Finding Extrema (Maximum & Minimum)

Def: Let $\begin{cases} \bullet\ f: A \to \mathbb{R}, & A \subset \mathbb{R}^n \text{ (may not be open)} \\ \bullet\ \vec{a} \in A \end{cases}$

(1) f is said to have a <u>global (absolute) maximum at $\vec{a}$</u>

if $\quad f(\vec{a}) \geq f(\vec{x}) \quad \forall\, \vec{x} \in A$

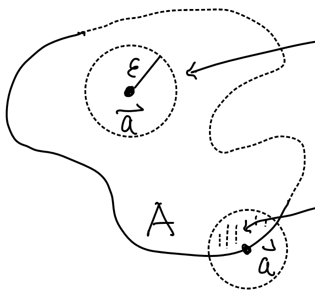
(2) f is said to have a <u>local (relative) maximum at $\vec{a}$</u>

if $\quad f(\vec{a}) \geq f(\vec{x}) \quad \forall\, \vec{x} \in A$ "near" $\vec{a}$

(i.e. $\exists\, \varepsilon > 0$ s.t. $f(\vec{a}) \geq f(\vec{x})$, $\forall\, \vec{x} \in A \cap B_\varepsilon(\vec{a})$)

(3) Similar definitions for <u>global (absolute) minimum</u> and <u>local (relative) minimum</u> by changing the inequality to $f(\vec{a}) \leq f(\vec{x})$.

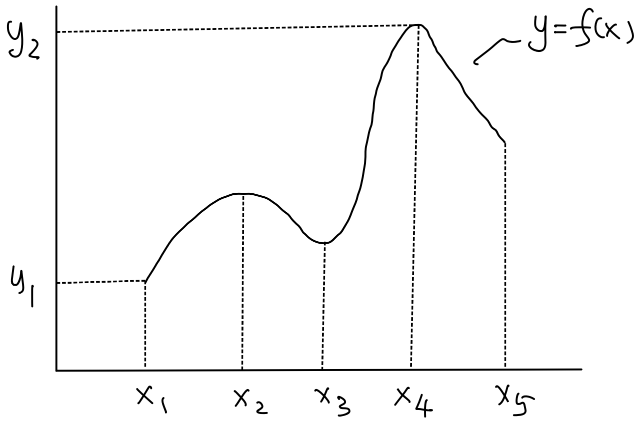
global max at $\vec{a}$ means $f(\vec{a}) \geq f(\vec{x})$ on the whole A



local max at $\vec{a}$ means $f(\vec{a}) \geq f(\vec{x})$ only holds inside the ball.

Remark: Global <u>Extremum</u> (max/min) is also a local extremum.

eg1 $f : [x_1, x_5] \longrightarrow \mathbb{R}$



Global: max : x_4, min : x_1

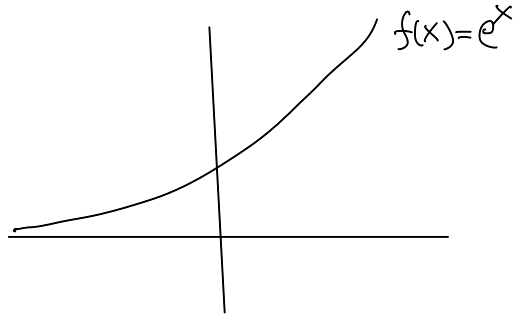
Local: max : x_2, x_4, min : x_1, x_3, x_5

max value : y_2

min value : y_1

eg 2 (NOT every function has global max/min)

(i) $f(x) = e^x$ on $\mathbb{R}$ (Domain unbounded)



No global min : $\lim_{x \to -\infty} f(x) = 0$

$\left(\Rightarrow \forall a \in \mathbb{R},\ f(a) = e^a > 0.\ \text{The limit} \Rightarrow \exists x \in \mathbb{R} \text{ s.t. } f(x) < f(a) \right)$

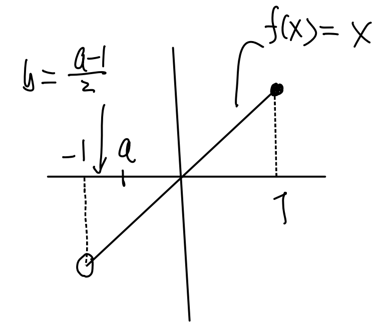
No global max : $\lim_{x \to +\infty} f(x) = +\infty$

$(\Rightarrow \forall a \in \mathbb{R},\ \exists x \in \mathbb{R} \text{ s.t. } f(x) > f(a) = e^a)$

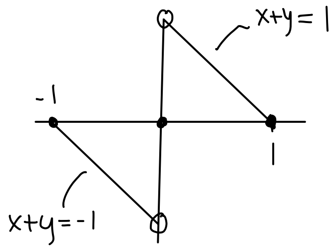
(ii) $f(x) = x$ on $(-1, 1]$ (Domain not closed)

Has global max at $x = 1$ (with max value $= f(1) = 1$), (Easy)

but no global min, $\forall a \in (-1, 1]$, $\exists b = \frac{a-1}{2} \in (-1, 1]$ s.t. $f(b) < f(a)$



(iii) $f = \begin{cases} 1-x, & 0<x\leq 1 \\ 0, & x=0 \\ -1-x, & -1\leq x<0 \end{cases}$ (discontinuous)



No global max and global min

(Ex! Similar argument as in (ii))

Extreme Value Thm (EVT)

Let
- $A \subseteq \mathbb{R}^n$ be closed and bounded,
- $f: A \to \mathbb{R}$ be continuous

Then f has global max and min.

(Proof: Omitted)

Remarks: (1) "Compact" = closed and bounded

(2) The Thm is a sufficient, but not a necessary condition.

Def: Let
- $f: A \to \mathbb{R}$, $A \subseteq \mathbb{R}^n$ (not necessary open)
- $\vec{a} \in \text{Int}(A)$

Then $\vec{a}$ is called a critical point of f if

either (1) $\vec{\nabla} f(\vec{a})$ DNE (does not exist)

or (2) $\vec{\nabla} f(\vec{a}) = \vec{0}$

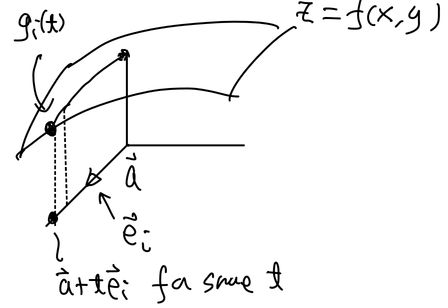
(i.e. either "$\frac{\partial f}{\partial x_i}(\vec{a})$ DNE for some $i=1,\cdots,n$" or "$\frac{\partial f}{\partial x_i}(\vec{a})=0$ for all $i=1,\cdots,n$")

Thm (First Derivative Test)

Suppose $f: A \to \mathbb{R}$ $(A \subset \mathbb{R}^n)$ attains a local extremum at $\vec{a} \in \mathrm{Int}(A)$, then $\vec{a}$ is a critical point of f.

Pf: Suppose f has a local extremum at $\vec{a} \in \mathrm{Int}(A)$.

If $\vec{\nabla} f(\vec{a})$ DNE, then $\vec{a}$ is a critical point.



If $\vec{\nabla} f(\vec{a})$ exists, then

$$\frac{\partial f}{\partial x_i}(\vec{a}), \quad \forall i = 1, \cdots, n$$

$\forall i$, consider the 1-variable function

$$g_i(t) = f(\vec{a} + t\vec{e}_i) \qquad (i = 1, \cdots, n) \qquad \vec{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i\text{th component}$$

Then $g_i(0) = f(\vec{a}) \begin{smallmatrix} \le \\ \text{or} \\ \ge \end{smallmatrix} f(\vec{a}+t\vec{e}_i)$ for $|t|$ small enough s.t. $\vec{a}+t\vec{e}_i$ is "near" $\vec{a}$

$\Rightarrow$ $t = 0$ is a local max/min

1-variable theory $\Rightarrow$

$$0 = g_i'(0) = \left.\frac{d}{dt}\right|_{t=0} f(\vec{a} + t\vec{e}_i)$$

$$= \vec{\nabla} f(\vec{a}) \cdot \vec{e}_i$$

$$= \frac{\partial f}{\partial x_i}(\vec{a}) \quad (\forall i)$$

$$\Rightarrow \quad \vec{\nabla} f(\vec{a}) = \vec{0}$$

※

Strategy for finding Extrema

$f : A \to \mathbb{R}$

(1) Find critical points of f in $\mathrm{Int}(A)$.

(2) Study f on boundary ∂A :
find max/min of f on ∂A.

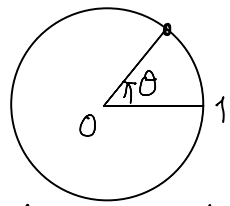
(3) Compare values of f at points found in steps (1) & (2).

<u>eg1</u> Find global max / min of

cts closed & bdd

$$f(x,y) = x^2 + 2y^2 - x + 3 \quad \text{for} \quad x^2+y^2 \le 1 \quad \left(A = \{x^2+y^2 \le 1\} \right)$$

<u>Soln</u> : <u>Step 1</u>



closed unit disk

Critical points in $\mathrm{Int}(A)$

$$= \mathrm{Int}\{x^2+y^2 \le 1\}$$

$$= \{x^2+y^2 < 1\}$$

(f is a polynomial $\Rightarrow \vec{\nabla} f$ always exist)

$$\vec{\nabla} f = (2x-1, 4y)$$

$$\vec{\nabla} f = \vec{0} \iff \begin{cases} 2x-1=0 \\ 4y=0 \end{cases} \iff (x,y) = (\tfrac{1}{2}, 0)$$

the only critical pt. in $\mathrm{Int}(A)$.

$\Rightarrow f(\frac{1}{2}, 0) = (\frac{1}{2})^2 + 0 - \frac{1}{2} + 3 = \frac{11}{4}$ (check!)

Step 2 Study f on $\partial\{x^2+y^2 \leq 1\} = \{x^2+y^2=1\}$

parametrize the boundary $\{x^2+y^2=1\}$ by angle θ:

$$\begin{cases} x = \cos\theta \\ y = \sin\theta \end{cases} \qquad \theta \in [0, 2\pi]$$

Therefore on $\partial\{x^2+y^2 \leq 1\}$,

$$f(\theta) = f(\cos\theta, \sin\theta) = \cos^2\theta + 2\sin^2\theta - \cos\theta + 3$$

$$= -\cos^2\theta - \cos\theta + 5$$

$$= -(\cos\theta + \tfrac{1}{2})^2 + \frac{21}{4} \quad \text{(check!)}$$

max value of f on $\partial A = \frac{21}{4}$ (at $\cos\theta = -\frac{1}{2} \Rightarrow (x,y) = (-\frac{1}{2}, \pm\frac{\sqrt{3}}{2})$)

min value of f on $\partial A = -(1+\frac{1}{2})^2 + \frac{21}{4} = 3$ (check!)

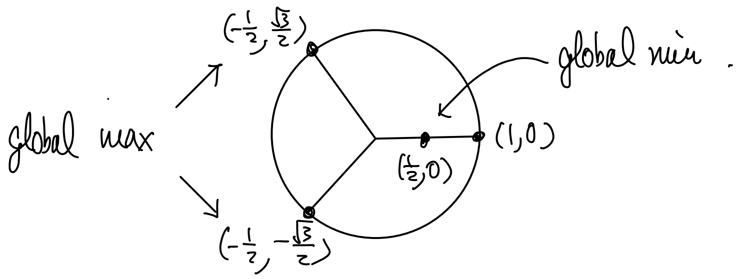
(at $\cos\theta = 1 \Rightarrow (x,y) = (1,0)$)

Step 3 Compare values of f at points from steps (1) & (2)

$f(\frac{1}{2}, 0) = \frac{11}{4}$ (value of critical pt. in Int(A))

$f(-\frac{1}{2}, \pm\frac{\sqrt{3}}{2}) = \frac{21}{4}$ (max value of f on ∂A)

$f(1, 0) = 3$ (min value of f on ∂A



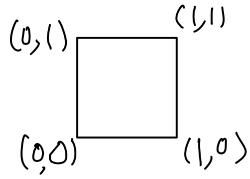
$\Rightarrow$ $\begin{cases} \text{max value} = \frac{21}{4} \text{ at the global max points } (-\frac{1}{2}, \pm\frac{\sqrt{3}}{2}) \\ \text{min value} = \frac{11}{4} \text{ at the global min point } (\frac{1}{2}, 0) \end{cases}$

$\#$

eg2 Find global max & min and values of

$$f(x,y) = 6xy - 4x^3 - 3y^2$$

on the set $\{(x,y) \in \mathbb{R}^2 : 0 \le x \le 1,\ 0 \le y \le 1\}$



Soln Step 1: Interior extremum

(f is a polynomial, $\vec{\nabla} f$ always exist)

$$\begin{cases} 0 = \frac{\partial f}{\partial x} = 6y - 12x^2 \\ 0 = \frac{\partial f}{\partial y} = 6x - 6y \end{cases}$$

$\Rightarrow$ $y = 2x^2$ & $x = y$

$\Rightarrow$ $x = 2x^2$

$\Rightarrow$ $x = 0$ or $x = \frac{1}{2}$

Hence $\underset{\text{on boundary}}{(x,y) = (0,0)}$ & $\underset{\text{interior.}}{(x,y) = (\frac{1}{2}, \frac{1}{2})}$

values $\quad f(0,0)=0$ (boundary)

$$f\left(\tfrac{1}{2},\tfrac{1}{2}\right) = 6\cdot\tfrac{1}{2}\cdot\tfrac{1}{2} - 4\left(\tfrac{1}{2}\right)^3 - 3\left(\tfrac{1}{2}\right)^2$$

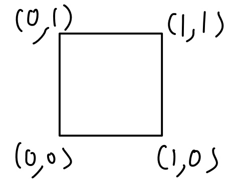
$$= \tfrac{1}{4}$$ (check!)

(To be cont'd)

(Cont'd) $f(x,y) = 6xy - 4x^3 - 3y^2$

Interior critical point $(\frac{1}{2}, \frac{1}{2})$ with value $f(\frac{1}{2},\frac{1}{2}) = \frac{1}{4}$.

<u>Boundary</u>



(1) $\{x=0,\ 0 \le y \le 1\}$: $f(0,y) = -3y^2$

$\Rightarrow f(0,1) = -3 \le f(0,y) \le 0 = f(0,0)$

(2) $\{y=0,\ 0 \le x \le 1\}$: $f(x,0) = -4x^3$

$\Rightarrow -4 = f(1,0) \le f(x,0) \le 0 = f(0,0)$

(3) $\{x=1,\ 0 \le y \le 1\}$: $f(1,y) = 6y - 4 - 3y^2 = -3(y-1)^2 - 1$

$f(1,0) = -4 \le f(1,y) \le -1 = f(1,1)$ $\left(\begin{array}{c} -1 \le y-1 \le 0 \\ \Rightarrow (y-1)^2 \le 1 \end{array} \right)$

(4) $\{y=1,\ 0 \le x \le 1\}$: $f(x,1) = 6x - 4x^3 - 3$

$0 = \frac{d}{dx} f(x,1) = 6 - 12x^2 \Rightarrow x = \frac{1}{\sqrt{2}}$ $\left(-\frac{1}{\sqrt{2}} \text{ rejected as } 0 \le x \le 1 \right)$

$f(\frac{1}{\sqrt{2}}, 1) = \frac{6}{\sqrt{2}} - 4\left(\frac{1}{\sqrt{2}}\right)^3 - 3 = -(3 - 2\sqrt{2}) < 0$ (check!)

$f(0,1) = -3$, $f(1,1) = -1$ $(< -(3-2\sqrt{2}))$

$\Rightarrow -3 = f(0,1) \le f(x,1) \le f(\frac{1}{\sqrt{2}}, 1) = -(3 - 2\sqrt{2})$

<u>Compare values</u>

Global max. pt. $(\frac{1}{2}, \frac{1}{2})$ with value $\frac{1}{4}$

Global min. pt. $(1,0)$ with value -4 ※

Matrix form for 2^{nd} order Taylor Polynomial

Def: Let $f: \Omega \to \mathbb{R}$ be C^2 ($\Omega \subseteq \mathbb{R}^n$, open).
Then the Hessian matrix of f at $\vec{a} \in \Omega$ is

$$Hf(\vec{a}) = \begin{bmatrix} f_{x_1x_1}(\vec{a}) & \cdots & f_{x_1x_n}(\vec{a}) \\ \vdots & & \vdots \\ f_{x_nx_1}(\vec{a}) & \cdots & f_{x_nx_n}(\vec{a}) \end{bmatrix}$$

Remarks (1) $Hf(\vec{a})$ is $n \times n$ symmetric (by Clairaut's Thm)

(2) In Textbook, Hessian of $f = \det(Hf(\vec{a}))$

So we emphasize our definition is a matrix (More common in advanced level math)

eg: $f(x,y)$ at $(0,0)$

$$Hf(0,0) = \begin{bmatrix} f_{xx}(0,0) & f_{xy}(0,0) \\ f_{yx}(0,0) & f_{yy}(0,0) \end{bmatrix} \qquad (f_{xy} = f_{yx})$$

$$[x, y] \begin{bmatrix} f_{xx}(0,0) & f_{xy}(0,0) \\ f_{yx}(0,0) & f_{yy}(0,0) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \underbrace{f_{xx}(0,0)x^2 + 2f_{xy}(0,0)xy + f_{yy}(0,0)y^2}_{2^{nd} \text{ order term in Taylor polynomials (up to a factor } \frac{1}{2!})}$$

2^{nd} order Taylor polynomial of f at $\vec{a}$ in matrix form

$$\boxed{P_2(\vec{x}) = f(\vec{a}) + \vec{\nabla} f(\vec{a})(\vec{x}-\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^T Hf(\vec{a})(\vec{x}-\vec{a})}$$

where $\vec{\nabla} f(\vec{a})$ regarded as row vector $[f_{x_1}(\vec{a}) \cdots f_{x_n}(\vec{a})]$,

$\vec{x}-\vec{a}$ regarded as column vector $\begin{bmatrix} x_1-a_1 \\ \vdots \\ x_n-a_n \end{bmatrix}$

& $(\vec{x}-\vec{a})^T$ is the transpose $[x_1-a_1 \ \cdots \ x_n-a_n]$
(row vector)

eg $g(x,y) = \frac{\ln x}{1-y}$. Find $P_2(x,y)$ at $(1,0)$ using matrix form.

Soln: $g(1,0) = 0$

$$\vec{\nabla} g = [\tfrac{\partial g}{\partial x} \ \ \tfrac{\partial g}{\partial y}] = [\tfrac{1}{x(1-y)} \ \ \tfrac{\ln x}{(1-y)^2}]$$

$$Hg = \begin{bmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{bmatrix} = \begin{bmatrix} -\frac{1}{x^2(1-y)} & \frac{1}{x(1-y)^2} \\ \frac{1}{x(1-y)^2} & \frac{\ln x}{(1-y)^3} \end{bmatrix}$$

$$\Rightarrow \vec{\nabla} g(1,0) = [1,0], \quad Hg(1,0) = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\therefore P_2(x,y) = g(1,0) + \vec{\nabla} g(1,0)\begin{bmatrix} x-1 \\ y \end{bmatrix} + \frac{1}{2}[x-1 \ \ y]\, Hg(1,0)\begin{bmatrix} x-1 \\ y \end{bmatrix}$$

$$= 0 + [1 \ 0]\begin{bmatrix} x-1 \\ y \end{bmatrix} + \frac{1}{2}[x-1 \ \ y]\begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x-1 \\ y \end{bmatrix}$$

$$= (x-1) - \frac{1}{2}(x-1)^2 + (x-1)y \quad \text{(Check!)}$$

※

Application to local max/min

If f is C^2, and $\vec{a}$ is a critical point of f.

Then $\vec{\nabla} f(\vec{a}) = \vec{0}$

$$\Rightarrow \quad f(\vec{x}) \simeq P_2(\vec{x}) = f(\vec{a}) + \overset{\vec{0}}{\cancel{\vec{\nabla} f(a)}}(\vec{x}-\vec{a}) + \tfrac{1}{2}(\vec{x}-\vec{a})^T Hf(a)(\vec{x}-\vec{a})$$

$$f(\vec{x}) - f(\vec{a}) \simeq \tfrac{1}{2}(\vec{x}-\vec{a})^T Hf(a)(\vec{x}-\vec{a})$$

$\therefore$ If $(\vec{x}-\vec{a})^T Hf(a)(\vec{x}-\vec{a}) < 0 \quad \forall \vec{x}$ near $\vec{a}$

then $f(\vec{x}) < f(\vec{a}) \quad \forall \vec{x}$ near $\vec{a}$

$\Rightarrow$ $\vec{a}$ is a <u>local max</u>

If $(\vec{x}-\vec{a})^T Hf(a)(\vec{x}-\vec{a}) > 0 \quad \forall \vec{x}$ near $\vec{a}$

then $f(\vec{x}) > f(\vec{a}) \quad \forall \vec{x}$ near $\vec{a}$

$\Rightarrow$ $\vec{a}$ is a <u>local min</u>

So we need to study when is a sym matrix H satisfies

$$\vec{v}^T H \vec{v} > 0 \quad \forall \text{ vector } \vec{v} \neq \vec{0}$$

and

$$\vec{v}^T H \vec{v} < 0 \quad \forall \text{ vector } \vec{v} \neq \vec{0}$$

Hence we make the following

Def: Let H be a symmetric $n \times n$ matrix.
Then H is said to be

(1) positive definite if $\vec{x}^T H \vec{x} > 0$
for all column vector $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$

(2) negative definite if $\vec{x}^T H \vec{x} < 0$
for all column vector $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$

(3) indefinite if $\exists$ column vectors $\vec{x}, \vec{y} \in \mathbb{R}^n \setminus \{\vec{0}\}$
such that $\vec{x}^T H \vec{x} > 0$ and $\vec{y}^T H \vec{y} < 0$

Remark: These are not all possibilities : $\exists$ sym. matrix which is not positive definite, negative definite, nor indefinite.

Then the discussion above implies

Thm (Second Derivative Test)

Let $\begin{cases} \bullet\ f: \Omega \to \mathbb{R} \text{ be } C^2,\ \Omega \subseteq \mathbb{R}^n \text{, open} \\ \bullet\ \vec{a} \in \Omega \text{ such that } \vec{\nabla} f(\vec{a}) = \vec{0} \end{cases}$

Then

$$Hf(\vec{a}) \text{ is } \begin{cases} \text{positive definite} & \Rightarrow \vec{a} \text{ is a local } \underline{\text{min}} \\ \text{negative definite} & \Rightarrow \vec{a} \text{ is a local } \underline{\text{max}} \\ \text{indefinite} & \Rightarrow \vec{a} \text{ is a } \underline{\text{saddle}} \text{ point} \end{cases}$$

Remark: A critical point which is neither local max nor local min is called a saddle point.

In particular for 2-variable, $\vec{v}^T H \vec{v}$ is of the form

$$g(x,y) = [x\ y]\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2$$

eg (1) $[x,y]\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = x^2 + 4y^2 > 0, \ \forall \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \setminus \{\vec{0}\}$

$\therefore \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$ is positive definite.

(2) $[x,y]\begin{bmatrix} -1 & 0 \\ 0 & -4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = -x^2 - 4y^2 < 0, \ \forall \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \setminus \{\vec{0}\}$

$\therefore \begin{bmatrix} -1 & 0 \\ 0 & -4 \end{bmatrix}$ is negative definite.

(3) $[x,y]\begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = -x^2 + 4y^2$

If $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, then $[x,y]\begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = -0^2 + 4(1)^2 = 4 > 0$

If $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, then $[x,y]\begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = -(1)^2 + 4(0)^2 = -1 < 0$

$\therefore \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$ is indefinite.

(4) $[x,y]\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = x^2 \geq 0, \ \forall \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \setminus \{\vec{0}\}$

But $[0\ 1]\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \Rightarrow$ not positive definite

$\therefore \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is not positive definite, negative definite, nor indefinite.

(5) $[x, y]\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = x^2 + 4xy + 5y^2$

$$= x^2 + 4xy + 4y^2 + y^2$$

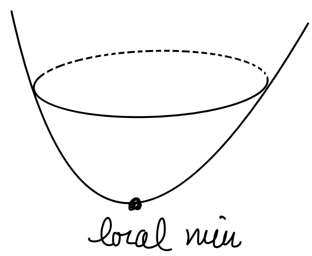
$$= (x+2y)^2 + y^2$$

$$> 0 \qquad \forall \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \setminus \{\vec{0}\}$$

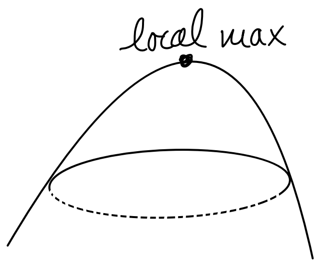
$\Rightarrow \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$ is positive definite.

<u>Geometrically</u> (locally near the <u>critical point</u>)

(1) $Hf(\vec{a})$ positive definite



(2) $Hf(\vec{a})$ negative definite



(3) $Hf(\vec{a})$ is indefinite