,b) + f_x(a,b)(x - a) + f_y(a,b)(y - b).
\]
Differentiability of $f(x,y)$

Let $f : X \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ be a function of two variables. We say that $f$ is **differentiable at** $a = (a,b) \in X$ if the partial derivatives $f_x(a,b)$, $f_y(a,b)$ exist and if

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

If $f$ is differentiable for every $a \in X$ then we say that $f$ is **differentiable**.

In words, $f$ is differentiable at $(a,b)$ if $L(x,y)$ provides a ‘good’ approximation of $f(x,y)$ near to $(a,b)$.

**Remark:**

1. Analytically, ‘good’ means that $f(x) - L(x)$ goes to 0 faster than $|x - a|$.
2. This definition of differentiability extends to scalar-valued functions of $n$ variables $f : X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$.

**Example:** Consider the function

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x,y) \mapsto 10 - x^2 - y^2$$

Then, the linearisation of $f$ at $a = (a,b)$ is

$$L(x,y) = 10 - a^2 - b^2 - 2a(x - a) - 2b(y - b)$$

We have

$$f(x,y) - L(x,y) = a^2 - x^2 + b^2 - y^2 + 2a(x - a) + 2b(y - b)$$

$$= -(x - a)^2 - (y - b)^2$$

Then,

$$\frac{f(x,y) - L(x,y)}{|x - a|} = -\left( \frac{(x-a)^2 + (y-b)^2}{\sqrt{(x-a)^2 + (y-b)^2}} \right) = -\sqrt{(x-a)^2 + (y-b)^2}$$

It is now not too difficult to see that

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

Hence, $f$ is differentiable.

**Exercise:** show that

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

using $\epsilon - \delta$ definition.
**Remark:** Geometrically, \( f \) is differentiable if its graph does not have any ‘corners’.

**Sufficient Condition for differentiability**

Let \( f : X \subseteq \mathbb{R}^2 \rightarrow \mathbb{R} \) be a function of two variables, \((a, b) \in X\). If the partial derivatives \( f_x(x, y) \) and \( f_y(x, y) \) are continuous in a sufficiently small disk centred at \((a, b)\) then \( f \) is differentiable at \((a, b)\).

**Necessary Condition for differentiability**

Let \( f : X \subseteq \mathbb{R}^2 \rightarrow \mathbb{R} \) be a function of two variables, \((a, b) \in X\). If \( f \) is differentiable at \((a, b)\) then \( f \) is continuous at \((a, b)\).

**Example:**

1. Consider the function \( f(x, y) = 2xy + \cos(y^2 + x^2) \). Then,

\[
    f_x(x, y) = 2y - 2x \sin(y^2 + x^2),
\]

\[
    f_y(x, y) = 2x - 2y \cos(y^2 + x^2).
\]

Both the partial derivatives are continuous - use the Algebraic Properties of Continuous Functions (p.111 of Colley). Hence, \( f \) is differentiable.

2. Consider the function \( f(x, y) = \frac{x^3 + 5y^4}{1 + x^2 + y^2} \), defined for all \((x, y) \in \mathbb{R}^2\). Then,

\[
    f_x(x, y) = \frac{3x^2 + x^4 + 3(xy)^2 - 10xy^4}{(1 + x^2 + y^2)^2}
\]

\[
    f_y(x, y) = \frac{20y^2 + 20x^2y^3 + 10y^5 - 2yx^3}{(1 + x^2 + y^2)^2}
\]

Both of those functions are continuous - they are rational functions whose numerator/denominator are polynomial functions are continuous. Hence, \( f(x, y) \) is differentiable.

3. Consider the function

\[
    f(x, y) = \begin{cases} 
        \frac{x^2y^2}{x^4+y^4}, & (x, y) \neq (0, 0) \\
        0, & (x, y) = (0, 0)
    \end{cases}
\]

The limit does not exist as \((x, y) \rightarrow (0, 0)\) (Exercise!). Hence, \( f(x, y) \) can’t be continuous at \((0, 0)\).
However, the partial derivatives
\[
\frac{\partial f}{\partial x}(0, 0) = \lim_{h \to 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \to 0} 0 = 0
\]
and
\[
\frac{\partial f}{\partial y}(0, 0) = \lim_{h \to 0} \frac{f(0, h) - f(0, 0)}{h} = \lim_{h \to 0} 0 = 0
\]
do exist. Hence, we see that we require a stronger condition than existence of partial derivatives to ensure differentiability of \( f \) at \((0, 0)\).

**Higher order partial derivatives**

It’s possible to ‘mix and match’ partial derivatives: given a function \( f : X \subset \mathbb{R}^n \to \mathbb{R} \) we know how to compute \( \frac{\partial f}{\partial x_i} \). We may now compute the partial derivative of this function with respect to any of the \( n \) variables \( x_1, \ldots, x_n \).

For example, if \( f(x, y, z) = x^2 - 2yz^3 + \frac{3xy^2}{z} \). Then,
\[
\frac{\partial f}{\partial x} = 2x + \frac{3y^2}{z}, \quad \frac{\partial f}{\partial y} = -2z^3 + \frac{6xy}{z}, \quad \frac{\partial f}{\partial z} = -6yz^2 - \frac{3y^2}{z^2}
\]
We may now compute the partial derivatives of each of these functions with respect to \( x, y, z \) (to obtain a total of nine new functions). We call these functions **second order (or mixed) partial derivatives** of \( f \):

\[
\frac{\partial^2 f}{\partial x^2} \overset{\text{def}}{=} \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial x} \right) = 2
\]
\[
\frac{\partial^2 f}{\partial x \partial y} \overset{\text{def}}{=} \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) = \frac{6y}{z}
\]
\[
\frac{\partial^2 f}{\partial x \partial z} \overset{\text{def}}{=} \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial z} \right) = -\frac{3y^2}{z^2}, \quad \cdots
\]

**Check your understanding**

Compute three of the remaining six second order partial derivatives.

---

**Clairaut’s Theorem**

Let \( f : X \subset \mathbb{R}^n \to \mathbb{R} \) be a function of \( n \) variables, \((a,b) \in X\). If all first order and second order partial derivatives exist and are continuous then, for any \( i, j = 1, \ldots, n \),
\[
\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}
\]
In words, **partial differentiation commutes**.
The Derivative

Let \( f : X \subseteq \mathbb{R}^2 \to \mathbb{R} \) be a function of two variables. The \textbf{gradient of \( f \) at \( a \)} is the (row) vector

\[
\nabla f(a) = \begin{bmatrix} f_x(a) & f_y(a) \end{bmatrix}
\]

\textbf{Observation:} we can write (1) as

\[
z = f(a) + \nabla f(a)(x - a), \quad x = \begin{bmatrix} x \\ y \end{bmatrix}
\]

(1*)

The product here is multiplication of the \( 1 \times 2 \) matrix \( \nabla f(a) \) with the \( 2 \times 1 \) matrix \( x - a \).

\textbf{Remark:}

1. Note the analogy with the equation of a tangent line of the graph of a single variable function:

\[
y = f(a) + f'(a)(x - a).
\]

2. If we consider the change of coordinates

\[
\hat{x} = x - a, \quad \hat{y} = y - b, \quad \hat{z} = z - f(a, b)
\]

then (1*) becomes

\[
\hat{z} = \nabla f(a)\hat{x}, \quad \hat{x} = \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix}
\]

3. The above remarks generalise to scalar-valued functions \( f : X \subseteq \mathbb{R}^n \to \mathbb{R} \), where we define the \textbf{gradient of \( f \) at \( a \)} to be the \( 1 \times n \) row vector

\[
\nabla f(a) = \begin{bmatrix} f_x(a) & f_x(a) & \cdots & f_x(a) \\ f_y(a) & f_y(a) & \cdots & f_y(a) \\ \vdots & \vdots & \ddots & \vdots \\ f_n(a) & f_n(a) & \cdots & f_n(a) \end{bmatrix}
\]

Suppose that \( f : X \subseteq \mathbb{R}^n \to \mathbb{R}^m \) is a \textbf{vector-valued} function, \( f(x) = (f_1(x), \ldots, f_m(x)) \), with each \( f_1, \ldots, f_m : X \subseteq \mathbb{R}^n \to \mathbb{R} \) a scalar-valued function.

Define the \textbf{matrix of partial derivatives of \( f \) at \( a \) in \( X \)}, or the \textbf{Jacobian matrix of \( f \) at \( a \)}, to be the \( m \times n \) matrix \( Df(a) \) having \( i \)-th row \( \nabla f_i(a) \):

\[
Df(a) = \begin{bmatrix}
\frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\
\frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_n}(a) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a)
\end{bmatrix}
\]

We write \( Df(x) \), or simple \( Df \), for the \( m \times n \) matrix whose \( ij \)-entry is \( \frac{\partial f_i}{\partial x_j}(x) \), and call it the \textbf{Jacobian of \( f \)}.

Define the \textbf{linearisation of \( f \) at \( a \) in \( X \)} to be the function

\[
L(x) = f(a) + Df(a)(x - a), \quad x \in \mathbb{R}^n
\]

The product here is multiplication of the \( m \times n \) matrix with the \( n \times 1 \) matrix \( x - a \). In particular, \( L(x) \in \mathbb{R}^m \).
Example: Consider the function
\[ f : \mathbb{R}^2 \to \mathbb{R}^3, (x, y) \mapsto (x^2 + y, 2xy, x + y^2) \]
Then,
\[ Df(x) = \begin{bmatrix} 2x & 1 \\ 2y & 2x \\ 1 & 2y \end{bmatrix} \]

**Differentiability of \( f(x) \)**
Let \( f : X \subseteq \mathbb{R}^n \to \mathbb{R}^m \) be a vector-valued function. We say that \( f \) is differentiable at \( a \in X \) if all partial derivatives \( f_{x_i}(a) \) exist and if
\[
\lim_{x \to a} \frac{f(x) - L(x)}{|x - a|} = 0
\]
If \( f \) is differentiable for every \( a \in X \) then we say that \( f \) is differentiable.

There are analogous results as for the two variable case.

**Sufficient Condition for differentiability**
Let \( f : X \subseteq \mathbb{R}^n \to \mathbb{R}^m, a \in X \). If all partial derivatives \( f_{x_i}(x) \) are continuous nearby to \( a \) then \( f \) is differentiable at \( a \).

**Necessary Condition for differentiability**
Let \( f : X \subseteq \mathbb{R}^n \to \mathbb{R}^m, a \in X \). If \( f \) is differentiable at \( a \) then \( f \) is continuous at \( a \).

Moreover, we can reduce differentiability of vector-valued functions to the differentiability of its component functions

Let \( f : X \subseteq \mathbb{R}^n \to \mathbb{R}^m, f(x) = (f_1(x), \ldots, f_m(x)), a \in X \). If \( f_1, \ldots, f_m \) are differentiable at \( a \) then \( f \) is differentiable at \( a \).

**What is the derivative?**
Observe the similarity between the linearisation of \( f \) at \( a \)
\[
L(x) = f(a) + Df(a)(x - a)
\]
and function whose graph is the tangent line of a single variable function \( f(x) \):
\[
L(x) = f(a) + f'(a)(x - a)
\]
Define the ‘multiplication by \( Df(a) \)’ linear map
\[
T_{Df(a)} : \mathbb{R}^n \to \mathbb{R}^m, \ x \mapsto Df(a)x
\]
then the linear map \( T_{Df(a)} \) plays the role of the derivative.

The derivative of a vector-valued function \( f \) of several variables is the linear map defined by the Jacobian of \( f \).