LECTURE 7 — GRA 6035

EIVIND ERIKSEN, OCT 3RD 2011

REVIEW + EXTRA MATERIAL: CONVEX/CONCAVE FUNCTIONS

Assume that $f(\underline{x}) = f(x_1, x_2, \ldots, x_n)$ is defined on a convex set $S = D_f$ in $\mathbb{R}^n$.

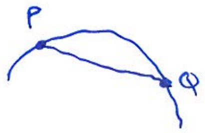
S must be convex set, otherwise defn. of convex/concave functions will not make sense

$\mathbb{R}^n$ = all points $\underline{x} = (x_1, \ldots, x_n)$ in n-dim. space

DEFINITION:

f convex on S $\iff$ for any points P, Q on the graph of f the line segment [P, Q] is never under the graph

f concave on S $\iff$ for any points P, Q on the graph of f, the line segment [P, Q] is never over the graph



Ex: f is convex

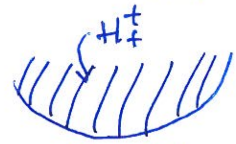
$f = \underbrace{x^2 + y^2}_{\text{convex}} + \underbrace{x + y}_{\text{convex}}$

Properties:

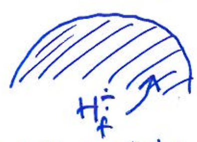
(a) f convex $\iff$ $-f$ concave

(b) f, g convex, $a, b \geq 0$ $\Rightarrow$ $af + bg$ convex
f, g concave, $a, b \geq 0$ $\Rightarrow$ $af + bg$ concave

(c) f convex function $\iff$ H_f^+ is a convex set
f concave function $\iff$ H_f^- is a convex set



= all point over or on the graph of f



= all points under or on the graph of f

S = 

Useful result:

Let f be a function defined on a convex set S, and assume that f is continuous on S and C^2 on the interior of S. Then we have:

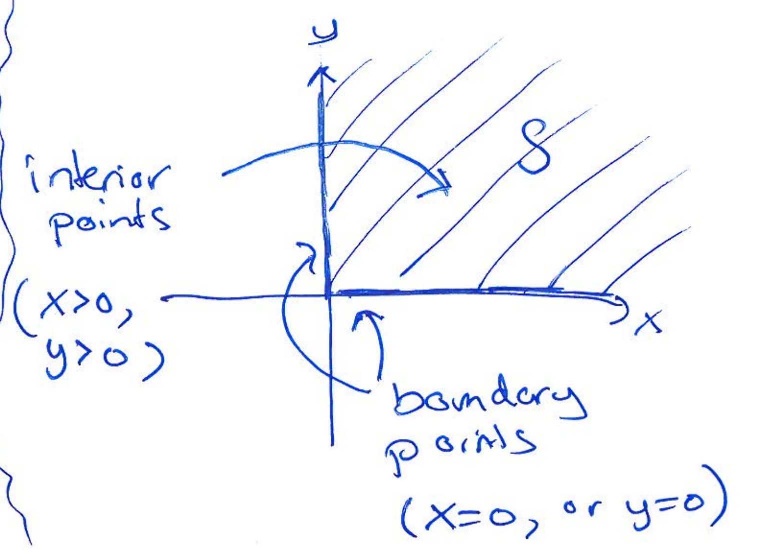
(a) f is convex on S $\iff$ $H(f)(\underline{x})$ is positive semidefinite for all interior pts $\underline{x}$ in S

(b) f is concave on S $\iff$ $H(f)(\underline{x})$ is negative semi-definite for all interior points $\underline{x}$ in S

interior point of S = point in S that is not on the boundary

Ex: $f(x,y) = xy$ defined on $S = \{(x,y): x \geq 0, y \geq 0\}$

Is f convex or concave?



$f'_x = y$

$f'_y = x$

$$H(f) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$\Delta_1 = 0, 0$

$\Delta_2 = -1$

<u>Condition for convexity:</u> $\Delta_1 \geq 0, \Delta_2 \geq 0$ for all points $x>0, y>0$

<u>not</u> convex

<u>— || — concavity:</u> $\Delta_1 \leq 0$, $\Delta_2 \geq 0$ for all points $x>0, y>0$

<u>not</u> concave

Ex: $f(x,y) = xyz$ defined on $S = \{(x,y,z): x \geq 0, y \geq 0, z \geq 0\}$

$$\left.\begin{aligned} f'_x &= yz \\ f'_y &= xz \\ f'_z &= xy \end{aligned}\right\} \qquad H(f) = \begin{vmatrix} 0 & z & y \\ z & 0 & x \\ y & x & 0 \end{vmatrix}$$

$\Delta_1 = 0, 0, 0$

$\Delta_2 = -z^2, -x^2, -y^2$

$\Delta_3 = -z(-xy) + y(zx)$

$\quad = 2xyz$

<u>Convex:</u> $\Delta_1 \geq 0$, $\Delta_2 \geq 0$, $\Delta_3 \geq 0$ for all $x>0, y>0, z>0$

$-z^2 < 0 \Rightarrow$ not convex

<u>Concave:</u> $\Delta_1 \leq 0$, $\Delta_2 \geq 0$, $\Delta_3 \cancel{\geq} 0$ — || —

$-z^2 < 0 \Rightarrow$ not concave

Plan: Lecture 7

① Extremal points
② Lagrange problems

} [FMEA] 3.1–3.3 (except envelope thm.)

Keep in mind: We want to solve the following problems:

Ⓐ max/min $f(x_1,\dots,x_n)$ — Unconstrained optimization ($S = D_f = \mathbb{R}^n$)

Ⓑ max/min $f(x_1,\dots,x_n)$ subject to { one or more equality/ineq. constraints — Constrained optimization (S is the subset of $\mathbb{R}^n$ of points that satisfy the constraints)

---

① Extremal points:

We assume that all functions are C^2.

Definition: $f(x_1,\dots,x_n)$ a function defined on S

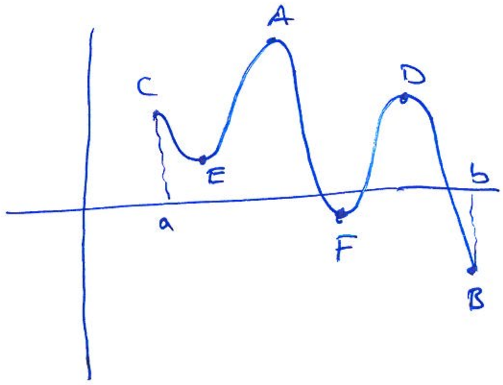
$\underline{x}^*$ is a global maximum for f if
$\underline{x}^*$ in S and $f(\underline{x}^*) \geq f(\underline{x})$ for all $\underline{x}$ in S

$\underline{x}^*$ is a global minimum for f if
$\underline{x}^*$ in S and $f(\underline{x}^*) \leq f(\underline{x})$ for all $\underline{x}$ in S

$\underline{x}^*$ is a local maximum for f if
$\underline{x}^*$ in S and $f(\underline{x}^*) \geq f(\underline{x})$ for all $\underline{x}$ in S close to $\underline{x}^*$

$\underline{x}^*$ is a local minimum for f if
$\underline{x}^*$ in S and $f(\underline{x}^*) \leq f(\underline{x})$ for all $\underline{x}$ in S close to $\underline{x}^*$

Ex: f defined on $[a,b]$
with graph



A : global max
B : global min

C, A, D : local max
E, F, B : local min

Definition: $f(x_1, ..., x_n)$ defined on S

A stationary point $\underline{x}^*$ is a point in the interior of S such that

$$f_1'(\underline{x}^*) = f_2'(\underline{x}^*) = \dots = f_n'(\underline{x}^*) = 0$$

Ex: $f(x,y) = 2x - y - x^2 + xy - y^2$ defined on $S = \mathbb{R}^2$

$f_x' = 2 - 2x + y = 0$ $\quad$ $2x - y = 2$
$f_y' = -1 + x - 2y = 0$ $\quad$ $-x + 2y = -1$ $\Big]\,2\cdot$ $\quad$ $3y = 0 \Rightarrow y = 0$, $x = 1$

Stationary pts: $(1,0)$

$(1,0)$ is an interior point of $S = \mathbb{R}^2$

Thm: $f(x_1, ..., x_n)$ defined on S

If $\underline{x}^*$ is a (local or global) extremal point in the interior of S, then $\underline{x}^*$ is a stationary point.

(local or global) extremal point = (local or global) min or max

Conclusion:

A local or global extremal point of f is either

i) stationary point (in the interior)
ii) boundary point

Ex: $f(x,y) = 2x - y - x^2 + xy - y^2$ defn. on $S = \mathbb{R}^2$

- Boundary pts: none
- ~~Stationary~~ pts: $\underline{x} = (1,0)$

Convex/concave functions: $f(\underline{x})$ defined on S

$\underline{x}^*$ $\left\{ \begin{array}{l} \text{interior} \\ \text{stationary} \\ \text{point} \end{array} \right.$

~~Ex~~

f convex $\Rightarrow$ $\underline{x}^*$ global min

f concave $\Rightarrow$ $\underline{x}^*$ global max

Ex: $\left. \begin{array}{l} f'_x = 2 - 2x + y \\ f'_y = -1 + x - 2y \end{array} \right\}$ $H(x) = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$

$\left. \begin{array}{l} D_1 = -2 < 0 \\ D_2 = 3 > 0 \end{array} \right\}$ negative definite for all $\underline{x}$

$\Downarrow$

f is concave

Conclusion: $\underline{x}^* = (1,0)$ is a global maximum

Classification of local extremal points

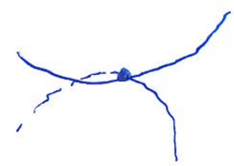
Thm: f defined on S, $\underline{x}^*$ interior stationary point

$$\left\{\begin{array}{lll} H(f)(\underline{x}^*) & \text{pos. definite} & \Longrightarrow \underline{x}^* \text{ local minimum} \\ H(f)(\underline{x}^*) & \text{neg. definite} & \Longrightarrow \underline{x}^* \text{ local maximum} \\ H(f)(\underline{x}^*) & \text{indefinite} & \Longrightarrow \underline{x}^* \text{ is neither local min. nor local max.} \\ & & \quad \text{(saddle point)} \end{array}\right.$$

second derivative test

Remark:

* A <u>saddle point</u> is a stationary point that is neither local min nor local max.



* If $H(f)(\underline{x}^*)$ is ① pos. semidefinite but not positive definite

  or

  ② neg. semidefinite but not negative definite

then the test is <u>inconclusive</u>.

Ex: $f = x^2y - xy^2 - x + y$

$$f'_x = 2xy - y^2 - 1 = 0 \qquad 2xy = y^2 + 1$$

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

$$x^2 - y^2 = 0$$

$$(x-y)(x+y) = 0$$

$$y = x \text{ or } y = -x$$

$\underline{y=x}$: $\left.\begin{array}{l} 2x^2 = x^2 + 1 \\ x^2 = 1 \quad x = \pm 1 \end{array}\right\}$ $(1,1)$, $(-1,-1)$

$\underline{y=-x}$: $\left.\begin{array}{l} -2x^2 = x^2 + 1 \\ -3x^2 = 1 \quad x^2 = -1/3 \end{array}\right\}$ no soln.

Conclusión: f has two stationary pts: $(1,1)$, $(-1,-1)$.

Classification:

$$\left.\begin{aligned} f'_x &= 2xy - y^2 - 1 \\ f'_y &= x^2 - 2xy + 1 \end{aligned}\right\} \Rightarrow H(f) = \begin{pmatrix} 2y & 2x-2y \\ 2x-2y & -2x \end{pmatrix}$$

In $(1,1)$: $H(f)(1,1) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$ $\left.\begin{aligned} D_1 &= 2 \\ D_2 &= -4 \end{aligned}\right\}$ indefinite

$\Rightarrow$ $(1,1)$ is saddle point

In $(-1,-1)$: $H(f)(-1,-1) = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$ $\left.\begin{aligned} D_1 &= -2 \\ D_2 &= -4 \end{aligned}\right\}$ indefinite

$\Rightarrow$ $(-1,-1)$ is saddle point

Ex: $f(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3$ defined on $\mathbb{R}^3$

$$\left.\begin{aligned} f'_{x_1} &= 3x_1^2 = 0 \\ f'_{x_2} &= 3x_2^2 = 0 \\ f'_{x_3} &= 3x_3^2 = 0 \end{aligned}\right\}$$

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

$$H(f) = \begin{pmatrix} 6x_1 & 0 & 0 \\ 0 & 6x_2 & 0 \\ 0 & 0 & 6x_3 \end{pmatrix}$$

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

positive semidefinite
negative semidefinite

$\Rightarrow$ test inconclusive

$$\left.\begin{aligned} f(a,0,0) = a^3 &> 0 \text{ if } a>0 \\ a^3 &< 0 \text{ if } a<0 \end{aligned}\right\} \Rightarrow (0,0,0) \text{ saddle point.}$$

② Lagrange problems

Std. form of a Lagrange problem:

$$\max/\min\ \underset{f(\underline{x})}{\overset{\parallel}{f(x_1,\dots,x_n)}} \quad \text{subject to} \quad \begin{cases} g_1(x_1,\dots,x_n) = b_1 \\ g_2(x_1,\dots,x_n) = b_2 \\ \vdots \\ g_m(x_1,\dots,x_n) = b_m \end{cases}$$

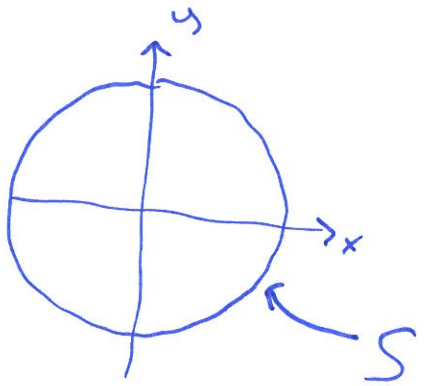
<u>Alterative formulation:</u>

Constrained optimization problem for the function $f(\underline{x})$ defined on the set

$$S = \{ \underline{x} \text{ in } \mathbb{R}^n : g_1(\underline{x}) = b_1,\ g_2(\underline{x}) = b_2, \dots, g_m(\underline{x}) = b_m \}$$

The points in S are called <u>admissible points</u>.

<u>Ex</u>: max $f(x,y) = x + 3y$ subject to $g(x,y) = x^2 + y^2 = 10$.



<u>Admissible points</u>: $S = \{ (x,y) : x^2 + y^2 = 10 \}$
the points on the circle

<u>Extreme Value Thm:</u>
f defined on S with S closed and bounded $\Rightarrow$ f has global min/max.

<u>Conclusion:</u>
there is a max and a min for $f(x,y) = x + 3y$ on the circle.
(S is closed + bounded).

How to solve Lagrange-problems:

Problem:

$$\max/\min\ f(x_1,\dots,x_n) \quad \text{subject to} \quad \begin{cases} g_1(x_1,\dots x_n) = b_1 \\ \vdots \\ g_m(x_1,\dots,x_n) = b_m \end{cases}$$

i) Form the Lagrangian function

$$\mathcal{L}(x_1,\dots,x_n;\underbrace{\lambda_1,\dots,\lambda_m}_{\substack{\text{Lagrange} \\ \text{multipliers}}}) = f(x_1,\dots,x_n) - \lambda_1 \cdot g_1(x_1,\dots,x_n) - \lambda_2 \cdot g_2(x_1,\dots,x_n) - \dots - \lambda_m \cdot g_m(x_1,\dots,x_n)$$

ii) Solve the Lagrange equations:

$$\left.\begin{array}{l} \left.\begin{array}{l} \mathcal{L}'_{x_1} = 0 \\ \mathcal{L}'_{x_2} = 0 \\ \vdots \\ \mathcal{L}'_{x_n} = 0 \end{array}\right\} \text{first order equations (n equations)} \\ \\ \left.\begin{array}{l} g_1(x_1,\dots x_n) = b_1 \\ \vdots \\ g_m(x_1,\dots,x_n) = b_m \end{array}\right\} \text{constraints (m equations)} \end{array}\right\} \begin{array}{c}(m+n) \text{ equations} \\ \text{in} \\ (m+n) \text{ variables}\end{array}$$

Ex: max/min $f(x,y) = x+3y$ subj. to $g(x,y) = x^2+y^2 = 10$

$$\mathcal{L}(x,y;\lambda) = x+3y-\lambda\cdot(x^2+y^2)$$

$$\boxed{\begin{aligned} \mathcal{L}'_x &= 1-\lambda\cdot 2x = 0 \\ \mathcal{L}'_y &= 3-\lambda\cdot 2y = 0 \\ & x^2+y^2 = 10 \end{aligned}}$$

3 eqn. in 3 var's (not linear)

$x = \frac{1}{2\lambda}$ ($\lambda \neq 0$ since $\lambda = 0$ does not give solutions)

$y = \frac{3}{2\lambda}$ — " —

$$x^2+y^2 = \left(\frac{1}{2\lambda}\right)^2 + \left(\frac{3}{2\lambda}\right)^2 = \frac{1+9}{(2\lambda)^2} = 10$$

$$\frac{10}{(2\lambda)^2} = 10 \Rightarrow (2\lambda)^2 = 1 \Rightarrow 2\lambda = \pm 1$$
$$\Rightarrow \underline{\lambda = \pm 1/2}$$

Solutions:

$\lambda = \frac{1}{2}$ gives $x=1, y=3$ $\Rightarrow (x,y;\lambda) = \underline{(1,3;1/2)}$

$\lambda = -\frac{1}{2}$ gives $x=-1, y=-3$ $\quad (x,y;\lambda) = \underline{(-1,-3;-1/2)}$

How can we use these solutions to solve the max/min-problem?

Thm:

If $(\underline{x}^*;\underline{\lambda}^*)$ is a solution of the Lagrange equations, then we may fix the specific $\underline{\lambda}^*$ and consider

$$\mathcal{L}(x_1,\dots,x_n;\lambda_1^*,\dots,\lambda_m^*) \text{ as a function of } x_1,x_2,\dots,x_n$$

Then we have:

$\mathcal{L}(\underline{x};\underline{\lambda}^*)$ convex function in $\underline{x}$ $\Rightarrow$ $\underline{x}^*$ solves the min-problem

$\mathcal{L}(\underline{x};\underline{\lambda}^*)$ concave function in $\underline{x}$ $\Rightarrow$ $\underline{x}^*$ solves the max-problem

Ex: We look at the solutions $(x,y;\lambda) = (1,3;1/2), (-1,-3;-1/2)$.

For $(x^*,y^*;\lambda^*) = (1,3;1/2)$: We fix $\lambda^* = \frac{1}{2}$ and let x,y vary

$\mathcal{L}(x,y) = x + 3y - \frac{1}{2}(x^2+y^2)$ is concave ; $\mathcal{L}'' = \begin{pmatrix} -1/2 & 0 \\ 0 & -1/2 \end{pmatrix}$ Hessian

$\Downarrow$

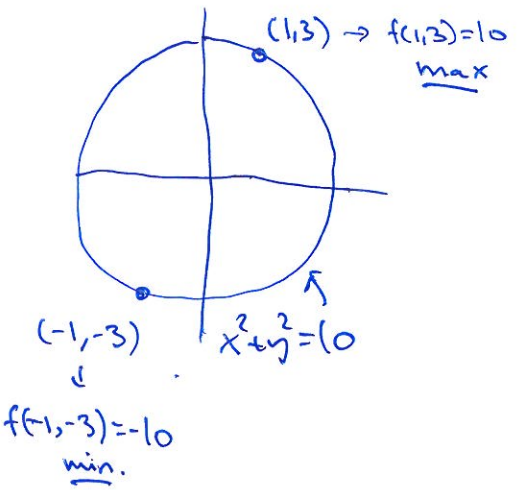
$\underline{x}^* = (x^*,y^*) = (1,3)$ solves the max-problem ; $f(1,3) = \underline{10}$ max value

For $(x^*,y^*;\lambda^*) = (-1,-3;-1/2)$: We fix $\lambda^* = -1/2$ and let x,y vary

$$\mathcal{L}(x,y) = x + 3y - (-1/2)(x^2+y^2)$$
$$= x + 3y + \frac{1}{2}(x^2+y^2)$$ is convex ; $\mathcal{L}'' = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$ Hessian

$\Downarrow$

$\underline{x}^* = (x^*,y^*) = (-1,-3)$ solves the min-problem ; $f(-1,-3) = \underline{-10}$ min value



Conclusion:

max $x+3y$ subj. to $x^2+y^2=10$ is 10 at (1,3)
min $x+3y$ subj. to $x^2+y^2=10$ is −10 at (−1,−3)

What if the functions are not convex or concave?

Consider the NDCQ (non-degenerate constraint qualification) for a point $\underline{x}^*$ that is admissible:

NDCQ:
$$\text{rk} \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\underline{x}^*) & \frac{\partial g_1}{\partial x_2}(\underline{x}^*) & \cdots & \frac{\partial g_1}{\partial x_n}(\underline{x}^*) \\ \vdots & & & \\ \frac{\partial g_m}{\partial x_1}(\underline{x}^*) & \frac{\partial g_m}{\partial x_2}(\underline{x}^*) & \cdots & \frac{\partial g_m}{\partial x_n}(\underline{x}^*) \end{pmatrix} = m$$

Ex: max/min $x+3y$ subj. to $x^2+y^2=10$

NDCQ: $\text{rk}\begin{pmatrix}2x & 2y\end{pmatrix} = 1$

Is NDCQ satisfied for all admissible points?

(x,y) admissible $\Leftrightarrow x^2+y^2=10$

$\text{rk}\begin{pmatrix}2x & 2y\end{pmatrix} = \begin{cases} 1, & x\neq 0 \text{ or } y\neq 0 \\ 0, & x=0 \text{ and } y=0 \end{cases}$

Since $(x,y)=(0,0)$ is not admissible $(0^2+0^2\neq 10)$, $\text{rk}\begin{pmatrix}2x & 2y\end{pmatrix}=1$ for all admissible points.

Thm:

If $\underline{x}^*$ solves the min/max Lagrange problem, then one of the following conditions hold:

i) $(\underline{x}^*;\underline{\lambda})$ solves the Lagrange equations for some $\underline{\lambda}=\lambda_1,\dots,\lambda_m$. (and $\underline{\lambda}$ is unique)

ii) $\underline{x}^*$ does not satisfy NDCQ

Conclusion:

We can make a list of possible candidates for max/min in a Lagrange problem, where we include both solutions of the Lagrange equations and admissible points that do not satisfy NDCQ.

Ex: In the example, if we choose not to use the method with convexity/concavity, we would:

(a) Make a list of candidate points: $(1,3)$, $(-1,-3)$

* $(1,3)$, $(-1,-3)$ from Lagrange equations
* no extra candidates since all admissible pts. satisfy NDCQ

(b) Argue that there must be max and min because of Extreme value thm.

(c) Compute $f(1,3)=10$ and $f(-1,-3)=-10$, and conclude that $(1,3)$ is max. and $(-1,-3)$ is min.

I did not have time to go through the interpretation of the Lagrange multiplier. This will be done next week.