Local extrema

$f : \mathbb{R}^n \longrightarrow \mathbb{R}$.

$$\begin{cases} x_0 \text{ is a maximum if } f(x) \leq f(x_0), \text{ for any } x \sim x_0 \\ x_0 \text{ is a minimum if } f(x) \geq f(x_0), \text{ for any } x \sim x_0 \end{cases}$$

stationary point or critical point

$$\nabla f(x_0) = 0$$

Theorem

$\Omega \subset \mathbb{R}^n$, $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ differentiable.

and x_0 local extrema $\Longrightarrow$ x_0 is a critical point

$$\nabla f(x_0) = 0$$

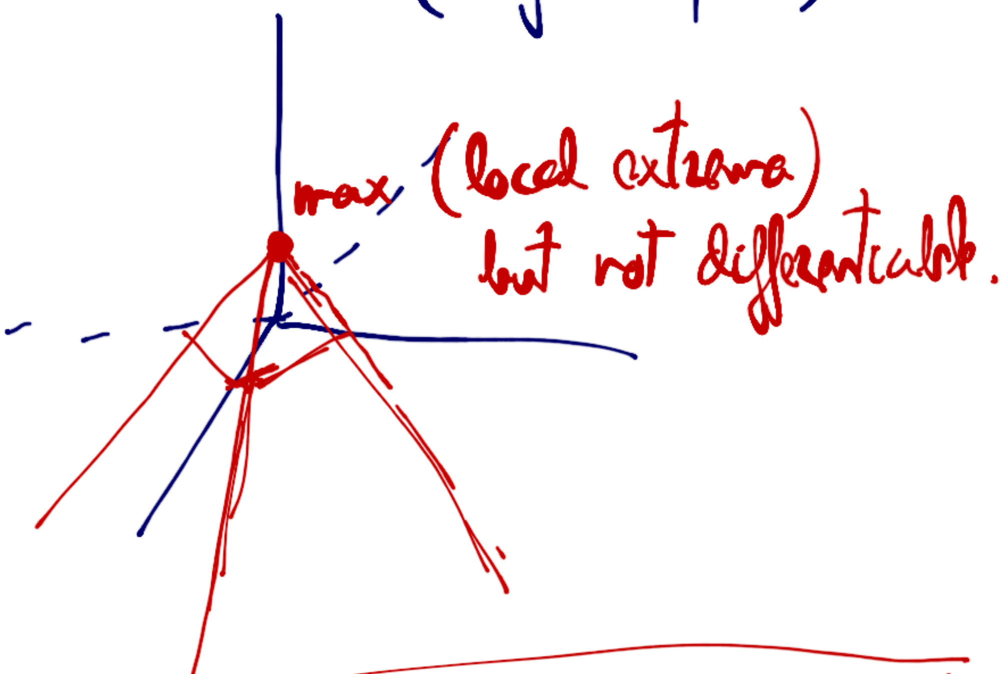
Remark

- f is differentiable and $\nabla f(x_0) = 0$, then all directional derivatives are zero (horizontal tangent plane)

If one of the partial derivatives at x_0 does not exist, f is not differentiable

There is not tangent plane.

but we might have a local extrema (angular point)



How to find local extrema?

We might find those local extrema just applying algebraic computations.

Example: $f(x,y) = 3x^2 + y^2 - 6x - 4y + 8$

$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ differentiable in $\mathbb{R}^2$

Find stationary points $\nabla f = (0,0)$

$$\frac{\partial f(x,y)}{\partial x} = 6x - 6 = 0 \Rightarrow x = 1$$

$$\frac{\partial f(x,y)}{\partial y} = 2y - 4 = 0 \Rightarrow y = 2$$

Stationary point $(1,2)$

$$f(1,2) = 1 = 3 + 4 - 6 - 8 + 8$$

$$f(x,y) = 3x^2 + y^2 - 6x - 4y + 8 = 3x^2 - 6x + y^2 - 4y + 8$$
$$= 3(x^2 - 2x + 1 \boxed{-1}) + (y^2 - 4y + 4 - 4) + 8$$
$$= 3(x-1)^2 + (y-2)^2 + 1 \geq 1 \Rightarrow (1,2) \text{ minimum}$$

Example: Find the critical points for

$$f(x,y) = 1 - \sqrt{x^2+y^2}$$

- We look for point where $\nabla f = (0,0)$
- We look for non-differentiable points.

$$\frac{\partial f(x,y)}{\partial x} = \frac{-x}{\sqrt{x^2+y^2}} \qquad \frac{\partial f(x,y)}{\partial y} = \frac{-y}{\sqrt{x^2+y^2}}$$

A possible local extrema will be $(0,0)$

$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ are not defined at $(0,0)$

$$\left.\begin{array}{l} f(0,0) = 1 \\ f(x,y) = 1 - \sqrt{x^2+y^2} \leq 1 \end{array}\right| \Rightarrow \begin{array}{l} f(x,y) \leq f(0,0) \\ \forall (x,y) \in Dom f \end{array}$$

Therefore, $(0,0)$ is global max.

* Another method to identify critical points

Intersection by vertical planes

$\Downarrow$

existence of saddle points

Example: $f(x,y) = 1 - x^2 + y^2$, $f: \mathbb{R}^2 \to \mathbb{R}$.

Critical points: $\dfrac{\partial f(x,y)}{\partial x} = -2x = 0 \Rightarrow x = 0$

$$\frac{\partial f(x,y)}{\partial y} = 2y = 0 \Rightarrow y = 0$$

One critical point $(0,0)$

$$f(0,0) = 1$$

In this direction f is bigger than 1 $\left\{\begin{array}{l} f(x,y) = 1 - x^2 + y^2 \\ x = 0 \end{array}\right\} \Rightarrow f(0,y) = 1 + y^2 \geq 1$

$(0,0)$ minimum

f is smaller than 1. $\left\{\begin{array}{l} f(x,y) = 1 - x^2 + y^2 \\ y = 0 \end{array}\right\} \Rightarrow f(x,0) = 1 - x^2 \leq 1$

$(0,0)$ maximum.

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

- Algebraic computations seem complicated
More systematic method to classify critical points:
  - Quadratic forms
  - Hessian matrix ("second derivatives")

Definition - Quadratic forms [Linear Algebra - Lay]

$Q : \mathbb{R}^N \longrightarrow \mathbb{R}$ quadratic form

$$Q(x) = \sum_{i,j=1}^{N} a_{ij}\, x_i\, x_j = x^T A x$$

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \qquad \boxed{A = \left(a_{ij}\right)_{i,j=1,\dots N}}$$

- $Q(x)$ is positive definite $\boxed{\text{if } Q(x) > 0}$ $\forall x \in \mathbb{R}^N$
- $Q(x)$ is negative definite if $\boxed{Q(x) < 0}$ $\forall x \in \mathbb{R}^N$
- Indefinite $\boxed{Q(x) < 0,\ Q(y) > 0 \text{ for some } x, y \in \mathbb{R}^N}$

If A is symmetric, there exists $\{u_1 \dots u_N\}$ an orthogonal basis such that $\quad P = [u_1 \dots u_N]$

change of basis.

$$x^T A x = (Py)^T A\, Py = y^T \underbrace{P^T A P}_{D}\, y = y^T D\, y$$

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}$$

$u_i \equiv$ eigenvectors

$\lambda_i \equiv$ eigenvalues.

Then,

$$Q(x) = \underline{\underline{x^T A x}} = y^T D y = \underbrace{\lambda_1 y_1^2 + \dots + \lambda_N y_N^2}_{\text{Spectral theorem.}}$$

Ex: Quadratic form

$$Q(x,y) = \overbrace{3x^2 + y^2} - 6x - 4y + 8$$

$$= (x,y)\, A \begin{pmatrix} x \\ y \end{pmatrix} + 8$$

Hessian matrix

$f: \mathbb{R}^N \longrightarrow \mathbb{R}$ scalar function

$$Hf = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_N \partial x_1} \\ \vdots & & & \\ \frac{\partial^2 f}{\partial x_1 \partial x_N} & \frac{\partial^2 f}{\partial x_2 \partial x_N} & \cdots & \frac{\partial^2 f}{\partial x_N^2} \end{pmatrix}$$

If $f \in C^2$ (all second derivatives are continuous) the Hessian matrix is symmetric (Schwarz's Th)

Remark

We will use the Hessian matrix and the associated quadratic form to classify critical points.

Hessian criteria in $\mathbb{R}^2$

$$Hf(x,y) = \begin{pmatrix} \dfrac{\partial^2 f(x,y)}{\partial x^2} & \dfrac{\partial^2 f(x,y)}{\partial y\,\partial x} \\ \dfrac{\partial^2 f(x,y)}{\partial x\,\partial y} & \dfrac{\partial^2 f(x,y)}{\partial y^2} \end{pmatrix} \qquad f \in C^2$$

$f : \mathbb{R}^2 \to \mathbb{R}$, $(x_0, y_0) \in \mathrm{Dom}\, f$ with

$$\boxed{\frac{\partial f(x_0, y_0)}{\partial x} = 0 = \frac{\partial f(x_0, y_0)}{\partial y}}$$

a) If $\det Hf(x_0, y_0) > 0$ and $\boxed{\dfrac{\partial^2 f(x_0, y_0)}{\partial x^2} > 0}$

then (x_0, y_0) is a local minimum.

$$\det Hf(x_0, y_0) = \frac{\partial^2 f(x_0,y_0)}{\partial x^2} \cdot \frac{\partial^2 f(x_0,y_0)}{\partial y^2} - \left(\frac{\partial^2 f(x_0,y_0)}{\partial x\,\partial y}\right)^2 > 0$$

so that

$$\frac{\partial^2 f(x_0,y_0)}{\partial x^2} \cdot \frac{\partial^2 f(x_0,y_0)}{\partial y^2} > \left(\frac{\partial^2 f(x_0,y_0)}{\partial x\,\partial y}\right)^2 > 0$$

So $\frac{\partial^2 f(x_0,y_0)}{\partial x^2} > 0$ and $\frac{\partial^2 f(x_0,y_0)}{\partial y^2} > 0$

f is concave upwards in both directions.

$\boxed{b)}$ $\det Hf(x_0,y_0) > 0$ and $\frac{\partial^2 f(x_0,y_0)}{\partial x^2} < 0$

<u>$(x_0 y_0)$ local maximum</u>

$\boxed{c)}$ <u>$\det Hf(x_0 y_0) < 0$ saddle point</u>

d) $\det Hf(x_0,y_0) = 0$ not conclusive !!

degenerate critical point.

Example: $f(x,y) = x^3 - 6xy + 3y^2 - 1$

First find critical points

$$\left.\begin{aligned} \frac{\partial f(x,y)}{\partial x} &= 3x^2 - 6y = 0 \\[2ex] \frac{\partial f(x,y)}{\partial y} &= -6x + 6y = 0 \end{aligned}\right\} \Rightarrow x = y$$

so that $\quad 3x^2 - 6x = 0 \Rightarrow x(3x-6) = 0$

$$x = 0, \; x = 2$$

Two critical points $(0,0), (2,2)$

Hessian matrix

$$Hf(x,y) = \begin{pmatrix} 6x & -6 \\ -6 & 6 \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

$$\det Hf(0,0) = \det \begin{vmatrix} 0 & -6 \\ -6 & 6 \end{vmatrix} = -36 < 0$$

This means $(0,0)$ is a saddle point

$$\det Hf(2,2) = \det \begin{vmatrix} 12 & -6 \\ -6 & 6 \end{vmatrix} = 72 - 36 = 36 > 0$$

$$\frac{\partial^2 f(2,2)}{\partial x^2} = 12 > 0$$

$(2,2)$ is a local minimum

Theorem - Taylor's polynomial (of second order)

$f: \mathbb{R}^2 \longrightarrow \mathbb{R}$ with continuous partial derivatives $f \in C^2$. Then, $(x_0 y_0) \in \operatorname{Dom} f$.

$$f(x,y) = \underbrace{f(x_0 y_0) + \frac{\partial f(x_0 y_0)}{\partial x}(x-x_0) + \frac{\partial f(x_0 y_0)}{\partial y}(y-y_0)}_{\text{linear part.}}$$

$$\underbrace{+ \frac{1}{2} \sum_{i,j=1}^{2} \frac{\partial^2 f(x_0 y_0)}{\partial x_i \, \partial x_j}(x-x_0)(y-y_0)}_{\text{quadratic form.}} + R$$

$$\frac{R}{\|(x,y)-(x_0 y_0)\|^2} \longrightarrow 0$$

$$= f(x_0 y_0) + \nabla f(x_0 y_0) \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}$$

$$+ \frac{1}{2}(x-x_0 \;\; y-y_0)\, Hf(x_0 y_0) \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}$$

Example: $f(x,y) = e^x \cos y$

Taylor's polynomial around $(x_0, y_0) = (0,0)$

$$f(0,0) = 1$$

$$\frac{\partial f(x,y)}{\partial x} = e^x \cos y \Rightarrow \frac{\partial f(0,0)}{\partial x} = 1$$

$$\frac{\partial f(x,y)}{\partial y} = -e^x \sin y \Rightarrow \frac{\partial f(0,0)}{\partial y} = 0$$

$$\nabla f(0,0) = (1,0)$$

$$\frac{\partial^2 f(x,y)}{\partial x^2} = e^x \cos y \Rightarrow \frac{\partial^2 f(0,0)}{\partial x^2} = 1$$

$$\frac{\partial^2 f(x,y)}{\partial x \partial y} = -e^x \sin y = \frac{\partial^2 f(x,y)}{\partial y \partial x}$$

$$Hf(0,0) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\frac{\partial^2 f(0,0)}{\partial x \partial y} = 0, \quad \frac{\partial^2 f(x,y)}{\partial y^2} = -e^x \cos y \Rightarrow \frac{\partial^2 f(0,0)}{\partial y^2} = -1$$

Then,

$$f(x,y) = 1 + (1,0)\begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2}(x\ y)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + R$$

$$= 1 + x + \frac{1}{2}(x, -y)\begin{pmatrix} x \\ y \end{pmatrix} + R$$

$$= \underbrace{1 + x + \frac{(x^2 - y^2)}{2}}_{\text{Taylor's polynomial of second order for } f(x,y) \text{ around } (0,0)} + R$$

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

Computations of local extrema

Assume $f : \mathbb{R}^N \to \mathbb{R}$ is of class C^2

a) Local extrema belong to the set of critical points

b) The type of local extrema is given by a quadratic form.

$$\underline{Q_H(x_1, \dots x_N)} = x^T \underbrace{Hf}_{\text{Hessian matrix.}} x = y^T D y = \underline{\lambda_1 y_1^2 + \dots + \lambda_N y_N^2}$$

$\{y_1, \dots y_N\}$ is an orthogonal basis.

Then, a) If all eigenvalues λ_i are positive $\Rightarrow$ Q_H positive definite.

b) If all eigenvalues λ_i are negative $\Rightarrow$ Q_H negative definite.

c) If we have positive and negative eigenvalues then Q_H indefinite. $Q_H = \lambda_1 y_1^2 + \dots + \lambda_N y_N^2$

The idea is to apply the quadratic form to get local extrema using Taylor's polynomial.

For example in $\mathbb{R}^2$

$$f(x,y) = f(x_0 y_0) + \nabla f(x_0 y_0) \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} + \frac{1}{2}(x-x_0 \; y-y_0) \mathbb{H} f(x_0 y_0) \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}$$

If $(x_0 y_0)$ is a critical point $\nabla f(x_0, y_0) = 0$

So

$$f(x,y) - f(x_0 y_0) = \overbrace{\frac{1}{2}(x-x_0 \; y-y_0) \mathbb{H} f(x_0 y_0) \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}}^{\text{quadratic form.}}$$

If the quadratic form is positive $\Rightarrow f(x,y) - f(x_0 y_0) > 0$

Then $f(x,y) > f(x_0, y_0)$

(x_0, y_0) is a local minimum.

Writing the quadratic form

$$\underline{Q_H(x,y)} = \underline{\lambda_1 y_1^2 + \lambda_2 y_2^2} \underline{> 0}$$

$$\Downarrow \quad \Uparrow$$

$$\lambda_1, \lambda_2 > 0$$

This is equivalent to

$$\boxed{\det \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} > 0} \quad \text{and} \quad \lambda_1 > 0, \lambda_2 > 0$$

$$\left[\det H f(x_0, y_0) > 0 \quad \text{and} \quad \frac{\partial^2 f(x_0, y_0)}{\partial x^2} > 0 \right]$$

If $Q_H < 0$

$$f(x\,y) < f(x_0\,y_0) \Rightarrow (x_0\,y_0) \text{ local max.}$$