MATH 237 Calculus 3 for Honours Mathematics

Keven Qiu

Contents

1 Graphs of Scalar Function

1.1 Scalar Functions

Definition - Scalar Function: A scalar function $f(x_1, \ldots, x_n)$ of n variables is a function whose domain is a subset of $\mathbb{R}^n$ and whose range is a subset of $\mathbb{R}$.

1.2 Geometric Interpretation of $z = f(x, y)$

Definition - Level Curves: The level curves of a function $f(x, y)$ are the curves

$$f(x, y) = k$$

where the values of k come from the range of f.

Definition - Cross Sections: A cross section of a surface $z = f(x, y)$ is the intersection of $z = f(x, y)$ with a vertical plane.

Definition - Level Surfaces: A level surface of a scalar function $f(x, y, z)$ is defined by

$$f(x, y, z) = k, \ \ k \in R(f)$$

Definition - Level Sets: A level set of a scalar function $f(\vec{x})$, $\vec{x} \in \mathbb{R}^n$ is defined by

$$\{\vec{x} \in \mathbb{R}^n \mid f(\vec{x}) = k\}, \text{ for } k \in R(f)$$

Function Inventory:

Function	General Form	Level Curves	Cross-Sections
Plane	$f(x, y) = ax + by + c$	Parallel lines	Parallel lines
Parabolic cylinder	$f(x, y) = ax^2$	Parallel lines	Vertical lines or parabolas
Elliptic paraboloid	$f(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ where $a, b \in \mathbb{R}$ have same sign	Circles or ellipses	Parabolas
Hyperbolic paraboloid	$f(x, y) = \frac{x^2}{a^2} - \frac{y^2}{b^2}$ where $a, b \in \mathbb{R}$ have different signs	Hyperbolas	Parabolas

2 Limits

2.1 Definition of Limit

Definition - Neighbourhood: An r-neighbourhood of a point $(a, b) \in \mathbb{R}^2$ is a set

$$N_r(a, b) = \{(x, y) \in \mathbb{R}^2 \mid \|(x, y) - (a, b)\| < r\}, \ \ r \in \mathbb{R}$$

Essentially the area of the circle excluding the perimeter. Note that if $r < 0$, N_r is the empty set.

Remark: Recall that $\|(x,y) - (a,b)\|$ is the Euclidean distance in $\mathbb{R}^2$. That is,

$$\|(x,y) - (a,b)\| = \sqrt{(x-a)^2 + (y-b)^2}$$

Definition - Limit: Assume $f(x,y)$ is defined in a neighbourhood of (a,b), except possibly at (a,b). If, for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$0 < \|(x,y) - (a,b)\| < \delta \quad \text{implies} \quad |f(x,y) - L| < \epsilon$$

then

$$\lim_{(x,y)\to(a,b)} f(x,y) = L$$

2.2 Limit Theorems

Limit Theorem 1: If $\lim\limits_{(x,y)\to(a,b)} f(x,y)$ and $\lim\limits_{(x,y)\to(a,b)} g(x,y)$ both exist, then

a. $\lim\limits_{(x,y)\to(a,b)} [f(x,y) + g(x,y)] = \lim\limits_{(x,y)\to(a,b)} f(x,y) + \lim\limits_{(x,y)\to(a,b)} g(x,y)$.

b. $\lim\limits_{(x,y)\to(a,b)} [f(x,y)g(x,y)] = \left[\lim\limits_{(x,y)\to(a,b)} f(x,y)\right]\left[\lim\limits_{(x,y)\to(a,b)} g(x,y)\right]$.

c. $\lim\limits_{(x,y)\to(a,b)} \frac{f(x,y)}{g(x,y)} = \frac{\lim_{(x,y)\to(a,b)} f(x,y)}{\lim_{(x,y)\to(a,b)} g(x,y)}$, provided $\lim\limits_{(x,y)\to(a,b)} g(x,y) \neq 0$.

Limit Theorem 2: If $\lim\limits_{(x,y)\to(a,b)} f(x,y)$ exists, then the limit is unique.

2.3 Proving a Limit Does Not Exist

To prove a limit does not exist, show that the limit along two different paths give different values. We can approach the limit along infinitely many lines or smooth curves at the same time by introducing an arbitrary coefficient m. If the value depends on the value of m, then it is not unique and the limit does not exist.

Caution: Make sure that all lines or curves you use approach the limit point.

2.4 Proving a Limit Exists

Theorem 1 - Squeeze Theorem: If there exists a function $B(x,y)$ such that

$$|f(x,y) - L| \leq B(x,y), \quad \text{for all } (x,y) \neq (a,b)$$

in some neighbourhood of (a,b) and $\lim\limits_{(x,y)\to(a,b)} B(x,y) = 0$, then

$$\lim_{(x,y)\to(a,b)} f(x,y) = L$$

2.5 Appendix: Inequalities and Absolute Values

Trichotomy Property: For any real numbers a and b, one and only one of the following holds:

$$a = b, \ a < b, \ b < a$$

Transitivity Property: If $a < b$ and $b < c$, then $a < c$.

Addition Property: If $a < b$, then for all c, $a + c < b + c$.

Multiplication Property: If $a < b$ and $c < 0$, then $bc < ac$.

Multiplication Inverse Property: If $ab > 0$ with $a < b$, then $\frac{1}{b} < \frac{1}{a}$. Note the change in order.

Absolute Value Properties (Commonly used for Squeeze Theorem):

1. $|a| = \sqrt{a^2}$.
2. $|a| < b$ if and only if $-b < a < b$.
3. Triangle Inequality: $|a + b| \leq |a| + |b|$ for all $a, b \in \mathbb{R}$.
4. If $c > 0$, then $a < a + c$.
5. Cosine Inequality: $2\,|x|\,|y| \leq x^2 + y^2$.

3 Continuous Functions

3.1 Definition of a Continuous Function

Definition - Continuous: A function $f(x, y)$ is continuous at (a, b) if and only if

$$\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b)$$

Additionally, if f is continuous at every point in a set $D \subset \mathbb{R}^2$, then we say that f is continuous on D.

Remark: Just like single variable calculus, there are three requirements in this definition:

1. $\lim\limits_{(x,y)\to(a,b)} f(x, y)$ exists,
2. f is defined at (a, b), and
3. $\lim\limits_{(x,y)\to(a,b)} = f(a, b)$.

3.2 The Continuity Theorems

Basic Functions (Continuous):

1. $f(x, y) = k$

2. $f(x,y) = x^n,\ f(x,y) = y^n$
3. $\ln(\cdot)$
4. $e^{(\cdot)}$
5. $\sin(\cdot), \cos(\cdot)$, etc.
6. $\arcsin(\cdot), \arccos(\cdot)$, etc.
7. $|\cdot|$

Operations on Functions: If $f(x,y)$ and $g(x,y)$ are scalar functions and $f(x,y) \in D(f) \cap D(g)$, then:

1. The sum $f+g$ is defined by

$$(f+g)(x,y) = f(x,y) + g(x,y)$$

2. The product fg is defined by

$$(fg)(x,y) = f(x,y)g(x,y)$$

3. The quotient $\frac{f}{g}$ is defined by

$$\left(\frac{f}{g}\right)(x,y) = \frac{f(x,y)}{g(x,y)},\ g(x,y) \neq 0$$

Composite Function: For scalar functions $g(t)$ and $f(x,y)$ the composite functions $g \circ f$ is defined by

$$(g \circ f)(x,y) = g(f(x,y))$$

for all $(x,y) \in D(f)$ for which $f(x,y) \in D(g)$.

Continuity Theorem 1: If f and g are both continuous at (a,b), then $f+g$ and fg are continuous at (a,b).

Continuity Theorem 2: If f and g are both continuous at (a,b) and $g(a,b) \neq 0$, then the quotient $\frac{f}{g}$ is continuous at (a,b).

Continuity Theorem 3: If $f(x,y)$ is continuous at (a,b) and $g(t)$ is continuous at $f(a,b)$, then the composition $g \circ f$ is continuous at (a,b).

4 The Linear Approximation and Partial Derivatives

4.1 Partial Derivatives

Definition - Partial Derivatives: The partial derivatives of $f(x,y)$ are defined by

$$\frac{\partial f}{\partial x}(x,y) = f_x(x,y) = \lim_{h \to 0} \frac{f(x+h,y) - f(x,y)}{h}$$
$$\frac{\partial f}{\partial y}(x,y) = f_y(x,y) = \lim_{h \to 0} \frac{f(x,y+h) - f(x,y)}{h}$$

provided that these limits exist.

Remark: It is sometimes convenient to use operator notation $D_1 f$ and $D_2 f$ for the partial derivatives of $f(x, y)$. $D_1 f$ means to differentiate f with respect to the variable in first position. We also sometimes relax the notation and use $\frac{\partial f}{\partial x}$ for $\frac{\partial f}{\partial x}(x, y)$.

Generalization of Partial Derivatives: To take the partial derivative of f with respect to its i-th variable, we hold all the other variables constant, and differentiate with respect to the i-th variable.

4.2 Higher-Order Partial Derivatives

Second-Order Partial Derivative:

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= f_{xx} = D_1^2 f = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \\
\frac{\partial^2 f}{\partial y \partial x} &= f_{xy} = D_2 D_1 f = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \\
\frac{\partial^2 f}{\partial x \partial y} &= f_{yx} = D_1 D_2 f = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \\
\frac{\partial^2 f}{\partial y^2} &= f_{yy} = D_2^2 f = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)
\end{aligned}$$

Theorem 1 - Clairaut's Theorem: If f_{xy} and f_{yx} are defined in some neighborhood of (a, b) and are both continuous at (a, b), then

$$f_{xy}(a, b) = f_{yx}(a, b)$$

Clairaut's Theorem also extends to higher-order partial derivatives:

$$f_{i_1 \ldots i_k} = f_{j_1 \ldots j_k}$$

Third-Order Partial Derivatives:

$$f_{xxx}, f_{xxy}, f_{xyx}, f_{xyy}, f_{yxx}, f_{yxy}, f_{yyx}, f_{yyy}$$

If the k-th partial derivatives of $f(x_1, \ldots, x_n)$ are continuous, then we write

$$f \in C^k$$

and say f is in class C^k.

4.3 The Tangent Plane

Definition - Tangent Plane: The tangent plane to $z = f(x, y)$ at the point $(a, b, f(a, b))$ is

$$z = f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b)$$

4.4 Linear Approximation for $z = f(x, y)$

Definition - Linearization and Linear Approximation: For a function $f(x, y)$ we define the linearization $L_{(a,b)}(x, y)$ of f at (a, b) by

$$L_{(a,b)}(x, y) = f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b)$$

We call the approximation

$$f(x, y) \approx L_{(a,b)}(x, y)$$

the linear approximation of $f(x, y)$ at (a, b)

Increment Form: Approximation for the change Δf in $f(x, y)$ due to a change $(\Delta x, \Delta y)$ away from the point (a, b):

$$\Delta f \approx \frac{\partial f}{\partial x}(a, b)\Delta x + \frac{\partial f}{\partial y}(a, b)\Delta y$$

4.5 Linear Approximation in Higher Dimensions

Definition - Gradient: Suppose that $f(x, y, z)$ has partial derivatives at $\vec{a} \in \mathbb{R}^3$. The gradient of f at $\vec{a}$ is defined by

$$\nabla f(\vec{a}) = (f_x(\vec{a}), f_y(\vec{a}), f_z(\vec{a}))$$

Definition - Linearization and Linear Approximation: Suppose that $f(\vec{x}), \vec{x} \in \mathbb{R}^3$, has partial derivatives at $\vec{a} \in \mathbb{R}^3$.
The linearization of f at $\vec{a}$ is defined by

$$L_{\vec{a}}(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a})$$

The linear approximation of f at $\vec{a}$ is

$$f(\vec{x}) \approx f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a})$$

Linear Approximation in $\mathbb{R}^n$: For an arbitrary vector $\vec{a} \in \mathbb{R}^n$, we have

$$\nabla \vec{x} = \vec{x} - \vec{a} = (x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n)$$

and define the gradient of f at $\vec{a}$ to be

$$\nabla f(\vec{a}) = (D_1 f(\vec{a}), D_2 f(\vec{a}), \ldots, D_n f(\vec{a}))$$

Then, the increment form of the linear approximation for $f(\vec{x})$ is

$$\nabla f \approx \nabla f(\vec{a}) \cdot \nabla \vec{x}$$

5 Differentiable Functions

5.1 Definition of Differentiability

Theorem 1: If $g'(a)$ exists, then $\lim\limits_{x \to a} \frac{|R_{1,a}(x)|}{|x-a|}n = 0$ where

$$R_{1,a}(x) = g(x) - L_a(x) = g(x) - g(a) - g'(a)(x - a)$$

Definition - Differentiable: A function $f(x, y)$ is differentiable at (a, b) if

$$\lim_{(x,y)\to(a,b)} \frac{\left|R_{1,(a,b)}(x, y)\right|}{\|(x, y) - (a, b)\|} = 0$$

where

$$R_{1,(a,b)}(x, y) = f(x, y) - L_{(a,b)}(x, y)$$

Theorem 2: If a function $f(x, y)$ satisfies

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

for some constants c and d then $c = f_x(a, b)$ and $d = f_y(a, b)$.

Remark: For the linear approximation to exist at (a, b), both partial derivatives must exist. However, both partial derivatives existing does not guarantee f will be differentiable.

Definition - Tangent Plane: Consider a function $f(x, y)$ which is differentiable at (a, b). The tangent plane of the surface $z = f(x, y)$ at $(a, b, f(a, b))$ is the graph of the linearization. That is, the tangent plane is given by

$$z = f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b)$$

5.2 Differentiability and Continuity

Theorem 1: If $f(x, y)$ is differentiable at (a, b), then f is continuous at (a, b).

5.3 Continuous Partial Derivatives and Differentiability

Theorem 1 - The Mean Value Theorem: If $f(x)$ is continuous on the closed interval $[x_1, x_2]$ and f is differentiable on the open interval (x_1, x_2), then there exists $x_0 \in (x_1, x_2)$ such that

$$f(x_2) - f(x_1) = f'(x_0)(x_2 - x_1)$$

Theorem 2: If the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are both continuous at (a, b), then $f(x, y)$ is differentiable at (a, b).

Definition - Differentiability for $f : \mathbb{R}^n \to \mathbb{R}$ A function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at a point $\vec{a} = (a_1, \ldots, a_n)$ if

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

where $L : \mathbb{R}^n \to \mathbb{R}$ is a linear transformation.

Theorem 1 for functions of n variables: If $f(x_1, \ldots, x_n)$ is differentiable at $\vec{a} = (a_1, \ldots, a_n)$, then f is continuous at $\vec{a}$.

Theorem 2 for functions of n variables: If $\frac{\partial f}{\partial x_1}, \ldots \frac{\partial f}{\partial x_n}$ are continuous at $\vec{a} = (a_1, \ldots, a_n)$, then $f(x_1, \ldots, x_n)$ is differentiable at $\vec{a}$.

5.4 Linear Approximation Revisited

$$f(x,y) = f(a,b) + \nabla f(a,b) \cdot (x-a, y-b) + R_{1,(a,b)}(x,y) \ (*)$$
$$f(x,y) \approx f(a,b) + \nabla f(a,b) \cdot (x-a, y-b) \ (**)$$

In this case, the approximation $(**)$ is reasonable for (x,y) sufficiently close to (a,b). In general, we have no information about $R_{1,(a,b)}(x,y)$.

Approximation is a recurring theme in calculus, and the equation

$$f(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) + R_{1,\vec{a}}(\vec{x})$$

is of fundamental importance.

6 Chain Rule

6.1 Basic Chain Rule in Two Dimensions

Theorem 1 - Chain Rule: Let $G(t) = f(x(t), y(t))$, and let $a = x(t_0)$ and $b = y(t_0)$. If f is differentiable at (a,b) and $x'(t_0)$ and $y'(t_0)$ exist, then $G'(t_0)$ exists and is given by

$$G'(t_0) = f_x(a,b)x'(t_0) + f_y(a,b)y'(t_0)$$

Vector Form of Basic Chain Rule:

$$\begin{aligned} \frac{dT}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\ &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \cdot \left(\frac{dx}{dt}, \frac{dy}{dt}\right) \\ &= \nabla f \cdot \frac{d\vec{x}}{dt} \end{aligned}$$

$$\frac{d}{dt} f(\vec{x}(t)) = \nabla f(\vec{x}(t)) \cdot \frac{d\vec{x}}{dt}(t)$$

6.2 Extensions of the Basic Chain Rule

Algorithm: To write the Chain Rule from a dependence diagram we do the following:

1. Identify all of the variables.
2. Take all possible paths from the differentiated variable to the differentiating variable.
3. For each link in a given path, differentiate the upper variable with respect to the lower variable being careful to consider if this is a derivative or a partial derivative. Multiply all such derivatives in that path.
4. Add the products from step 3 together to complete the Chain Rule.

Generalized Chain Rule: Let $w = f(x_1, \dots, x_m)$ be a differentiable function of m independent variables and for $1 \le i \le m$ let $x_i = x_i(t_1, \dots, t_n)$ be a differentiable function of n independent variables. Then

$$\frac{\partial w}{\partial t_j} = \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_j} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_j} + \cdots + \frac{\partial w}{\partial x_m}\frac{\partial x_m}{\partial t_j}$$

7 Directional Derivatives and the Gradient Vector

7.1 Directional Derivatives

Definition - Directional Derivative: The directional derivative of $f(x, y)$ at a point (a, b) in the direction of a unit vector $\vec{u} = (u_1, u_2)$ where $\|\vec{u}\| = 1$ is defined by

$$D_{\vec{u}}f(a, b) = \frac{d}{ds}f(a + su_1, b + su_2)\bigg|_{s=0}$$

provided that the derivative exists. An alternate way is

$$D_{\vec{u}}f(a, b) = \lim_{t \to 0} \frac{f((a, b) + t\vec{u}) - f(a, b)}{t}$$

Theorem 1 - Directional Derivative (DD) Theorem: If $f(x, y)$ is differentiable at (a, b) and $\vec{u} = (u_1, u_2)$ where $\|\vec{u}\| = 1$ is a unit vector (otherwise you must normalize), then

$$D_{\vec{u}}f(a, b) = \nabla f(a, b) \cdot \vec{u}$$

where $\cdot$ represents the dot product.

If we choose $\vec{u} = \hat{\imath} = (1, 0)$ or $\vec{u} = \hat{\jmath} = (0, 1)$, then the directional derivative is equal to the partial derivatives f_x and f_y respectively.

7.2 The Gradient Vector in Two Dimensions

Theorem 1 - The Greatest Rate of Change (GRC) Theorem: If $f(x, y)$ is differentiable at (a, b) and $\nabla f(a, b) \neq (0, 0)$, then the largest value of $D_{\vec{u}}f(a, b)$ is $\|\nabla f(a, b)\|$, and occurs when $\vec{u}$ is in the direction of $\nabla f(a, b)$. This applies to any dimension.

Theorem 2 - Orthogonality Theorem: If $f(x, y) \in C^1$ in a neighbourhood of (a, b) and $\nabla f(a, b) \neq (0, 0)$, then $\nabla f(a, b)$ is orthogonal to the level curve $f(x, y) = k$ through (a, b).

7.3 The Gradient Vector in Three Dimensions

Theorem 1 - Orthogonality Theorem in Three Dimensions: If $f(x, y, z) \in C^1$ in a neighbourhood of (a, b, c) and $\nabla f(a, b, c) \neq (0, 0, 0)$, then $\nabla f(a, b, c)$ is orthogonal to the level surfaces $f(x, y, z) = k$ through (a, b, c).

$$\nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) = 0$$

This is the equation of the tangent plane. In component form,

$$f_x(a, b, c)(x - a) + f_y(a, b, c)(y - b) + f_z(a, b, c)(z - c) = 0$$

8 Taylor Polynomials and Taylor's Theorem

8.1 The Taylor Polynomial of Degree 2

Definition - 2nd degree Taylor Polynomial: Let f be a function of two variables. The second degree Taylor polynomial $P_{2,(a,b)}$ of $f(x,y)$ at (a,b) is given by

$$\begin{aligned} P_{2,(a,b)} &= f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b) \\ &+ \frac{1}{2}\left[f_{xx}(a,b)(x-a)^2 + 2f_{xy}(a,b)(x-a)(y-b) + f_{yy}(a,b)(y-b)^2\right] \end{aligned}$$

Second degree Taylor polynomial approximates: $f(x,y) \approx P_{2,(a,b)}(x,y)$.

Definition - Hessian Matrix: The Hessian matrix of $f(x,y)$, denoted by $Hf(x,y)$, is defined as

$$Hf(x,y) = \begin{bmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{yx}(x,y) & f_{yy}(x,y) \end{bmatrix}$$

8.2 Taylor's Formula with Second Degree Remainder

Theorem 1 - Taylor Remainder for Single Variable Functions: If $f''(x)$ exists on $[a,x]$, then there exists a number c between a and x such that

$$f(x) = f(a) + f'(a)(x-a) + R_{1,a}(x)$$

where

$$R_{1,a}(x) = \frac{1}{2}f''(c)(x-a)^2$$

Theorem 2 - Taylor Remainder for Functions of Two Variables: If $f(x,y) \in C^2$ in some neighbourhood $N(a,b)$ of (a,b), then for all $(x,y) \in N(a,b)$ there exists a point (c,d) on the line segment joining (a,b) and (x,y) such that

$$f(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b) + R_{1,(a,b)}(x,y)$$

where

$$R_{1,(a,b)}(x,y) = \frac{1}{2}[f_{xx}(c,d)(x-a)^2 + 2f_{xy}(c,d)(x-a)(y-b) + f_{yy}(c,d)(y-b)^2]$$

Remark: Like the one variable case, Taylor's formula tells us the existence of (c,d) but not how to find it.

Corollary: If $f(x,y) \in C^2$ in some closed neighbourhood $N(a,b)$ of (a,b), then there exists a positive constant M such that

$$\left|R_{1,(a,b)}(x,y)\right| \leq M \left\|(x,y) - (a,b)\right\|^2, \quad \text{for all } (x,y) \in N(a,b)$$

8.3 Generalizations of the Taylor Polynomial

Multi-Index Notation: If $f \in C^k$ is a function of n variables, we can write a k-th order partial derivative of $f(x_1, \ldots, x_n)$ as

$$\partial^\alpha f = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \left(\frac{\partial}{\partial x_2}\right)^{\alpha_2} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} f$$

where α is a multi-index; that is, $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ with $\alpha_i \in \mathbb{N}$. The sum $\alpha_1+\alpha_2+\cdots+\alpha_n = k$ is called the order of α and is sometimes denoted $|\alpha|$. We also define $\alpha! = \alpha_1!\alpha_2!\cdots\alpha_n!$. Given a multi-index of order k, $\partial^\alpha f$ is a partial derivative of order k of f.

In addition to using the multi-index notation for k-th partial derivatives, we can also use it as follows:
Let $\vec{x} = (x_1, \ldots, x_n)$, $\vec{a} = (a_1, \ldots, a_n)$, and $\alpha = (\alpha_1, \ldots, \alpha_n)$, then

$$(\vec{x} - \vec{a})^\alpha = (x_1 - a_1)^{\alpha_1}(x_2 - a_2)^{\alpha_2} \cdots (x_n - a_n)^{\alpha_n}$$

Definition - k-th degree Taylor Polynomial: The k-th degree Taylor polynomial of a function $f(x, y)$ is

$$P_{k,(a,b)}(x, y) = \sum_{|\alpha| \leq k} \partial^\alpha f(a, b) \frac{[(x, y) - (a, b)]^\alpha}{\alpha!}$$

Theorem 1 - Taylor's Theorem of order k: If $f(x, y) \in C^{k+1}$ in some neighbourhood $N(a, b)$ of (a, b), then for all $(x, y) \in N(a, b)$ there exists a point (c, d) on the line segment between (a, b) and (x, y) such that

$$f(x, y) = P_{k,(a,b)}(x, y) + R_{k,(a,b)}(x, y)$$

where

$$R_{k,(a,b)}(x, y) = \sum_{|\alpha| = k+1} \partial^\alpha f(c, d) \frac{[(x, y) - (a, b)]^\alpha}{\alpha!}$$

Corollary: If $f(x, y) \in C^k$ in some neighbourhood of (a, b), then

$$\lim_{(x,y)\to(a,b)} \frac{\left|f(x, y) - P_{k,(a,b)}(x, y)\right|}{\|(x, y) - (a, b)\|^k} = 0$$

Corollary: If $f(x, y) \in C^{k+1}$ in some closed neighbourhood $N(a, b)$ of (a, b), then there exists a constant $M > 0$ such that

$$\left|f(x, y) - P_{k,(a,b)}(x, y)\right| \leq M \|(x, y) - (a, b)\|^{k+1}$$

for all $(x, y) \in N(a, b)$.

Definition - Taylor Polynomial of Degree k for Functions of n Variables: The Taylor polynomial of degree k for functions of n variables is

$$P_{k,\vec{a}}(\vec{x}) = \sum_{\alpha \leq k} \partial^\alpha f(\vec{a}) \frac{(\vec{x} - \vec{a})^\alpha}{\alpha!}$$

9 Critical Points

9.1 Local Extrema and Critical Points

Definition - Local Maximum and Minimum: A point (a, b) is a local maximum point of f if $f(x, y) \leq f(a, b)$ for all (x, y) in some neighbourhood of (a, b).

A point (a, b) is a local minimum point of f if $f(x, y) \geq f(a, b)$ for all (x, y) in some neighbourhood of (a, b).
Points which are either local maxima or local minima are sometimes referred to as local extrema.

Theorem 1: If (a, b) is a local maximum or minimum point of f, then each partial derivative is either equal to zero or does not exist.

Definition - Critical Point: A point (a, b) in the domain of $f(x, y)$ is called a critical point of f if

$$\frac{\partial f}{\partial x}(a, b) = 0 \text{ or } \frac{\partial f}{\partial x}(a, b) \text{ does not exist,}$$
$$\frac{\partial f}{\partial y}(a, b) = 0 \text{ or } \frac{\partial f}{\partial y}(a, b) \text{ does not exist.}$$

Definition - Saddle Point: A critical point (a, b) of $f(x, y)$ is called a saddle point of f if in every neighbourhood of (a, b) there exist points (x_1, y_1) and (x_2, y_2) such that

$$f(x_1, y_1) > f(a, b) \text{ and } f(x_2, y_2) < f(a, b)$$

9.2 The Second Derivative Test

Definition - Quadratic Form: A function Q of the form

$$Q(u, v) = a_{11}u^2 + 2a_{12}uv + a_{22}v^2$$

where a_{11}, a_{12} and a_{22} are constants, is called a quadratic form on $\mathbb{R}^2$. Matrix form of Quadratic form:

$$Q(u, v) = \begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

Quadratic forms on $\mathbb{R}^2$ fall into four main classes:

1. If $Q(u, v) > 0$ for all $(u, v) \neq (0, 0)$, then $Q(u, v)$ is positive definite.
2. If $Q(u, v) < 0$ for all $(u, v) \neq (0, 0)$, then $Q(u, v)$ is negative definite.
3. If $Q(u, v) < 0$ for some (u, v) and $Q(w, z) > 0$ for some (w, z), then $Q(u, v)$ is indefinite.
4. If $Q(u, v)$ does not belong to those, then it is semidefinite.
   a. If $Q(u, v) \geq 0$ for all $(u, v) \neq (0, 0)$, then $Q(u, v)$ is positive semidefinite.
   b. If $Q(u, v) \leq 0$ for all $(u, v) \neq (0, 0)$, then $Q(u, v)$ is negative semidefinite.

Proposition - Determinant and Quadratic Forms: A quadratic form $Q(u, v) = a_{11}u^2 + 2a_{12}uv + a_{22}v^2$ on $\mathbb{R}^2$ is

1. Positive definite if $\det A > 0$ and $a_{11} > 0$
2. Negative definite if $\det A > 0$ and $a_{11} < 0$
3. Indefinite if $\det A < 0$

4. Semidefinite if $\det A = 0$

Theorem 1 - Second Partial Derivatives Test: Suppose that $f(x, y) \in C^2$ in some neighbourhood of (a, b) and that

$$f_x(a, b) = 0 = f_y(a, b)$$

1. If $Hf(a, b)$ is positive definite, then (a, b) is a local minimum point of f.
2. If $Hf(a, b)$ is negative definite, then (a, b) is a local maximum point of f.
3. If $Hf(a, b)$ is indefinite, then (a, b) is a saddle point of f.
4. If $Hf(a, b)$ is semidefinite. This means that the critical point (a, b) is degenerate. Check the sign of $f(x, y) - f(a, b)$.
   a. If $f(x, y) - f(a, b)$ is always positive, then (a, b) is a local minimum.
   b. If $f(x, y) - f(a, b)$ is always negative, then (a, b) is a local maximum.
   b. If $f(x, y) - f(a, b)$ assumes both positive and negative, then (a, b) is a saddle point.

Remark: Another way of classifying the Hessian matrix is by finding its eigenvalues. A symmetric matrix is positive definite if all eigenvalues are positive, negative definite if all eigenvalues are negative, and indefinite if it has both positive and negative eigenvalues.

Generalizations: The Hessian matrix of f at $\vec{a}$ is the $n \times n$ symmetric matrix given by

$$Hf(\vec{a}) = \left[\tfrac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a})\right]$$

where $i, j = 1, 2, \ldots, n$. The Hessian matrix can classified positive definite, negative definite, indefinite or semidefinite by considering the quadratic form in $\mathbb{R}^n$:

$$Q(\vec{u}) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a}) u_i u_j$$

It can be justified heuristically by using second degree Taylor polynomial approximation,

$$f(\vec{x}) \approx P_{2,\vec{a}}(\vec{x})$$

which leads to

$$f(\vec{x}) - f(\vec{a}) \approx \frac{1}{2!} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a})(x_i - a_i)(x_j - a_j)$$

9.3 Convex Functions

Definition - Convex and strictly convex functions of one variable: A twice differentiable function $f(x)$ is convex if $f''(x) \geq 0$ for all x and f is strictly convex if $f''(x) > 0$ for all x. Convex means concave up.

Theorem 1 - Properties of convex functions of one variable: If $f(x) \in C^2$ and is strictly convex, then

1. $f(x) > L_a(x) = f(a) + f'(x)(x - a)a$ for all $x \neq a$, for any $a \in \mathbb{R}$.

2. For $a < b$, $f(x) < f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$ for all $x \in (a, b)$.

Definition - Convex and strictly convex functions of two variables: Let $f(x, y)$ have continuous second partial derivatives. We say that f is convex if $Hf(x, y)$ is positive semidefinite for all (x, y) and that f is strictly convex if $Hf(x, y)$ is positive definite for all (x, y).

Theorem 2 - Properties of convex functions of two variables: If $f(x, y)$ has continuous second partial derivatives and is strictly convex, then

1. $f(x, y) > L_{(a,b)}(x, y)$ for all $(x, y) \neq (a, b)$, and
2. $f(a_1 + t(b_1 - a_1), a_2 + t(b_2 - a_2)) < f(a_1, a_2) + t[f(b_1, b_2) - f(a_1, a_2)]$ for $0 < t < 1$, $(a_1, a_2) \neq (b_1, b_2)$.

Theorem 3 - Critical points of convex and strictly convex functions: If $f(x, y) \in C^2$ is convex, then every critical point (c, d) satisfies $f(x, y) \geq f(c, d)$ for all $(x, y) \neq (c, d)$.
If $f(x, y) \in C^2$ is strictly convex and has a critical point (c, d), then $f(x, y) \geq f(c, d)$ for all $(x, y) \neq (c, d)$ and f has no other critical point.

9.4 Proof of the Second Partial Derivative Test

Lemma 1: Let $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ be a positive definite matrix. If $|\tilde{a} - a|$, $\left|\tilde{b} - b\right|$ and $|\tilde{c} - c|$ are sufficiently small, then $\begin{bmatrix} \tilde{a} & \tilde{b} \\ \tilde{b} & \tilde{c} \end{bmatrix}$ is positive definite. The lemma is also true if "positive definite" is replaced by negative definite or indefinite.

10 Optimization Problems

10.1 The Extreme Value Theorem

Definition - Absolute Maximum and Minimum: Given a function $f(x, y)$ and a set $S \subseteq \mathbb{R}^2$,

1. a point $(a, b) \in S$ is an absolute maximum point of f on S if

$$f(x, y) \leq f(a, b) \quad \text{for all } (x, y) \in S$$

   The value $f(a, b)$ is called the absolute maximum value of f on S.

2. a point $(a, b) \in S$ is an absolute minimum point of f on S if

$$f(x, y) \geq f(a, b) \quad \text{for all } (x, y) \in S$$

   The value $f(a, b)$ is called the absolute minimum value of f on S.

Definition - Bounded Set: A set $S \subset \mathbb{R}^2$ is said to be bounded if and only if it is contained in some neighbourhood of the origin.

Definition - Boundary Point: Given a set $S \subset \mathbb{R}^2$, a point $(a, b) \in \mathbb{R}^2$ is aid to be a boundary point of S if and only if every neighbourhood of (a, b) contains at least one point in S and one point not in S.

Definition - Boundary of S: The set $B(S)$ of all boundary points of S is called the boundary of S.

Theorem 2 - Extreme Value Theorem (EVT) for Functions of Two Variables: If $f(x, y)$ is continuous on a closed and bounded set $S \subset \mathbb{R}^2$, then there exist points $(a, b), (c, d) \in S$ such that

$$f(a, b) \leq f(x, y) \leq f(c, d) \quad \text{for all } (x, y) \in S$$

Remark: A function f may have an absolute max/min on a set $S \in \mathbb{R}^2$ even if conditions of EVT are not satisfied.

10.2 Algorithm for Extreme Values

Algorithm: First, check to see if the given set $S \subset \mathbb{R}^2$ is closed and bounded.
Next, check to see if the given function $f(x, y)$ is continuous on S.
If both conditions above are satisfied, then to find the maximum and/or minimum value of a function $f(x, y)$:

1. Find all critical points of f that are contained in S.
2. Evaluate f at each such point.
3. Find the maximum and minimum values of f on the boundary $B(S)$.
4. The maximum value of f on S is the largest value of the function found in steps 2 and 3. The minimum value of f on S is the smallest value of the function found in steps 2 and 3.

10.3 Optimization with Constraints

Lagrange Multiplier Algorithm: Assume that $f(x, y)$ is a differentiable function and $g \in C^1$. To find the maximum value and minimum value of f subject to the constraint $g(x, y) = k$, evaluate $f(x, y)$ at all points (a, b) which satisfy one of the following conditions.

1. $\nabla f(a, b) = \lambda \nabla g(a, b)$ and $g(a, b) = k$
2. $\nabla g(a, b) = (0, 0)$ and $g(a, b) = k$
3. (a, b) is an end point of the curve $g(x, y) = k$

The maximum/minimum value of $f(x, y)$ is the largest/smallest value of f obtained at the points found in conditions 1-3. To find the points (a, b) in condition 1 we have to solve the system of 3 equations in 3 unknowns

$$\begin{aligned} f_x(x, y) &= \lambda g_x(x, y) \\ f_y(x, y) &= \lambda g_y(x, y) \\ g(x, y) &= k \end{aligned}$$

The variable λ, called the Lagrange multiplier, is not required for our purposes.

Remark:

1. Observe that condition 2 must be included since we assumed that $\nabla g(a,b) \neq (0,0)$ in the derivation.
2. Condition 3 will only arise if there are restrictions on the domain of $g(x,y)$ that result in the curve $g(x,y)=k$ having end points. For instance, if $g(x,y)=x^2+y^2$, then the curve $x^2+y^2=1$ (the unit circle) does not have end points if there are no restrictions on x and y. However, if we restrict the domain of g to $\{(x,y) : y \geq 0\}$, then there will be two end points: $(1,0)$ and $(1,0)$.
3. If the curve $g(x,y)=k$ is unbounded, then we must consider $\lim\limits_{\|(x,y)\|\to\infty} f(x,y)$ for (x,y) satisfying $g(x,y)=k$. We will not see cases like this in this course.

Method of Lagrange Multipliers for Three Variables: To find the maximum/minimum value of a differentiable function $f(x,y,z)$ subject to $g(x,y,z)=k$ such that $g \in C^1$, we evaluate $f(x,y,z)$ at all points (a,b,c) which satisfy one of the following:

1. $\nabla f(a,b,c) = \lambda \nabla g(a,b,c)$ and $g(a,b,c)=k$.
2. $\nabla g(a,b,c) = (0,0,0)$ and $g(a,b,c)=k$.
3. (a,b,c) is an edge point of the surface $g(x,y,z)=k$

The maximum/minimum value of $f(x,y,z)$ is the largest/smallest value of f obtained from all points found in conditions 1-3.

Generalization: The method of Lagrange multipliers can be generalized to functions of n variables $f(\vec{x}), \vec{x} \in \mathbb{R}^n$ and with r constraints of the form

$$g_1(\vec{x}) = 0, g_2(\vec{x}) = 0, \ldots, g_r(\vec{x}) = 0 \ (*)$$

In order to find the possible maximum and minimum points of f subject to the constraints $(*)$, we have to find all the points $\vec{a}$ such that

$$\nabla f(\vec{a}) = \lambda_1 \nabla g_1(\vec{a}) + \cdots + \lambda_r \nabla g_r(\vec{a}), \text{ and } g_i(\vec{a}) = 0, 1 \leq i \leq r$$

The scalars $\lambda_1, \ldots, \lambda_r$ are the Lagrange multipliers. When $r=1$ and $n=2$ or 3, this reduces to the previous algorithms.

11 Coordinate Systems

11.1 Polar Coordinates

Relationship to Cartesian Coordinates: $x = r\cos\theta$, $y = r\sin\theta$, $r = \sqrt{x^2+y^2}$, $\tan\theta = \frac{y}{x}$.

Area in Polar Coordinates:

$$A = \lim_{\|\Delta\theta_i\|\to 0} \sum_{i=0}^{n-1} \frac{1}{2}[f(\theta_i^*)]^2 \Delta\theta = \int_a^b \frac{1}{2}[f(\theta)]^2 \, d\theta$$

Algorithm - Area Between Curves in Polar Coordinates: To find the area between two curves in polar coordinates, we use the same method we used for doing this in Cartesian coordinates.

1. Find the points of intersections.
2. Graph the curves and split the desired region into easily integrable regions.
3. Integrate.

11.2 Cylindrical Coordinates

We add another axis called the axis of symmetry to polar coordinates.

$$(r, \theta, z)$$

Relationship to Cartesian Coordinates: $x = r\cos\theta,\ y = r\sin\theta,\ z = z,\ r = \sqrt{x^2+y^2}$, $\tan\theta = \frac{y}{x}$, $z = z$.

11.3 Spherical Coordinates

Let P be any point in 3-dimensional space. We will represent P by the coordinates (ρ, φ, θ) where $\rho \geq 0$ is the length of the line OP, θ is the same angle as in polar and cylindrical coordinates, and φ is the angle between the positive z-axis and the line OP.

Relationship to Cartesian Coordinates:

$$\begin{aligned} x &= \rho\sin\varphi\cos\theta & \rho &= \sqrt{x^2+y^2+z^2} \\ y &= \rho\sin\varphi\sin\theta & \tan\theta &= \frac{y}{x} \\ z &= \rho\cos\varphi & \cos\varphi &= \frac{z}{\sqrt{x^2+y^2+z^2}} \end{aligned}$$

12 Mapping of $\mathbb{R}^2$ into $\mathbb{R}^2$

12.1 The Geometry of Mappings

Definition - Vector-Valued Function: A function whose domain is a subset of $\mathbb{R}^n$ and whose codomain is $\mathbb{R}^m$ is called a vector-valued function.

Definition - Mapping: A vector-valued function whose domain is a subset of $\mathbb{R}^n$ and whose codomain is a subset of $\mathbb{R}^n$ is called a mapping (or transformation).
A pair of equations

$$\begin{aligned} u &= f(x, y) \\ v &= g(x, y) \end{aligned}$$

associates with each point $(x, y) \in \mathbb{R}^2$ a single point $(u, v) \in \mathbb{R}^2$, and thus defines a vector-valued function

$$(u, v) = F(x, y) = (f(x, y), g(x, y))$$

The scalar functions f and g are called the component functions of the mapping.

12.2 The Linear Approximation of a Mapping

Definition - Derivative Matrix: The derivative matrix of a mapping defined by

$$F(x,y) = (f(x,y), g(x,y))$$

is denoted DF and defined by

$$DF = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix}$$

If we introduce the column vectors

$$\Delta\vec{u} = \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix},\ \Delta\vec{x} = \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

then the increment form of the linear approximation for mappings becomes

$$\Delta\vec{u} \approx DF(a,b)\Delta\vec{x}$$

for $\Delta\vec{x}$ sufficiently small. Thus, the linear approximation for mappings is

$$F(x,y) \approx F(a,b) + DF(a,b)\Delta\vec{x}$$

Generalization: A mapping F from $\mathbb{R}^n$ to $\mathbb{R}^m$ is defined by a set of m component functions:

$$\begin{aligned} u_1 &= f_1(x_1, \ldots, x_n) \\ &\vdots \\ u_m &= f_m(x_1, \ldots, x_n) \end{aligned}$$

Or, in vector notation

$$\vec{u} = F(\vec{x}) = (f_1(\vec{x}), \cdots, f_m(\vec{x})),\ \vec{x} \in \mathbb{R}^n$$

If we assume that F has continuous partial derivatives, then the derivative matrix of F is the $m \times n$ matrix defined by

$$DF(\vec{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

The linear approximation for F at $\vec{a}$ is

$$F(\vec{x}) \approx F(\vec{a}) + DF(\vec{a})\Delta\vec{x}$$

where

$$\Delta\vec{u} = \begin{bmatrix} \Delta u_1 \\ \vdots \\ \Delta u_m \end{bmatrix} \in \mathbb{R}^m,\ \Delta\vec{x} = \begin{bmatrix} \Delta x_1 \\ \vdots \\ \Delta x_n \end{bmatrix} \in \mathbb{R}^n$$

12.3 Composite Mappings and the Chain Rule

Theorem 1 - Chain Rule in Matrix Form for mappings from $\mathbb{R}^2$ to $\mathbb{R}^2$: Let F and G be mappings from $\mathbb{R}^2$ to $\mathbb{R}^2$. If G has continuous partial derivatives at (x,y) and F has continuous partial derivatives at $(u,v) = G(x,y)$, then the composite mapping $F \circ G$ has continuous partial derivatives at (x,y) and

$$D(F \circ G)(x,y) = DF(u,v)DG(x,y)$$

13 Jacobians and Inverse Mappings

13.1 The Inverse Mapping Theorem

Definition - Invertible Mapping and Inverse Mapping: Let F be a mapping from a set D_{xy} onto a set D_{uv}. If there exists a mapping F^{-1}, called the inverse of F which maps D_{uv} to D_{xy} such that

$$(x, y) = F^{-1}(u, v) \text{ if and only if } (u, v) = F(x, y)$$

then F is invertible on D_{xy}.

Definition - One-to-One: A mapping F from $\mathbb{R}^2$ to $\mathbb{R}^2$ is said to be one-to-one (or injective) on a set D_{xy} if and only if $F(a, b) = F(c, d)$ implies $(a, b) = (c, d)$, for all $(a, b), (c, d) \in D_{xy}$.

Theorem 1 - One-to-One Implies Invertible: Let F be a mapping from a set D_{xy} onto a set D_{uv}. If F is one-to-one on D_{xy}, then F is invertible on D_{xy}.

Theorem 2 - Inverse of the Derivative Matrix: Consider a mapping F which maps D_{xy} onto D_{uv}. If F has continuous partial derivatives at $\vec{x} \in D_{xy}$ and there exists an inverse mapping F^{-1} of F which has continuous partial derivatives at $\vec{u} = F(\vec{x}) \in D_{uv}$, then

$$DF^{-1}(\vec{u})DF(\vec{x}) = I$$

Definition - The Jacobian: The Jacobian of a mapping

$$(u, v) = F(x, y) = (u(x, y), v(x, y))$$

is denoted $\frac{\partial(u,v)}{\partial(x,y)}$, and is defined by

$$\frac{\partial(u, v)}{\partial(x, y)} = \det[DF(x, y)] = \det \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}$$

Corollary 3: Consider a mapping defined by

$$(u, v) = F(x, y) = (f(x, y), g(x, y))$$

which maps a subset D_{xy} onto a subset D_{uv}. Suppose that f and g have continuous partial derivatives on D_{xy}. If F has an inverse mapping F^{-1}, with continuous partial derivatives on D_{uv}, then the Jacobian of F is non-zero:

$$\frac{\partial(u, v)}{\partial(x, y)} \neq 0, \text{ on } D_{xy}$$

Corollary 4 - Inverse Property of the Jacobian: Consider a mapping F which maps D_{xy} onto D_{uv}. If F has continuous partial derivatives at $\vec{x} \in D_{xy}$ and there exists an inverse mapping F^{-1} of F which has continuous partial derivatives at $\vec{u} = F(\vec{x}) \in D_{uv}$, then

$$\frac{\partial(x, y)}{\partial(u, v)} = \frac{1}{\frac{\partial(u,v)}{\partial(x,y)}}$$

Theorem 5 - The Inverse Mapping Theorem: If a mapping $(u, v) = F(x, y)$ has continuous partial derivatives in some neighbourhood of (a, b) and $\frac{\partial(u,v)}{\partial(x,y)} \neq 0$ at (a, b), then there is a neighbourhood of (a, b) in which F has an inverse mapping $(x, y) = F^{-1}(u, v)$ which has continuous partial derivatives.

13.2 Geometrical Interpretation of the Jacobian

We can think of the Jacobian as the extent to which F increases/decreases areas. Jacobian is a scaling factor for (very small) areas that are mapped by F.

Area of a rectangle transformation:

$$\Delta A_{uv} \approx \left| \frac{\partial(u,v)}{\partial(x,y)} \right| \Delta A_{xy}$$

Definition - The Jacobian General Form: For a mapping defined by

$$\vec{u} = F(\vec{x}) = (f_1(\vec{x}), \ldots, f_n(\vec{x}))$$

where $\vec{u} = (u_1, \ldots, u_n)$ and $\vec{x} = (x_1, \ldots, x_n)$, the Jacobian of F is

$$\frac{\partial(u_1, \ldots, u_n)}{\partial(x_1, \ldots, x_n)} = \det[DF(\vec{x})] = \det \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

14 Double Integrals

14.1 Definition of Double Integrals

Definition - Integrable Function: Let $D \subset \mathbb{R}^2$ be closed and bounded. Let P be a partition of D as described above, and let $|\Delta P|$ denote the length of the longest side of all rectangles in the partition P. A function $f(x,y)$ which is bounded on D is integrable on D if all Riemann sums approach the same value as $|\Delta P| \to 0$.

Definition - Double Integral: If $f(x,y)$ is integrable on a closed bounded set D, then we define the double integral of f on D as

$$\iint_D f(x,y)\,dA = \lim_{\Delta P \to 0} \sum_{i=1}^{n} f(x_i, y_i) \Delta A_i$$

Interpretations of the Double Integral:
Double Integral as Area: The simplest interpretation is when we specialize f to be the constant function with value one:

$$f(x,y) = 1, \text{ for all } (x,y) \in D$$

Then the Riemann sum $\sum_{i=1}^{n} f(x_i, y_i) \Delta A_i$ simply sums the areas of all rectangles in D, and the double integral serves to define the area $A(D)$ of the set D:

$$A(D) = \iint_D 1\,dA$$

Double Integral as Volume: If $f(x,y) \geq 0$ for all $(x,y) \in D$, then the double integral

$$\iint_D f(x,y)\,dA$$

can be interpreted as the volume $V(S)$ of the region defined by

$$S = \{(x, y, z) \mid 0 \leq z \leq f(x, y), (x, y) \in D\}$$

Double Integral as Mass:

$$M = \iint\limits_D f(x, y)\, dA$$

Double Integral as Probability:

$$P((X, Y) \in D) = \iint\limits_D f(x, y)\, dA$$

Average Value of a Function:

$$f_{av} = \frac{1}{A(D)} \iint\limits_D f(x, y)\, dA$$

Theorem 1 - Linearity: If $D \subset \mathbb{R}^2$ is a closed and bounded set and f and g are two integrable functions on D, then for any constant c:

1. $$\iint\limits_D (f + g)\, dA = \iint\limits_D f\, dA + \iint\limits_D g\, dA$$

2. $$\iint\limits_D cf\, dA = c \iint\limits_D f\, dA$$

Theorem 2 - Basic Inequality: If $D \subset \mathbb{R}^2$ is a closed and bounded set and f and g are two integrable functions on D such that $f(x, y) \leq g(x, y)$ for all $(x, y) \in D$, then

$$\iint\limits_D f\, dA \leq \iint\limits_D g\, dA$$

Theorem 3 - Absolute Value Inequality: If $D \subset \mathbb{R}^2$ is a closed and bounded set and f is an integrable function on D, then

$$\left| \iint\limits_D f\, dA \right| \leq \iint\limits_D |f|\, dA$$

Theorem 4 - Decomposition: Let $D \subset \mathbb{R}^2$ be a closed and bounded set and let f be an integrable function on D. If D is decomposed into two closed and bounded subsets D_1 and D_2 by a piecewise smooth curve C, then

$$\iint\limits_D f\, dA = \iint\limits_{D_1} f\, dA + \iint\limits_{D_2} f\, dA$$

14.2 Iterated Integrals

Theorem 1 - Iterated Integrals: Let $D \subset \mathbb{R}^2$ be defined by

$$y_l(x) \leq y \leq y_u(x), \text{ and } x_l \leq x \leq x_u$$

where $y_l(x)$ and $y_u(x)$ are continuous for $x_l \leq x \leq x_u$. If $f(x, y)$ is continuous on D, then

$$\iint\limits_D f(x, y)\, dA = \int\limits_{x_l}^{x_u} \int\limits_{y_l(x)}^{y_u(x)} f(x, y)\, dy\, dx$$

Remark: Although the parentheses around the inner integral are usually omitted, we must evaluate it first. Moreover, as in our interpretation of volume above, when evaluating the inner integral, we are integrating with respect to y while holding x constant. That is, we are using partial integration.

14.3 The Change of Variable Theorem

Theorem 1 - Change of Variable Theorem: Let each of D_{uv} and D_{xy} be a closed bounded set whose boundary is a piecewise-smooth closed curve. Let

$$(x, y) = F(u, v) = (f(u, v), g(u, v))$$

be a one-to-one mapping of D_{uv} onto D_{xy}, with $f, g \in C^1$, and $\frac{\partial(x,y)}{\partial(u,v)} \neq 0$ except for possibly on a finite collection of piecewise-smooth curves in D_{uv}. If $G(x, y)$ is continuous on D_{xy}, then

$$\iint\limits_{D_{xy}} G(x, y)\, dx\, dy = \iint\limits_{Duv} G(f(u, v), g(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du\, dv$$

Double Integral of Mapping from Cartesian to Polar Coordinates:

$$(x, y) = f(r, \theta) = (r \cos \theta, r \sin \theta)$$

which has Jacobian,

$$\frac{\partial(x, y)}{\partial(r, \theta)} = r$$

15 Triple Integrals

15.1 Definition of Triple Integrals

Definition - Integrable: A function $f(x, y, z)$ which is bounded on a closed bounded set $D \subset \mathbb{R}^3$ is said to be integrable on D if and only if all Riemann sums approach the same value as $\Delta P \to 0$.

Definition - Triple Integral: If $f(x, y, z)$ is integrable on a closed bounded set D, then we define the triple integral of f over D as

$$\iiint\limits_D f(x, y, z)\, dV = \lim_{\Delta P \to 0} \sum_{i=1}^{n} f(x_i, y_i, z_i) \Delta V_i$$

Interpretations of the Triple Integral:

Triple Integral as Volume: The simplest interpretation is when we specialize f to be the constant function with value one:

$$f(x,y,z) = 1, \text{ for all } (x,y,z) \in D$$

Then the Riemann sum $\sum_{i=1}^{n} f(x_i, y_i, z_i)\Delta V_i$ simply sums the volumes of all rectangular blocks in D, and the triple integral over D serves to define the volume $V(D)$ of the set D:

$$V(D) = \iiint\limits_D 1\, dV$$

Triple Integral as Mass:

$$M = \iiint\limits_D f(x,y,z)\, dV$$

Definition - Average Value of a Function: Let $D \subset \mathbb{R}^3$ be closed and bounded with volume $V(D) \neq 0$, and let $f(x,y,z)$ be a bounded and integrable function on D. The average value of f over D is defined by

$$f_{avg} = \frac{1}{V(D)} \iiint\limits_D f(x,y,z)\, dV$$

Theorem 1 - Linearity: If $D \subset \mathbb{R}^3$ is a closed and bounded set, c is a constant, and f and g are two integrable functions on D, then

1. $$\iiint\limits_D (f+g)\, dV = \iiint\limits_D f\, dV + \iiint\limits_D g\, dV$$

2. $$\iiint\limits_D cf\, dA = c\iiint\limits_D f\, dV$$

Theorem 2 - Basic Inequality: If $D \subset \mathbb{R}^3$ is a closed and bounded set and f and g are two integrable functions on D such that $f(x,y,z) \leq g(x,y,z)$ for all $(x,y,z) \in D$, then

$$\iiint\limits_D f\, dV \leq \iiint\limits_D g\, dV$$

Theorem 3 - Absolute Value Inequality: If $D \subset \mathbb{R}^3$ is a closed and bounded set and f is an integrable function on D, then

$$\left| \iiint\limits_D f\, dV \right| \leq \iiint\limits_D |f|\, dV$$

Theorem 4 - Decomposition: Let $D \subset \mathbb{R}^3$ be a closed and bounded set and let f be an integrable function on D. If D is decomposed into two closed and bounded subsets D_1 and D_2 by a piecewise smooth curve C, then

$$\iiint\limits_D f\, dV = \iiint\limits_{D_1} f\, dV + \iiint\limits_{D_2} f\, dV$$

15.2 Iterated Integrals

Theorem 1 - Iterated Integrals: Let $D \subset \mathbb{R}^2$ be defined by

$$z_l(x,y) \le z \le z_u(x,y), \text{ and } (x,y) \in D_{xy}$$

where z_l and z_u are continuous on D_{xy}, and D_{xy} is a closed bounded subset in $\mathbb{R}^2$, whose boundary is a piecewise smooth closed curve. If $f(x,y,z)$ is continuous on D, then

$$\iiint\limits_D f(x,y,z)\,dV = \iint\limits_{D_{xy}} \int\limits_{z_l(x,y)}^{z_u(x,y)} f(x,y,z)\,dz\,dA$$

15.3 The Change of Variable Theorem

Theorem 1 - Change of Variable Theorem: Let

$$x = f(u,v,w), y = g(u,v,w), z = h(u,v,w)$$

be a one-to-one mapping of D_{uvw} onto D_{xyz}, with f, g, h having continuous partials, and

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} \ne 0 \text{ on } D_{uvw}$$

If $G(x,y,z)$ is continuous on D_{xyz}, then

$$\iiint\limits_{D_{xyz}} G(x,y,z)\,dV = \iiint\limits_{D_{uvw}} G(f(u,v,w), g(u,v,w), h(u,v,w)) \left| \frac{\partial(x,y,z)}{\partial(u,v,w)} \right| dV$$

Triple Integrals of Mapping from Cartesian to Cylindrical Coordinates:

$$\left| \frac{\partial(x,y,z)}{\partial(r,\theta,z)} \right| = r$$

$$\iiint\limits_{D_{xyz}} G(x,y,z)\,dx\,dy\,dz = \iiint\limits_{D_{r\theta z}} G(r\cos\theta, r\sin\theta, z) r\,dr\,d\theta\,dz$$

Triple Integrals of Mapping from Cartesian to Spherical Coordinates:

$$\left| \frac{\partial(x,y,z)}{\partial(\rho,\theta,\varphi)} \right| = \rho^2 \sin\varphi$$

$$\iiint\limits_{D_{xyz}} G(x,y,z)\,dx\,dy\,dz = \iiint\limits_{D_{\rho\theta\varphi}} G(\rho\sin\varphi\cos\theta, \rho\sin\varphi\sin\theta, \rho\cos\varphi)\rho^2 \sin\varphi\,d\rho\,d\theta\,d\varphi$$