Higher Partial derivatives and the Hessian.

Let $f(x,y) = \begin{cases} xy\,\dfrac{x^2-y^2}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$

Then $\dfrac{\partial f}{\partial x \partial y}(0,0) \neq \dfrac{\partial f}{\partial y \partial x}(0,0)$

A sufficient condition for the equality of mixed partial derivatives is

Thm Let $f \in C^2(\Omega)$ $f : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^n$ (ie $\dfrac{\partial f}{\partial x^i}$, $\dfrac{\partial f}{\partial x^i \partial x^j}$ are continous). Then

$$\frac{\partial f}{\partial x^i \partial x^j} = \frac{\partial f}{\partial x^j \partial x^i} \qquad 1 \leq i, j \leq n.$$

Def^n let $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$

The $n \times n$ Matrix $H(f; x_0) = \left(D_{ij} f(x_0)\right)_{i,j=1}^{n}$,

the matrix whose (i,j) entry is $\frac{\partial f}{\partial x^i \partial x^j}(x_0)$,

is called the Hessian Matrix

First order Taylor Formula

$f : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^n$. f is differentiable at x_0. Then

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \underbrace{R(|x - x_0|)}_{o(|x - x_0|)}$$

where $\lim\limits_{x \to x_0} \left(\frac{R(x; x_0)}{|x - x_0|} \right) = 0$

The polynomial $T_1(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0)$ is the first order approximation to f near x_0.

Thm. Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ with continuous second order partial derivatives. Then $\forall y \in \mathbb{R}^n$

$$f(x+y) = f(x) + \nabla f(x) \cdot y + \frac{1}{2!}\, y\, H_f(x)\, y^t + R_2(x,y)$$

where $\dfrac{R_2(x,y)}{\|x-y\|^2} \to 0$ as $x \to y$.

or equivalently

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2}(x-x_0)\, H_f(x_0)\, (x-x_0)^t + R_2(x, x_0)$$

where $\dfrac{R_2(x,x_0)}{\|x-x_0\|^2} \to 0$ as $x \to x_0$.

$$f(x^1, x^2) = f(x_0^1, x_0^2) + \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} (x^i - x_0^i) +$$

$$\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial f}{\partial x^i \partial x^j} (x_0)(x^i - x_0^i)(x^j - x_0^j).$$

<u>Rk</u>. Sometimes Hessian Matrix is also denoted by

$\nabla^2 f$!

There is a general Taylor formula which looks as follows.

Thm let $f \in C^k(\Omega)$, $x, x_0 \in \Omega$. Then

$$f(x) = f(x_0) + \nabla f(x_0)\cdot(x - x_0) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(x^i - x_0^i)(x^j - x_0^j)$$

$$+ \cdots + \frac{1}{k!}\sum_{i_1, i_2, \ldots, i_k = 1}^{n} \frac{\partial^k f}{\partial x^{i_1}\partial x^{i_2}\cdots\partial x^{i_k}}(x_0)\prod_{j=1}^{k}(x^{i_j} - x_0^{i_j})$$

$$+ R_k(x, x_0))$$

such that $\lim \dfrac{R_k(x - x_0)}{\|x - x_0\|^k} = 0$ as $x \to x_0$.

Ex: $f(x,y) = e^{x+y}\cos x$ at pt $(0,0)$

$e^x \cdot e^y \cdot \cos x$

Find the Taylor exponsion of degree 3 around $(0,0)$

$f(0,0) = 1.$

$$\frac{\partial f}{\partial x} = e^x e^y \cos x - e^{x+y}\sin x \Big|_{(0,0)} = 1.$$

$$\frac{\partial f}{\partial y} = e^x \cos x \, e^y \Big|_{(0,0)} = 1.$$

$$\frac{\partial^2 f}{\partial x^2} = e^y \left[e^x(\cos x - \sin x) + e^x(-\sin x - \cos x) \right] \Big|_{0,0} = 0.$$

$$\frac{\partial^2 f}{\partial x \partial y} = e^y e^x (\cos x - \sin x) \Big|_{(0,0)} = 1.$$

$$\partial^2 f / \partial y^2 = e^y e^x \cos x \Big|_{(0,0)} = 1.$$

$$\frac{\partial^3 f}{\partial x^3} = -2e^y(e^x\sin x + e^x\cos x)\Big|_{(0,0)} = -2$$

$$\frac{\partial^3 f}{\partial y^3} = e^y e^x \cos x \Big|_{(0,0)} = 1.$$

$$\frac{\partial^3 f}{\partial x^2 \partial y} = e^y e^x (-2\sin x)\Big|_{(0,0)} = 0 = \frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial y \partial^2 x} \cdots$$

$$\frac{\partial^3 f}{\partial x \partial^2 y}\Big|_{(0,0)} = 1.$$

Hence

$$e^{x+y}\cos x = 1 + \nabla f(0,0)\cdot(x,y) + \cdots$$

$$= 1 + x + y + \frac{1}{2}\left(2xy + 1y^2\right) \quad (+\,0\cdot x^2)$$

$2\frac{\partial^2 f}{\partial x \partial y}$, $\frac{\partial^2 f}{\partial y^2}$, $\frac{\partial^2 f}{\partial x^2}$

$$+ \frac{1}{6}\left(-2x^3 + 3xy^2 + y^3\right) + o(|x-y|^3)$$

$\frac{\partial^3 f}{\partial x^3}$, $3\frac{\partial^3 f}{\partial x \partial^2 y}$.

$$T_3(x,y) = 1 + x + y + \frac{1}{2}(2xy + y^2) + \frac{1}{6}(-2x^3 + 3xy^2 + y^3).$$

7

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \qquad \text{conv for } \underline{\text{all}}\ x.$$

$$\frac{1}{1-x} = \sum x^k, \qquad |x| < 1$$

$$\frac{1}{1-y^2} = \sum y^{2k}, \qquad |y^2| < 1$$

Lokal Extrema of functions $f: \mathbb{R}^n \to \mathbb{R}$.

For $f: \mathbb{R} \to \mathbb{R}$ we have seen that local extrema of f occur at critical points (ie points for which $f'(x)=0$)

The nature of extrema is determined by $f''(x)$.

If x_0 is a critical pt, Then

i) x_0 is a local min if $f''(x_0) > 0$

ii) x_0 " " max if $f''(x_0) < 0$

(iii) x_0 is a saddle point if $f''(x_0) = 0$.

For $f: \mathbb{R}^n \to \mathbb{R}$, the role of $f'(x)$ is taken over by ∇f

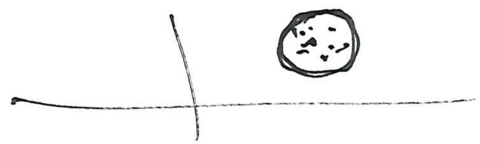
" " $f''(x)$ is " " by $\mathrm{Hes}(f)$.

Defn let $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ a scalar field.

A point $a = (a^1, a^2 \ldots a^n)$ is a <u>local maximum</u> <u>(resp loc min)</u> if we can find a neighborhood $B_a(r) := \{x \in \mathbb{R}^n \mid |x-a| < r\}$ completely contained in Ω

such that $\forall x \in B_a(r)$, $f(x) \leq f(a)$

(resp. $f(x) \geq f(a)$).



A local extrema is a point $a \in \Omega$ which is a local maximum or local minimum.

Thm. If f is a differentiable scalar field and $a = (a^1 \dots a^n)$ is a local extremum then $\nabla f(a) = 0$.

Defn A stationary point or a critical point of $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is a point $a \in \Omega$ where $\nabla f(a) = 0$.

It is called a saddle point if it is not a local extrema

Rk- If a is a saddle point then every nbhd $B_a(r)$ contains points x with $f(x) \geq f(a)$ and points x' such that $f(x') \leq f(a)$.

Ex: $f(x,y) = x^2 - y^2$

$\nabla f(x,y) = (2x, -2y)$

$\nabla f(x,y) = (0,0) = (2x, -2y) \implies x=0, y=0.$

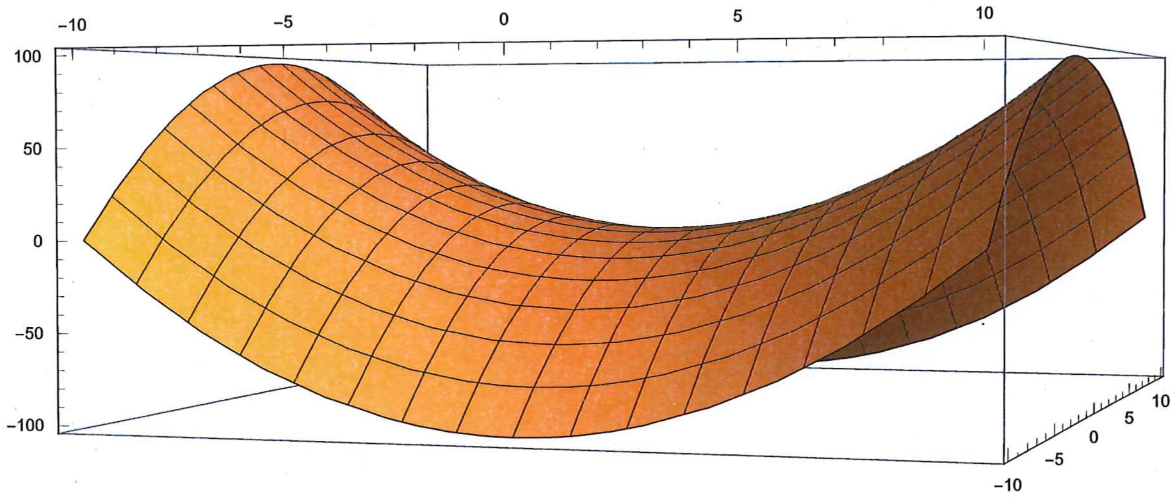
$(0,0)$ is the only critical point

It is a sadde a point. Because: if we move away from it along the y-axis the function is $f(0,y) = -y^2 < 0 = f(0,0)$

If we move along the x-axis, the function is $f(x,0) = x^2 > 0 = f(0,0)$

In[1]:= `Plot3D[x^2 - y^2, {x, -10, 10}, {y, -10, 10}]`

Out[1]=



To classify critical points we will use the Hessian matrix..

Assume $f \in C^2$ so that 2nd order mixed partials are equal and hence the Hessian Matrix is symmetric.

Defn. A symmetric Matrix $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ is called

1) positive definite iff $xAx^t = \sum a_{ij} x^i x^j > 0$ $\forall x = (x^1 \dots x^n) \in \mathbb{R}^n \setminus \{0\}$.

iff all its eigenvalues $\lambda_1, \dots, \lambda_n > 0$.

2) negative definite iff $\underset{1\times n}{x}\underset{n\times n}{A}x^t < 0$ $\forall x \in \mathbb{R}^n \setminus \{0\}$

iff $\lambda_1, \dots \lambda_n < 0$.

3) otherwise it is called indefinite (it has both positive and negative e.values or zero e.v.).

In the case $n=2$ of 2×2 matrices, the (in)definiteness is easily determined using the following thm.

<u>**Thm**</u> A symmetric 2×2 Matrix $A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ with $\det A \neq 0$ is

1) positive definite $\Leftrightarrow$ $a_{11}>0$ and $\det A>0$
2) negative definite $\Leftrightarrow$ $a_{11}<0$, $\det A>0$
3) indefinite $\Leftrightarrow$ $\det A<0$.

$$\begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdot & m_{2n} \\ \vdots & & & \\ m_{n1} & \cdots & & m_{nn} \end{pmatrix}$$

For the local extrema of $f: \mathbb{R}^n \to \mathbb{R}$.

Thm Let $f: \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^n$ be a scalar field. Assume $f \in C^2$. Let a be a critical point of f (ie $\nabla f(a) = 0$) with $\det(\operatorname{Hess} f(a)) \neq 0$

1) If $\operatorname{Hess}(f(a)) > 0$ (positive definite) then a is a local minimum.

2) If $\operatorname{Hess}(f(a)) < 0$ (negative defin) then a is a local maxim.

3) otherwise it is a saddle point.

Note If $\det(\operatorname{Hess}) = 0$, ie Hess is a singular Matrix the test does not say anything.

In the special case $f: \mathbb{R}^2 \to \mathbb{R}$ we have

Thm If $f \in C^2$, a is a critical point ($f:(x,y)$)

Then 1) If $\frac{\partial^2 f}{\partial x^2}(a) < 0$ and $\det(\text{Hess} f(a)) > 0$

then a is a loc. max of f.

2) If $\frac{\partial^2 f}{\partial x^2}(a) > 0$, $\det(\text{Hess} f) > 0$.

then a is a loc. min of f

3) If $\det(Hf(a)) < 0$ then a is a saddle point.

Ex. 1) $f(x,y,z) = (x-1)^2 + (y+2)^2 + (z+1)^2$

$\nabla f = \left(2(x-1),\ 2(y+2),\ 2(z+1)\right) = (0,0,0)$

$\Rightarrow x = 1$ $\Rightarrow (1,-2,-1)$ is the only critical point.
$y = -2$
$z = -1$

$$\mathrm{Hess}(f) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} > 0 \text{ positive defin.}$$

$(1,-2,-1)$ is a local min.

(2) $f(x,y) = \cos(x+2y) + \cos(2x+3y)$

$$\nabla f = \left(-\sin(x+2y) - 2\sin(2x+3y),\ -2\sin(x+2y) - 3\sin(2x+3y)\right).$$

$$= (0,0)$$

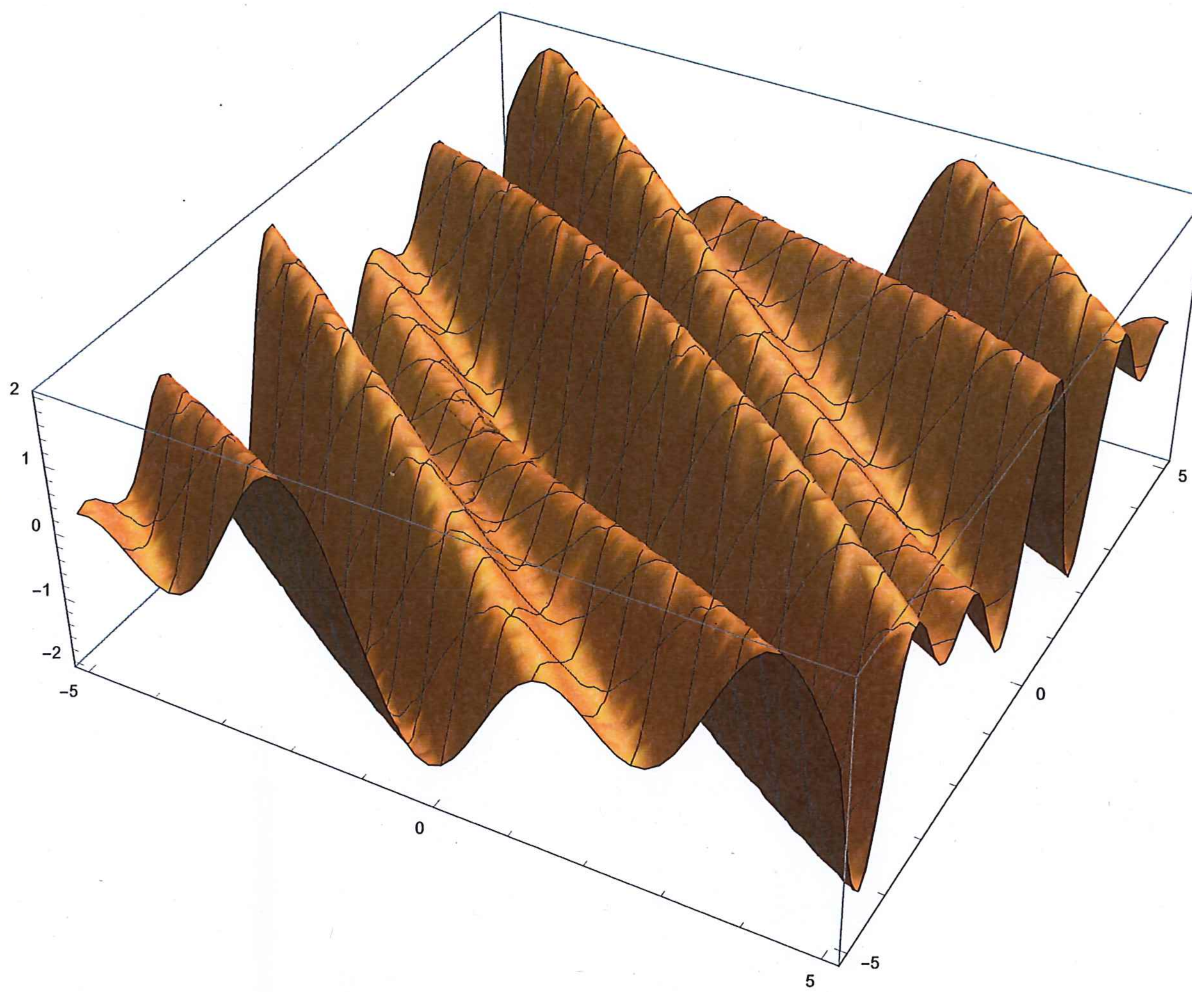
$$\Longrightarrow \left.\begin{array}{l} -\sin(x+2y) - 2\sin(2x+3y) = 0 \\ -2\sin(x+2y) - 3\sin(2x+3y) = 0 \end{array}\right\} \Longrightarrow \begin{array}{l} \sin(2x+3y) = 0. \\ \sin(\cancel{x} + 2y) = 0. \end{array}$$

$$\Longrightarrow \begin{array}{l} x+2y = k\pi \\ 2x+3y = l\pi \end{array} \Longrightarrow x \in k\pi,\ y \in l\pi \text{ with } k, l \in \mathbb{Z}.$$

crit pts $= \{ (\pi k, \pi l) \mid l, k \in \mathbb{Z} \}$.

In[5]:= Plot3D[Cos[x + 2 y] + Cos[2 x + 3 y], {x, -5, 5}, {y, -5, 5}]

Out[5]=



$19\frac{1}{2}$

$$\frac{\partial f}{\partial x \partial y} = -2\cos(x+2y) - 6\cos(2x+3y)$$

$$\frac{\partial^2 f}{\partial x^2} = -\cos(x+2y) - 4\cos(2x+3y)$$

$$\frac{\partial^2 f}{\partial y^2} = -4\cos(x+2y) - 9\cos(2x+3y)$$

In pt $(0,0)$ $\quad \nabla^2 f(0,0) = \underbrace{\begin{pmatrix} -5 & -8 \\ -8 & -13 \end{pmatrix}}_{\text{negative def.}}$ $\quad$ det $= 1 > 0$.

$\frac{\partial^2 f}{\partial x^2}(0,0) = -5 < 0$

Hence $(0,0)$ is a local maximum.

Similarly for all points $(2\pi k, 2\pi l)$ by periodicity.

Analog: at (π, π) $\frac{\partial^2 f}{\partial x^2} = 5 > 0$

$\frac{\partial^2 f}{\partial y^2} = 13$ $\frac{\partial^2 f}{\partial x \partial y} = 8.$, $Hess(f) = \begin{pmatrix} 5 & 8 \\ 8 & 13 \end{pmatrix} > 0$

$\det > 0$

$\frac{\partial^2 f}{\partial x^2} > 0$

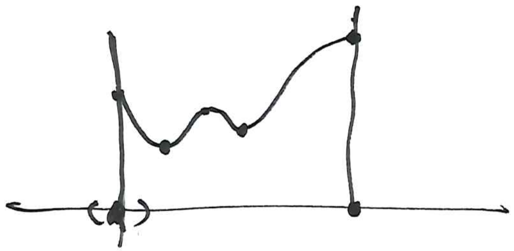
(π, π) is a loc min.

Similarly by periodicity also $((2k+1)\pi, (2l+1)\pi)$ are all loc. minima.

Check: $(0, \pi)$ and $(\pi, 0)$ Hence also $(2k\pi, (2l+1)\pi)$ $((2k+1)\pi, 2l\pi)$ are saddle points.

Global Extrema.

Recall $f: [a, b] \longrightarrow \mathbb{R}$, we saw that the global extrema of f is either at a critical point inside (a, b) or at the boundary points.



Thm Let f be differentiable in the interior of Ω. Then every global extremum of f is either at a critical point in the interior of Ω or on the boundary of Ω.