LECTURE 6 (F)

GRA 6035 MATHEMATICS

EIVIND ERIKSEN

SEP 29, 2014

Plan:

1. Unconstrained optimization
2. Partial derivatives and Hessians
3. First order conditions
4. Second order conditions
5. Convex and concave functions

Reading:

[ME] 14.1–14.4, 14.8
17.1–17.5

---

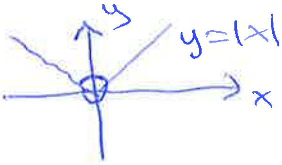
1. Unconstrained optimization problems

max/min $f(x_1, \ldots, x_n) = f(\underline{x})$ when $\underline{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is any point

We assume that f is "nice": polynomials, rational, exponentials, logarithms, ...

f is C^1: all partial derivatives of f exist and are continuous

Ex: $f(x) = |x|$ is not C^1



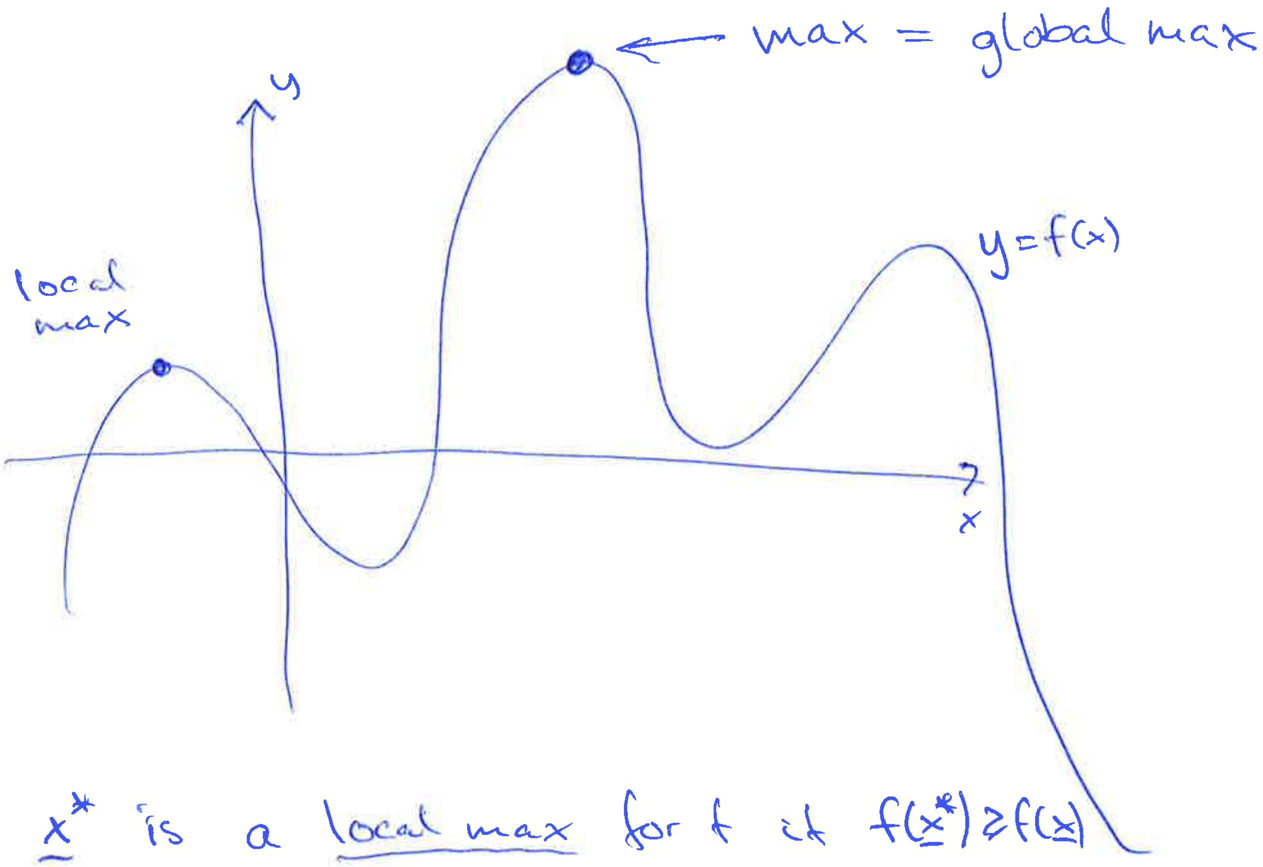
$$y = |x| = \begin{cases} x, & x \geq 0 \\ -x, & x < 0 \end{cases}$$

f is C^2: all second order partial derivatives exist and are continuous

max/min:

$\underline{x}^*$ is a <u>maximum for f</u> if $f(\underline{x}^*) \geq f(\underline{x})$ at all points $\underline{x}$ in $\mathbb{R}^n$

$\underline{x}^*$ is a <u>minimum for f</u> if $f(\underline{x}^*) \leq f(\underline{x})$ at all points $\underline{x}$ in $\mathbb{R}^n$.



$\underline{x}^*$ is a <u>local max</u> for f if $f(\underline{x}^*) \geq f(\underline{x})$ for all points $\underline{x}$ close to $\underline{x}^*$

$\underline{x}^*$ is a <u>local min</u> for f if $f(\underline{x}^*) \leq f(\underline{x})$ for all points $\underline{x}$ close to $\underline{x}^*$

2 Partial derivatives

Ex: $f(x,y) = x^3 + xy - y^3$

$$f'_x = \frac{\partial f}{\partial x} = \underline{3x^2 + y}$$
$$f'_y = \frac{\partial f}{\partial y} = \underline{x - 3y^2}$$

first order partial derivatives

$$f''_{xx} = 6x$$
$$f''_{xy} = 1$$
$$f''_{yx} = 1$$
$$f''_{yy} = -6y$$

Second order partial derivatives

Hessian matrix:

$$H(f)(x,y) = \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{pmatrix} = \begin{pmatrix} 6x & 1 \\ 1 & -6y \end{pmatrix}$$

Young's lemma:

If f is a C^2 function, then $f''_{x_i x_j} = f''_{x_j x_i}$.

Notice that

$$f''_{x_i x_j} = f''_{x_j x_i} \iff H(f) \text{ is symmetric}$$

Ex: $f(x_1, x_2, x_3) = x_1^2 - 6x_1x_2 + 2x_2^2 + 10x_2x_3 + 4x_3^2$

$$f'_{x_1} = 2x_1 - 6x_2$$

$$f'_{x_2} = -6x_1 + 4x_2 + 10x_3$$

$$f'_{x_3} = 10x_2 + 8x_3$$

$$f''_{11} = 2 \qquad f''_{12} = -6 \qquad f''_{13} = 0$$

$$f''_{21} = -6 \qquad f''_{22} = 4 \qquad f''_{23} = 10$$

$$f''_{31} = 0 \qquad f''_{32} = 10 \qquad f''_{33} = 8$$

$$H(f) = \begin{pmatrix} 2 & -6 & 0 \\ -6 & 4 & 10 \\ 0 & 10 & 8 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & -3 & 0 \\ -3 & 2 & 5 \\ 0 & 5 & 4 \end{pmatrix}$$

Symm. matrix of the quadr. form

$$2A = H(f)$$

Ex: $f(x,y) = \ln(1+x^2+y^2)$

$$f'_x = \frac{1}{1+x^2+y^2} \cdot 2x = \frac{2x}{1+x^2+y^2}$$

$$f'_y = \frac{1}{1+x^2+y^2} \cdot 2y = \frac{2y}{1+x^2+y^2}$$

$$f''_{xx} = \left(\frac{2x}{1+x^2+y^2}\right)'_x = \frac{2\cdot(1+x^2+y^2) - 2x\cdot 2x}{(1+x^2+y^2)^2}$$

$$= \boxed{\frac{2-2x^2+2y^2}{u^2}} \qquad u = 1+x^2+y^2$$

$$f''_{xy} = \left(\frac{2x}{1+x^2+y^2}\right)'_y = \frac{0 - 2x\cdot 2y}{u^2} = \boxed{\frac{-4xy}{u^2}}$$

$$f''_{yy} = \left(\frac{2y}{1+x^2+y^2}\right)'_y = \frac{2\cdot(1+x^2+y^2) - 2y\cdot 2y}{(1+x^2+y^2)^2}$$

$$= \boxed{\frac{2+2x^2-2y^2}{u^2}}$$

$$H(f)(x,y) = \begin{pmatrix} \frac{2-2x^2+2y^2}{(1+x^2+y^2)^2} & \frac{-4xy}{(1+x^2+y^2)^2} \\ \frac{-4xy}{(1+x^2+y^2)^2} & \frac{2+2x^2-2y^2}{(1+x^2+y^2)^2} \end{pmatrix}$$

$f = \ln(1+x^2+y^2)$

3 First order conditions

Defn: A stationary pt. for f is a pt. that satisfies

FOC = first order conditions $\longrightarrow$ $\boxed{\frac{\partial f}{\partial x_1} = 0, \ \frac{\partial f}{\partial x_2} = 0, \ \ldots, \ \frac{\partial f}{\partial x_n} = 0}$

Ex: $f(x,y) = x^3 + 9xy - y^3$

Stationary pts:
$y=0 \Rightarrow x=0$
$y=-3 \Rightarrow x=3$

$$\boxed{\begin{aligned} &\text{I} \quad f'_x = 3x^2 + 9y = 0 \\ &\text{II} \quad f'_y = 9x - 3y^2 = 0 \end{aligned}} \quad \text{FOC}$$

II: $x = \frac{3y^2}{9} = \frac{y^2}{3}$

I: $3\cdot(y^2/3)^2 + 9y = 0$

$\frac{1}{3}y^4 + 9y = 0 \quad |\cdot 3$

$y^4 + 27y = 0$

$y(y^3+27) = 0 \Rightarrow \underline{y=0}$ or $y = \sqrt[3]{-27} = \underline{-3}$

Conclusion: $f(x,y) = x^3 + 9xy - y^3$ has two stationary pts:

$(x,y) = \underline{(0,0)}$ and $(x,y) = \underline{(3,-3)}$

A critical point for f is a pt. where either i) $\frac{\partial f}{\partial x_1} = \dots = \frac{\partial f}{\partial x_n} = 0$ or ii) at least one of the partial derivatives are not defined.

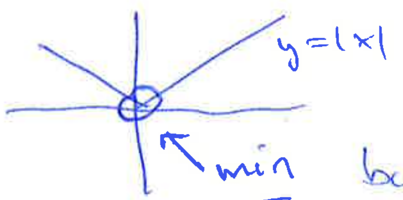
When f is C^1, critical points = stationary pts.

Fact: When f is a C^2 function,

$$\boxed{\underline{x}^* \text{ is max/min for } f \implies \underline{x}^* \text{ is stationary pt. for } f}$$

This means that the stationary pts are the candidates for max/min.

Ex: $f(x) = |x|$ $\qquad$ f is not C^1



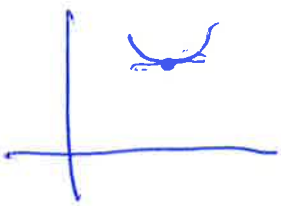
min but not stationary pt (it is a critical pt)

④ Second order conditions

Fact: If $\underline{x}^*$ is a stationary pt. for f and f is C^2, then

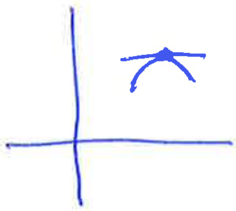
$H(f)(\underline{x}^*)$ positive definite	$\Rightarrow$	$\underline{x}^*$ local min
—"— negative definite	$\Rightarrow$	$\underline{x}^*$ local max
—"— indefinite	$\Rightarrow$	$\underline{x}^*$ saddle pt.

If $H(f)(\underline{x}^*)$ is positive/negative semidefinite, but not positive/negative definite, then the test is inconclusive.



$H(f)(\underline{x}^*)$ is positive definite

local min



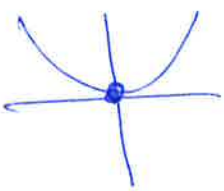
$H(f)(\underline{x}^*)$ is negative definite

local max

Saddle point = stationary point that is not local max, not local min.

Ex:

$f(x) = x^4$

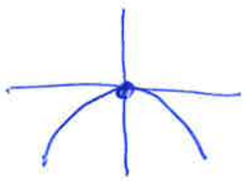


$f' = 4x^3 = 0$

$x = 0$

$f'' = 12x^2 \quad f''(0) = 0.$

$f(x) = -x^4$



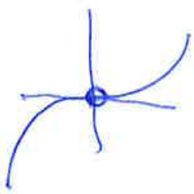
$f' = -4x^3 = 0$

$x = 0$

$f'' = -12x^2$

$f''(0) = 0$

$f(x) = x^3$



$f' = 3x^2 = 0$

$x = 0$

$f'' = 6x$

$f''(0) = 0$

Ex: $f(x,y) = x^3 + 9xy - y^3$

$f'_x = 3x^2 + 9y$

$f'_y = 9x - 3y^2$

Stationary pts:

$(0,0)$ $f(0,0) = \underline{0}$

$(3,-3)$ $f(3,-3) = \underline{-27}$

(candidates for max/min)

$$H(f) = \begin{pmatrix} 6x & 9 \\ 9 & -6y \end{pmatrix}$$

Classify $(0,0)$: $H(f)(0,0) = \begin{pmatrix} 0 & 9 \\ 9 & 0 \end{pmatrix}$

$D_1 = 0$

$D_2 = -81$

$D_2 < 0$ means $H(f)(0,0)$ is indefinite

$(0,0)$ is a saddle pt

(f has no max)

Classify $(3,-3)$: $H(f)(3,-3) = \begin{pmatrix} 18 & 9 \\ 9 & 18 \end{pmatrix}$

$D_1 = 18 > 0$

$D_2 = 18^2 - 9^2 > 0$

$D_1 > 0, D_2 > 0$ means that $H(f)(3,-3)$ is pos. defn.

$(3,-3)$ is local min

it may be the min. of f

Question: Is $(3,-3)$ a global min or just a local min?

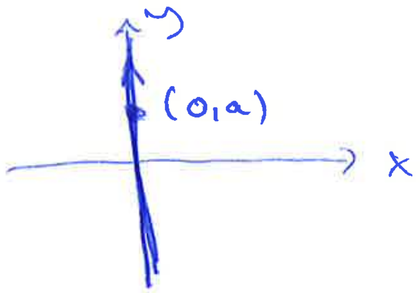
$f(x,y) = x^3 + 9xy - y^3$

$x = 0,\ y = a$:

$f(0,a) = 0^3 + 9 \cdot 0 \cdot a - a^3$

$= -a^3 \to -\infty$

when $a \to \infty$



Conclusion: f has no min.

Comment: $f(x,y) = --$

$f'_x = --$ $f''_{xx} = --$ $f''_{xy} = --$

$f'_y = --$ $f''_{yy} = --$

$$H(f)(\underline{x}^*) = \begin{pmatrix} A & B \\ B & C \end{pmatrix} \quad \begin{cases} A = f''_{xx}(\underline{x}^*) \\ B = f''_{xy}(\underline{x}^*) \\ C = f''_{yy}(\underline{x}^*) \end{cases}$$

$D_1 = A$

$D_2 = AC - B^2$

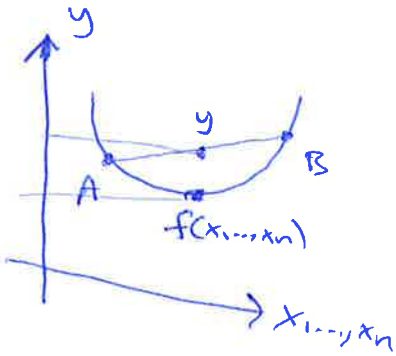
$A > 0,\ AC - B^2 > 0$

$\Downarrow$

$\underline{x}^*$ local min

(5) Convex and concave functions

$f(\underline{x}) = f(x_1, \dots, x_n)$

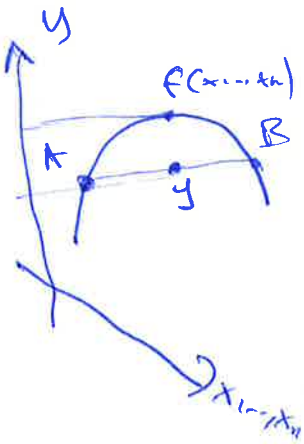


__Defn:__ f is <u>convex</u> if the following condition holds:

For all pts. A and B on the graph of f, any pt $(x_1, \dots, x_n, y)$ on the line segment $[A, B]$ satisfies

$$y \geq f(x_1, \dots, x_n)$$

("the points lies on/over the graph")



f is <u>concave</u> if the following condition holds:

For all pts A and B on the graph of f, any pt. $(x_1, \dots, x_n, y)$ on the line segment $[A, B]$ satisfies

$$y \leq f(x_1, \dots, x_n)$$

("the point lies on/under the graph")

Condition for convex/concave functions:

f convex $\Leftrightarrow$ $H(f)(\underline{x})$ is positive semidefinite for all $\underline{x}$

f concave $\Leftrightarrow$ $H(f)(\underline{x})$ is negative semidefinite for all $\underline{x}$

Ex: $f(x,y) = x^3 + 9xy - y^3$

$$H(f) = \begin{pmatrix} 6x & 9 \\ 9 & -6y \end{pmatrix} \qquad \begin{array}{l} D_1 = 6x \\ D_2 = -36xy - 81 \end{array}$$

f convex $\Leftrightarrow$ positive semidefn. for all x,y $\Leftrightarrow$ $D_1 = 6x \geq 0$ $\quad \Delta_1 = -6y \geq 0$

$D_2 = -36xy - 81 \geq 0$

for all x,y

f is not convex

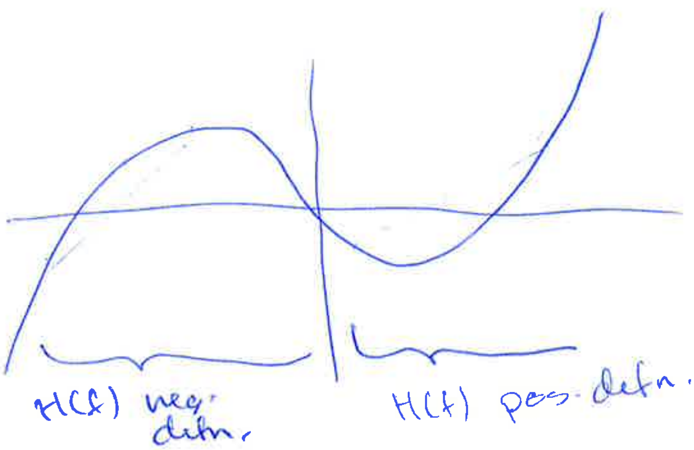
f concave $\Leftrightarrow$ $H(f)$ neg. semidefn for all x,y $\Leftrightarrow$ $D_1 = 6x \leq 0$ for all x,y

$D_2 = \ldots$

$\ldots$

f is not concave

Ex:



$H(f)(x)$ pos. semidefn. for $x \geq 0$

not convex

not concave

Fact:

If f is convex, any stationary pt. is global min
If f is concave, any stat. pt is global max

Ex: $f(x,y,z) = x^2 + 6xy + 2y^2 + 4z^2$

$f'_x = 2x + 6y$

$f'_y = 6x + 4y$

$f'_z = 8z$

$$H(t) = \begin{pmatrix} 2 & 6 & 0 \\ 6 & 4 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

$D_1 = 2 > 0$

$D_2 = 2 \cdot 4 - 6^2 < 0$

$D_3 = 8 \cdot (2 \cdot 4 - 6^2) < 0$

indefinite for all (x,y,z)

not convex
not concave

Ex: $f(x,y) = e^{2x+3y}$

$f'_x = e^{2x+3y} \cdot 2$

$f'_y = e^{2x+3y} \cdot 3$

$$H(t) = \begin{pmatrix} 4e^{2x+3y} & 6e^{2x+3y} \\ 6e^{2x+3y} & 9e^{2x+3y} \end{pmatrix}$$

$H(t)$ is positive semidefin. for all (x,y).
$\left\{ \begin{array}{l} \Delta_1 = 9e^{2x+3y} > 0 \end{array} \right.$

$D_1 = 4e^{2x+3y} > 0$

$D_2 = 36\left(e^{2x+3y}\right)^2 - 36\left(e^{2x+3y}\right)^2$
$= 0$

f is convex

Fact:

f convex $\Longleftrightarrow$ $\Delta_1 \geq 0$, $\Delta_2 \geq 0$, ..., $\Delta_n \geq 0$
for all principal minors
and all pts $(x_1, \dots, x_n)$.

f concave $\Longleftrightarrow$ $\Delta_1 \leq 0$, $\Delta_2 \geq 0$, ..., $(-1)^n \Delta_n \geq 0$
for all principal minors
at all pts. $(x_1, \dots, x_n)$

Method for finding global max/min:

① Find stationary pts using FOC

② Classify stationary pts in ① as local min, local max or saddle pts. using $H(f)(x^*)$.

Try to say something about global max/min using either

i) f is convex or concave, or

ii) by considering the candidate point, ~~[illegible]~~ their type and their values

This can in some cases be difficult.

Ex: $f(x,y) = xy^2 + x^3y - xy$

$f'_x = y^2 + 3x^2y - y = 0 \qquad y(y+3x^2-1)=0$

$f'_y = 2xy + x^3 - x = 0 \qquad x(2y+x^2-1)=0$

Stationary pts:

$x=0,\ y=0 \qquad \Rightarrow (0,0)$

or

$x=0,\ y+3x^2-1=0 \Rightarrow y=1 \Rightarrow (0,1)$

or

$y=0,\ 2y+x^2-1=0 \Rightarrow x=\pm 1 \Rightarrow (\pm 1, 0)$

or

$y+3x^2-1=0,\ 2y+x^2-1=0 \Rightarrow y = 1-3x^2$

$2(1-3x^2)+x^2-1=0$

$1-5x^2=0$

$x^2 = 1/5 \quad x = \pm\sqrt{1/5},\ y = 1-3/5 = 2/5$

$(\pm\sqrt{1/5}, 2/5)$

Candidates for max/min:

$(0,0)$, $(0,1)$, $(\pm 1, 0)$, $(\pm\sqrt{1/5}, 2/5)$

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

$$f(\sqrt{1/5}, 2/5) = \frac{4}{25}\cdot\sqrt{\frac{1}{5}} + \frac{2}{25}\sqrt{\frac{1}{5}} - \frac{2}{5}\sqrt{\frac{1}{5}} = -\frac{4}{25}\cdot\sqrt{\frac{1}{5}}$$

$$f\left(-\sqrt{\frac{1}{5}}, \frac{2}{5}\right) = -\frac{4}{25}\sqrt{\frac{1}{5}} - \frac{2}{25}\sqrt{\frac{1}{5}} + \frac{2}{5}\sqrt{\frac{1}{5}} = \frac{4}{25}\sqrt{\frac{1}{5}}$$

The function values means that

$(-\sqrt{1/5}, 2/5)$ candidates for **max**

$(\sqrt{1/5}, 2/5)$ — " — **min**

$$H(f) = \begin{pmatrix} 6xy & 2y+3x^2-1 \\ 2y+3x^2-1 & 2x \end{pmatrix}$$

$D_1 = 6xy$ can be both pos. and neg.

$\Downarrow$

f not concave, not convex

$$H(f)(\sqrt{1/5}, 2/5) = \begin{pmatrix} 12/5\cdot\sqrt{1/5} & 2/5 \\ 2/5 & 2\sqrt{1/5} \end{pmatrix}$$

$D_1 = \frac{12}{5}\cdot\sqrt{\frac{1}{5}} > 0$

$D_2 = \frac{24}{25} - \frac{4}{25} = \frac{20}{25} > 0$

pos. defn. $\Downarrow$ **local min**

$$H(f)(-\sqrt{1/5}, 2/5) = \begin{pmatrix} -12/5\sqrt{1/5} & 2/5 \\ 2/5 & -\sqrt{1/5} \end{pmatrix}$$

$D_1 = -\frac{12}{5}\sqrt{\frac{1}{5}} < 0$

$D_2 = \frac{24}{25} - \frac{4}{25} = \frac{20}{25} > 0$

neg. defn. $\Downarrow$ **local max**

Conclusion:

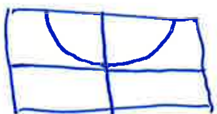
$(\sqrt{1/5}, 2/5)$ is local min, $f = -\frac{4}{25}\sqrt{\frac{1}{5}}$

$(-\sqrt{\frac{1}{5}}, 2/5)$ is local max, $f = \frac{4}{25}\sqrt{\frac{1}{5}}$

but difficult to determine if these points are global max/min.

$x=1, y=a$: $f(x,y) = f(1,a) = a^2 + a - a = a^2$

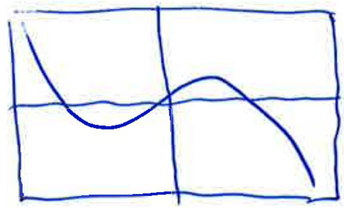
$$\Longrightarrow \lim_{\substack{x=1 \\ y\to\infty}} f(x,y) = \lim_{a\to\infty} a^2 = \infty$$



one slice of graph

there is no max

$x = a,\ y=-1$: $\quad f(x,y) = f(a,-1) = a + a^3\cdot(-1) - (-a) = 2a - a^3$



another slice of graph

$\lim_{\substack{x\to\infty \\ y=-1}} f(x,y) = \lim_{a\to\infty} 2a - a^3 = -\infty$

there is no min

In other words,

$f(1,1000) = 1000^2 = 1000000 > \frac{4}{25}\sqrt{\frac{1}{5}}$

$f(1000,-1) = 2\cdot 1000 - 1000^3$

$= -999.998.000 < -\frac{4}{25}\sqrt{\frac{1}{5}}$

Conclusion: When you have found local min/max that is best candidate for global min/max, there are two possibilities

i) it is global min/max

ii) there are points with even smaller/bigger values, and there is no min/max