LECTURE 6 (B)

GRA 6035

EIVIND ERIKSEN

SEP 25, 2014

MATHEMATICS

Plan:

1. Unconstrained optimization problems
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 $\underline{x}$ is any point in $\mathbb{R}^n$

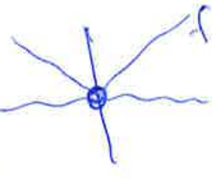
(that is, $\underline{x} = (x_1, \ldots, x_n)$ is an n-tuple of numbers)

We assume that f is "nice" (polynomials, rational functions, exp. / logarithms)

C^1: all partial derivatives exist and are continuous

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

Ex: $f(x) = |x| = \begin{cases} x, & x \geq 0 \\ -x, & x < 0 \end{cases}$



$f(x) = |x|$

example of a function that is not C^1

max/min:

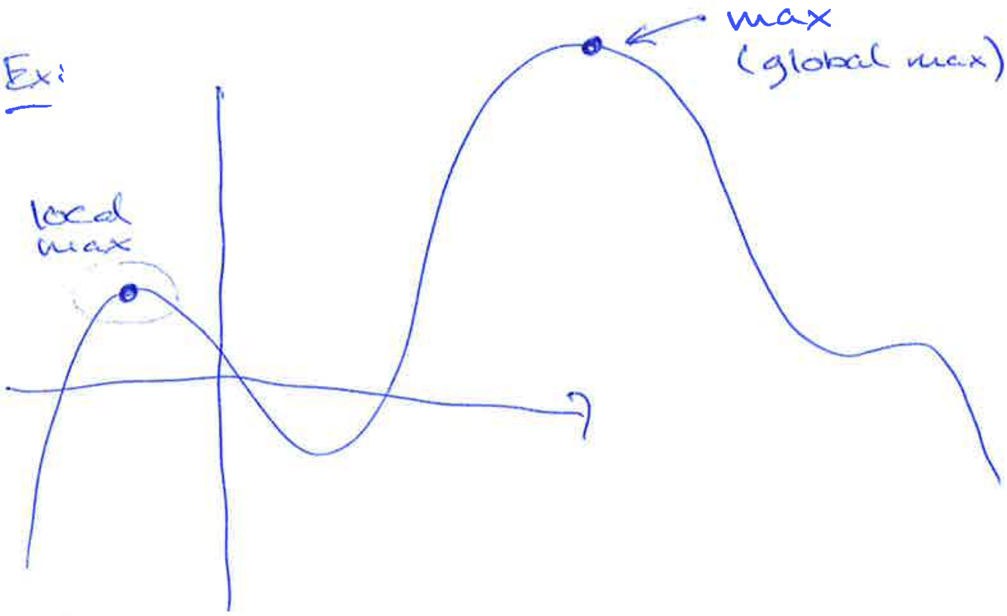
$\underline{x}^*$ is a <u>max</u> for f if $f(\underline{x}^*) \geq f(\underline{x})$ for all pts. $\underline{x}$

— " — <u>min</u> — " — $f(\underline{x}^*) \leq f(\underline{x})$ f— " —

global max = max

global min = min

Ex:



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

$\underline{x}^*$ — " — <u>local min</u> — " — $f(\underline{x}^*) \leq f(\underline{x})$ for all $\underline{x}$ close to $\underline{x}^*$

(2) Partial derivatives

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

$$\left.\begin{aligned} f'_x &= \frac{\partial f}{\partial x} = \underline{3x^2+y} \\ f'_y &= \underline{x - 3y^2} \end{aligned}\right\} \text{ first order partial derivatives}$$

$$\left.\begin{aligned} f''_{xx} &= \underline{6x} \\ f''_{xy} &= \underline{1} \\ f''_{yx} &= \underline{1} \\ f''_{yy} &= \underline{-6y} \end{aligned}\right\} \text{ second order partial derivatives}$$

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 $H(f)$ is a symmetric matrix. In other words,

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} \quad \text{or} \quad f''_{x_i x_j} = f''_{x_j x_i}$$

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

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

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

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

$f''_{x_1x_1} = 2 \quad f''_{x_1x_2} = -6 \quad f''_{x_1x_3} = 0$

$f''_{x_2x_1} = -6 \quad f''_{x_2x_2} = 4 \quad f''_{x_2x_3} = 10$

$f''_{x_3x_1} = 0 \quad f''_{x_3x_2} = 10 \quad f''_{x_3x_3} = 8$

Symm. matrix A of the quadratic form

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

For a quadratic form, $H(f)$ is a constant matrix and equal to $2A$.

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

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

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

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

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

$f''_{yy} = \frac{2+2x^2-2y^2}{u^2}$

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

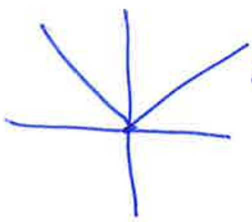
3 First order conditions

Defn: A stationary pt for f is a point where

$$f'_{x_1} = f'_{x_2} = \cdots = f'_{x_n} = 0$$

(A critical pt for f is a point where either i) $f'_{x_1} = \cdots = f'_{x_n} = 0$ or ii) at least one of the partial derivatives does not exist.)

Ex:



$f(x) = |x|$

* no stationary pts
* critical pt at $x = 0$

Fact:

$$\boxed{x^* \text{ max/min for } f \Longrightarrow x^* \text{ stationary pt for } f}$$

(If f is not C^1, we must replace stationary pt with critical pt)

When we find the stationary pts for f, we find candidates for max/min.

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

$f'_x = 3x^2 + 9y = 0$

$f'_y = 9x - 3y^2 = 0$

} FOC (first order conditions = conditions for stationary pts)

$9y = -3x^2$

$y = -\frac{3x^2}{9} = -x^2/3$

$9x - 3 \cdot (-x^2/3)^2 = 0$

$9x - \frac{1}{3}x^4 = 0 \quad | \cdot 3$

$27x - x^4 = 0$

$x(27 - x^3) = 0$

$\underline{x = 0}$ or $x^3 = 27$

$\Downarrow$ $\quad x = \sqrt[3]{27} = \underline{3}$

$y = 0$ $\quad \Downarrow$

$y = -3$

Stationary pts = solution of FOC's

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

and

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

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

$f(3,-3) = 27 - 81 + 27 = \underline{-27}$

Candidate for max: $(x,y) = (0,0)$ $\quad f = 0$

— " — min: $(x,y) = (3,-3)$ $\quad f = -27$

④ Second order conditions

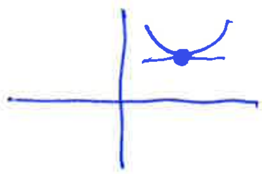
Fact:

If $\underline{x}^*$ is stationary pt. for f, then

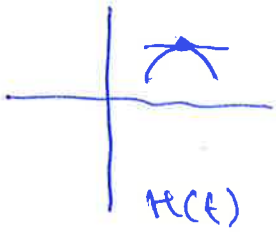
$H(f)(\underline{x}^*)$ positive definite	$\Rightarrow$ $\underline{x}^*$	local min
$H(f)(\underline{x}^*)$ negative definite	$\Rightarrow$ $\underline{x}^*$	local max
$H(f)(\underline{x}^*)$ indefinite	$\Rightarrow$ $\underline{x}^*$	saddle pt

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

Explanation:



$H(f)$ positive definite

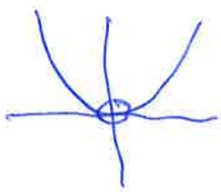


$H(f)$ negative definite

saddle pt = stationary pt that isn't local max/min

Ex:

$f(x) = x^4$



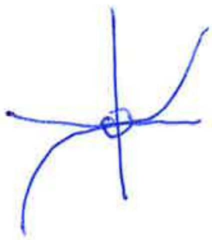
$f' = 4x^3 = 0$

$x = 0$

$f'' = 12x^2$

$H(f)(0) = 0$

$f(x) = x^3$



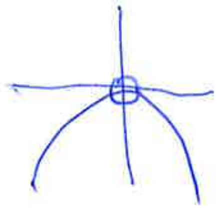
$f' = 3x^2 = 0$

$x = 0$

$f'' = 6x$

$H(f)(0) = 0$

$f(x) = -x^4$



$f' = -4x^3$

$x = 0$

$f'' = -12x^2$

$H(f)(0) = 0$

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

Stationary pts: $(0,0)$, $(3,-3)$

$$H(f)(x,y) = \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{xy} & f''_{yy} \end{pmatrix} = \begin{pmatrix} 6x & 9 \\ 9 & -6y \end{pmatrix}$$

$(x,y)=(0,0)$: Indefinite $\Rightarrow$ saddle pt.

$$H(f)(0,0) = \begin{pmatrix} 0 & 9 \\ 9 & 0 \end{pmatrix} \qquad D_1 = 0, \quad D_2 = -81 < 0$$

$(x,y)=(3,-3)$: Pos. defn. $\Rightarrow$ local min

$$H(f)(3,-3) = \begin{pmatrix} 18 & 9 \\ 9 & 18 \end{pmatrix} \qquad D_1 = 18 > 0, \quad D_2 = 18^2 - 9^2 > 0$$

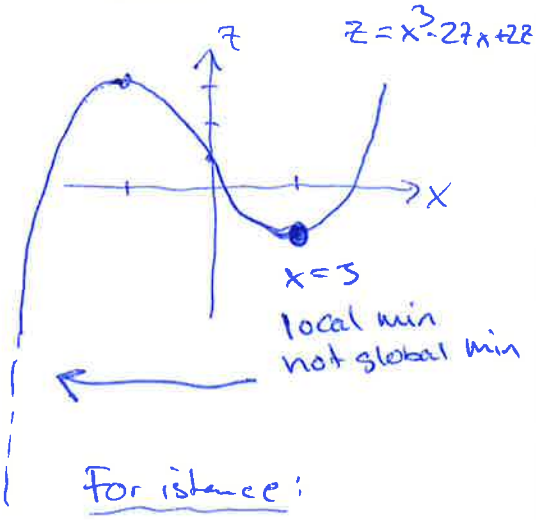
No max for f Local min for f at $(3,-3)$

In fact: $(3,-3)$ is not global min:

Method: "cut" through $y=-3$, where $f(x,-3) = x^3 + 9x(-3) - (-3)^3 = x^3 - 27x + 27$

FOC: $3x^2 - 27 = 0$, $x^2 = 9$, $x = \pm 3$

Hessian: $6x$



For instance:
$f(-10,-3) = -1000 + 270 + 27 = -703$
which is less than
$f(3,-3) = -27$

Comment:

$$\left.\begin{aligned} A &= f''_{xx}(\underline{x}^*) \\ B &= f''_{xy}(\underline{x}^*) \\ C &= f''_{yy}(\underline{x}^*) \end{aligned}\right\} \quad A > 0,\ C > 0,\ AC - B^2 > 0$$

This is the two-variable case of positive definite

$$\left.\begin{aligned} \Delta_1 &= A, C \\ \Delta_2 &= AC - B^2 \end{aligned}\right\} \quad H(f)(x^*) = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

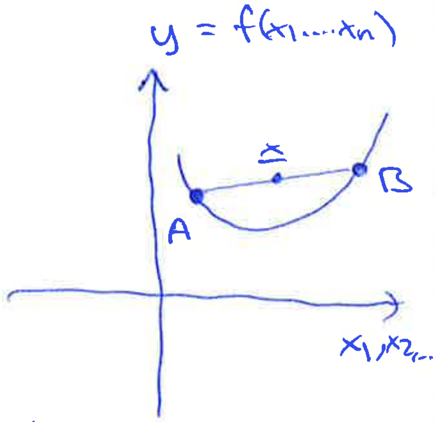
$A > 0$, $C > 0$, $AC - B^2 > 0$
$\Rightarrow$ local min

(5) Convex and concave functions

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

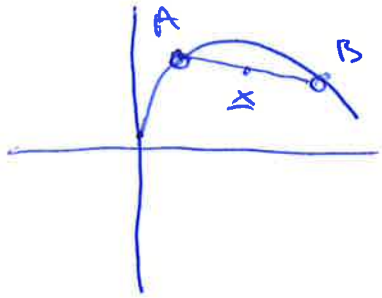
<u>Defn:</u> f is <u>convex</u> if the following condition holds:

For any pts A, B on the graph of f, all pts. $\underline{x}$ on the line segment [A, B] lies on or <u>over</u> the graph.



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

For any pts A, B on the graph of f, all pts on the line segment [A, B] lies on or <u>under</u> the graph



We use the Hessian matrix to determine if a function is convex or concave:

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

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

Ex: $f(x) = x^3 - 9x$

$f'(x) = 3x^2 - 9$

$f''(x) = 6x$ $\qquad\qquad H(f) = (6x)$

$D_1 = 6x$

$x > 0$: $H(f)$ is positive definite

$x < 0$: $H(f)$ is negative definite

$x > 0$: $H(f)$ is pos. defn.

$x < 0$: $H(f)$ is neg. defn.

f is not convex, not concave

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

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

If f is convex, then $H(f)$ is pos. semidef. for all (x,y), and in particular $\begin{cases} D_1 \geq 0 \\ D_2 \geq 0 \end{cases}$ for all x,y. This is not the case.

$\Rightarrow$ f not convex

If f is concave, then $D_1 \leq 0$, $D_2 \geq 0$ for all x,y. $\Rightarrow$ f is not concave.

Fact: If f is convex, then any stationary pt. is a global min.

If f is concave, then any stationary pt. is a global max.

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

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

$D_1 = 2$

$D_2 = 16 - 36 = -20$

indefinite

f not convex, not concave

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

f concave $\iff \Delta_1 \leq 0, \Delta_2 \geq 0, \Delta_3 \leq 0, \dots, (-1)^n \cdot \Delta_n \geq 0$ for all principal minors and all $(x_1, \dots, x_n)$.

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

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

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

$$H(f) = \begin{pmatrix} e^{x+y} & e^{x+y} \\ e^{x+y} & e^{x+y} \end{pmatrix}$$

$D_1 = e^{x+y} > 0$ for all x,y

$D_2 = (e^{x+y})^2 - (e^{x+y})^2 = 0$ for all x,y

$\Delta_1 = e^{x+y}, e^{x+y} > 0$ for all x,y

$\Delta_2 = 0 \geq 0$ for all x,y

f is convex

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, ~~...~~ 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:

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

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

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

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

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

$\Big\}$ pos. defn. $\Downarrow$ local min

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

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

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

$\Big\}$ neg. defn. $\Downarrow$ local max

Conclusion:

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

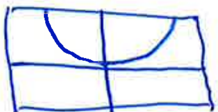
$(-\sqrt{\frac{1}{5}}, 2/5)$ is local max, $f = \frac{4}{25}\sqrt{\frac{1}{5}}$

$\Big\}$ but difficult to determine if these points are global max/min.

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

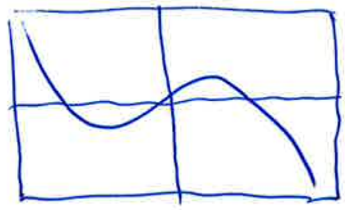
$$\Longrightarrow \lim_{\substack{x=1 \\ y\to\infty}} f(x,y) = \lim_{a\to\infty} a^2 = \infty$$

there is no max



one slice of graph

$\underset{\sim\sim\sim}{x=a,\ y=-1:}$ $f(x,y) = f(a,-1) = a + a^3\cdot(-1) - (-a) = 2a - a^3$



another slice of graph

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

<u>there is no min</u>

<u>In other words,</u> $f(1,1000) = 1000^2 = 1000000 > \frac{4}{25}\sqrt{\frac{1}{5}}$

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

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

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

i) it <u>is</u> global min/max

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