MA102: Multivariable Calculus

Arup Chattopadhyay and Subhamay Saha
Department of Mathematics
IIT Guwahati

Differentiability of $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$

Definition: Let $U \subset \mathbb{R}^n$ be open. Then $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $X_0 \in U$ if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{H \to 0} \frac{\|f(X_0 + H) - f(X_0) - L(H)\|}{\|H\|} = 0.$$

Differentiability of $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$

Definition: Let $U \subset \mathbb{R}^n$ be open. Then $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $X_0 \in U$ if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{H \to 0} \frac{\|f(X_0 + H) - f(X_0) - L(H)\|}{\|H\|} = 0.$$

The linear map L is called the derivative of f at X_0 and is denoted by $\mathrm{D}f(X_0)$, that is, $L = \mathrm{D}f(X_0)$.

Other notations: $f'(X_0)$, $\frac{\mathrm{d}f}{\mathrm{d}X}(X_0)$.

Characterization of differentiability

Theorem: Consider $f : \mathbb{R}^n \to \mathbb{R}^m$ with $f(X) = (f_1(X), \ldots, f_m(X))$, where $f_i : \mathbb{R}^n \to \mathbb{R}$. Then f is differentiable at $X_0 \in \mathbb{R}^n \iff f_i$ is differentiable at X_0 for $i = 1, 2, \ldots, m$. Further

$$\mathrm{D}f(X_0)(H) = (\nabla f_1(X_0) \bullet H, \ldots, \nabla f_m(X_0) \bullet H).$$

Characterization of differentiability

Theorem: Consider $f : \mathbb{R}^n \to \mathbb{R}^m$ with $f(X) = (f_1(X), \ldots, f_m(X))$, where $f_i : \mathbb{R}^n \to \mathbb{R}$. Then f is differentiable at $X_0 \in \mathbb{R}^n \iff f_i$ is differentiable at X_0 for $i = 1, 2, \ldots, m$. Further

$$\mathrm{D}f(X_0)(H) = (\nabla f_1(X_0) \bullet H, \ldots, \nabla f_m(X_0) \bullet H).$$

The matrix of $\mathrm{D}f(X_0)$ is called the Jacobian matrix of f at X_0 and is denoted by $J_f(X_0)$.

Jacobian matrix of $f : \mathbb{R}^n \to \mathbb{R}^m$

$J_f(X_0)$ is an $m \times n$ matrix with (i, j)-th entry $a_{ij} := \partial_j f_i(X_0)$.

Jacobian matrix of $f : \mathbb{R}^n \to \mathbb{R}^m$

$J_f(X_0)$ is an $m \times n$ matrix with (i, j)-th entry $a_{ij} := \partial_j f_i(X_0)$.

$$J_f(X_0) = \begin{bmatrix} \nabla f_1(X_0) \\ \vdots \\ \nabla f_m(X_0) \end{bmatrix}_{m\times n} = \left[\frac{\partial f_i}{\partial x_j}(X_0)\right]_{m\times n}.$$

Jacobian matrix of $f : \mathbb{R}^n \to \mathbb{R}^m$

$J_f(X_0)$ is an $m \times n$ matrix with (i,j)-th entry $a_{ij} := \partial_j f_i(X_0)$.

$$J_f(X_0) = \begin{bmatrix} \nabla f_1(X_0) \\ \vdots \\ \nabla f_m(X_0) \end{bmatrix}_{m\times n} = \left[\frac{\partial f_i}{\partial x_j}(X_0)\right]_{m\times n}.$$

- $f(x,y) = (f_1(x,y), f_2(x,y), f_3(x,y))$

$$J_f(a,b) = \begin{bmatrix} \partial_x f_1(a,b) & \partial_y f_1(a,b) \\ \partial_x f_2(a,b) & \partial_y f_2(a,b) \\ \partial_x f_3(a,b) & \partial_y f_3(a,b) \end{bmatrix}$$

Jacobian matrix of $f : \mathbb{R}^n \to \mathbb{R}^m$

$J_f(X_0)$ is an $m \times n$ matrix with (i,j)-th entry $a_{ij} := \partial_j f_i(X_0)$.

$$J_f(X_0) = \begin{bmatrix} \nabla f_1(X_0) \\ \vdots \\ \nabla f_m(X_0) \end{bmatrix}_{m\times n} = \left[\frac{\partial f_i}{\partial x_j}(X_0)\right]_{m\times n}.$$

- $f(x,y) = (f_1(x,y), f_2(x,y), f_3(x,y))$

$$J_f(a,b) = \begin{bmatrix} \partial_x f_1(a,b) & \partial_y f_1(a,b) \\ \partial_x f_2(a,b) & \partial_y f_2(a,b) \\ \partial_x f_3(a,b) & \partial_y f_3(a,b) \end{bmatrix}$$

- $f(x,y,z) = (f_1(x,y,z), f_2(x,y,z))$

$$J_f(a,b,c) = \begin{bmatrix} \partial_x f_1(a,b,c) & \partial_y f_1(a,b,c) & \partial_z f_1(a,b,c) \\ \partial_x f_2(a,b,c) & \partial_y f_2(a,b,c) & \partial_z f_2(a,b,c) \end{bmatrix}$$

Examples

- If $f(x,y) = (xy, e^x y, \sin y)$ then

$$J_f(x,y) = \begin{bmatrix} y & x \\ e^x y & e^x \\ 0 & \cos y \end{bmatrix}$$

Examples

- If $f(x,y) = (xy, e^x y, \sin y)$ then
$$J_f(x,y) = \begin{bmatrix} y & x \\ e^x y & e^x \\ 0 & \cos y \end{bmatrix}$$

- If $f(x,y,z) = (x+y+z, xyz)$ then
$$J_f(x,y,z) = \begin{bmatrix} 1 & 1 & 1 \\ yz & xz & xy \end{bmatrix}$$

Chain rule

Theorem-A: Let $X : \mathbb{R} \to \mathbb{R}^n$ be differentiable at t_0 and $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable at $X_0 := X(t_0)$. Then $f \circ X$ is differentiable at t_0 and

$$\frac{\mathrm{d}}{\mathrm{d}t} f(X(t))|_{t=t_0} = \nabla f(X_0) \bullet X'(t_0) = \sum_{i=1}^{n} \partial_i f(X_0) \frac{\mathrm{d}x_i(t_0)}{\mathrm{d}t}.$$

Chain rule

Theorem-A: Let $X : \mathbb{R} \to \mathbb{R}^n$ be differentiable at t_0 and $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable at $X_0 := X(t_0)$. Then $f \circ X$ is differentiable at t_0 and

$$\frac{\mathrm{d}}{\mathrm{d}t} f(X(t))|_{t=t_0} = \nabla f(X_0) \bullet X'(t_0) = \sum_{i=1}^{n} \partial_i f(X_0) \frac{\mathrm{d}x_i(t_0)}{\mathrm{d}t}.$$

Proof: Since f is differentiable at X_0, therefore

$$f(X_0 + H) = f(X_0) + \nabla f(X_0) \bullet H + E(H)\|H\| \cdots\cdots (*)$$

and $E(H) \to 0$ as $H \to 0$. Put $H := X(t) - X(t_0)$ in $(*)$ to complete the proof.

Chain rule for partial derivatives

Theorem-B: If $X : \mathbb{R}^2 \to \mathbb{R}^n, (u, v) \mapsto (x_1(u, v), \ldots, x_n(u, v))$ has partial derivatives at (a, b) and $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $Y := X(a, b)$ then $F(u, v) := f(X(u, v))$ has partial derivatives at (a, b) and

$$\begin{aligned}
\partial_u F(a, b) &= \nabla f(Y) \bullet \partial_u X(a, b) = \sum_{j=1}^{n} \frac{\partial f(Y)}{\partial x_j} \frac{\partial x_j(a, b)}{\partial u}, \\
\partial_v F(a, b) &= \nabla f(Y) \bullet \partial_v X(a, b) = \sum_{j=1}^{n} \frac{\partial f(Y)}{\partial x_j} \frac{\partial x_j(a, b)}{\partial v}.
\end{aligned}$$

Chain rule for partial derivatives

Case n=2:

If $x = x(u, v)$ and $y = y(u, v)$ have first order partial derivatives at the point (u, v), and if $z = f(x, y)$ is differentiable at the point $(x(u, v), y(u, v))$, then $z = f(x(u, v), y(u, v))$ has first order partial derivatives at (u, v) given by

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u} \text{ and } \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v} .$$

Example: Find $\partial w/\partial u$ and $\partial w/\partial v$ when $w = x^2 + xy$ and $x = u^2 v, y = uv^2$.

Graph and level set

Let $f : \mathbb{R}^n \to \mathbb{R}$. Then $G(f) := \{(X, f(X)) : X \in \mathbb{R}^n\} \subset \mathbb{R}^{n+1}$ is the graph of f. $G(f)$ represents a hyper-surface in $\mathbb{R}^{n+1}$.

Graph and level set

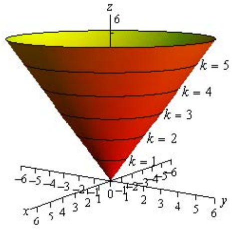
Let $f : \mathbb{R}^n \to \mathbb{R}$. Then $G(f) := \{(X, f(X)) : X \in \mathbb{R}^n\} \subset \mathbb{R}^{n+1}$ is the graph of f. $G(f)$ represents a hyper-surface in $\mathbb{R}^{n+1}$.

The set $S(f, \alpha) := \{X \in \mathbb{R}^n : f(X) = \alpha\}$ is called a level set of f and represents a hyper-surface in $\mathbb{R}^n$.

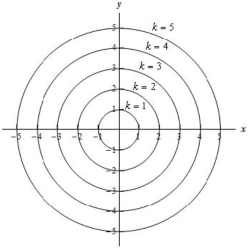
Graph and level set

Let $f : \mathbb{R}^n \to \mathbb{R}$. Then $G(f) := \{(X, f(X)) : X \in \mathbb{R}^n\} \subset \mathbb{R}^{n+1}$ is the graph of f. $G(f)$ represents a hyper-surface in $\mathbb{R}^{n+1}$.

The set $S(f, \alpha) := \{X \in \mathbb{R}^n : f(X) = \alpha\}$ is called a level set of f and represents a hyper-surface in $\mathbb{R}^n$.



(e) Graph of $f(x, y) := \sqrt{x^2 + y^2}$



(f) Level curve $\sqrt{x^2 + y^2} = k$

Level sets and gradients

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable at $X_0 \in \mathbb{R}^n$. Suppose that X_0 is point on the hyper-surface $f(X) = \alpha$.

Level sets and gradients

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable at $X_0 \in \mathbb{R}^n$. Suppose that X_0 is point on the hyper-surface $f(X) = \alpha$.

Let $X : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ be a curve on the hyper-surface $f(X) = \alpha$ passing through X_0, i.e, $X(0) = X_0$ and $f(X(t)) = \alpha$ for $t \in (-\varepsilon, \varepsilon)$.

Level sets and gradients

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable at $X_0 \in \mathbb{R}^n$. Suppose that X_0 is point on the hyper-surface $f(X) = \alpha$.

Let $X : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ be a curve on the hyper-surface $f(X) = \alpha$ passing through X_0, i.e, $X(0) = X_0$ and $f(X(t)) = \alpha$ for $t \in (-\varepsilon, \varepsilon)$.

Suppose that $X(t)$ is differentiable at 0. Then

$$0 = \frac{\mathrm{d}f(X(t))}{\mathrm{d}t}|_{t=0} = \nabla f(X_0) \bullet X'(0) \Rightarrow \nabla f(X_0) \perp X'(0)$$

Level sets and gradients

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable at $X_0 \in \mathbb{R}^n$. Suppose that X_0 is point on the hyper-surface $f(X) = \alpha$.

Let $X : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ be a curve on the hyper-surface $f(X) = \alpha$ passing through X_0, i.e, $X(0) = X_0$ and $f(X(t)) = \alpha$ for $t \in (-\varepsilon, \varepsilon)$.

Suppose that $X(t)$ is differentiable at 0. Then

$$0 = \frac{\mathrm{d}f(X(t))}{\mathrm{d}t}\Big|_{t=0} = \nabla f(X_0) \bullet X'(0) \Rightarrow \nabla f(X_0) \perp X'(0)$$

Since the line $X_0 + t\,X'(0)$ is tangent to the curve $X(t)$ at X_0, $\nabla f(X_0)$ is normal to the hyper-surface $f(X) = \alpha$ at X_0.

Tangent Plane and Normal Line to a Level Surface

$f : E \subseteq \mathbb{R}^3 \to \mathbb{R}$. $X_0 = (x_0,\ y_0,\ z_0)$ is a point on the level surface $S = \{(x,\ y,\ z) \in \mathbb{R}^3\ :\ f(x,\ y,\ z) = c\}$ where c is a fixed real number. Let f be differentiable at X_0.

The tangent plane to S at X_0 is the plane passing through X_0 and normal to the gradient vector ∇f at X_0. Its equation is

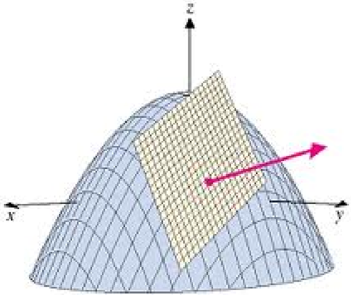
$$f_x(X_0)(x - x_0) + f_y(X_0)(y - y_0) + f_z(X_0)(z - z_0) = 0\ .$$

The normal line to S at X_0 is the line perpendicular to the tangent plane and parallel to $\nabla f(X_0)$, given by equations

$$x = x_0 + f_x(X_0)t,\ \ y = y_0 + f_y(X_0)t,\ \ z = z_0 + f_z(X_0)t$$

for $t \in \mathbb{R}$.

Normal Line and Tangent Plane to a Surface



Example

Find the tangent plane and normal line to the surface $x^2 + y^2 + z^2 = 3$ at the point $(-1, 1, 1)$.
The given surface can be written as a level surface $f(x, y, z) = 3$ where $f(x, y, z) = x^2 + y^2 + z^2$.
$f_x(x, y, z) = 2x$, $f_y(x, y, z) = 2y$ and $f_z(x, y, z) = 2z$
$f_x(-1, 1, 1) = -2$, $f_y(-1, 1, 1) = 2$, $f_z(-1, 1, 1) = 2$.
Equation of tangent plane:

$$-2(x - (-1)) + 2(y - 1) + 2(z - 1) = 0$$

or equivalently, $-x + y + z = 3$.
The normal line to the surface at $(-1, 1, 1)$ is given by

$$x = -1 - 2t, \;\; y = 1 + 2t, \;\; z = 1 + 2t \qquad \text{for } t \in \mathbb{R}\,.$$

Tangent Plane and Normal line for a Surface $z = f(x,\ y)$

The equation for a surface $S:\ z = f(x,\ y)$ can be written in the form $f(x,\ y) - z = 0$.
Hence, the surface $z = f(x,\ y)$ is also the level surface $F(x,\ y,\ z) = 0$ of the function $F(x,\ y,\ z) = f(x,\ y) - z$.
If $X_0 = (x_0,\ y_0,\ z_0)$ is a point on the surface $z = f(x,\ y)$ and f_x and f_y are continuous at $(x_0,\ y_0)$ then the tangent plane to the surface S at X_0 is the plane

$$f_x(x_0,\ y_0)(x - x_0) + f_y(x_0,\ y_0)(y - y_0) - (z - z_0) = 0$$

or equivalently

$$z = z_0 + f_x(x_0,\ y_0)(x - x_0) + f_y(x_0,\ y_0)(y - y_0).$$

Tangent Plane and Normal line for a Surface $z = f(x, y)$

The normal line to the surface S at P_0 is the line

$$x = x_0 + f_x(x_0, y_0)t, \quad y = y_0 + f_y(x_0, y_0)t, \quad z = z_0 - t \text{ for } t \in \mathbb{R}.$$

Example: Find equations for the tangent plane and normal line to the surface $z = 9 - x^2 - y^2$ at the point $P_0 = (1, 2, 4)$.
Observe that $f_x(1, 2) = -2$ and $f_y(1, 2) = -4$. Therefore, the equation for the tangent plane is

$$(-2)(x-1) + (-4)(y-2) - (z-4) = 0 \implies 2x + 4y + z = 14 .$$

The equation for the normal line is

$$x = 1 - 2t, \quad y = 2 - 4t, \quad z = 4 - t \text{ for } t \in \mathbb{R}.$$

MA102: Multivariable Calculus

Arup Chattopadhyay and Subhamay Saha
Department of Mathematics
IIT Guwahati

Partial Derivatives of Higher Order

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x, y) = x^4 y + y^3 x$ for $(x, y) \in \mathbb{R}^2$. Then

$$\frac{\partial f}{\partial x} = 4x^3 y + y^3 \qquad \text{and} \qquad \frac{\partial f}{\partial y} = x^4 + 3y^2 x \,.$$

Now, these first order partial derivatives are again functions from $\mathbb{R}^2$ to $\mathbb{R}$. Again, we can do partial differentiation with respect to the variables x and y. For example,

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}\left(4x^3 y + y^3\right) = 12x^2 y = \frac{\partial^2 f}{\partial x^2}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial y}\left(4x^3 y + y^3\right) = 4x^3 + 3y^2 = \frac{\partial^2 f}{\partial y\, \partial x}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial x}\left(x^4 + 3y^2 x\right) = 4x^3 + 3y^2 = \frac{\partial^2 f}{\partial x\, \partial y}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}\left(x^4 + 3y^2 x\right) = 6xy = \frac{\partial^2 f}{\partial y^2}$$

Notations for Partial Derivatives of Higher Order

Let $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ where S is an open set in $\mathbb{R}^n$.

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) = f_{x_i x_j} = \partial_{ij} f = \partial_{x_i x_j} f$$

$$\frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i} = \frac{\partial}{\partial x_k}\left(\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)\right) = \partial_{ijk} f = f_{x_i x_j x_k}$$

and so on.

Note: In the notation $f_{x_i x_j x_k}$, the variable x_i which is close to f is first and then next variable x_j and so on.

Mixed Partial Derivatives:

The partial derivatives ∂_{ij} and ∂_{ji} with $i \neq j$ are called the mixed (partial) derivatives.

$\partial_{ij} f(X_0) \neq \partial_{ji} f(X_0)$ is possible

Let $f(x, y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & \text{if} \quad (x, y) \neq (0, 0), \\ 0 & \text{if} \quad (x, y) = (0, 0). \end{cases}$

Prove that

- f, f_x, f_y are continuous in $\mathbb{R}^2$.
- $\partial_{12} f$ and $\partial_{21} f$ exist at every point of $\mathbb{R}^2$, and are continuous except at (0, 0).
- $\partial_{12} f(0, 0) = 1$ and $\partial_{21} f(0, 0) = -1$ (Second Order Mixed Partial Derivatives are not equal at origin).
- If $X_0 = (x_0, y_0) \neq (0, 0)$ then $\partial_{12} f(X_0) = \partial_{21} f(X_0)$.

When mixed partial derivatives are equal?

Theorem

Let $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ where S is an open set in $\mathbb{R}^n$. Let $X_0 \in S$. Assume that $\partial_{ij} f$ and $\partial_{ji} f$ exist in a neighborhood of the point X_0.

If the mixed partial derivatives $\partial_{ij} f$ and $\partial_{ji} f$ are continuous at X_0 then

$$\partial_{ij} f(X_0) = \partial_{ji} f(X_0) \ .$$

If $\partial_{ij} f$ and $\partial_{ji} f$ are not continuous at a point then they may or may not be equal. Let $f(x,\, y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & \text{if} \quad (x,\, y) \neq (0,\, 0)\,, \\ 0 & \text{if} \quad (x,\, y) = (0,\, 0)\,. \end{cases}$

Then we have

$$f_{xy}(x,y) = \begin{cases} \frac{x^6+9x^4y^2-9x^2y^4-y^6}{(x^2+y^2)^3} & (x,y) \neq (0,0), \\ -1 & (x,y) = (0,0). \end{cases}$$

$$f_{yx}(x,y) = \begin{cases} \frac{x^6+9x^4y^2-9x^2y^4-y^6}{(x^2+y^2)^3} & (x,y) \neq (0,0), \\ 1 & (x,y) = (0,0). \end{cases}$$

Here f_{xy} and f_{yx} are not continuous at $(0,0)$ (put $y = mx$).

On the other hand, let $f(x,\, y) = \begin{cases} \frac{x^2y^2}{x^2+y^2} & \text{if} \quad (x,\, y) \neq (0,\, 0)\,, \\ 0 & \text{if} \quad (x,\, y) = (0,\, 0)\,. \end{cases}$

Then we have

$$f_{xy}(x,y) = f_{yx}(x,y) = \begin{cases} \frac{x^6+9x^4y^2-9x^2y^4-y^6}{(x^2+y^2)^3} & (x,y) \neq (0,0), \\ 0 & (x,y) = (0,0). \end{cases}$$

Here again f_{xy} and f_{yx} are not continuous at $(0,0)$ (put $y = mx$), but they are equal at $(0,0)$.

Continuous partial derivatives

Let S be an open subset of $\mathbb{R}^n$ and $f : S \subset \mathbb{R}^n \to \mathbb{R}$.

Suppose that $\partial_i f(X)$ exists for all $X \in S$ and $i = 1, \ldots, n$. Then each $\partial_i f$ defines a function on S.

If $\partial_i f : S \to \mathbb{R}, X \mapsto \partial_i f(X)$ is continuous for $i = 1, \ldots, n$ then f is said to be continuously differentiable on S (in short, C^1).

Continuous partial derivatives

Let S be an open subset of $\mathbb{R}^n$ and $f : S \subset \mathbb{R}^n \to \mathbb{R}$.

Suppose that $\partial_i f(X)$ exists for all $X \in S$ and $i = 1, \dots, n$. Then each $\partial_i f$ defines a function on S.

If $\partial_i f : S \to \mathbb{R}, X \mapsto \partial_i f(X)$ is continuous for $i = 1, \dots, n$ then f is said to be continuously differentiable on S (in short, C^1).

Fact: f is $C^1 \iff \nabla f : S \subset \mathbb{R}^n \to \mathbb{R}^n, X \mapsto \nabla f(X)$ is continuous.

Recall: f is $C^1 \Rightarrow f$ is differentiable $\not\Rightarrow f$ is C^1.

Examples:

- Consider $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = x^2 + e^{xy} + y^2$. Then f is C^1.

- Consider $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(0, 0) = 0$ and $f(x, y) := (x^2 + y^2)\sin(1/(x^2 + y^2))$ if $(x, y) \neq (0, 0)$. Then f is differentiable but NOT C^1.

Continuous partial derivatives

Let S be an open subset of $\mathbb{R}^n$. Let $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ be differentiable so that $\partial_i f : S \to \mathbb{R}$ for $i = 1, \ldots, n$.

If the partial derivatives of $\partial_j f$ exist at $X_0 \in S$ for $j = 1, \ldots, n$, that is, $\partial_i \partial_j f(X_0)$ exists for $i, j = 1, 2, \ldots, n$, then f is said to have second order partial derivatives at X_0.

Continuous partial derivatives

Let S be an open subset of $\mathbb{R}^n$. Let $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ be differentiable so that $\partial_i f : S \to \mathbb{R}$ for $i = 1, \ldots, n$.

If the partial derivatives of $\partial_j f$ exist at $X_0 \in S$ for $j = 1, \ldots, n$, that is, $\partial_i \partial_j f(X_0)$ exists for $i, j = 1, 2, \ldots, n$, then f is said to have second order partial derivatives at X_0.

f is said to be C^2 (twice continuously differentiable) if $\partial_i \partial_j f(X)$ exists for $X \in S$ and $\partial_i \partial_j f : S \to \mathbb{R}$ is continuous for $i, j = 1, 2, \ldots, n$.

- p-th order partial derivatives of f are defined similarly.

Continuous partial derivatives

Fact: $f : \mathbb{R}^n \to \mathbb{R}$ is $C^2 \Rightarrow \nabla f : \mathbb{R}^n \to \mathbb{R}^n$ is differentiable.

Hessian:

Suppose that $f(x, y)$ has second order partial derivatives at $X_0 = (a, b)$. Then the matrix

$$H_f(X_0) := \begin{bmatrix} \partial_x \partial_x f(X_0) & \partial_y \partial_x f(X_0) \\ \partial_x \partial_y f(X_0) & \partial_y \partial_y f(X_0) \end{bmatrix} = \begin{bmatrix} f_{xx}(X_0) & f_{xy}(X_0) \\ f_{yx}(X_0) & f_{yy}(X_0) \end{bmatrix}$$

is called the Hessian of f at X_0.

Hessian

Fact: Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is C^2 and $X_0 \in \mathbb{R}^n$. Then the Hessian

$$H_f(X_0) := \begin{bmatrix} \partial_1\partial_1 f(X_0) & \cdots & \partial_n\partial_1 f(X_0) \\ \vdots & \cdots & \vdots \\ \partial_1\partial_n f(X_0) & \cdots & \partial_n\partial_n f(X_0) \end{bmatrix}$$

is symmetric.
Also $H_f(X_0) = J_{\nabla f}(X_0) =$ Jacobian of ∇f at X_0.

Hessian

Fact: Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is C^2 and $X_0 \in \mathbb{R}^n$. Then the Hessian

$$H_f(X_0) := \begin{bmatrix} \partial_1\partial_1 f(X_0) & \cdots & \partial_n\partial_1 f(X_0) \\ \vdots & \cdots & \vdots \\ \partial_1\partial_n f(X_0) & \cdots & \partial_n\partial_n f(X_0) \end{bmatrix}$$

is symmetric.
Also $H_f(X_0) = J_{\nabla f}(X_0) =$ Jacobian of ∇f at X_0.

Example: Consider $f(x, y) = x^2 - 2xy + 2y^2$. Then

$$H_f(x, y) = \begin{bmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}.$$

Mean Value Theorem

Let $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ where S is an open and convex set in $\mathbb{R}^n$. If f is differentiable in S then for any two points X_1 and X_2 in S, there exists a point X_0 on the line segment L joining X_1 and X_2 such that

$$f(X_2) - f(X_1) = \nabla f(X_0) \bullet (X_2 - X_1).$$

Proof: Let $\Phi : \mathbb{R} \to \mathbb{R}^n$ be defined by

$$\Phi(t) = (1-t)X_1 + tX_2 \qquad \text{for } t \in \mathbb{R}\ .$$

Consider the function $g(t) = f(\Phi(t))$ for $t \in [0,\ 1]$ and invoke chain rule and apply the mean value theorem of single variable calculus to g.

MA102: Multivariable Calculus

Arup Chattopadhyay and Subhamay Saha
Department of Mathematics
IIT Guwahati

Taylor's Theorem

We can also write the MVT in the following way:
Let $f : S \to \mathbb{R}$ be differentiable. Let $X_0 \in S$. Then there exists $0 < \theta < 1$ such that

$$f(X_0 + H) \;=\; f(X_0) + \sum_{i=1}^{n} \partial_i f(X_0 + \theta H) h_i.$$

Let $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ where S is an open set in $\mathbb{R}^n$. Let $X_0 = (x_1, \cdots, x_n) \in S$ and let $H = (h_1, \cdots, h_n) \in \mathbb{R}^n$ be such that the line segment L joining X_0 and $X_0 + H$ lies inside S.
Suppose that f and its partial derivatives through order $(n+1)$ are continuous in S.

Taylor's Theorem

Then, for some $0 < \theta < 1$, we have

$$f(X_0 + H) = f(X_0) + \sum_{k=1}^{n} \frac{1}{k!} \left(h_1\partial_1 + \cdots + h_n\partial_n\right)^k f\Big|_{X_0}$$

$$+ \frac{1}{(n+1)!} \left(h_1\partial_1 + \cdots + h_n\partial_n\right)^{n+1} f\Big|_{X_0+\theta H}.$$

Case n=2: Let $f : \mathbb{R}^2 \to \mathbb{R}$. Set the differential operators as: $F' = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)$,

$F'' = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 = h^2\frac{\partial^2}{\partial x^2} + 2hk\frac{\partial^2}{\partial x \partial y} + k^2\frac{\partial^2}{\partial y^2}$ and

$F^{(n)} = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n$ (Expand it by binomial theorem).

Taylor's Formula:

Suppose $f(x,\ y)$ and its partial derivatives through order $(n+1)$ are continuous throughout an open rectangular region R centered at $(a,\ b)$ in $\mathbb{R}^2$. Then, throughout R,

$$f(a+h,\ b+k) = f(a,\ b)+\sum_{r=1}^{n}\frac{1}{r!}\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^{r} f|_{(a,\ b)}+\text{Remainder term}$$

The *Remainder term* is

$$\frac{1}{(n+1)!}\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^{n+1} f|_{(a+ch,\ b+ck)}$$

where $((a+ch),\ (b+ck))$ for some c $(0 < c < 1)$, is a point on the line segment joining $(a,\ b)$ and $(a+h,\ b+k)$.

Taylor's formula and Polynomial Approximations

Taylor's formula provides polynomial approximations of f near the point (a, b). The first n derivative terms yield the n-the degree polynomial approximation to f and the last term (remainder term) gives the approximation error. Therefore, an upper bound of the remainder term is called as an upper bound for error term while approximating f by the n-th degree polynomial.
If $n = 1$, we get linear approximation of f near (a, b).
If $n = 2$, we get quadratic approximation of f near (a, b).
If $n = 3$, we get cubic approximation of f near (a, b).

Example

Let $f(x, y) = \frac{1}{xy}$ if $xy \neq 0$. Let $(a, b) = (1, -1)$. Compute the first two terms in the Taylor's formula of f near $(1, -1)$.

$f_x = -x^{-2}y^{-1}$, $f_y = -x^{-1}y^{-2}$, $f_{xx} = 2x^{-3}y^{-1}$, $f_{xy} = x^{-2}y^{-2}$ and $f_{yy} = 2x^{-1}y^{-3}$.

$$\frac{1}{(1+h)(-1+k)} = -1 + (h-k) + \left(\frac{h^2}{(1+\theta h)^3(-1+\theta k)} + \frac{hk}{(1+\theta h)^2(-1+\theta k)^2} + \frac{k^2}{(1+\theta h)(-1+\theta k)^3}\right)$$

where $0 < \theta < 1$.

Maxima/Minima

Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$.

If f is continuous on E and E is a closed & bounded, then f attains its maximum and minimum value on E.

How to find the (extremum) points at which f attains the maximum value or the minimum value on E?

Before finding answer to this question, we formally define maximum and minimum of f.

Local minimum/maximum and Global minimum/maximum

Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$.
A point $X^* \in E$ is said to be a point of relative/local minimum of f if there exists a $r > 0$ such that $f(X^*) \leq f(X)$ for all $X \in E$ with $\|X - X^*\| < r$. In such case, the value $f(X^*)$ is called the relative/local minimum of f.

Local minimum/maximum and Global minimum/maximum

Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$.
A point $X^* \in E$ is said to be a point of relative/local minimum of f if there exists a $r > 0$ such that $f(X^*) \leq f(X)$ for all $X \in E$ with $\|X - X^*\| < r$. In such case, the value $f(X^*)$ is called the relative/local minimum of f.

A point $X^* \in E$ is said to be a point of relative/local maximum of f if there exists a $r > 0$ such that $f(X) \leq f(X^*)$ for all $X \in E$ with $\|X - X^*\| < r$. In such case, the value $f(X^*)$ is called the relative/local maximum of f.

Local minimum/maximum and Global minimum/maximum

Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$.
A point $X^* \in E$ is said to be a point of relative/local minimum of f if there exists a $r > 0$ such that $f(X^*) \leq f(X)$ for all $X \in E$ with $\|X - X^*\| < r$. In such case, the value $f(X^*)$ is called the relative/local minimum of f.

A point $X^* \in E$ is said to be a point of relative/local maximum of f if there exists a $r > 0$ such that $f(X) \leq f(X^*)$ for all $X \in E$ with $\|X - X^*\| < r$. In such case, the value $f(X^*)$ is called the relative/local maximum of f.

A point $X^* \in E$ is said to be a point of absolute/ global minimum of f if $f(X^*) \leq f(X)$ for all $X \in E$.

Local minimum/maximum and Global minimum/maximum

Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$.
A point $X^* \in E$ is said to be a point of relative/local minimum of f if there exists a $r > 0$ such that $f(X^*) \leq f(X)$ for all $X \in E$ with $\|X - X^*\| < r$. In such case, the value $f(X^*)$ is called the relative/local minimum of f.

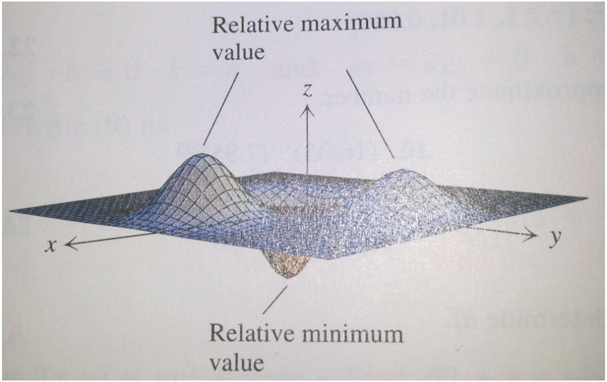
A point $X^* \in E$ is said to be a point of relative/local maximum of f if there exists a $r > 0$ such that $f(X) \leq f(X^*)$ for all $X \in E$ with $\|X - X^*\| < r$. In such case, the value $f(X^*)$ is called the relative/local maximum of f.

A point $X^* \in E$ is said to be a point of absolute/ global minimum of f if $f(X^*) \leq f(X)$ for all $X \in E$.

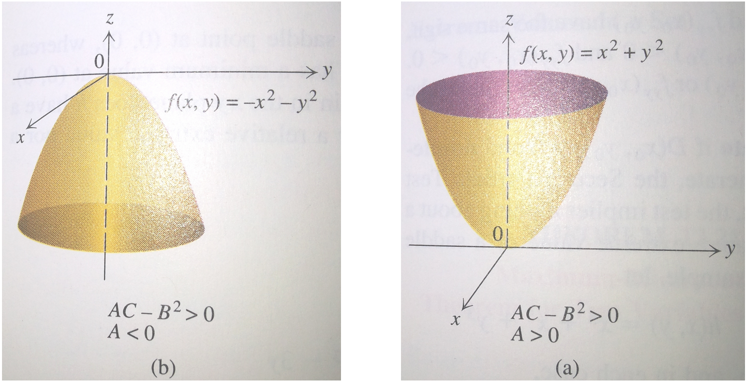
A point $X^* \in E$ is said to be a point of absolute/ global maximum of f if $f(X) \leq f(X^*)$ for all $X \in E$.

Extremum Points and Extremum Values

A point $X^* \in E$ is said to be a point of extremum of f if it is either (local/global) minimum point or maximum point of f. The function value $f(X^*)$ at the extremum point X^* is called an extremum value of f.

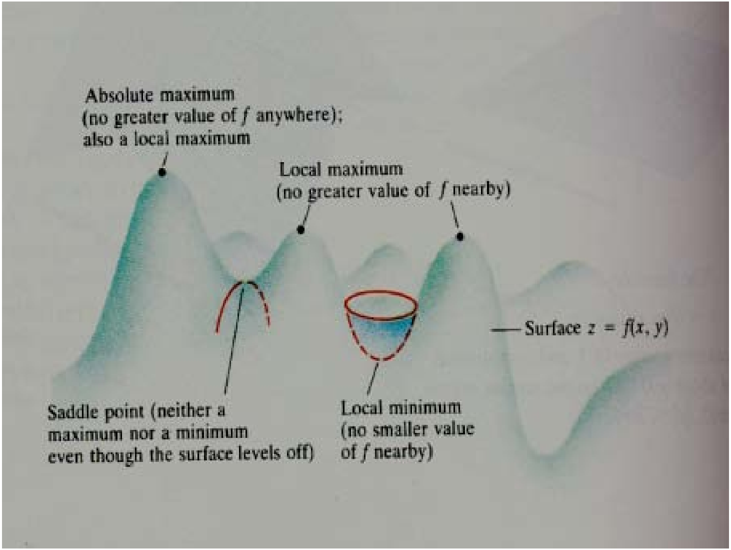


Absolute/Global Extremum



$f(x, y) = -x^2 - y^2$ has an absolute maximum at $(0, 0)$ in $\mathbb{R}^2$.
$f(x, y) = x^2 + y^2$ has an absolute minimum at $(0, 0)$ in $\mathbb{R}^2$.

Relative/Local Extremum and Saddle Point



Critical Points

Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$. A point $X^* \in E$ is said to be a critical point of f if

- either

$$\frac{\partial f}{\partial x_1}(X^*) = \frac{\partial f}{\partial x_2}(X^*) = \cdots = \frac{\partial f}{\partial x_n}(X^*) = 0 \ ,$$

- or at least one of the first order partial derivatives of f does not exist.

The function value $f(X^*)$ at the critical point X^* is called a critical value of f.

Critical Points: Examples

The point (0, 0) is the critical point of the function $f(x, y) = x^2 + y^2$ and the critical value is 0 corresponding to this critical point.

The point (0, 0) is a critical point of the function

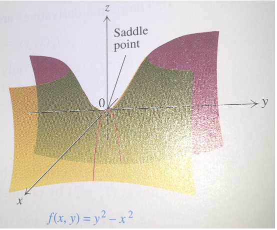
$$h(x, y) = \begin{cases} x \sin(1/x) + y & \text{if} \quad x \neq 0\,, \\ y & \text{if} \quad x = 0\,. \end{cases}$$

Here, $h_x(0,0)$ does not exist and $h_y(0,0) = 1$.

Saddle Points

Let X^* be a critical point of f. If every neighborhood $N(X^*)$ of the point X^* contains points at which f is strictly greater than $f(X^*)$ and also contains points at which f is strictly less than $f(X^*)$. Such X^* is said to be a Saddle point of f. That is, f attains neither relative maximum nor relative minimum at the critical point X^*.

Example: Let $f(x,\ y) = y^2 - x^2$ for $(x,\ y) \in \mathbb{R}^2$. Then $(0,\ 0)$ is a saddle point of f.



To find the extremum points of f, where to look for?

If $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$, then where to look in E for extremum values of f?
The maxima and minima of f can occur only at

- boundary points of E,
- critical points of E
  - interior point of E where all the first order partial derivatives of f are zero,
  - interior point of E where at least one of the first order partial derivatives of f does not exist.

Example:
In the closed rectangle
$R = \{(x, y) \in \mathbb{R}^2 \; : \; -1 \leq x \leq 1 \text{ and } -1 \leq y \leq 1\}$, the function $f(x, y) = x^2 + y^2$ attains

- its minimum value at $(0, 0)$,
- its maximum value at $(\pm 1, \pm 1)$.

Necessary Condition for Extremum

Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$. If an interior point X^* of E is a point of relative/absolute extremum of f, and if the first order partial derivatives of f at X^* exists then

$$\partial_1 f(X^*) = \cdots = \partial_n f(X^*) = 0.$$

That is, the gradient vector at X^* is the zero vector.

Further, the directional derivative of f at X^* in all directions is zero, if f is differentiable at X^*.

Quadratic Forms

Definition

Let $H = [a_{ij}]$ be an $n \times n$ symmetric matrix. A function of the form

$$Q(X) = X^T\ H\ X \text{ for } X^T = (x_1,\ x_2,\ \cdots,\ x_n) \in \mathbb{R}^n$$

from $\mathbb{R}^n$ into $\mathbb{R}$ is called a quadratic form (or bilinear form).

Examples of Quadratic Forms:

$$Q(x_1,\ x_2) = (x_1,\ x_2) \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = ax_1^2 + bx_1x_2 + cx_2^2\ .$$

$$Q(x_1,\ x_2,\ x_3) = (x_1,\ x_2,\ x_3) \begin{bmatrix} a & d/2 & f/2 \\ d/2 & b & e/2 \\ f/2 & e/2 & c \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= ax_1^2 + bx_2^2 + cx_3^2 + dx_1x_2 + ex_2x_3 + fx_3x_1\ .$$

Classification of quadratic forms Q

- If $Q(X) > 0$ for all $X \neq 0$, then Q is said to be positive definite.
- If $Q(X) < 0$ for all $X \neq 0$, then Q is said to be negative definite.
- If $Q(X) > 0$ for some X and $Q(X) < 0$ for some other X, then Q is said to be indefinite.
- If $Q(X) \geq 0$ for all X and $Q(X) = 0$ for some $X \neq 0$, then Q is said to be positive semidefinite.
- If $Q(X) \leq 0$ for all X and $Q(X) = 0$ for some $X \neq 0$, then Q is said to be negative semidefinite.

All the above terms used to describe quadratic forms Q can also be applied to the corresponding symmetric matrices H.

Examples

- Positive Definite: $Q(x_1, x_2) = x_1^2 + x_2^2$
- Negative Definite: $Q(x_1, x_2) = -x_1^2 - x_2^2$
- Indefinite: $Q(x_1, x_2) = x_1^2 - x_2^2$,
  Reasons: $Q(1, 0) = 1 > 0$ and $Q(0, 1) = -1 < 0$
- Positive Semidefinite: $Q(x_1, x_2) = x_2^2$,
  Reasons: $Q(X) \geq 0$ for all X and $Q(1, 0) = 0$.
- Negative Semidefinite: $Q(x_1, x_2) = -x_2^2$,
  Reasons: $Q(X) \leq 0$ for all X and $Q(1, 0) = 0$.

Classifying Quadratic Forms from the nature of Eigenvalues

Theorem

Let $Q(X) = X^T\ H\ X$ for $X^T = (x_1,\ x_2,\ \cdots,\ x_n) \in \mathbb{R}^n$ where H is a $n \times n$ symmetric matrix.

- If all the eigenvalues of H are positive, then Q (and H) is positive definite.
- If all the eigenvalues of H are negative, then Q (and H) is negative definite.
- If H has both positive and negative eigenvalues, then Q (and H) is indefinite.
- If all eigenvalues of H are non-negative (≥ 0), then H is positive semidefinite.
- If all eigenvalues of H are non-positive (≤ 0), then H is negative semidefinite.

Examples

- Positive Definite: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, Eigenvalues are 1, 2, 3.
- Negative Definite: $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$,

  Eigenvalues are -1, -2, -3.
- Indefinite: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$, Eigenvalues are 1, -2, -3.

Examples (Continuation)

- Positive Semidefinite: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, Eigenvalues are 1, 2, 0.
- Negative Semidefinite: $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix}$,

  Eigenvalues are -1, -2, 0.

Results for 2×2 Real Symmetric Matrices

Let $H = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ be a 2×2 symmetric matrix. Then H is

- positive definite if $\det H = ac - b^2 > 0$ and $a > 0$;
- negative definite if $\det H = ac - b^2 > 0$ and $a < 0$;
- indefinite if $\det H = ac - b^2 < 0$;

Example:

- Positive Definite: $\begin{bmatrix} 2 & 3 \\ 3 & 8 \end{bmatrix}$, $\det H = 7 > 0$ and $a = 2 > 0$.
- Negative Definite: $\begin{bmatrix} -2 & 3 \\ 3 & -8 \end{bmatrix}$, $\det H = 7 > 0$ and $a = -2 < 0$.
- Indefinite: $\begin{bmatrix} 2 & -3 \\ -3 & 1 \end{bmatrix}$, $\det H = -7 < 0$.

MA102: Multivariable Calculus

Arup Chattopadhyay and Subhamay Saha
Department of Mathematics
IIT Guwahati

Second Derivative Test

Motivation/Idea: When $f : \mathbb{R}^2 \to \mathbb{R}$, the Taylors Theorem (upto 2-nd derivative term) takes the form

$$f(X_0 + X) - f(X_0) = (xf_x(X_0) + yf_y(X_0))$$
$$+(x^2 f_{xx}(P) + 2xy\, f_{xy}(P) + y^2 f_{yy}(P)) \text{ for } X \in N(X_0).$$

Here, blue color terms are: $Q(X) = X^T HX$

Nature of Hessian matrix $\Longrightarrow$ Nature of Critical Point

Theorem

Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$ and let X_0 be an interior point of E. Suppose that all the second order partial derivatives of f exist and continuous at X_0 and X_0 is a critical point of f. Let H denote the Hessian matrix of f.

- If H is positive definite, then f has a local minimum at X_0.
- If H is negative definite, then f has a local maximum at X_0.
- If H is indefinite, then f has a saddle point at X_0.
- If H is semidefinite, then the test is inconclusive.

Case $n = 2$: $f : E \subseteq \mathbb{R}^2 \to \mathbb{R}$

Let $f : E \subseteq \mathbb{R}^2 \to \mathbb{R}$ and let X_0 be an interior point of E. Suppose that all the second order partial derivatives of f exist and continuous at X_0 and X_0 is a critical point of f. Then, the Hessian matrix of f is given by $H = \begin{bmatrix} f_{xx}(X_0) & f_{xy}(X_0) \\ f_{xy}(X_0) & f_{yy}(X_0) \end{bmatrix}$.

- If H is positive definite (that is, $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at X_0), then f has a local minimum at X_0.
- If H is negative definite (that is, $f_{xx} < 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at X_0), then f has a local maximum at X_0.
- If H is indefinite (that is, $f_{xx}f_{yy} - f_{xy}^2 < 0$ at X_0), then f has a saddle point at X_0.
- If H is semidefinite (that is, $f_{xx}f_{yy} - f_{xy}^2 = 0$ at X_0), then the test is inconclusive.

Example 1

Let $f(x,\ y) = x^2 - 2xy + \frac{y^3}{3} - 3y$. Determine at which points f have relative extremum values and at which points f has saddle points.

Step 1: Finding Critical Points:

Observe that $f_x(x,\ y) = 2x - 2y$ and $f_y(x,\ y) = -2x + y^2 - 3$. The critical points of f are the solutions of $f_x(x,\ y) = 0$ and $f_y(x,\ y) = 0$.

$$f_x(x,\ y) = 2x - 2y = 0 \text{ and } \ f_y(x,\ y) = -2x + y^2 - 3 = 0$$

$$\implies (x,\ y) = (3,\ 3) \text{ or } (-1,\ -1)\ .$$

Thus, the points $(3,\ 3)$ and $(-1,\ -1)$ are critical points of f.

Step 2: Determining Nature of Each Critical Point:

At $(3, 3)$, $f_{xx}f_{yy} - f_{xy}^2 = 8 > 0$ and $f_{xx} = 2 > 0$. Therefore, f has a relative minimum value at the critical point $(3,\ 3)$.

At $(-1, -1)$, $f_{xx}f_{yy} - f_{xy}^2 = -8 < 0$. Therefore, the function f has a saddle point at the critical point $(-1,\ -1)$.

Example 2

Let $f(x, y, z) = 1 - x^2 - 2y^2 - 3z^2$ for (x, y, z). Determine all critical points of f in $\mathbb{R}^3$ and find its nature.
Step 1: Finding Critical Points of f
$f_x = -2x = 0$, $f_y = -4y = 0$ and $f_z = -6z = 0$ implies that $(0, 0, 0)$ is the critical point of f.

Step 2: Determining the nature of each Critical Point
At $(0, 0, 0)$, the Hessian matrix $H = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -6 \end{bmatrix}$. The eigenvalues of H are -2, -4 and -6. Since all eigenvalues of H are negative, H is negative definite. Hence $(0, 0, 0)$ is a local maximum point of f.

Example 3

Find the maxima, minima and saddle points of $f(x, y) := (x^2 - y^2)e^{-(x^2+y^2)/2}$.

Example 3

Find the maxima, minima and saddle points of $f(x,y) := (x^2 - y^2)e^{-(x^2+y^2)/2}$. We have

$$\begin{aligned} f_x &= [2x - x(x^2 - y^2)]e^{-(x^2+y^2)/2} = 0, \\ f_y &= [-2y - y(x^2 - y^2)]e^{-(x^2+y^2)/2} = 0, \end{aligned}$$

so the critical points are $(0,0), (\pm\sqrt{2}, 0)$ and $(0, \pm\sqrt{2})$.

Point	f_{xx}	f_{xy}	f_{yy}	$\mathbf{det}(H)$	Type —
$(0,0)$	2	0	-2	-4	saddle
$(\sqrt{2},0)$	$-4/e$	0	$-4/e$	$16/e^2$	maximum
$(-\sqrt{2},0)$	$-4/e$	0	$-4/e$	$16/e^2$	maximum
$(0,\sqrt{2})$	$4/e$	0	$4/e$	$16/e^2$	minimum
$(0,-\sqrt{2})$	$4/e$	0	$4/e$	$16/e^2$	minimum

Global extrema

Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$ be continuous.
If E is a closed & bounded, then it is known that f attains its maximum and minimum value on E.

Find global maximum and global minimum of the function $f : [-2, 2] \times [-2, 2] \to \mathbb{R}$ given by $f(x, y) := 4xy - 2x^2 - y^4$.

Global extrema

Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$ be continuous.
If E is a closed & bounded, then it is known that f attains its maximum and minimum value on E.

Find global maximum and global minimum of the function $f : [-2, 2] \times [-2, 2] \to \mathbb{R}$ given by $f(x, y) := 4xy - 2x^2 - y^4$.

To find global extrema, find extrema of f in the interior and then on the boundary.

Global extrema

Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$ be continuous.
If E is a closed & bounded, then it is known that f attains its maximum and minimum value on E.

Find global maximum and global minimum of the function $f : [-2, 2] \times [-2, 2] \to \mathbb{R}$ given by $f(x, y) := 4xy - 2x^2 - y^4$.

To find global extrema, find extrema of f in the interior and then on the boundary.

Solving $f_x = 4y - 4x = 0$ and $f_y = 4x - 4y^3 = 0$ we obtain the critical points $(0, 0), (1, 1)$ and $(-1, -1)$. We have $f(1, 1) = f(-1, -1) = 1$.

$(0, 0)$ is a saddle point.

Global extrema

For the boundary, consider $f(x,2), f(x,-2), f(2,y), f(-2,y)$ and find their extrema on $[-2,2]$.

The global minimum is attained at $(2,-2)$ and $(-2,2)$ with $f(2,-2) = -40$.

The global maximum is attained at $(1,-1)$ and $(-1,1)$.

Example: Find the absolute maxima and minima of $f(x,y) = 2 + 2x + 2y - x^2 - y^2$ on the triangular region in the first quadrant bounded by the lines $x = 0, y = 0, y = 9 - x$.

We have $f_x = 2 - 2x = 0,\ f_y = 2 - 2y = 0$ which implies that $x = 1, y = 1$ is the only critical point (local maxima).

We next study the behaviour of f on the boundary which is a triangle.

Case 1. On the segment $y = 0$, $\varphi(x) := f(x, 0) = 2 + 2x - x^2$ is defined on $I = [0, 9]$. We have $\varphi(0) = f(0, 0) = 2, \varphi(9) = f(9, 0) = -61$. On $(0, 9)$, $\varphi'(x) = 2 - 2x = 0$ gives $x = 1$. Thus, $x = 1$ is the only critical point of $\varphi(x)$ in $(0, 9)$ and $\varphi(1) = f(1, 0) = 3$.

Case 2. On the segment $x = 0$, $\psi(y) = f(0, y) = 2 + 2y - y^2$ and $\psi'(y) = 2 - 2y = 0$ implies $y = 1$ and $\psi(1) = f(0, 1) = 3$.

Case 3. On the segment $y = 9 - x$, we have

$$f(x, 9 - x) = -61 + 18x - 2x^2$$

and the critical point is $x = 9/2$. At this point $f(9/2, 9/2) = -41/2$.

Finally, $f(1, 1) = 4$ (Global Maximum), $f(9, 0) = f(0, 9) = -61$ (Global Minimum)

Constrained extrema of $f : \mathbb{R}^n \to \mathbb{R}$

In the earlier two examples, the boundary of the domains are lines in $\mathbb{R}^2$. We now discuss a method to deal with more general boundary.

Let $U \subset \mathbb{R}^n$ be open and $f, g : U \subset \mathbb{R}^n \to \mathbb{R}$ be continuous. Then

Maximize or Minimize $f(X)$
Subject to the constraint $g(X) = \alpha$.

Constrained extrema of $f : \mathbb{R}^n \to \mathbb{R}$

In the earlier two examples, the boundary of the domains are lines in $\mathbb{R}^2$. We now discuss a method to deal with more general boundary.

Let $U \subset \mathbb{R}^n$ be open and $f, g : U \subset \mathbb{R}^n \to \mathbb{R}$ be continuous. Then

$$\text{Maximize or Minimize } f(X)$$
$$\text{Subject to the constraint } g(X) = \alpha.$$

Example: Find the extreme values of $f(x, y) = x^2 - y^2$ on the circle $x^2 + y^2 = 1$.

It turns out that f attains minimum at $(0, \pm 1)$ and maximum at $(\pm 1, 0)$ although $\nabla f(0, \pm 1) \neq 0$ and $\nabla f(\pm 1, 0) \neq 0$.

Test for constrained extrema of $f : \mathbb{R}^2 \to \mathbb{R}$

Theorem: Let $f, g : U \subset \mathbb{R}^2 \to \mathbb{R}$ be C^1. Suppose that f has an extremum at $(a, b) \in U$ such that $g(a, b) = \alpha$ and that $\nabla g(a, b) \neq (0, 0)$. Then there is a $\lambda \in \mathbb{R}$, called Lagrange multiplier, such that $\nabla f(a, b) = \lambda \nabla g(a, b)$.

Test for constrained extrema of $f : \mathbb{R}^2 \to \mathbb{R}$

Theorem: Let $f, g : U \subset \mathbb{R}^2 \to \mathbb{R}$ be C^1. Suppose that f has an extremum at $(a, b) \in U$ such that $g(a, b) = \alpha$ and that $\nabla g(a, b) \neq (0, 0)$. Then there is a $\lambda \in \mathbb{R}$, called Lagrange multiplier, such that $\nabla f(a, b) = \lambda \nabla g(a, b)$.

Proof: Let $r(t)$ be a local parametrization of the curve $g(x, y) = \alpha$ such that $r(0) = (a, b)$. Then $f(r(t))$ has an extremum at $t = 0$. Therefore

$$\frac{\mathrm{d}f(r(t))}{\mathrm{d}t}\Big|_{t=0} = \nabla f(a, b) \bullet r'(0) = 0.$$

Test for constrained extrema of $f : \mathbb{R}^2 \to \mathbb{R}$

Theorem: Let $f, g : U \subset \mathbb{R}^2 \to \mathbb{R}$ be C^1. Suppose that f has an extremum at $(a, b) \in U$ such that $g(a, b) = \alpha$ and that $\nabla g(a, b) \neq (0, 0)$. Then there is a $\lambda \in \mathbb{R}$, called Lagrange multiplier, such that $\nabla f(a, b) = \lambda \nabla g(a, b)$.

Proof: Let $r(t)$ be a local parametrization of the curve $g(x, y) = \alpha$ such that $r(0) = (a, b)$. Then $f(r(t))$ has an extremum at $t = 0$. Therefore

$$\frac{\mathrm{d}f(r(t))}{\mathrm{d}t}\Big|_{t=0} = \nabla f(a, b) \bullet r'(0) = 0.$$

Now $g(r(t)) = \alpha \Rightarrow \nabla g(a, b) \bullet r'(0) = 0$. This shows that $r'(0) \perp \nabla g(a, b)$ and $r'(0) \perp \nabla f(a, b)$. Hence $\nabla f(a, b) = \lambda \nabla g(a, b)$ for some $\lambda \in \mathbb{R}$.

Method of Lagrange multipliers for $f : \mathbb{R}^2 \to \mathbb{R}$

To find extremum of f subject to the constraint $g(x, y) = \alpha$, define $L(x, y, \lambda) := f(x, y) - \lambda(g(x, y) - \alpha)$ and solve the equations

Method of Lagrange multipliers for $f : \mathbb{R}^2 \to \mathbb{R}$

To find extremum of f subject to the constraint $g(x, y) = \alpha$, define $L(x, y, \lambda) := f(x, y) - \lambda(g(x, y) - \alpha)$ and solve the equations

$$
\begin{aligned}
\frac{\partial L}{\partial x} &= 0 \Rightarrow \frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}, \\
\frac{\partial L}{\partial y} &= 0 \Rightarrow \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}, \\
\frac{\partial L}{\partial \lambda} &= 0 \Rightarrow g(x, y) = \alpha.
\end{aligned}
$$

Method of Lagrange multipliers for $f : \mathbb{R}^2 \to \mathbb{R}$

To find extremum of f subject to the constraint $g(x,y) = \alpha$, define $L(x,y,\lambda) := f(x,y) - \lambda(g(x,y) - \alpha)$ and solve the equations

$$\begin{aligned}
\frac{\partial L}{\partial x} &= 0 \Rightarrow \frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}, \\
\frac{\partial L}{\partial y} &= 0 \Rightarrow \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}, \\
\frac{\partial L}{\partial \lambda} &= 0 \Rightarrow g(x,y) = \alpha.
\end{aligned}$$

• The auxiliary function $L(x,y,\lambda)$ is called Lagrangian which converts constrained extrema to unconstrained extrema.

Method of Lagrange multipliers for $f : \mathbb{R}^2 \to \mathbb{R}$

To find extremum of f subject to the constraint $g(x, y) = \alpha$, define $L(x, y, \lambda) := f(x, y) - \lambda(g(x, y) - \alpha)$ and solve the equations

$$\begin{aligned}
\frac{\partial L}{\partial x} &= 0 \Rightarrow \frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}, \\
\frac{\partial L}{\partial y} &= 0 \Rightarrow \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}, \\
\frac{\partial L}{\partial \lambda} &= 0 \Rightarrow g(x, y) = \alpha.
\end{aligned}$$

• The auxiliary function $L(x, y, \lambda)$ is called Lagrangian which converts constrained extrema to unconstrained extrema. • Critical points of L are eligible solutions for constrained extrema.

Example

Find the extreme values of $f(x, y) = x^2 - y^2$ on the circle $x^2 + y^2 = 1$.

Example

Find the extreme values of $f(x, y) = x^2 - y^2$ on the circle $x^2 + y^2 = 1$.

The equations $f_x = \lambda g_x, f_y = \lambda g_y$ and $g(x, y) = 1$ give $2x = \lambda 2x, -2y = \lambda 2y$ and $x^2 + y^2 = 1$. The first equation shows either $x = 0$ or $\lambda = 1$.

Example

Find the extreme values of $f(x, y) = x^2 - y^2$ on the circle $x^2 + y^2 = 1$.

The equations $f_x = \lambda g_x, f_y = \lambda g_y$ and $g(x, y) = 1$ give $2x = \lambda 2x, -2y = \lambda 2y$ and $x^2 + y^2 = 1$. The first equation shows either $x = 0$ or $\lambda = 1$.

If $x = 0$ then $y = \pm 1 \Rightarrow \lambda = -1$. Thus $(x, y, \lambda) := (0, \pm 1, -1)$ are possible solutions.

Example

Find the extreme values of $f(x, y) = x^2 - y^2$ on the circle $x^2 + y^2 = 1$.

The equations $f_x = \lambda g_x, f_y = \lambda g_y$ and $g(x, y) = 1$ give $2x = \lambda 2x, -2y = \lambda 2y$ and $x^2 + y^2 = 1$. The first equation shows either $x = 0$ or $\lambda = 1$.

If $x = 0$ then $y = \pm 1 \Rightarrow \lambda = -1$. Thus $(x, y, \lambda) := (0, \pm 1, -1)$ are possible solutions.

If $\lambda = 1$ then $y = 0 \Rightarrow x = \pm 1$. Thus $(x, y, \lambda) := (\pm 1, 0, 1)$ are also possible solutions.

Example

Find the extreme values of $f(x, y) = x^2 - y^2$ on the circle $x^2 + y^2 = 1$.

The equations $f_x = \lambda g_x, f_y = \lambda g_y$ and $g(x, y) = 1$ give $2x = \lambda 2x, -2y = \lambda 2y$ and $x^2 + y^2 = 1$. The first equation shows either $x = 0$ or $\lambda = 1$.

If $x = 0$ then $y = \pm 1 \Rightarrow \lambda = -1$. Thus $(x, y, \lambda) := (0, \pm 1, -1)$ are possible solutions.

If $\lambda = 1$ then $y = 0 \Rightarrow x = \pm 1$. Thus $(x, y, \lambda) := (\pm 1, 0, 1)$ are also possible solutions.

Now $f(0, 1) = f(0, -1) = -1$ and $f(1, 0) = f(-1, 0) = 1$ so that minimum and maximum values are -1 and 1.

Finding global extrema of $f : \mathbb{R}^2 \to \mathbb{R}$

Let $f : \mathbb{R}^2 \to \mathbb{R}$ and $U \subset \mathbb{R}^2$ be a region with smooth closed boundary curve C. To find global emtremum of f in U :

Finding global extrema of $f : \mathbb{R}^2 \to \mathbb{R}$

Let $f : \mathbb{R}^2 \to \mathbb{R}$ and $U \subset \mathbb{R}^2$ be a region with smooth closed boundary curve C. To find global emtremum of f in U :

- Locate all critical points of f in U.

Finding global extrema of $f : \mathbb{R}^2 \to \mathbb{R}$

Let $f : \mathbb{R}^2 \to \mathbb{R}$ and $U \subset \mathbb{R}^2$ be a region with smooth closed boundary curve C. To find global emtremum of f in U :

- Locate all critical points of f in U.
- Find eligible global extremum of f on the curve C by using Lagrange multipliers or parametrization.

Finding global extrema of $f : \mathbb{R}^2 \to \mathbb{R}$

Let $f : \mathbb{R}^2 \to \mathbb{R}$ and $U \subset \mathbb{R}^2$ be a region with smooth closed boundary curve C. To find global emtremum of f in U :

- Locate all critical points of f in U.
- Find eligible global extremum of f on the curve C by using Lagrange multipliers or parametrization.
- Choose points among eligible solutions in C and the critical points at which f attains extreme values. These extreme values are global extremum.

Example: global extrema

Find global maximum and global minimum of the function $f(x, y) := (x^2 + y^2)/2$ such that $x^2/2 + y^2 \leq 1$.

Example: global extrema

Find global maximum and global minimum of the function $f(x, y) := (x^2 + y^2)/2$ such that $x^2/2 + y^2 \leq 1$.

Since $f_x = x$ and $f_y = y$, $(0, 0)$ is the only critical point.

Example: global extrema

Find global maximum and global minimum of the function $f(x,y) := (x^2+y^2)/2$ such that $x^2/2+y^2 \leq 1$.

Since $f_x = x$ and $f_y = y$, $(0,0)$ is the only critical point.

Next, consider $L(x,y,\lambda) := (x^2+y^2)/2 - \lambda(x^2/2+y^2-1)$. Then Lagrange multiplier equations are

$$x = \lambda x, \quad y = 2\lambda y, \quad x^2/2+y^2 = 1.$$

If $x=0$ then $y=\pm 1$ and $\lambda = 1/2$. If $y=0$ then $x=\pm\sqrt{2}$ and $\lambda = 1$. If $xy \neq 0$ then $\lambda = 1$ and $\lambda = 1/2$ -which is not possible.

Example: global extrema

Find global maximum and global minimum of the function $f(x, y) := (x^2 + y^2)/2$ such that $x^2/2 + y^2 \leq 1$.

Since $f_x = x$ and $f_y = y$, $(0, 0)$ is the only critical point.

Next, consider $L(x, y, \lambda) := (x^2 + y^2)/2 - \lambda(x^2/2 + y^2 - 1)$. Then Lagrange multiplier equations are

$$x = \lambda x, \quad y = 2\lambda y, \quad x^2/2 + y^2 = 1.$$

If $x = 0$ then $y = \pm 1$ and $\lambda = 1/2$. If $y = 0$ then $x = \pm\sqrt{2}$ and $\lambda = 1$. If $xy \neq 0$ then $\lambda = 1$ and $\lambda = 1/2$ -which is not possible.

Thus $(0, \pm 1)$ and $(\pm\sqrt{2}, 0)$ are eligible solutions for the boundary curve. We have $f(0, \pm 1) = 1/2$, $f(\pm\sqrt{2}, 0) = 1$ and $f(0, 0) = 0$.

Test for constrained extrema of $f : \mathbb{R}^n \to \mathbb{R}$

Theorem: Let $f, g : U \subset \mathbb{R}^n \to \mathbb{R}$ be C^1. Suppose that f has an extremum at $X_0 \in U$ such that $g(X_0) = \alpha$ and $\nabla g(X_0) \neq (0, \ldots, 0)$. Then there is a $\lambda \in \mathbb{R}$ such that $\nabla f(X_0) = \lambda \nabla g(X_0)$.

Test for constrained extrema of $f : \mathbb{R}^n \to \mathbb{R}$

Theorem: Let $f, g : U \subset \mathbb{R}^n \to \mathbb{R}$ be C^1. Suppose that f has an extremum at $X_0 \in U$ such that $g(X_0) = \alpha$ and $\nabla g(X_0) \neq (0, \ldots, 0)$. Then there is a $\lambda \in \mathbb{R}$ such that $\nabla f(X_0) = \lambda \nabla g(X_0)$.

If $g(X) = \alpha$ is a closed surface then global extremum is obtained by finding all points where $\nabla f(X) = \lambda \nabla g(X)$ and choosing those where f is largest or smallest.

Test for constrained extrema of $f : \mathbb{R}^n \to \mathbb{R}$

Theorem: Let $f, g : U \subset \mathbb{R}^n \to \mathbb{R}$ be C^1. Suppose that f has an extremum at $X_0 \in U$ such that $g(X_0) = \alpha$ and $\nabla g(X_0) \neq (0, \ldots, 0)$. Then there is a $\lambda \in \mathbb{R}$ such that $\nabla f(X_0) = \lambda \nabla g(X_0)$.

If $g(X) = \alpha$ is a closed surface then global extremum is obtained by finding all points where $\nabla f(X) = \lambda \nabla g(X)$ and choosing those where f is largest or smallest.

The Lagrangian is given by $L(X, \lambda) := f(X) - \lambda(g(X) - \alpha)$. So, the multiplier equations are $\nabla L(X, \lambda) = (0, 0, \ldots, 0)$.

Example

Maximize the function $f(x, y, z) := x + z$ subject to the constraint $x^2 + y^2 + z^2 = 1$.

Example

Maximize the function $f(x, y, z) := x + z$ subject to the constraint $x^2 + y^2 + z^2 = 1$.

The multiplier equations are

$$1 = 2x\lambda, \quad 0 = 2y\lambda, \quad 1 = 2z\lambda, \quad x^2 + y^2 + z^2 = 1.$$

From first and 3rd equation, $\lambda \neq 0$. Thus, by second equation, $y = 0$. By first and 3rd equations $x = z \Rightarrow x = z = \pm 1/\sqrt{2}$.

Example

Maximize the function $f(x, y, z) := x + z$ subject to the constraint $x^2 + y^2 + z^2 = 1$.

The multiplier equations are

$$1 = 2x\lambda, \quad 0 = 2y\lambda, \quad 1 = 2z\lambda, \quad x^2 + y^2 + z^2 = 1.$$

From first and 3rd equation, $\lambda \neq 0$. Thus, by second equation, $y = 0$. By first and 3rd equations $x = z \Rightarrow x = z = \pm 1/\sqrt{2}$.

Hence $X_0 := (1/\sqrt{2}, 0, 1/\sqrt{2})$ and $X_1 := (-1/\sqrt{2}, 0, -1/\sqrt{2})$ are eligible solutions. This shows that $f(X_0) = \sqrt{2}$ and $f(X_1) = -\sqrt{2}$.

Riemann sum for double integral

Consider the rectangle $\mathcal{R} := [a, b] \times [c, d]$ and a bounded function $f : \mathcal{R} \to \mathbb{R}$.

Let P be a partition of $\mathcal{R}$ into mn sub-rectangles R_{ij} and $c_{ij} \in R_{ij}$ for $i = 1, 2, \ldots, m, j = 1, 2, \ldots, n$. Also let

$$\Delta A_{ij} = \mathrm{area}(R_{ij}) = \Delta x_i \Delta y_j \text{ and } \mu(P) := \max_{ij} \Delta A_{ij}.$$

Consider the Riemann sum

$$S(P, f) := \sum_{i=1}^{m} \sum_{j=1}^{n} f(c_{ij}) \Delta A_{ij} = \sum_{i=1}^{m} \sum_{j=1}^{n} f(c_{ij}) \Delta x_i \Delta y_j.$$

Double integral

Definition: If $\lim_{\mu(P)\to 0} S(P,f)$ exists then f is said to be Riemann integrable and the (double) integral of f over $\mathcal{R}$ is given by

$$\iint_{\mathcal{R}} f(x,y)dA = \iint_{\mathcal{R}} f(x,y)dxdy = \lim_{\mu(P)\to 0} S(P,f).$$

- If $f(x,y) \geq 0$ then $\iint_{\mathcal{R}} f(x,y)dA$ gives the volume of the region bounded by $\mathcal{R}$ and the graph of f.

Theorem: If $f : \mathcal{R} \to \mathbb{R}$ is continuous then f is Riemann integrable over $\mathcal{R}$.

Theorem: Let $f, g : \mathcal{R} \to \mathbb{R}$ be Riemann integrable. Then

- $f + \alpha g$ is Riemann integrable for $\alpha \in \mathbb{R}$ and

$$\iint_{\mathcal{R}} (f + \alpha g) dA = \iint_{\mathcal{R}} f dA + \alpha \iint_{\mathcal{R}} g dA$$

- $|f|$ is Riemann integrable and

$$\left| \iint_{\mathcal{R}} f(x, y) dA \right| \leq \iint_{\mathcal{R}} |f(x, y)| dA.$$

- $\iint_{\mathcal{R}} dA = \mathrm{Area}(\mathcal{R})$.

- If $\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2$, where $\mathcal{R}_1$ and $\mathcal{R}_2$ are two disjoint rectangles then

$$\iint_{\mathcal{R}} f(x, y) dA = \iint_{\mathcal{R}_1} f(x, y) dA + \iint_{\mathcal{R}_2} f(x, y) dA.$$