Plan:

Review : Lecture 6

1. Convex and concave functions
2. Constrained optimization and admissible points
3. Lagrange problems
4. Kuhn-Tucker problems

} relevant for midterm exam

Note:
1. Plenary Session II on Mond. 17-20.
2. Midterm exam: Fri Oct 12 15-16
3. Extended office hours: (B4-032) Mond 10-17, Wed 10-17, Thu 10-17

Review: Lecture 6

Unconstrained optimization : max/min $f(x_1, \dots, x_n) = f(\underline{x})$

Stationary pts of f: $\boxed{f'_{x_1} = f'_{x_2} = \dots = f'_{x_n} = 0}$ FOC

Classification:

$$H(f)(\underline{x}) = \begin{pmatrix} f''_{x_1x_1} & f''_{x_1x_2} & \dots \\ \vdots & & \end{pmatrix}$$

Hessian of f

n×n symm. matrix

Local classification: Second derivative test

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

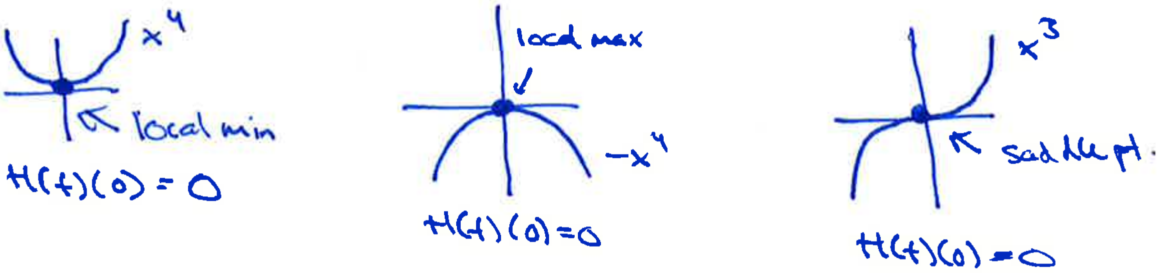
$H(f)(\underline{x}^*)$ pos. defn. $\Rightarrow$ $\underline{x}^*$ is local min

— " — neg. defn. $\Rightarrow$ " is local max

— " — indefn. $\Rightarrow$ " is saddle pt.

If $H(f)(\underline{x}^*)$ is pos./neg. semi-defn (but not defn) then there is no conclusion

Ex:



Classification (globally): the ultimate goal!

- convex/concave optimization
- ad hoc methods

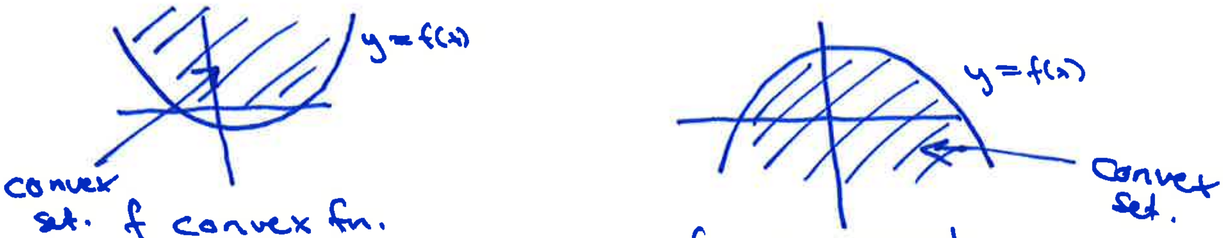
① Convex and concave functions

Convex/concave optimization:

If $\underline{x}^*$ is a stationary pt. for f and f convex $\Rightarrow$ $\underline{x}^*$ is global min

— " — f concave $\underline{x}^*$ is global max

Ex: $f(x)$ fn. in one variable



f convex fn. $f''(x) \geq 0$ for all x

f concave fn. $f''(x) \leq 0$ for all x

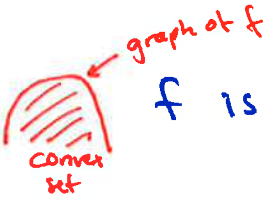
Defn: A set D is convex if the following condition holds:

P, Q points in D $\Rightarrow$ the line segment $[P, Q]$ is contained in D.

D convex $\Updownarrow$ "no holes" in D

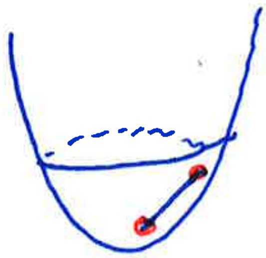
Defn: $f = f(x_1, \ldots, x_n)$ fn. in n variables.

f is <u>convex fn.</u> if the set of points lying <u>over</u> the graph of f is a <u>convex set</u>

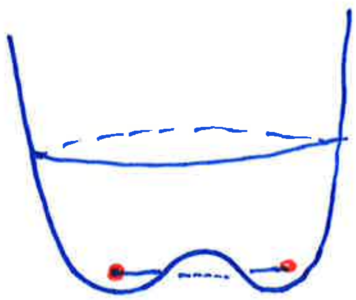


f is <u>concave fn.</u> if the set of points lying <u>under</u> the graph of f is a <u>convex set</u>

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



f convex



f not <u>convex</u>

Computational method:

f convex fn. $\iff$ $H(f)(\underline{x})$ is pos. semidefn. for all pts $\underline{x}$

f concave fn. $\iff$ " neg. semidefn. — " —

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

$f'_x = 4x + 2y$

$f'_y = 2x + 12y$

$$H(f)(x,y) = \begin{pmatrix} 4 & 2 \\ 2 & 12 \end{pmatrix} \qquad D_1 = 4 > 0, \quad D_2 = 44 > 0$$

pos. defn. for all (x,y)

$\Downarrow$

pos. semidefn — " —

$(0,0)$ global min $\Leftarrow$ f <u>convex</u>

Remember: A is positive semidefn. if $\underline{x}^T A \underline{x} \geq 0$ for all $\underline{x}$
" positive defn. if $\underline{x}^T A \underline{x} > 0$ for all $\underline{x} \neq \underline{0}$

$$\boxed{\begin{array}{c} \underline{x}^T A \underline{x} > 0 \\ \Downarrow \\ \underline{x}^T A \underline{x} \geq 0 \end{array}}$$

Therefore, pos. defn. matrices are pos. semidefinite
Similarly, neg. defn. matrices " neg. semidefn.

pos./neg. definite is a special case of pos./neg. semidefinite!

Note: A is both pos. semidefn. and neg. semidefn.

$$\Updownarrow$$

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

Therefore, a polynomial function of degree ≤ 1 is both convex and concave.

Ex: $f = x - y + z$ is convex and concave

$f'_x = 1$
$f'_y = -1$
$f'_z = 1$

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

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

$f'_x = 3x^2 - 3y$

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

$H(f) = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}$ " $H(f)(x,y)$

$D_1 = 6x$

$D_2 = 36xy - 9$

D_1 can be both pos. and neg.

$\Rightarrow$ f is not concave, not concave

One variable:

$f(x) = x^3 - 3x$

$f'' \geq 0$

$f'' = H(f) \leq 0$

concave on $x \leq 0$ | convex on $x \geq 0$

$H(f)(2,2) = \begin{pmatrix} 12 & -3 \\ -3 & 12 \end{pmatrix}$ $D_1 = 12$, $D_2 = 135$ pos. defin.

$H(f)(0,0) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$ $D_1 = 0$, $D_2 = -9$ indefin.

$H(f)(-2,-2) = \begin{pmatrix} -12 & -3 \\ -3 & -12 \end{pmatrix}$ $D_1 = -12$, $D_2 = 135$ neg. defin.

Note:

f,g convex	$\Rightarrow$	f+g convex
f,g concave		f+g concave
f convex, $a>0$	$\Rightarrow$	$a \cdot f$ convex
" $a<0$		$a \cdot f$ concave
f concave, $a>0$		$a \cdot f$ concave
" $a<0$		$a \cdot f$ convex

$H(f+g) = H(f) + H(g)$

$H(a \cdot f) = a \cdot H(f)$

$\boxed{f \text{ concave} \Leftrightarrow -f \text{ convex}}$

Ex: $f(x,y,z) = \underbrace{x^4 + y^4 + z^4}_{f_1} + \underbrace{x^2 - xy + y^2 + yz + z^2}_{f_2}$

$A = \begin{pmatrix} 1 & -1/2 & 0 \\ -1/2 & 1 & 1/2 \\ 0 & 1/2 & 1 \end{pmatrix}$

$H(f_1) = \begin{pmatrix} 12x^2 & 0 & 0 \\ 0 & 12y^2 & 0 \\ 0 & 0 & 12z^2 \end{pmatrix}$

f_1 convex

$H(f) = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \leftarrow H(f) = 2A$

$D_1 = 1$, $D_2 = 3/4$, $D_3 = 3/4 - \frac{1}{2} \cdot \frac{1}{2} = 1/2$

f_2 convex

$f = f_1 + f_2$ convex

(2) Constrained optimization

max/min $f(x_1, \dots, x_n)$ when $g(x_1, \dots, x_n) = a$
$g(x_1, \dots, x_n) \leq a$

one constraint

Ex: min $\underbrace{f(x,y) = x^2+y^2}_{\text{objective fn.}}$ when $\underbrace{xy = 1}_{\text{constraints}}$

$=$ Equality constraints: Lagrange problem

$\leq, \geq$ Closed inequality constraints: Kuhn-Tucker problem

Admissible points: A point that satisfies all constraints

D = set of admissible points

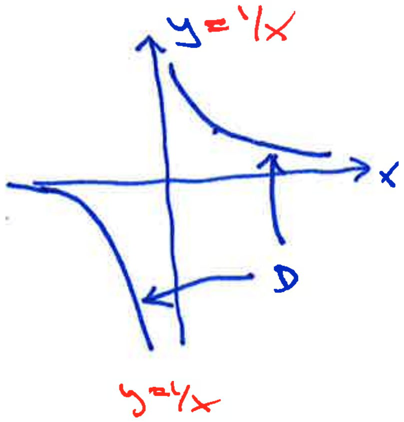
Ex: min $f(x,y) = \underline{x^2+y^2}$ when $xy=1$

Adm. pts: $xy = 1$

$D = \{(x,y) : xy = 1\}$

all pts. (x,y) such that $xy = 1$

$xy = 1 \iff y = 1/x$



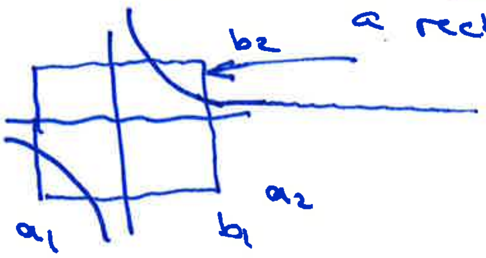
Defn: D is <u>closed</u> when the constraints are given by equality ($=$) or closed inequality ($\leq$ $\geq$)

D is <u>bounded</u> if there are constants $a_1, \ldots, a_n, b_1, \ldots, b_n$ such that

$$\boxed{\begin{array}{c} a_1 \leq x_1 \leq b_1 \\ a_2 \leq x_2 \leq b_2 \\ \vdots \\ a_n \leq x_n \leq b_n \end{array}}$$

for all pts $(x_1, \ldots, x_n)$ in D.

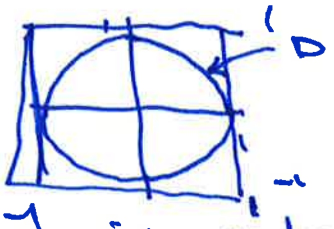
Ex: $D: xy = 1$ is <u>bounded</u> if there is a rectangle containing all pts of D



$D: xy=1$

<u>No</u>, D is <u>not</u> bounded

Ex: $D: x^2+y^2=1$



circle, center $(0,0)$ radius $=1$

$D: x^2+y^2=1$

Yes, D is <u>bounded</u>

<u>Check $x^2+y^2=1$ algebraically:</u>

$$\left.\begin{array}{l} -1 \leq x \leq 1 \\ -1 \leq y \leq 1 \end{array}\right\} \text{Yes, } D \text{ is bounded}$$

Defn: D is <u>compact</u> if it is closed and bounded

> **Thm:** Extreme value theorem (EVT)
> If f is a continuous fn. defined on a compact set D, then f attains a max and a min on D.

③ Lagrange problems

Ex: $\min \underbrace{f(x,y) = x^2+y^2}$ when $\underline{xy=1}$
$g(x,y)=a$

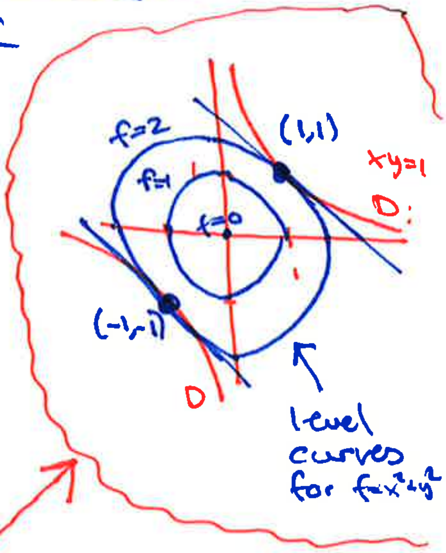
$$L(x,y;\lambda) = f(x,y) - \lambda \cdot g(x,y)$$
$$= x^2+y^2-\lambda \cdot xy$$

x,y: variables
λ: Lagrange multiplier

FOC: $\begin{cases} L'_x = & 2x - \lambda \cdot y = 0 \\ L'_y = & 2y - \lambda \cdot x = 0 \end{cases}$

C: $\{ \quad xy = 1$

Lagrange conditions: FOC + C



$$x = \frac{\lambda y}{2} \Rightarrow 2y - \lambda \cdot \left(\frac{\lambda y}{2}\right) = 0 \quad | \cdot 2$$
$$4y - \lambda^2 y = 0$$
$$y(4-\lambda^2) = 0$$

$\underline{y=0}$ or $4-\lambda^2=0$

$\lambda = 2$ or $\lambda = -2$

~~$y=0$, $x=0$, $0=1$ imp.~~	$\lambda=2$, $x=y$, $x^2=1$, $x=\pm 1$, $(x,y;\lambda) = (1,1;2)$, $(-1,-1;2)$	~~$\lambda=-2$, $x=-y$, $-x^2=1$, $x^2=-1$ imp.~~

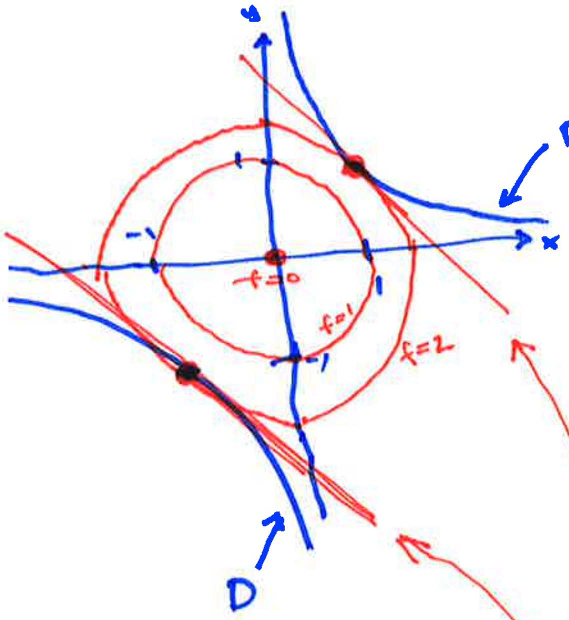
$(x,y;\lambda) =$
$(1,1;2)$ $f=2$
$(-1,-1;2)$ $f=2$

Candidates for min

Explanation of Lagrange conditions $L'_x = L'_y = 0$

Lagrange problem: min $f = x^2 + y^2$ when $xy = 1$

① Adm. pts: $\underbrace{xy}_{g(x,y)} = \underbrace{1}_{a}$



② Values of $f(x,y) = x^2 + y^2$ (objective fn).

Level curves: $f(x,y) = c$

c=0: $f(x,y) = 0$: $x^2 + y^2 = 0$
$(x,y) = (0,0)$

c=1: $f(x,y) = 1$ $x^2 + y^2 = 1$
circle, centr $(0,0)$
rad. 1

no adm. pts with $f=0$ or $f=1$

c=2: $f(x,y) = 2$ $x^2 + y^2 = 2$
circle, centr $(0,0)$
radius $\sqrt{2}$
— two adm. pts. $(1,1), (-1,-1)$ with $f = 2$

Minimum: D: $xy = 1$ and level curve $f(x,y) = c$ meet in a point where the tangent line of $g(x,y) = a$ and of $f(x,y) = c$ are the same

tangent lines have the same slope $\Rightarrow \; -\dfrac{f'_x}{f'_y} = -\dfrac{g'_x}{g'_y} \iff \begin{matrix} f'_x = \lambda \cdot g'_x \\ f'_y = \lambda \cdot g'_y \end{matrix} \iff \begin{cases} L'_x = 0 \\ L'_y = 0 \end{cases}$

General Lagrange problem: max/min $f(\underline{x}) = f(x_1, \ldots, x_n)$ when
$$\begin{cases} g_1(\underline{x}) = a_1 \\ g_2(\underline{x}) = a_2 \\ \vdots \\ g_m(\underline{x}) = a_m \end{cases}$$

n variables, m constraints

$$\mathcal{L}(x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_m) = f(x_1, \ldots, x_n) - \lambda_1 \cdot g_1(x_1, \ldots, x_n) - \ldots - \lambda_m g_m(x_1, \ldots, x_n)$$

Lagrange conditions: FOC + C

FOC: $\boxed{\begin{matrix} \mathcal{L}'_{x_1} = 0 \\ \mathcal{L}'_{x_2} = 0 \\ \vdots \\ \mathcal{L}'_{x_n} = 0 \end{matrix}}$ C: $\boxed{\begin{matrix} g_1(\underline{x}) = a_1 \\ g_2(\underline{x}) = a_2 \\ \vdots \\ g_m(\underline{x}) = a_m \end{matrix}}$

$(n+m)$ eqn's in $(n+m)$ unknown var's

Thm: Necessary condition for Lagrange problems

If $(x_1^*, x_2^*, \ldots, x_n^*) = \underline{x}^*$ is a max/min in the general Lagrange problem, and NDCQ is satisfied at $\underline{x}^*$, then there are Lagrange multipliers $\lambda_1^*, \ldots, \lambda_m^*$ such that $(\underline{x}^*; \underline{\lambda}^*)$ satisfies FOC + C.

$$\boxed{\underline{x}^* \text{ max/min} \xRightarrow[\text{NDCQ}]{} (\underline{x}^*; \lambda^*) \text{ satisfies FOC+C}}$$

Ex: min $x^2 + y^2$ when $xy = 1$

D: $xy = 1$ is not bounded

FOC+C: $\underbrace{\begin{matrix} (1,1;2) & f=2 \\ (-1,-1;2) & f=2 \end{matrix}}_{\text{cand. for min.}}$

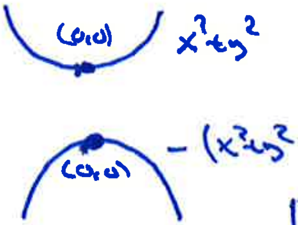
(4) Kuhn-Tucker problems

<u>Ex:</u> $\max f(x,y) = x^2 + y^2$ when $xy \leq 1$
$g(x,y) \leq a$

Kuhn-tucker problem in std. form:

$$\max f(x_1, \ldots, x_n) \quad \text{when} \quad \begin{cases} g_1(x_1, \ldots, x_n) \leq a_1 \\ g_2(x_1, \ldots, x_n) \leq a_2 \\ \vdots \\ g_m(x_1, \ldots, x_n) \leq a_m \end{cases}$$

<u>Ex:</u> $\boxed{\min\ x^2 + y^2 \text{ when } xy \geq 1}$ $\quad 1 \cdot (-1)$ KT problem, not in std. form



$\max -(x^2+y^2) \qquad -xy \leq -1$

same solution-pts.

$\boxed{\max\ -x^2 - y^2 \text{ when } -xy \leq -1}$ KT problem in std. form.

<u>Method:</u> Kuhn-tucker formulation (using std. form)

$$\mathcal{L} = f(x_1, \ldots, x_n) - \lambda_1 \cdot g_1(x_1, \ldots, x_n) - \ldots - \lambda_m \cdot g_m(x_1, \ldots, x_n)$$

<u>FOC:</u> $\boxed{\begin{array}{l} \mathcal{L}'_{x_1} = 0 \\ \mathcal{L}'_{x_2} = 0 \\ \vdots \\ \mathcal{L}'_{x_n} = 0 \end{array}}$

<u>C:</u> $\boxed{\begin{array}{l} g_1(\underline{x}) \leq a_1 \\ g_2(\underline{x}) \leq a_2 \\ \vdots \\ g_m(\underline{x}) \leq a_m \end{array}}$

<u>CSC:</u> complementary slackness condition $\boxed{\begin{array}{ll} \lambda_1 \geq 0, & \lambda_1(g_1(\underline{x}) - a_1) = 0 \\ \lambda_2 \geq 0, & \lambda_2(g_2(\underline{x}) - a_2) = 0 \\ \vdots & \vdots \\ \lambda_m \geq 0 & \end{array}}$

applies when the KT problem is in std. form

CSC: complementary slackness conditions

Ex: max x^2+y^2 when $xy \le 1$ ← std. form

$L = x^2 + y^2 - \lambda \cdot xy$

FOC: $L'_x = 2x - \lambda y = 0$

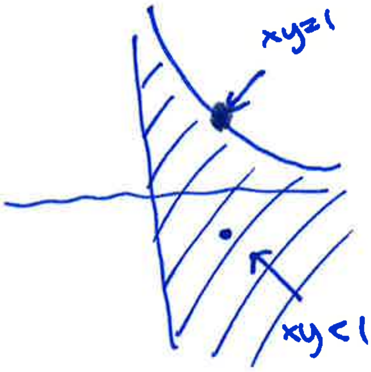
$L'_y = 2y - \lambda x = 0$

C: $xy \le 1$

CSC: $\lambda \ge 0$ and $\lambda \cdot (xy - 1) = 0$

$\Updownarrow$

$\lambda = 0$ or $xy = 1$



C:	$xy = 1$ (binding)	$xy < 1$ (non-binding)
CSC:	$\lambda \ge 0$	$\lambda = 0$
FOC:	$2x - \lambda y = 0$ $2y - \lambda x = 0$ $\Downarrow$ $x = \frac{\lambda y}{2}$ $2y = \lambda \cdot \left(\frac{\lambda y}{2}\right)$ $4y = \lambda^2 y$ $(4-\lambda^2)y = 0$ $y = 0$ or $\lambda^2 = 4$	$2x - \lambda y = 0$ $2y - \lambda x = 0$ $\Downarrow$ $2x = 0 \Rightarrow x = 0$ $2y = 0 \quad y = 0$ $x \cdot y = 0 \cdot 0 < 1$ ok.

(continued)

$xy = 1$
$\lambda \geq 0$
$2x - \lambda y = 0$
$2y - \lambda x = 0$

$x = \frac{\lambda y}{2}$

$y = 0$ or $\lambda^2 = 4$

~~$x = 0$, $xy = 0 = 1$ imp.~~

$\lambda = 2$ $(\lambda \geq 0)$
$x = y$
$x^2 = 1$
$(x, y; \lambda) = (1, 1; 2)$
$(-1, -1; 2)$

Candidates:

$(1, 1; 2)$ $f = 2$
$(-1, -1; 2)$ $f = 2$

$xy < 1$
$\lambda = 0$
$2x - \lambda y = 0$
$2y - \lambda x = 0$

$\lambda = 0 \Rightarrow x = 0, y = 0$

Candidates:

$(0, 0; 0)$ $f = 0$

best candidates for max since $2 > 0$

Thm: (Necessary condition for Lagrange problems)

If $\underline{x}^* = (x_1^*, \ldots, x_n^*)$ is a max in a Kuhn-Tucker problem in std. form, and NDCQ is satisfied at $\underline{x}^*$, then there are Lagrange multipliers $\lambda_1^*, \ldots, \lambda_m^*$ such that

$(\underline{x}^*; \lambda^*)$ satisfies $\underbrace{\text{FOC+C+CSC}}_{\text{Kuhn-Tucker conditions}}$

$$\underline{x}^* \text{ max } \xrightarrow[\text{NDCQ}]{} (\underline{x}^*; \underline{\lambda}^*) \text{ satisfies FOC+C+CSC}$$