Hi W02

global topic: optimization

today: F.O.C for inequality constraints

Sundaram "First course in optimization"

theorem 6.1

terminology and notation

$f(x_1, x_2, x_3 \dots x_n) \to \boxed{max}$

$\min\ x^2 + y^2$
s.t $x + y \le 4$
$\to$
$\max\ -x^2 - y^2$
s.t $x + y \le 4$

* we consider only maximization problems.

① Discuss the problem informally

$f(x_1, x_2 \dots x_n)$ $\quad f: \mathbb{R}^n \to \mathbb{R}$

max f

$$\begin{cases} h_1(x_1, \dots x_n) \ge 0 \\ h_2(x_1 \dots x_n) \ge 0 \\ \vdots \\ h_e(x_1 \dots x_n) \ge 0 \end{cases}$$

$5x + x^2 + y^2 \le 6 \longrightarrow 6 - 5x - x^2 - y^2 \ge 0$

$$h_1(x_1, \dots, x_n) \geq 0$$

satisfied → $h_1 = 0$: active, effective, binding

satisfied → $h_1 > 0$: non-active, non-eff., non-bind...

violated.

Ex
$x^2 + y^2 - 10 \geq 0$
$(x,y) = (5,7)$
sat.
not active.

Step A. Informal

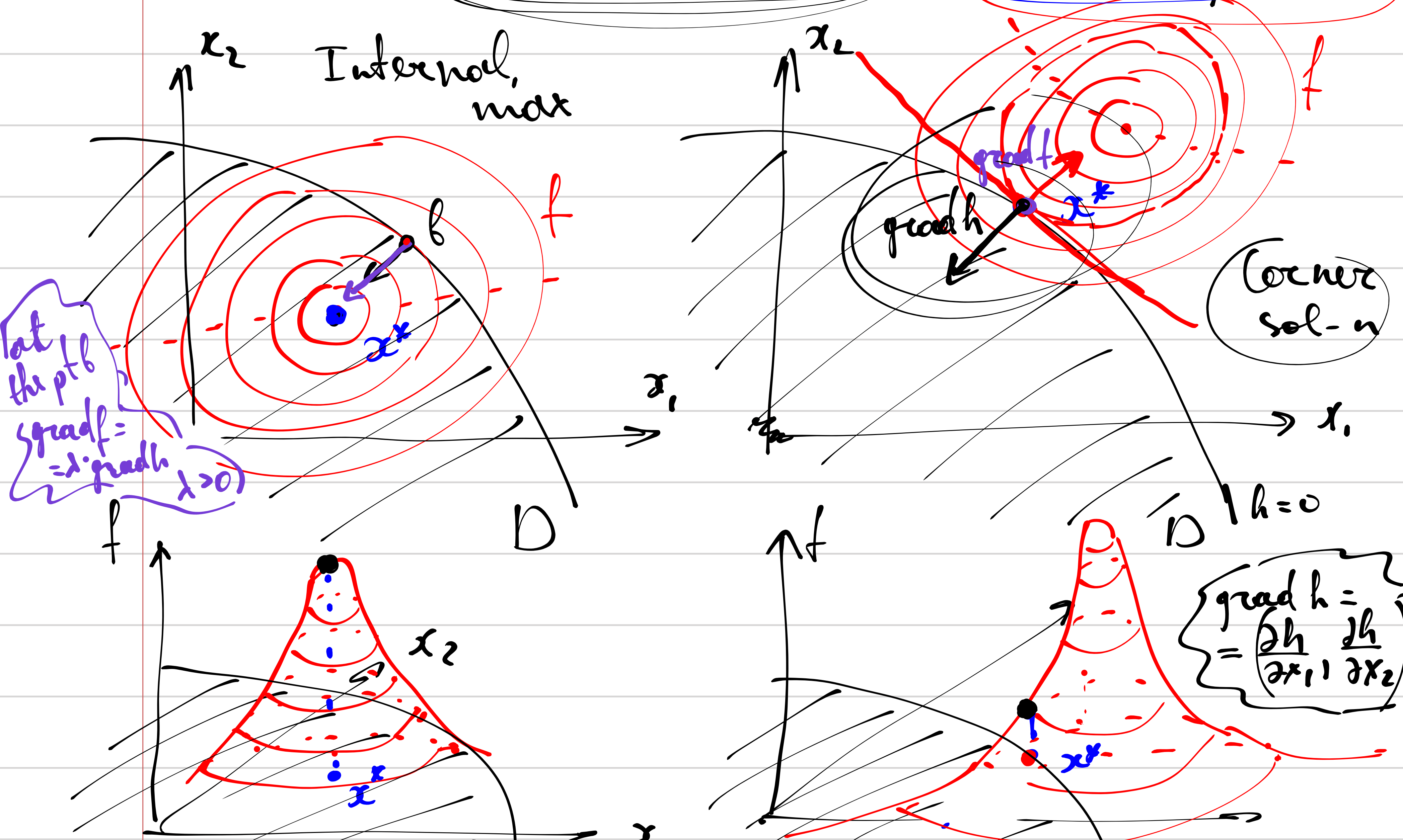
Step B. Formal statement with all det-s.

Example $x \in \mathbb{R}^2$

$$f(x_1, x_2) \to \max$$
$$h(x_1, x_2) \geq 0$$

$D \subset \mathbb{R}^2$

$D = \{h(x_1, x_2) \geq 0\}$



equations! $x^* \in D$ $h(x_1^*, x_2^*) \geq 0$

$$\left.\frac{\partial f}{\partial x_1}\right|_{x^*} = 0 \quad \left.\frac{\partial f}{\partial x_2}\right|_{x^*} = 0$$

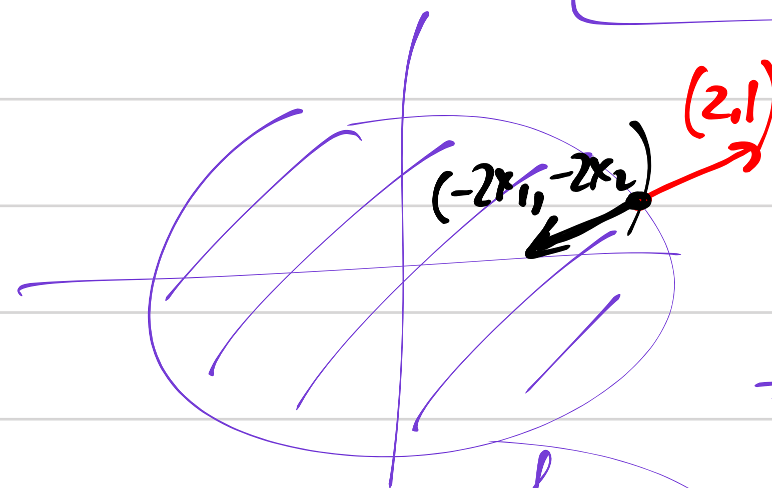
$h(x_1^*, x_2^*) = 0$

grad $f = -\lambda \cdot$ grad h

$\lambda > 0$

$x_1^2 + x_2^2 \le 1$

$\boxed{1 - x_1^2 - x_2^2 \ge 0}$



grad $h =$
$= (-2x_1, -2x_2)$

f

max $\boxed{2x_1 + x_2}$

grad $f = (2, 1)$

$$\boxed{\begin{array}{c} \lambda > 0 \\ \text{grad } f = -\lambda \cdot \text{grad } h \\ h(x^*) = 0 \end{array}} \begin{cases} (2, 1) = -\lambda \cdot (-2x_1, -2x_2) \\ 1 - x_1^2 - x_2^2 = 0 \end{cases}$$

$\lambda > 0$

$2 = \lambda \cdot 2x_1$

$1 = \lambda \cdot 2x_2$

$x_1 > 0 \quad x_2 > 0$

$x_1 = 2x_2$

Internal sol-n | Corner sol-n

equations! $x^* \in D$ $\quad h(x_1^*, x_2^*) \ge 0$

$$\left.\frac{\partial f}{\partial x_1}\right|_{x^*} = 0 \quad \left.\frac{\partial f}{\partial x_2}\right|_{x^*} = 0$$

cond C

fake: $\lambda = 0$

$h(x_1^*, x_2^*) = 0$

grad $f = -\lambda \cdot$ grad h

$\lambda > 0$

$$\boxed{\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right) = -\lambda \cdot \left(\frac{\partial h}{\partial x_1}; \frac{\partial h}{\partial x_2}\right)}$$

$\boxed{\text{cond A}}$

write unified system

$$\begin{cases} \partial f/\partial x_1 = -\lambda \cdot \partial h/\partial x_1 \\ \partial f/\partial x_2 = -\lambda\, \partial h/\partial x_2 \end{cases}$$

$$L(x_1, x_2) = f(x_1, x_2) + \lambda \cdot h(x_1, x_2)$$

$\boxed{\text{almost F.O.C}}$

$$\begin{cases} \dfrac{\partial L}{\partial x_1} = \dfrac{\partial f}{\partial x_1} + \lambda \cdot \dfrac{\partial h}{\partial x_1} = 0 \\ \dfrac{\partial L}{\partial x_2} = \dfrac{\partial f}{\partial x_2} + \lambda \cdot \dfrac{\partial h}{\partial x_2} = 0 \end{cases} \text{cond B}$$

complementary slackness condition

$\lambda \ge 0$ $\quad \boxed{h(x_1, x_2) = \frac{\partial L}{\partial \lambda} \ge 0}$ $\quad \lambda \cdot \frac{\partial L}{\partial \lambda} = 0$

One small but important detail.

Ex. $\left\{x_1^3 + x_2^3 \le 0\right\} = D \to \boxed{h = -x_1^3 - x_2^3 \ge 0}$

$$f(x_1, x_2) = -(1-x_1)^2 - (1-x_2)^2 \to \max$$

geometry:

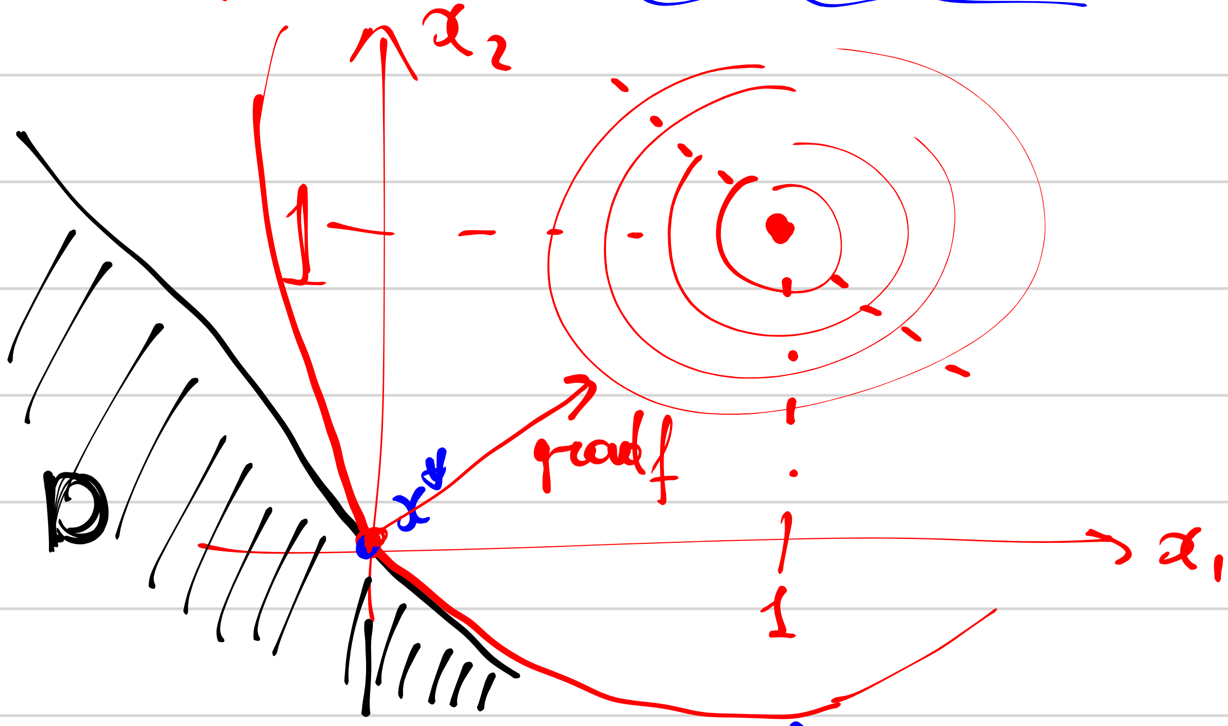
$x_1^3 \le -x_2^3$

$\Downarrow$

$x_1 \le -x_2$

$x_1 + x_2 \le 0$

$f(x_1, x_2) = -(1-x_1)^2 - (1-x_2)$



$x^* = (0,0)$

$$\operatorname{grad} f\Big|_{x^*} = \left(2(1-x_1),\ 2(1-x_2)\right)\Big|_{x^*=(0,0)} = (2,\ 2)$$

$$\operatorname{grad} h = (-3x_1^2;\ -3x_2^2)\Big|_{x^*=(0,0)} = (0,0)$$

$$(2, 2) = -\lambda \cdot (0, 0)$$ ‼ not possible

The Kuhn – Tucker theorem.

(1) If $f : \mathbb{R}^n \to \mathbb{R}$, $h_1, h_2 \dots h_\ell : \mathbb{R}^n \to \mathbb{R}$

(2) $f, h_1, h_2 \dots h_\ell$ are C^1 functions

(3) [cont-ous der-s]

(4)

(5) $D = U \cap \{ x \in \mathbb{R}^n \mid h_1(x) \geq 0, h_2(x) \geq 0 \dots$

(6) $h_\ell(x) \geq 0 \}$

(7) U – any open subset of $\mathbb{R}^n$.

⋮

(8) x^* is the maximum of the function f

(9) on the set D

(10)

NDCQ
non degenerate constraints qualification

$$\boxed{\operatorname{rank}\left(\begin{matrix}\text{Jacobian of active}\\ \text{constraints at } x^*\end{matrix}\right) = \begin{matrix}\text{number of}\\ \text{active constr-nts}\end{matrix}}$$

$$L(x) = f(x) + \sum_{i=1}^{\ell} \lambda_i \cdot h_i(x)$$

$i \in \{1 \dots \ell\}$
$j \in \{1 \dots n\}$

then:

$$\begin{cases} \dfrac{\partial L}{\partial x_j} = \dfrac{\partial f}{\partial x_j} + \displaystyle\sum_{i=1}^{\ell} \lambda_i \cdot \dfrac{\partial h_i}{\partial x_j} = 0 & (j \in \{1 \dots n\}) \\ \lambda_i \geq 0 & (i \in \{1 \dots \ell\}) \\ \dfrac{\partial L}{\partial \lambda_i} = \boxed{h_i(x) \geq 0} & (i \in \{1 \dots \ell\}) \\ \lambda_i \cdot h_i(x) = 0 & (i \in \{1 \dots \ell\}) \end{cases}$$

Old case with (=)
$\dfrac{\partial L}{\partial x_j} = 0 \qquad h_i(x) = 0$

6.1 in Sundaram
First course in optimization