

**Optimality conditions  
for  
constrained optimisation**

## **Recall:** Optimisation conditions for unconstrained problem with one variable function

- One variable function:  $f : R^1 \rightarrow R^1$

One variable function:  $f : R^1 \rightarrow R^1$

**Result:** Let  $c \in R^1$  such that

$$f'(c) = \frac{df(c)}{dx} = \lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c} \neq 0.$$

Then  $c$  is neither a local minimum nor a local maximum of  $f$ .

Furthermore,

i) if  $f'(c) > 0$ , There exists an interval  $(c - \delta, c + \delta)$  with  $\delta > 0$  such that

$$f(x) < f(c) \quad \forall x \in (c - \delta, c)$$

$$f(c) < f(x) \quad \forall x \in (c, c + \delta)$$

and  $f$  is **increasing** at  $c$

ii) if  $f'(c) < 0$ , there exists an interval  $(c - \delta, c + \delta)$  with  $\delta > 0$  such that

$$f(x) > f(c) \quad \forall x \in (c - \delta, c)$$

$$f(c) > f(x) \quad \forall x \in (c, c + \delta)$$

and  $f$  is **decreasing** at  $c$

### Result : (Second derivative test)

Let  $c \in \mathbb{R}^1$  such that  $f'(c) = 0$ .

Assume also that the second derivative  $f''(x)$  exists  $\forall x \in B_\varepsilon(c)$ .

i) If  $f''(c) < 0$ , then  $c$  is a local maximum of  $f$ .

ii) If  $f''(c) > 0$ , then  $c$  is a local minimum of  $f$ .

**Remark:**  $f''(c)$  denotes the second derivative of  $f$  at  $c$ :

$$f''(c) = \frac{df'}{dx}(c) = \lim_{x \rightarrow c} \frac{f'(x) - f'(c)}{x - c} = \frac{d^2 f}{dx^2}(c).$$

**Intuitive justification of ii):**

If  $f''(c) > 0$ , then  $f'$  is an increasing function at  $c$ , and it follows that  $c$  is a local minimum.

## Recall: Optimisation conditions for unconstrained problem with several variables function

Several variables function:  $f : R^n \rightarrow R^1$

To extend the preceding results obtained for one variable function:

$$f : R^1 \rightarrow R^1 \leftrightarrow f : R^n \rightarrow R^1$$

$$f'(x) \leftrightarrow \nabla f(x) = \left[ \frac{\delta f(x)}{\delta x_1}, \dots, \frac{\delta f(x)}{\delta x_n} \right]^T$$

$$f''(x) \leftrightarrow \nabla^2 f(x) = \begin{pmatrix} \frac{\delta^2 f(x)}{\delta x_1 \delta x_1} & \dots & \frac{\delta^2 f(x)}{\delta x_1 \delta x_n} \\ \vdots & \ddots & \vdots \\ \frac{\delta^2 f(x)}{\delta x_n \delta x_1} & \dots & \frac{\delta^2 f(x)}{\delta x_n \delta x_n} \end{pmatrix}$$

$$\psi(x) = x^T D x = [x_1, \dots, x_n] \begin{pmatrix} d_{11} & \dots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{n1} & \dots & d_{nn} \end{pmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

To extend the preceding results obtained for one variable function:

$$f: \mathbb{R}^1 \rightarrow \mathbb{R}^1 \leftrightarrow f: \mathbb{R}^n \rightarrow \mathbb{R}^1$$

$$f'(x) \leftrightarrow \nabla f(x) = \left[ \frac{\delta f(x)}{\delta x_1}, \dots, \frac{\delta f(x)}{\delta x_n} \right]^T$$

$$f''(x) \leftrightarrow \nabla^2 f(x) = \begin{pmatrix} \frac{\delta^2 f(x)}{\delta x_1 \delta x_1} & \dots & \frac{\delta^2 f(x)}{\delta x_1 \delta x_n} \\ \vdots & \ddots & \vdots \\ \frac{\delta^2 f(x)}{\delta x_n \delta x_1} & \dots & \frac{\delta^2 f(x)}{\delta x_n \delta x_n} \end{pmatrix}$$

**Definition:** The **quadratic form** associated with the a real valued  $n \times n$  matrix  $D$  is the fonction  $\psi: \mathbb{R}^n \rightarrow \mathbb{R}^1$  specified as follows

$$\psi(x) = x^T D x.$$

**Définition:** A real valued  $n \times n$  matrix  $D$  is **positive semi-definite** (**positive definite**) if  $\psi(x) \geq 0 \quad \forall x \in \mathbb{R}^n$  ( $\psi(x) > 0 \quad \forall x \in \mathbb{R}^n, x \neq 0$ ).

**Results:** A real valued  $n \times n$  matrix  $D$  is **positive semi-definite** (**positive definite**) if and only if all its eigenvalues are **non negative** (**positive**).

## Necessary conditions

**Lemma:** Let  $X \subset \mathbb{R}^n$  be an open set and  $f \in C^2 / X$  be twice continuously differentiable. If  $x \in X$  is a local minimum of  $f$  on  $X$ , then  $\nabla f(x) = 0$  and  $\nabla^2 f(x)$  is a **positive semi-definite** matrix.

## Sufficient conditions

**Lemma:** Let  $X \subset \mathbb{R}^n$  be an open set and  $f \in C^2 / X$  be twice continuously differentiable. If  $\nabla f(x^*) = 0$  and  $\nabla^2 f(x^*)$  is a **positive definite** matrix, then  $x^*$  is a local minimum of  $f$  on  $X$ .

**Conter-exemple:** The conditions  $\nabla f(x) = 0$  and  $\nabla^2 f(x)$  being a positive semi definite matrix are not sufficient to guarantee that  $x$  is a local minimum.

$$f(x, y) = x^3 + y^3$$

$$\nabla f(x, y) = [3x^2, 3y^2]^T$$

$$\nabla^2 f(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix}$$



**Conter-exemple:** The conditions  $\nabla f(x) = 0$  and  $\nabla^2 f(x)$  being a positive semi definite matrix are not sufficient to guarantee that  $x$  is a local minimum.

$$f(x, y) = x^3 + y^3$$

$$\nabla f(x, y) = [3x^2, 3y^2]^T$$

$$\nabla^2 f(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix}$$

At the point  $x = y = 0$

$$f(x, y) = 0$$

$$\nabla f(0, 0) = [0, 0]^T$$

$$\nabla^2 f(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ pos. semi def.}$$

**Conter-exemple:** The conditions  $\nabla f(x) = 0$  and  $\nabla^2 f(x)$  being a positive semi definite matrix are not sufficient to guarantee that  $x$  is a local minimum.

$$f(x, y) = x^3 + y^3$$

$$\nabla f(x, y) = [3x^2, 3y^2]^T$$

$$\nabla^2 f(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix}$$

At the point  $x = y = 0$

$$f(x, y) = 0$$

$$\nabla f(0, 0) = [0, 0]^T$$

$$\nabla^2 f(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ pos. semi def.}$$

Then the conditions are satisfied at  $x = y = 0$ .

**Conter-exemple:** The conditions  $\nabla f(x) = 0$  and  $\nabla^2 f(x)$  being a positive semi definite matrix are not sufficient to guarantee that  $x$  is a local minimum.

$$\begin{array}{ll} f(x, y) = x^3 + y^3 & \text{At the point } x = y = 0 \\ \nabla f(x, y) = [3x^2, 3y^2]^T & \nabla f(0, 0) = [0, 0]^T \\ \nabla^2 f(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix} & \nabla^2 f(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ pos. semi def.} \end{array}$$

Then the conditions are satisfied at  $x = y = 0$ .

But for an  $\varepsilon > 0$  sufficiently small,  $\left[-\frac{\varepsilon}{2}, -\frac{\varepsilon}{2}\right] \in B_\varepsilon(0, 0)$  and

$$f\left(-\frac{\varepsilon}{2}, -\frac{\varepsilon}{2}\right) = \left(-\frac{\varepsilon}{2}\right)^3 + \left(-\frac{\varepsilon}{2}\right)^3 = -\frac{2\varepsilon^3}{8} < 0 = f(0, 0),$$

**Conter-exemple:** The conditions  $\nabla f(x) = 0$  and  $\nabla^2 f(x)$  being a positive semi definite matrix are not sufficient to guarantee that  $x$  is a local minimum.

$$\begin{array}{ll} f(x, y) = x^3 + y^3 & \text{At the point } x = y = 0 \\ \nabla f(x, y) = [3x^2, 3y^2]^T & \nabla f(0, 0) = [0, 0]^T \\ \nabla^2 f(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix} & \nabla^2 f(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ pos. semi def.} \end{array}$$

Then the conditions are satisfied at  $x = y = 0$ .

But for an  $\varepsilon > 0$  sufficiently small,  $\left[-\frac{\varepsilon}{2}, -\frac{\varepsilon}{2}\right] \in B_\varepsilon(0, 0)$  and

$$f\left(-\frac{\varepsilon}{2}, -\frac{\varepsilon}{2}\right) = \left(-\frac{\varepsilon}{2}\right)^3 + \left(-\frac{\varepsilon}{2}\right)^3 = -\frac{2\varepsilon^3}{8} < 0 = f(0, 0),$$

and  $[0, 0]$  is not a local minimum even if the conditions are satisfied.

## Lagrangean multipliers

Consider the following mathematical programming problem

$$\begin{aligned} & \text{Min } f(x) \\ & \text{s.t. } f_i(x) = 0 \quad i = 1, \dots, m \\ & \quad \quad x \in X \end{aligned} \tag{1}$$

where  $X \subset R^n$  and the functions  $f : X \rightarrow R^1$ ,  $f_i : X \rightarrow R^1, i = 1, \dots, m$ .

## Lagrangean multipliers

Consider the following mathematical programming problem

$$\begin{aligned} \text{Min } & f(x) \\ \text{s.t. } & f_i(x) = 0 \quad i = 1, \dots, m \\ & x \in X \end{aligned} \tag{1}$$

where  $X \subset R^n$  and the functions  $f : X \rightarrow R^1$ ,  $f_i : X \rightarrow R^1, i = 1, \dots, m$ .

To obtain the **lagrangean** function of (1), we associate a **lagrangean multiplicateur**  $\lambda_i$  with each constraint fonction  $f_i$  :

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x).$$

# Optimisation conditions for constrained problem

Consider the following mathematical programming problem

$$\begin{aligned} \text{Min } & f(x) \\ \text{s.t. } & f_i(x) = 0 \quad i = 1, \dots, m \\ & x \in X \end{aligned} \tag{1}$$

where  $X \subset \mathbb{R}^n$  and the functions  $f : X \rightarrow \mathbb{R}^1$ ,  $f_i : X \rightarrow \mathbb{R}^1, i = 1, \dots, m$ .

To obtain the **lagrangean** function of (1), we associate a **lagrangean multiplicateur**  $\lambda_i$  with each constraint fonction  $f_i$  :

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x).$$

Without any assumption on  $X$  or on the fonctions  $f$  and  $f_i$ , we can derive sufficient conditions for a point  $x^*$  to be a global optimal solution for problem (1).

Consider the following mathematical programming problem

$$\begin{aligned} & \text{Min } f(x) \\ & \text{s.t. } f_i(x) = 0 \quad i = 1, \dots, m \\ & \quad \quad x \in X \end{aligned} \tag{1}$$



Consider the following mathematical programming problem

$$\begin{aligned} & \text{Min } f(x) \\ & \text{s.t. } f_i(x) = 0 \quad i = 1, \dots, m \\ & \quad x \in X \end{aligned} \tag{1}$$

**Theorem 1:** Assume that the lagrangean function of (1)

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

has a global minimum  $x^*$  on  $X$  when the multiplier vector  $\lambda = \lambda^*$ . If  $f_i(x^*) = 0$  for all  $i = 1, \dots, m$ , then  $x^*$  is a global optimal solution of (1).

Consider the following mathematical programming problem

$$\begin{aligned} & \text{Min } f(x) \\ & \text{s.t. } f_i(x) = 0 \quad i = 1, \dots, m \\ & \quad x \in X \end{aligned} \tag{1}$$

**Theorem 1:** Assume that the lagrangean function of (1)

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

has a global minimum  $x^*$  on  $X$  when the multiplier vector  $\lambda = \lambda^*$ . If  $f_i(x^*) = 0$  for all  $i = 1, \dots, m$ , then  $x^*$  is a global optimal solution of (1).

**Proof.** For contradiction, suppose that  $x^*$  is not a global optimal solution of (1). Then there exists another solution  $\bar{x} \in X$  such that  $f_i(\bar{x}) = 0$  for all  $i = 1, \dots, m$ , and  $f(\bar{x}) < f(x^*)$ .

**Theorem 1:** Assume that the lagrangean function of (1)

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

has a global minimum  $x^*$  on  $X$  when the multiplier vector  $\lambda = \lambda^*$ . If  $f_i(x^*) = 0$  for all  $i = 1, \dots, m$ , then  $x^*$  is a global optimal solution of (1).

**Proof.** For contradiction, suppose that  $x^*$  is not a global optimal solution of (1). Then there exists another solution  $\bar{x} \in X$  such that  $f_i(\bar{x}) = 0$  for all  $i = 1, \dots, m$ , and  $f(\bar{x}) < f(x^*)$ .

**Theorem 1:** Assume that the lagrangean function of (1)

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

has a global minimum  $x^*$  on  $X$  when the multiplier vector  $\lambda = \lambda^*$ . If  $f_i(x^*) = 0$  for all  $i = 1, \dots, m$ , then  $x^*$  is a global optimal solution of (1).

**Proof.** For contradiction, suppose that  $x^*$  is not a global optimal solution of (1). Then there exists another solution  $\bar{x} \in X$  such that  $f_i(\bar{x}) = 0$  for all  $i = 1, \dots, m$ , and  $f(\bar{x}) < f(x^*)$ .

Hence, for all  $\lambda$

$$\sum_{i=1}^m \lambda_i f_i(\bar{x}) = \sum_{i=1}^m \lambda_i f_i(x^*) = 0$$

**Theorem 1:** Assume that the lagrangean function of (1)

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

has a global minimum  $x^*$  on  $X$  when the multiplier vector  $\lambda = \lambda^*$ . If  $f_i(x^*) = 0$  for all  $i = 1, \dots, m$ , then  $x^*$  is a global optimal solution of (1).

**Proof.** For contradiction, suppose that  $x^*$  is not a global optimal solution of (1). Then there exists another solution  $\bar{x} \in X$  such that  $f_i(\bar{x}) = 0$  for all  $i = 1, \dots, m$ , and  $f(\bar{x}) < f(x^*)$ .

Hence, for all  $\lambda$

$$\sum_{i=1}^m \lambda_i f_i(\bar{x}) = \sum_{i=1}^m \lambda_i f_i(x^*) = 0$$

and

$$f(\bar{x}) + \sum_{i=1}^m \lambda_i f_i(\bar{x}) < f(x^*) + \sum_{i=1}^m \lambda_i f_i(x^*).$$

**Theorem 1:** Assume that the lagrangean function of (1)

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

has a global minimum  $x^*$  on  $X$  when the multiplier vector  $\lambda = \lambda^*$ . If  $f_i(x^*) = 0$  for all  $i = 1, \dots, m$ , then  $x^*$  is a global optimal solution of (1).

**Proof.** For contradiction, suppose that  $x^*$  is not a global optimal solution of (1). Then there exists another solution  $\bar{x} \in X$  such that  $f_i(\bar{x}) = 0$  for all  $i = 1, \dots, m$ , and  $f(\bar{x}) < f(x^*)$ .

Hence, for all  $\lambda$

$$\sum_{i=1}^m \lambda_i f_i(\bar{x}) = \sum_{i=1}^m \lambda_i f_i(x^*) = 0$$

and

$$f(\bar{x}) + \sum_{i=1}^m \lambda_i f_i(\bar{x}) < f(x^*) + \sum_{i=1}^m \lambda_i f_i(x^*).$$

If  $\lambda = \lambda^*$ , then the preceding relation contradicts the fact that  $x^*$  is a global minimum of the lagrangean on  $X$  when  $\lambda = \lambda^*$ . □

Consider the following mathematical programming problem

$$\begin{aligned} & \text{Min } f(x) \\ & \text{s.t. } f_i(x) \leq 0 \quad i = 1, \dots, m \\ & \quad x \in X \end{aligned} \tag{2}$$

Consider the following mathematical programming problem

$$\begin{aligned} & \text{Min } f(x) \\ & \text{s.t. } f_i(x) \leq 0 \quad i = 1, \dots, m \\ & \quad x \in X \end{aligned} \tag{2}$$

**Theorem 2:** Assume that the lagrangean function of (2)

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

has a global minimum  $x^*$  on  $X$  when the multiplier vector  $\lambda = \lambda^*$ . If  $f_i(x^*) \leq 0$ ,  $\lambda_i^* \geq 0$ , and  $\lambda_i^* f_i(x^*) = 0$  for all  $i = 1, \dots, m$ , then  $x^*$  is a global optimal solution of (2).



Consider the following mathematical programming problem

$$\begin{aligned} & \text{Min } f(x) \\ & \text{s.t. } f_i(x) \leq 0 \quad i = 1, \dots, m \\ & \quad x \in X \end{aligned} \tag{2}$$

**Theorem 2:** Assume that the lagrangean function of (2)

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

has a global minimum  $x^*$  on  $X$  when the multiplier vector  $\lambda = \lambda^*$ . If  $f_i(x^*) \leq 0$ ,  $\lambda_i^* \geq 0$ , and  $\lambda_i^* f_i(x^*) = 0$  for all  $i = 1, \dots, m$ , then  $x^*$  is a global optimal solution of (2).

**Proof.** For contradiction, suppose that  $x^*$  is not a global optimal solution of (2). Then there exists another solution  $\bar{x} \in X$  such that  $f_i(\bar{x}) \leq 0$  for all  $i = 1, \dots, m$ , and  $f(\bar{x}) < f(x^*)$ .

**Theorem 2:** Assume that the lagrangean function of (2)

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

has a global minimum  $x^*$  on  $X$  when the multiplier vector  $\lambda = \lambda^*$ . If  $f_i(x^*) \leq 0$ ,  $\lambda_i^* \geq 0$ , and  $\lambda_i^* f_i(x^*) = 0$  for all  $i = 1, \dots, m$ , then  $x^*$  is a global optimal solution of (2).

**Proof.** For contradiction, suppose that  $x^*$  is not a global optimal solution of (2). Then there exists another solution  $\bar{x} \in X$  such that  $f_i(\bar{x}) \leq 0$  for all  $i = 1, \dots, m$ , and  $f(\bar{x}) < f(x^*)$ .

**Theorem 2:** Assume that the lagrangean function of (2)

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

has a global minimum  $x^*$  on  $X$  when the multiplier vector  $\lambda = \lambda^*$ . If  $f_i(x^*) \leq 0$ ,  $\lambda_i^* \geq 0$ , and  $\lambda_i^* f_i(x^*) = 0$  for all  $i = 1, \dots, m$ , then  $x^*$  is a global optimal solution of (2).

**Proof.** For contradiction, suppose that  $x^*$  is not a global optimal solution of (2). Then there exists another solution  $\bar{x} \in X$  such that  $f_i(\bar{x}) \leq 0$  for all  $i = 1, \dots, m$ , and  $f(\bar{x}) < f(x^*)$ .

Hence, for  $\lambda^* \geq 0$

$$\sum_{i=1}^m \lambda_i^* f_i(\bar{x}) \leq 0 \quad \text{and} \quad \sum_{i=1}^m \lambda_i^* f_i(x^*) = 0$$

**Theorem 2:** Assume that the lagrangean function of (2)

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

has a global minimum  $x^*$  on  $X$  when the multiplier vector  $\lambda = \lambda^*$ . If  $f_i(x^*) \leq 0$ ,  $\lambda_i^* \geq 0$ , and  $\lambda_i^* f_i(x^*) = 0$  for all  $i = 1, \dots, m$ , then  $x^*$  is a global optimal solution of (2).

**Proof.** For contradiction, suppose that  $x^*$  is not a global optimal solution of (2). Then there exists another solution  $\bar{x} \in X$  such that  $f_i(\bar{x}) \leq 0$  for all  $i = 1, \dots, m$ , and  $f(\bar{x}) < f(x^*)$ .

Hence, for  $\lambda^* \geq 0$

$$\sum_{i=1}^m \lambda_i^* f_i(\bar{x}) \leq 0 \quad \text{and} \quad \sum_{i=1}^m \lambda_i^* f_i(x^*) = 0$$

and

$$f(\bar{x}) + \sum_{i=1}^m \lambda_i^* f_i(\bar{x}) < f(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*).$$

**Theorem 2:** Assume that the lagrangean function of (2)

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

has a global minimum  $x^*$  on  $X$  when the multiplier vector  $\lambda = \lambda^*$ . If  $f_i(x^*) \leq 0$ ,  $\lambda_i^* \geq 0$ , and  $\lambda_i^* f_i(x^*) = 0$  for all  $i = 1, \dots, m$ , then  $x^*$  is a global optimal solution of (2).

**Proof.** For contradiction, suppose that  $x^*$  is not a global optimal solution of (2). Then there exists another solution  $\bar{x} \in X$  such that  $f_i(\bar{x}) \leq 0$  for all  $i = 1, \dots, m$ , and  $f(\bar{x}) < f(x^*)$ .

Hence, for  $\lambda^* \geq 0$

$$\sum_{i=1}^m \lambda_i^* f_i(\bar{x}) \leq 0 \quad \text{and} \quad \sum_{i=1}^m \lambda_i^* f_i(x^*) = 0$$

and

$$f(\bar{x}) + \sum_{i=1}^m \lambda_i^* f_i(\bar{x}) < f(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*).$$

The preceding relation contradicts the fact that  $x^*$  is a global minimum of the lagrangean on  $X$  when  $\lambda = \lambda^*$ . □

# First order Karush-Kuhn-Tucker (KKT) optimality conditions

To have conditions easier to verify, we need additional assumptions on  $X$  and on the functions  $f$  and  $f_i$ .

If  $X$  is convex, and if  $f$  and  $f_i$  are differentiable and convex in problem (2)

$$\begin{aligned} \text{Min } & f(x) \\ \text{s.t. } & f_i(x) \leq 0 \quad i = 1, \dots, m \\ & x \in X \end{aligned} \tag{2}$$

and if  $\lambda_i \geq 0, i = 1, \dots, m,$

then the lagrangean  $L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$

is also a convex function of  $x$  on  $X$  since

$$\lambda_i \geq 0 \text{ et } f_i(x) \text{ convex} \Rightarrow \lambda_i f_i(x) \text{ convex}$$

$$f(x) + \sum_{i=1}^m \lambda_i f_i(x) \text{ sum of convex functions.}$$

If  $f$  and  $f_i$  are differentiable and convex, then the lagrangean

$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$  is also a differentiable and convex function in  $x$ ,

and hence  $x^*$  is a global minimum on  $X$  when  $\lambda = \lambda^*$  if

$$\nabla_x L(\lambda^*, x^*) = \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0.$$

If  $f$  and  $f_i$  are differentiable and convex, then the lagrangean

$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$  is also a differentiable and convex function in  $x$ ,

and hence  $x^*$  is a global minimum on  $X$  when  $\lambda = \lambda^*$  if

$$\nabla_x L(\lambda^*, x^*) = \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0.$$

Thus, theorem 2 can be written as:

**Theorem 2:** Assume that the lagrangean function of problem (2)

$$L(\lambda, x) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

has a global minimum  $x^*$  on  $X$

when the multiplier vector  $\lambda = \lambda^*$ .

If  $f_i(x^*) \leq 0$ ,  $\lambda_i^* \geq 0$ , and  $\lambda_i^* f_i(x^*) = 0$

for all  $i = 1, \dots, m$ ,

then  $x^*$  is a global optimal solution of (2).

**K-K-T**

If there exists a vector  $\lambda^*$  such that for  $x^* \in X$

$$\nabla_x L(\lambda^*, x^*) = \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, m$$

$$f_i(x^*) \leq 0 \quad i = 1, \dots, m$$

$$\lambda_i^* \geq 0 \quad i = 1, \dots, m$$



## Sufficiency of the K-K-T conditions

Referring to theorem 2, we can show that the K-K-T conditions are sufficient when in addition  $X$  is convex and the functions  $f$  and  $f_i$  are convex on  $X$ .

## Sufficiency of the K-K-T conditions

Referring to theorem 2, we can show that the K-K-T conditions are sufficient when in addition  $X$  is convex and the functions  $f$  and  $f_i$  are convex on  $X$ .

We can also use the gradient inequality to show the same result.

**Theorem 3:** Assume that  $X$  is convex and that the functions  $f$  et  $f_i$  are differentiable and convex. If the K-K-T conditions are verified at  $x^*$ , then  $x^*$  is a global minimum of the problem (2).

## Sufficiency of the K-K-T conditions

Referring to theorem 2, we can show that the K-K-T conditions are sufficient when in addition  $X$  is convex and the functions  $f$  et  $f_i$  are convex on  $X$ .

We can also use the gradient inequality to show the same result.

**Theorem 3:** Assume that  $X$  is convex and that the functions  $f$  and  $f_i$  are differentiable and convex. If the K-K-T conditions are verified at  $x^*$ , then  $x^*$  is a global minimum of the problem (2).

**Proof.** Since the lagrangean is convex (shown before), then it follows from the gradient inequality that for all  $x \in X$

$$f(x) + \sum_{i=1}^m \lambda_i^* f_i(x) \geq f(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \left[ \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) \right]^T (x - x^*)$$

**Theorem 3:** Assume that  $X$  is convex and that the functions  $f$  and  $f_i$  are differentiable and convex. If the K-K-T conditions are verified at  $x^*$ , then  $x^*$  is a global minimum of the problem (2).

**Proof.** Since the lagrangean is convex (shown before), then it follows from the gradient inequality that for all  $x \in X$

$$f(x) + \sum_{i=1}^m \lambda_i^* f_i(x) \geq f(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \left[ \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) \right]^T (x - x^*)$$

**Theorem 3:** Assume that  $X$  is convex and that the functions  $f$  and  $f_i$  are differentiable and convex. If the K-K-T conditions are verified at  $x^*$ , then  $x^*$  is a global minimum of the problem (2).

**Proof.** Since the lagrangean is convex (shown before), then it follows from the gradient inequality that for all  $x \in X$

$$f(x) + \sum_{i=1}^m \lambda_i^* f_i(x) \geq f(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \left[ \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) \right]^T (x - x^*)$$

K-K-T

$$\nabla_x L(\lambda^*, x^*) = \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, n$$

$$f_i(x^*) \leq 0 \quad i = 1, \dots, n$$

$$\lambda_i^* \geq 0 \quad i = 1, \dots, n$$

**Theorem 3:** Assume that  $X$  is convex and that the functions  $f$  and  $f_i$  are differentiable and convex. If the K-K-T conditions are verified at  $x^*$ , then  $x^*$  is a global minimum of the problem (2).

**Proof.** Since the lagrangean is convex (shown before), then it follows from the gradient inequality that for all  $x \in X$

$$f(x) + \sum_{i=1}^m \lambda_i^* f_i(x) \geq f(x^*) + \underbrace{\sum_{i=1}^m \lambda_i^* f_i(x^*)}_0 + \underbrace{\left[ \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) \right]^T}_{0} (x - x^*)$$

K-K-T

$$\nabla_x L(\lambda^*, x^*) = \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, n$$

$$f_i(x^*) \leq 0 \quad i = 1, \dots, n$$

$$\lambda_i^* \geq 0 \quad i = 1, \dots, n$$

**Theorem 3:** Assume that  $X$  is convex and that the functions  $f$  and  $f_i$  are differentiable and convex. If the K-K-T conditions are verified at  $x^*$ , then  $x^*$  is a global minimum of the problem (2).

**Proof.** Since the lagrangean is convex (shown before), then it follows from the gradient inequality that for all  $x \in X$

$$f(x) + \sum_{i=1}^m \lambda_i^* f_i(x) \geq f(x^*) + \underbrace{\sum_{i=1}^m \lambda_i^* f_i(x^*)}_0 + \underbrace{\left[ \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) \right]^T}_{0} (x - x^*)$$

Then, for all  $x \in X$

$$f(x^*) - f(x) \leq \sum_{i=1}^m \lambda_i^* f_i(x) \leq 0$$

K-K-T

$$\nabla_x L(\lambda^*, x^*) = \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, n$$

$$f_i(x^*) \leq 0 \quad i = 1, \dots, n$$

$$\lambda_i^* \geq 0 \quad i = 1, \dots, n$$

**Theorem 3:** Assume that  $X$  is convex and that the functions  $f$  and  $f_i$  are differentiable and convex. If the K-K-T conditions are verified at  $x^*$ , then  $x^*$  is a global minimum of the problem (2).

**Proof.** Since the lagrangean is convex (shown before), then it follows from the gradient inequality that for all  $x \in X$

$$f(x) + \sum_{i=1}^m \lambda_i^* f_i(x) \geq f(x^*) + \underbrace{\sum_{i=1}^m \lambda_i^* f_i(x^*)}_0 + \underbrace{\left[ \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) \right]^T}_{0} (x - x^*)$$

K-K-T

$$\nabla_x L(\lambda^*, x^*) = \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, n$$

$$f_i(x^*) \leq 0 \quad i = 1, \dots, n$$

$$\lambda_i^* \geq 0 \quad i = 1, \dots, n$$

Then, for all  $x \in X$

$$f(x^*) - f(x) \leq \sum_{i=1}^m \lambda_i^* f_i(x) \leq 0$$

and

$$f(x^*) \leq f(x). \quad \square$$



## Necessity of the K-K-T conditions

Assume that  $x^* \in X$ ,  $f_i(x^*) \leq 0$ ,  $i = 1, \dots, m$ , and is a local minimum of problem (2):

$$\begin{aligned} \text{Min } & f(x) \\ \text{s.t. } & f_i(x) \leq 0 \quad i = 1, \dots, m \\ & x \in X. \end{aligned} \tag{2}$$

Are the K-K-T satisfied; i.e.,  
can we find a multiplier vector  $\lambda^*$  such that

$$\begin{aligned} \nabla_x L(\lambda^*, x^*) &= \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0 \\ \lambda_i^* f_i(x^*) &= 0 \quad i = 1, \dots, n \\ f_i(x^*) &\leq 0 \quad i = 1, \dots, n \\ \lambda_i^* &\geq 0 \quad i = 1, \dots, n? \end{aligned}$$

Assume that  $x^* \in X$ ,  $f_i(x^*) \leq 0$ ,  $i = 1, \dots, m$ , and is a local minimum of problem (2):

$$\begin{aligned} \text{Min } & f(x) \\ \text{s.t. } & f_i(x) \leq 0 \quad i = 1, \dots, m \\ & x \in X. \end{aligned} \tag{2}$$

Are the K-K-T satisfied; i.e.,

can we find a multiplier vector  $\lambda^*$  such that

$$\nabla_x L(\lambda^*, x^*) = \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, n$$

$$f_i(x^*) \leq 0 \quad i = 1, \dots, n$$

$$\lambda_i^* \geq 0 \quad i = 1, \dots, n?$$

Assume that  $x^* \in X$ ,  $f_i(x^*) \leq 0$ ,  $i = 1, \dots, m$ , and is a local minimum of problem (2):

$$\begin{aligned} \text{Min } & f(x) \\ \text{s.t. } & f_i(x) \leq 0 \quad i = 1, \dots, m \\ & x \in X. \end{aligned} \tag{2}$$

Are the K-K-T satisfied; i.e.,  
can we find a multiplier vector  $\lambda^*$  such that

$$\begin{aligned} \nabla_x L(\lambda^*, x^*) &= \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0 \\ \lambda_i^* f_i(x^*) &= 0 \quad i = 1, \dots, n \\ f_i(x^*) &\leq 0 \quad i = 1, \dots, n \\ \lambda_i^* &\geq 0 \quad i = 1, \dots, n? \end{aligned}$$

It is more difficult to answer to this question than it was in studying the sufficiency.

The K-K-T are satisfied when the feasible domain of problem (2) (and hence the constraint functions  $f_i$ ,  $i = 1, \dots, n$ ) verifies some conditions.

Different sets of such conditions exist.

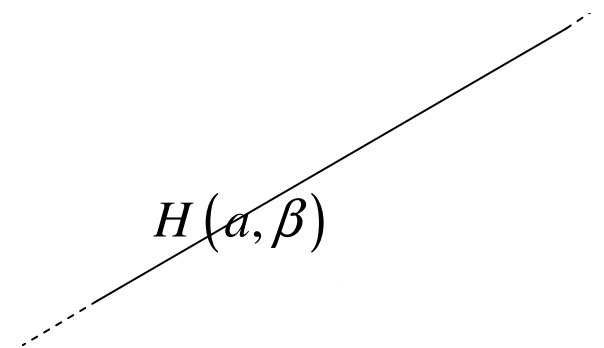
To analyse the necessity of the K-K-T conditions, additional notions and preliminary results related to theorems of alternatives are required

To analyse the necessity of the K-K-T conditions, additional notions and preliminary results related to theorems of alternatives are required

**Definitions.**

The **hyperplan** specified by a point  $a \in R^n$  and a scalar  $\beta$  is the following set in  $R^n$

$$H(a, \beta) = \{x \in R^n : a^T x = \beta\}.$$



To analyse the necessity of the K-K-T conditions, additional notions and preliminary results related to theorems of alternatives are required

### Definitions.

The **hyperplan** specified by a point  $a \in R^n$  and a scalar  $\beta$  is the following set in  $R^n$

$$H(a, \beta) = \{x \in R^n : a^T x = \beta\}.$$

The **half spaces** (closed) associated with the hyperplan  $H(a, \beta)$  are the following sets in  $R^n$  :

$$H^+[a, \beta] = \{x \in R^n : a^T x \geq \beta\}$$

$$H^-[a, \beta] = \{x \in R^n : a^T x \leq \beta\}.$$

$$H^-[a, \beta]$$

$$H(a, \beta)$$

$$H^+[a, \beta]$$

Remark: It is easy to verify that these sets are convex.

**Definition.** Separating hyperplan .

The hyperplan  $H(a, \beta)$  separates two non empty sets  $X$  and  $Y$  if

$$a^T x \geq \beta \text{ for all } x \in X \text{ (i.e., } X \subset H^+[a, \beta])$$

$$a^T y \leq \beta \text{ for all } y \in Y \text{ (i.e., } Y \subset H^-[a, \beta]).$$

The **separation** is **strict** if the inequalities in both preceding relations are strict.

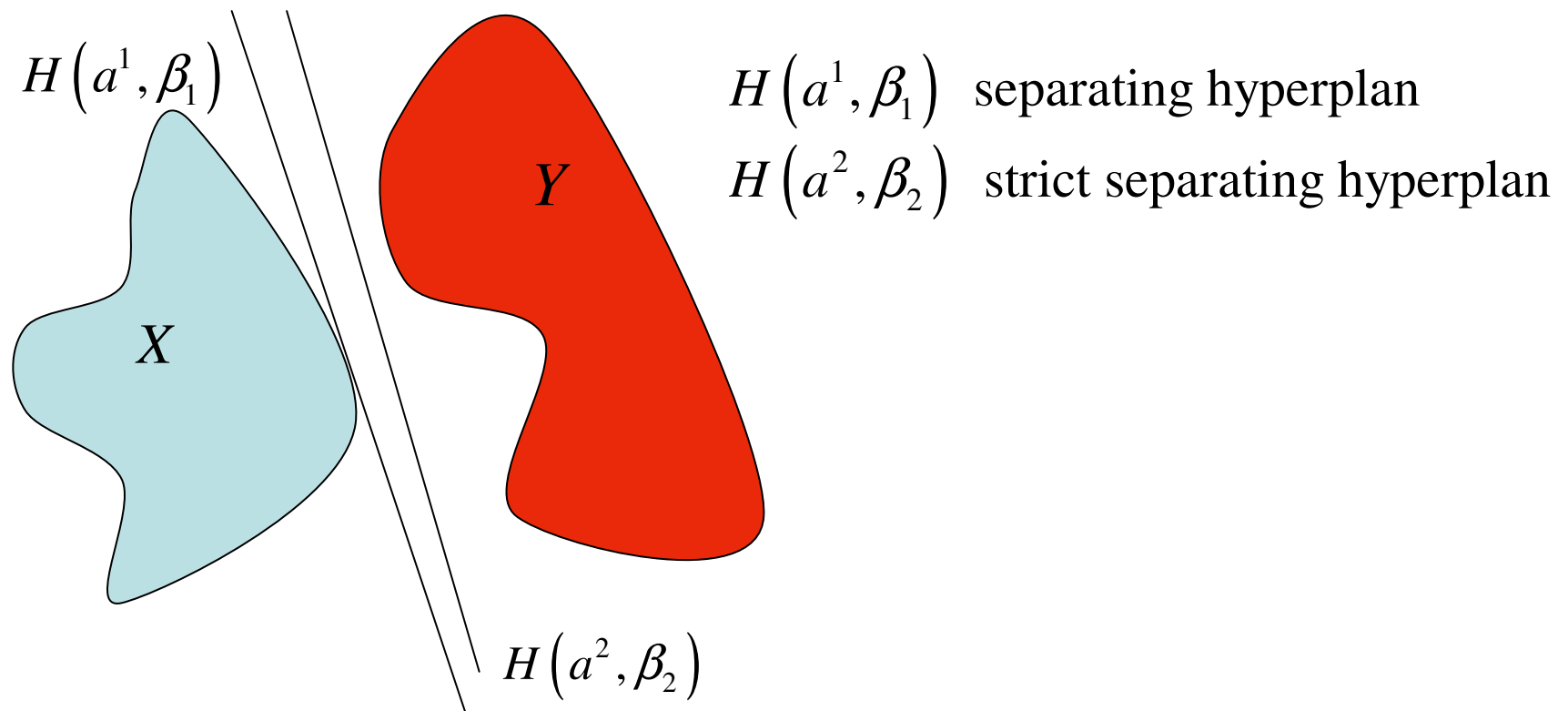
**Definition.** Separating hyperplan .

The hyperplan  $H(a, \beta)$  separates two non empty sets  $X$  and  $Y$  if

$$a^T x \geq \beta \text{ for all } x \in X \text{ (i.e., } X \subset H^+[a, \beta])$$

$$a^T y \leq \beta \text{ for all } y \in Y \text{ (i.e., } Y \subset H^-[a, \beta]).$$

The **separation** is **strict** if the inequalities in both preceding relations are strict.





**Theorem 4:** Let the vectors  $x, y, a \in R^n$ . If  $a^T y < a^T x$ , then for all  $\theta \in (0,1)$ ,

$$a^T y < a^T (\theta x + (1-\theta) y) < a^T x.$$

**Theorem 4:** Let the vectors  $x, y, a \in R^n$ . If  $a^T y < a^T x$ , then for all  $\theta \in (0, 1)$ ,

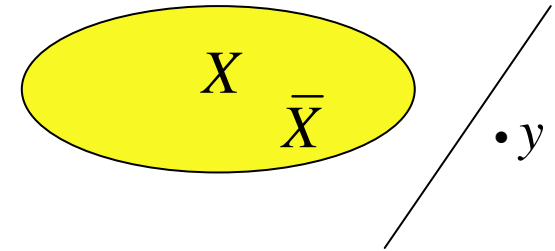
$$a^T y < a^T (\theta x + (1 - \theta) y) < a^T x.$$

**Proof.**  $a^T (\theta x + (1 - \theta) y) < a^T (\theta x + (1 - \theta) x) = a^T x.$

$$a^T (\theta x + (1 - \theta) y) > a^T (\theta y + (1 - \theta) y) = a^T y.$$

□

**Théorème 5: (Separating theorem)** If  $X \subset R^n$  is a non empty convex set and if  $y \notin \bar{X}$ , then there exists an hyperplan separating (strictly)  $\bar{X}$  and  $y$ .

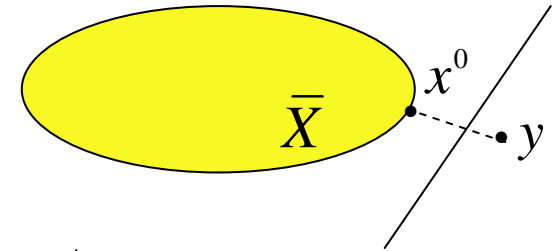


**Théorème 5: (Separating theorem)** If  $X \subset R^n$  is a non empty convex set and if  $y \notin \bar{X}$ , then there exists an hyperplan separating (strictly)  $\bar{X}$  and  $y$ .

**Proof.** There exists a point  $x^0 \in \bar{X}$  such that

$$\|x^0 - y\| = \min_{x \in \bar{X}} \|x - y\|$$

(where  $\|z\| = \sqrt{z^T z}$  denotes the euclidean norm of  $z$ ).



**Théorème 5: (Separating theorem)** If  $X \subset R^n$  is a non empty convex set and if  $y \notin \bar{X}$ , then there exists an hyperplan separating (strictly)  $\bar{X}$  and  $y$ .

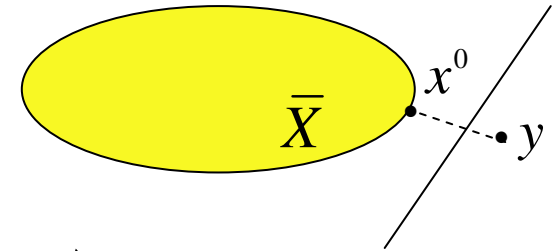
**Proof.** There exists a point  $x^0 \in \bar{X}$  such that

$$\|x^0 - y\| = \min_{x \in \bar{X}} \|x - y\|$$

(where  $\|z\| = \sqrt{z^T z}$  denotes the euclidean norm of  $z$ ).

It is easy to verify that  $\bar{X}$  is a convex set, and consequently the line segment  $\mathfrak{S}(x^0, x) \subset \bar{X}$  for all  $x \in \bar{X}$ . Then for all  $\theta \in [0, 1]$

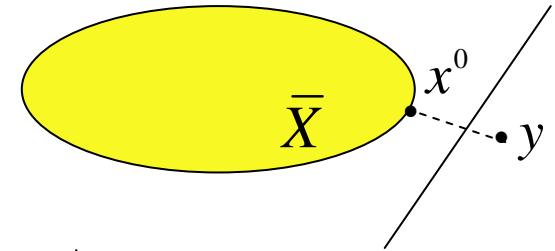
$$\|x^0 - y\| \leq \|(\theta x + (1 - \theta) x^0) - y\|.$$



**Théorème 5: (Separating theorem)** If  $X \subset R^n$  is a non empty convex set and if  $y \notin \bar{X}$ , then there exists an hyperplan separating (strictly)  $\bar{X}$  and  $y$ .

**Proof.** There exists a point  $x^0 \in \bar{X}$  such that

$$\|x^0 - y\| = \min_{x \in \bar{X}} \|x - y\|$$



(where  $\|z\| = \sqrt{z^T z}$  denotes the euclidean norm of  $z$ ).

It is easy to verify that  $\bar{X}$  is a convex set, and consequently the line segment  $\mathfrak{S}(x^0, x) \subset \bar{X}$  for all  $x \in \bar{X}$ . Then for all  $\theta \in [0, 1]$

$$\|x^0 - y\| \leq \|(\theta x + (1 - \theta)x^0) - y\|.$$

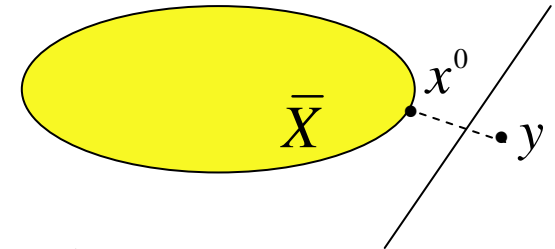
Thus for all  $\theta \in [0, 1]$

$$(x^0 - y)^T (x^0 - y) \leq ((\theta x + (1 - \theta)x^0) - y)^T ((\theta x + (1 - \theta)x^0) - y)$$

**Théorème 5: (Separating theorem)** If  $X \subset R^n$  is a non empty convex set and if  $y \notin \bar{X}$ , then there exists an hyperplan separating (strictly)  $\bar{X}$  and  $y$ .

**Proof.** There exists a point  $x^0 \in \bar{X}$  such that

$$\|x^0 - y\| = \min_{x \in \bar{X}} \|x - y\|$$



(where  $\|z\| = \sqrt{z^T z}$  denotes the euclidean norm of  $z$ ).

It is easy to verify that  $\bar{X}$  is a convex set, and consequently the line segment  $\mathfrak{S}(x^0, x) \subset \bar{X}$  for all  $x \in \bar{X}$ . Then for all  $\theta \in [0, 1]$

$$\|x^0 - y\| \leq \|(\theta x + (1 - \theta)x^0) - y\|.$$

Thus for all  $\theta \in [0, 1]$

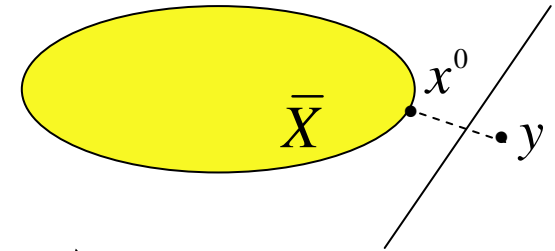
$$(x^0 - y)^T (x^0 - y) \leq ((\theta x + (1 - \theta)x^0) - y)^T ((\theta x + (1 - \theta)x^0) - y)$$

$$(x^0 - y)^T (x^0 - y) \leq ((x^0 - y) + \theta(x - x^0))^T ((x^0 - y) + \theta(x - x^0))$$

**Théorème 5: (Separating theorem)** If  $X \subset R^n$  is a non empty convex set and if  $y \notin \bar{X}$ , then there exists an hyperplan separating (strictly)  $\bar{X}$  and  $y$ .

**Proof.** There exists a point  $x^0 \in \bar{X}$  such that

$$\|x^0 - y\| = \min_{x \in \bar{X}} \|x - y\|$$



(where  $\|z\| = \sqrt{z^T z}$  denotes the euclidean norm of  $z$ ).

It is easy to verify that  $\bar{X}$  is a convex set, and consequently the line segment  $\mathfrak{S}(x^0, x) \subset \bar{X}$  for all  $x \in \bar{X}$ . Then for all  $\theta \in [0, 1]$

$$\|x^0 - y\| \leq \|(\theta x + (1 - \theta)x^0) - y\|.$$

Thus for all  $\theta \in [0, 1]$

$$(x^0 - y)^T (x^0 - y) \leq ((\theta x + (1 - \theta)x^0) - y)^T ((\theta x + (1 - \theta)x^0) - y)$$

$$(x^0 - y)^T (x^0 - y) \leq ((x^0 - y) + \theta(x - x^0))^T ((x^0 - y) + \theta(x - x^0))$$

$$(x^0 - y)^T (x^0 - y) \leq (x^0 - y)^T (x^0 - y) + 2\theta(x - x^0)^T (x^0 - y) + \theta^2(x - x^0)^T (x - x^0)$$



Thus for all  $\theta \in [0, 1]$

$$(x^0 - y)^T (x^0 - y) \leq ((\theta x + (1 - \theta)x^0) - y)^T ((\theta x + (1 - \theta)x^0) - y)$$

$$(x^0 - y)^T (x^0 - y) \leq ((x^0 - y) + \theta(x - x^0))^T ((x^0 - y) + \theta(x - x^0))$$

$$\cancel{(x^0 - y)^T (x^0 - y)} \leq \cancel{(x^0 - y)^T (x^0 - y)} + 2\theta(x - x^0)^T (x^0 - y) + \theta^2(x - x^0)^T (x - x^0).$$

Thus for all  $\theta \in [0, 1]$

$$(x^0 - y)^T (x^0 - y) \leq ((\theta x + (1 - \theta)x^0) - y)^T ((\theta x + (1 - \theta)x^0) - y)$$

$$(x^0 - y)^T (x^0 - y) \leq ((x^0 - y) + \theta(x - x^0))^T ((x^0 - y) + \theta(x - x^0))$$

$$\cancel{(x^0 - y)^T (x^0 - y)} \leq \cancel{(x^0 - y)^T (x^0 - y)} + 2\theta(x - x^0)^T (x^0 - y) + \theta^2(x - x^0)^T (x - x^0).$$

Consequently for all  $\theta \in [0, 1]$

$$2\theta(x - x^0)^T (x^0 - y) + \theta^2(x - x^0)^T (x - x^0) \geq 0.$$

Thus for all  $\theta \in [0, 1]$

$$(x^0 - y)^T (x^0 - y) \leq ((\theta x + (1 - \theta)x^0) - y)^T ((\theta x + (1 - \theta)x^0) - y)$$

$$(x^0 - y)^T (x^0 - y) \leq ((x^0 - y) + \theta(x - x^0))^T ((x^0 - y) + \theta(x - x^0))$$

$$\cancel{(x^0 - y)^T (x^0 - y)} \leq \cancel{(x^0 - y)^T (x^0 - y)} + 2\theta(x - x^0)^T (x^0 - y) + \theta^2(x - x^0)^T (x - x^0).$$

Consequently for all  $\theta \in [0, 1]$

$$2\theta(x - x^0)^T (x^0 - y) + \theta^2(x - x^0)^T (x - x^0) \geq 0.$$

But this implies  $(x - x^0)^T (x^0 - y) \geq 0$ .

Thus for all  $\theta \in [0, 1]$

$$(x^0 - y)^T (x^0 - y) \leq ((\theta x + (1 - \theta)x^0) - y)^T ((\theta x + (1 - \theta)x^0) - y)$$

$$(x^0 - y)^T (x^0 - y) \leq ((x^0 - y) + \theta(x - x^0))^T ((x^0 - y) + \theta(x - x^0))$$

$$\cancel{(x^0 - y)^T (x^0 - y)} \leq \cancel{(x^0 - y)^T (x^0 - y)} + 2\theta(x - x^0)^T (x^0 - y) + \theta^2(x - x^0)^T (x - x^0).$$

Consequently for all  $\theta \in [0, 1]$

$$2\theta(x - x^0)^T (x^0 - y) + \theta^2(x - x^0)^T (x - x^0) \geq 0.$$

But this implies  $(x - x^0)^T (x^0 - y) \geq 0$ . Indeed if

$(x - x^0)^T (x^0 - y) < 0$ , then for  $\tilde{\theta} > 0$  small enough

$$2\tilde{\theta} \left| (x - x^0)^T (x^0 - y) \right| > \tilde{\theta}^2 \left| (x - x^0)^T (x - x^0) \right|$$

Thus for all  $\theta \in [0, 1]$

$$(x^0 - y)^T (x^0 - y) \leq ((\theta x + (1 - \theta)x^0) - y)^T ((\theta x + (1 - \theta)x^0) - y)$$

$$(x^0 - y)^T (x^0 - y) \leq ((x^0 - y) + \theta(x - x^0))^T ((x^0 - y) + \theta(x - x^0))$$

$$\cancel{(x^0 - y)^T (x^0 - y)} \leq \cancel{(x^0 - y)^T (x^0 - y)} + 2\theta(x - x^0)^T (x^0 - y) + \theta^2(x - x^0)^T (x - x^0).$$

Consequently for all  $\theta \in [0, 1]$

$$2\theta(x - x^0)^T (x^0 - y) + \theta^2(x - x^0)^T (x - x^0) \geq 0.$$

But this implies  $(x - x^0)^T (x^0 - y) \geq 0$ . Indeed if

$(x - x^0)^T (x^0 - y) < 0$ , then for  $\tilde{\theta} > 0$  small enough

$$2\tilde{\theta} \left| (x - x^0)^T (x^0 - y) \right| > \tilde{\theta}^2 \left| (x - x^0)^T (x - x^0) \right|$$

and then we would have

$$2\tilde{\theta}(x - x^0)^T (x^0 - y) + \tilde{\theta}^2(x - x^0)^T (x - x^0) < 0.$$

We then have  $(x - x^0)^T (x^0 - y) \geq 0$  or

$$x^{0T} (x^0 - y) \leq x^T (x^0 - y). \quad (5.3)$$

We then have  $(x - x^0)^T (x^0 - y) \geq 0$  or

$$x^{0T} (x^0 - y) \leq x^T (x^0 - y). \quad (5.3)$$

Since  $y \notin \bar{X}$ ,  $\|x^0 - y\|^2 = (x^0 - y)^T (x^0 - y) > 0$ , or

$$y^T (x^0 - y) < x^{0T} (x^0 - y). \quad (5.4)$$

We then have  $(x - x^0)^T (x^0 - y) \geq 0$  or

$$x^{0T} (x^0 - y) \leq x^T (x^0 - y). \quad (5.3)$$

Since  $y \notin \bar{X}$ ,  $\|x^0 - y\|^2 = (x^0 - y)^T (x^0 - y) > 0$ , or

$$y^T (x^0 - y) < x^{0T} (x^0 - y). \quad (5.4)$$

Thus applying theorem 4 with the vectors  $(x^0 - y)$ ,

$x^0$ ,  $y$ , using  $\theta = \frac{1}{2}$ , and referring to (5.4)

$$y^T (x^0 - y) < \frac{1}{2} (x^0 + y)^T (x^0 - y) < x^{0T} (x^0 - y).$$



We then have  $(x - x^0)^T (x^0 - y) \geq 0$  or

$$x^{0T} (x^0 - y) \leq x^T (x^0 - y). \quad (5.3)$$

Since  $y \notin \bar{X}$ ,  $\|x^0 - y\|^2 = (x^0 - y)^T (x^0 - y) > 0$ , or

$$y^T (x^0 - y) < x^{0T} (x^0 - y). \quad (5.4)$$

Thus applying theorem 4 with the vectors  $(x^0 - y)$ ,

$x^0$ ,  $y$ , using  $\theta = \frac{1}{2}$ , and referring to (5.4)

$$y^T (x^0 - y) < \frac{1}{2} (x^0 + y)^T (x^0 - y) < x^{0T} (x^0 - y).$$

Combining with relation (5.3),

$$y^T (x^0 - y) < \frac{1}{2} (x^0 + y)^T (x^0 - y) < x^{0T} (x^0 - y) \leq x^T (x^0 - y).$$

Combining with relation (5.3),

$$y^T (x^0 - y) < \frac{1}{2} (x^0 + y)^T (x^0 - y) < x^{0T} (x^0 - y) \leq x^T (x^0 - y).$$

Combining with relation (5.3),

$$y^T (x^0 - y) < \frac{1}{2} (x^0 + y)^T (x^0 - y) < x^{0T} (x^0 - y) \leq x^T (x^0 - y).$$

Thus

$$y^T (x^0 - y) < \frac{1}{2} (x^0 + y)^T (x^0 - y) < x^T (x^0 - y).$$

Combining with relation (5.3),

$$y^T (x^0 - y) < \frac{1}{2} (x^0 + y)^T (x^0 - y) < x^{0T} (x^0 - y) \leq x^T (x^0 - y).$$

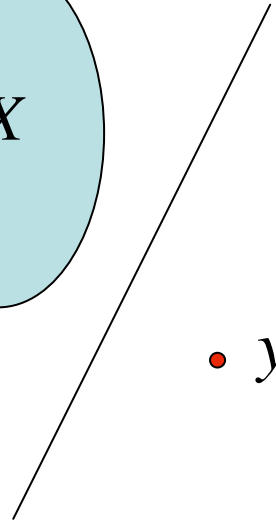
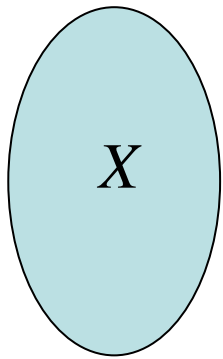
Thus

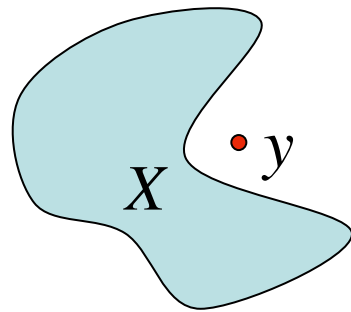
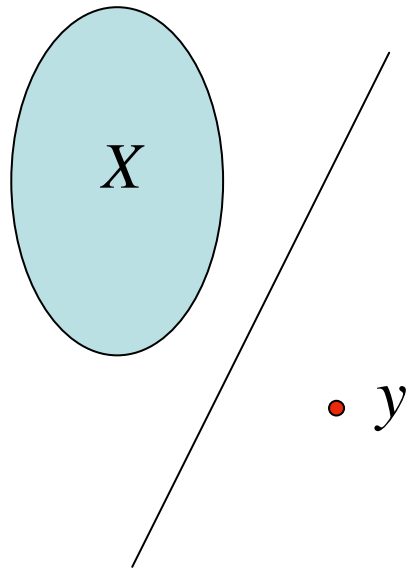
$$y^T (x^0 - y) < \frac{1}{2} (x^0 + y)^T (x^0 - y) < x^T (x^0 - y).$$

Since this relation holds for all  $x \in \bar{X}$ , it follows that

the hyperplan  $H(a, \beta)$  where  $a = (x^0 - y)$  and

$\beta = \frac{1}{2} (x^0 + y)^T (x^0 - y)$  is separating (strictly)  $\bar{X}$  and  $y$ . □





In theorem 5, the convexity of  $X$  is a sufficient condition insuring that an hyperplan exists to separate it (strictly) from  $y \notin \bar{X}$ .

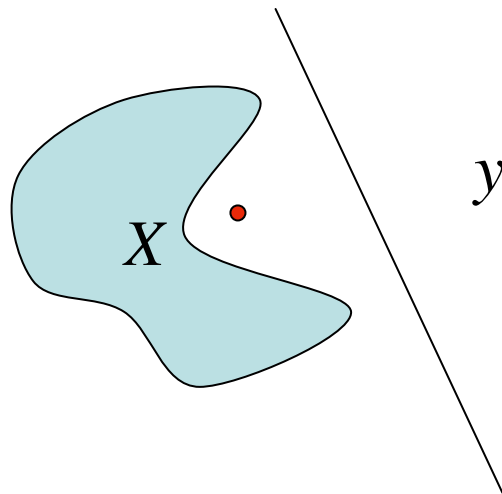
In theorem 5, the convexity of  $X$  is a sufficient condition insuring that an hyperplan exists to separate it (strictly) from  $y \notin \bar{X}$ .

But as illustrated in the following figure, the convexity of  $X$  is not a necessary condition.



In theorem 5, the convexity of  $X$  is a sufficient condition insuring that an hyperplan exists to separate it (strictly) from  $y \notin \bar{X}$ .

But as illustrated in the following figure, the convexity of  $X$  is not a necessary condition.



**Theorem 6:(Farkas Lemma)** Let the vectors

$a^1, \dots, a^n$  and  $b \in R^m$ . A sufficient condition for  $b$  to be a non negative linear combination of the  $a^j$  (i.e., a sufficient condition for the existence of non negative scalars  $x_1, \dots, x_n$  such that  $b = a^1 x_1 + \dots + a^n x_n$ ) is that each time there exists  $y \in R^m$  such that  $y^T a^j \geq 0$  for all  $j = 1, \dots, n$ , then necessarily it follows that  $y^T b \geq 0$ .

**Theorem 6:(Farkas Lemma)** Let the vectors

$a^1, \dots, a^n$  and  $b \in R^m$ . A sufficient condition for  $b$  to be a non negative linear combination of the  $a^j$  (i.e., a sufficient condition for the existence of non negative scalars  $x_1, \dots, x_n$  such that  $b = a^1 x_1 + \dots + a^n x_n$ ) is that each time there exists  $y \in R^m$  such that  $y^T a^j \geq 0$  for all  $j = 1, \dots, n$ , then necessarily it follows that  $y^T b \geq 0$ .

**Proof.** We have to show that

$$\left\{ \begin{array}{l} \text{for all } y \in R^m, \\ \text{if } y^T a^j \geq 0 \text{ for all } j = 1, \dots, n, \\ \text{then necessarily } y^T b \geq 0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \exists x_1, \dots, x_n \geq 0 \text{ such that} \\ b = a^1 x_1 + \dots + a^n x_n \end{array} \right\}$$

**Proof.** We have to show that

$$\left\{ \begin{array}{l} \text{for all } y \in R^m, \\ \text{if } y^T a^j \geq 0 \text{ for all } j = 1, \dots, n, \\ \text{then necessarily } y^T b \geq 0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \exists x_1, \dots, x_n \geq 0 \text{ such that} \\ b = a^1 x_1 + \dots + a^n x_n \end{array} \right\}$$

**Proof.** We have to show that

$$\left\{ \begin{array}{l} \text{for all } y \in R^m, \\ \text{if } y^T a^j \geq 0 \text{ for all } j = 1, \dots, n, \\ \text{then necessarily } y^T b \geq 0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \exists x_1, \dots, x_n \geq 0 \text{ such that} \\ b = a^1 x_1 + \dots + a^n x_n \end{array} \right\}$$

We will rather show the contrapose of the implication:

$$\neg \left\{ \begin{array}{l} \text{for all } y \in R^m, \\ \text{if } y^T a^j \geq 0 \text{ for all } j = 1, \dots, n, \\ \text{then necessarily } y^T b \geq 0 \end{array} \right\} \Leftarrow \neg \left\{ \begin{array}{l} \exists x_1, \dots, x_n \geq 0 \text{ such that} \\ b = a^1 x_1 + \dots + a^n x_n \end{array} \right\}$$

**Proof.** We have to show that

$$\left\{ \begin{array}{l} \text{for all } y \in R^m, \\ \text{if } y^T a^j \geq 0 \text{ for all } j = 1, \dots, n, \\ \text{then necessarily } y^T b \geq 0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \exists x_1, \dots, x_n \geq 0 \text{ such that} \\ b = a^1 x_1 + \dots + a^n x_n \end{array} \right\}$$

We will rather show the contrapose of the implication:

$$\neg \left\{ \begin{array}{l} \text{for all } y \in R^m, \\ \text{if } y^T a^j \geq 0 \text{ for all } j = 1, \dots, n, \\ \text{then necessarily } y^T b \geq 0 \end{array} \right\} \Leftarrow \neg \left\{ \begin{array}{l} \exists x_1, \dots, x_n \geq 0 \text{ such that} \\ b = a^1 x_1 + \dots + a^n x_n \end{array} \right\}$$

i.e.,

$$\left\{ \begin{array}{l} \nexists x_1, \dots, x_n \geq 0 \text{ such that} \\ b = a^1 x_1 + \dots + a^n x_n \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \exists y \in R^m, \text{ such that} \\ y^T a^j \geq 0 \text{ for all } j = 1, \dots, n, \\ \text{and } y^T b < 0 \end{array} \right\}$$

To show that

$$\left\{ \begin{array}{l} \nexists x_1, \dots, x_n \geq 0 \text{ such that} \\ b = a^1 x_1 + \dots + a^n x_n \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \exists y \in R^m, \text{ such that} \\ y^T a^j \geq 0 \text{ for all } j = 1, \dots, n, \\ \text{and } y^T b < 0 \end{array} \right\}$$

consider the following set

$$Z = \left\{ z \in R^m : \exists x_1, \dots, x_n \geq 0 \text{ such that } z = a^1 x_1 + \dots + a^n x_n \right\}.$$

To show that

$$\left\{ \begin{array}{l} \nexists x_1, \dots, x_n \geq 0 \text{ such that} \\ b = a^1 x_1 + \dots + a^n x_n \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \exists y \in R^m, \text{ such that} \\ y^T a^j \geq 0 \text{ for all } j = 1, \dots, n, \\ \text{and } y^T b < 0 \end{array} \right\}$$

consider the following set

$$Z = \left\{ z \in R^m : \exists x_1, \dots, x_n \geq 0 \text{ such that } z = a^1 x_1 + \dots + a^n x_n \right\}.$$

It is easy to show that  $Z$  is convex. It is also possible to show that  $Z$  is a close set (i.e.  $Z = \bar{Z}$ ).



To show that

$$\left\{ \begin{array}{l} \nexists x_1, \dots, x_n \geq 0 \text{ such that} \\ b = a^1 x_1 + \dots + a^n x_n \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \exists y \in R^m, \text{ such that} \\ y^T a^j \geq 0 \text{ for all } j = 1, \dots, n, \\ \text{and } y^T b < 0 \end{array} \right\}$$

consider the following set

$$Z = \left\{ z \in R^m : \exists x_1, \dots, x_n \geq 0 \text{ such that } z = a^1 x_1 + \dots + a^n x_n \right\}.$$

It is easy to show that  $Z$  is convex. It is also possible to show that  $Z$  is a close set (i.e.  $Z = \bar{Z}$ ).

$$z^1 = a^1 x_1^1 + \dots + a^n x_n^1$$

$$z^2 = a^1 x_1^2 + \dots + a^n x_n^2$$

$$\theta z^1 + (1 - \theta) z^2 = a^1 \left[ \theta x_1^1 + (1 - \theta) x_1^2 \right] + \dots + a^n \left[ \theta x_n^1 + (1 - \theta) x_n^2 \right]$$

$$\text{and then } \theta z^1 + (1 - \theta) z^2 \in Z$$

To show that

$$\left\{ \begin{array}{l} \nexists x_1, \dots, x_n \geq 0 \text{ such that} \\ b = a^1 x_1 + \dots + a^n x_n \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \exists y \in R^m, \text{ such that} \\ y^T a^j \geq 0 \text{ for all } j = 1, \dots, n, \\ \text{and } y^T b < 0 \end{array} \right\}$$

consider the following set

$$Z = \left\{ z \in R^m : \exists x_1, \dots, x_n \geq 0 \text{ such that } z = a^1 x_1 + \dots + a^n x_n \right\}.$$

It is easy to show that  $Z$  is convex. It is also possible to show that  $Z$  is a close set (i.e.  $Z = \bar{Z}$ ).

To show that

$$\left\{ \begin{array}{l} \nexists x_1, \dots, x_n \geq 0 \text{ such that} \\ b = a^1 x_1 + \dots + a^n x_n \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \exists y \in R^m, \text{ such that} \\ y^T a^j \geq 0 \text{ for all } j = 1, \dots, n, \\ \text{and } y^T b < 0 \end{array} \right\}$$

consider the following set

$$Z = \left\{ z \in R^m : \exists x_1, \dots, x_n \geq 0 \text{ such that } z = a^1 x_1 + \dots + a^n x_n \right\}.$$

It is easy to show that  $Z$  is convex. It is also possible to show that  $Z$  is a close set (i.e.  $Z = \bar{Z}$ ).

Since that an assumption of the contrapose is that  $b \notin Z = \bar{Z}$ , then by theorem 5, there exists an hyperplan  $H(p, \beta)$  separating strictly  $Z$  and  $b$ :

$$p^T b < \beta < p^T z \quad \text{for all } z \in Z. \quad (5.5)$$

$$Z = \left\{ z \in R^m : \exists x_1, \dots, x_n \geq 0 \text{ such that } z = a^1 x_1 + \dots + a^n x_n \right\}.$$

Since that an assumption of the contrapose is that  $b \notin Z = \bar{Z}$ , then by theorem 5, there exists an hyperplan  $H(p, \beta)$  separating strictly  $Z$  and  $b$ :

$$p^T b < \beta < p^T z \quad \text{for all } z \in Z. \quad (5.5)$$

$$Z = \left\{ z \in R^m : \exists x_1, \dots, x_n \geq 0 \text{ such that } z = a^1 x_1 + \dots + a^n x_n \right\}.$$

Since that an assumption of the contrapose is that  $b \notin Z = \bar{Z}$ , then by theorem 5, there exists an hyperplan  $H(p, \beta)$  separating strictly  $Z$  and  $b$ :

$$p^T b < \beta < p^T z \quad \text{for all } z \in Z. \quad (5.5)$$

But  $0 \in Z$ , implying that  $\beta < 0$ , and consequently

$$p^T b < 0.$$

$$Z = \left\{ z \in R^m : \exists x_1, \dots, x_n \geq 0 \text{ such that } z = a^1 x_1 + \dots + a^n x_n \right\}.$$

Since that an assumption of the contrapose is that  $b \notin Z = \bar{Z}$ , then by theorem 5, there exists an hyperplan  $H(p, \beta)$  separating strictly  $Z$  and  $b$ :

$$p^T b < \beta < p^T z \quad \text{for all } z \in Z. \quad (5.5)$$

But  $0 \in Z$ , implying that  $\beta < 0$ , and consequently

$$p^T b < 0.$$

Furthermore,  $p^T z \geq 0$  for all  $z \in Z$ .

$$Z = \left\{ z \in R^m : \exists x_1, \dots, x_n \geq 0 \text{ such that } z = a^1 x_1 + \dots + a^n x_n \right\}.$$

Since that an assumption of the contrapose is that  $b \notin Z = \bar{Z}$ , then by theorem 5, there exists an hyperplan  $H(p, \beta)$  separating strictly  $Z$  and  $b$ :

$$p^T b < \beta < p^T z \quad \text{for all } z \in Z. \quad (5.5)$$

But  $0 \in Z$ , implying that  $\beta < 0$ , and consequently

$$p^T b < 0.$$

Furthermore,  $p^T z \geq 0$  for all  $z \in Z$ . Indeed, if  $\tilde{z} \in Z$  were existing such that  $p^T \tilde{z} < 0$ , then since  $Z$  is a cone (i.e., if  $z \in Z$  then

$\lambda z \in Z$  for all  $\lambda \geq 0$ ), we would get that  $p^T (\lambda \tilde{z}) \xrightarrow{\lambda \rightarrow \infty} -\infty$

contradicting (5.5).

But  $0 \in Z$ , implying that  $\beta < 0$ , and consequently

$$p^T b < 0.$$

Furthermore,  $p^T z \geq 0$  for all  $z \in Z$ . Indeed, if  $\tilde{z} \in Z$  were existing such that  $p^T \tilde{z} < 0$ , then since  $Z$  is a cone (i.e., if  $z \in Z$  then  $\lambda z \in Z$  for all  $\lambda \geq 0$ ), we would get that  $p^T (\lambda \tilde{z}) \xrightarrow{\lambda \rightarrow \infty} -\infty$  contradicting (5.5).



But  $0 \in Z$ , implying that  $\beta < 0$ , and consequently

$$p^T b < 0.$$

Furthermore,  $p^T z \geq 0$  for all  $z \in Z$ . Indeed, if  $\tilde{z} \in Z$  were existing such that  $p^T \tilde{z} < 0$ , then since  $Z$  is a cone (i.e., if  $z \in Z$  then  $\lambda z \in Z$  for all  $\lambda \geq 0$ ), we would get that  $p^T (\lambda \tilde{z}) \xrightarrow{\lambda \rightarrow \infty} -\infty$  contradicting (5.5).

Since it is easy to verify that  $a^j \in Z$ ,  $j = 1, \dots, n$ ,

$$Z = \left\{ z \in R^m : \exists x_1, \dots, x_n \geq 0 \text{ such that } z = a^1 x_1 + \dots + a^n x_n \right\}$$

it follows that  $p^T a^j \geq 0$ ,  $j = 1, \dots, n$ .

But  $0 \in Z$ , implying that  $\beta < 0$ , and consequently

$$p^T b < 0.$$

Furthermore,  $p^T z \geq 0$  for all  $z \in Z$ . Indeed, if  $\tilde{z} \in Z$  were existing such that  $p^T \tilde{z} < 0$ , then since  $Z$  is a cone (i.e., if  $z \in Z$  then  $\lambda z \in Z$  for all  $\lambda \geq 0$ ), we would get that  $p^T (\lambda \tilde{z}) \xrightarrow{\lambda \rightarrow \infty} -\infty$  contradicting (5.5).

Since it is easy to verify that  $a^j \in Z$ ,  $j = 1, \dots, n$ ,

$$Z = \left\{ z \in R^m : \exists x_1, \dots, x_n \geq 0 \text{ such that } z = a^1 x_1 + \dots + a^n x_n \right\}$$

it follows that  $p^T a^j \geq 0$ ,  $j = 1, \dots, n$ .

We have shown the contrapose since  $p \in R^m$  is such that  $p^T a^j \geq 0$ ,  $j = 1, \dots, n$  and  $p^T b < 0$ . □

**Theorem 6:(Farkas Lemma)** Let the vectors

$a^1, \dots, a^n$  and  $b \in R^m$ . A sufficient condition for  $b$  to be a non negative linear combination of the  $a^j$  (i.e., a sufficient condition for the existence of non negative scalars  $x_1, \dots, x_n$  such that  $b = a^1 x_1 + \dots + a^n x_n$ ) is that each time there exists  $y \in R^m$  such that  $y^T a^j \geq 0$  for all  $j = 1, \dots, n$ , then necessarily it follows that  $y^T b \geq 0$ .

**Proof.** We have to show that

$$\left\{ \begin{array}{l} \text{for all } y \in R^m, \\ \text{if } y^T a^j \geq 0 \text{ for all } j = 1, \dots, n, \\ \text{then necessarily } y^T b \geq 0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \exists x_1, \dots, x_n \geq 0 \text{ such that} \\ b = a^1 x_1 + \dots + a^n x_n \end{array} \right\}$$

**Corollary 7: (Theorem of alternatives)** Let  $A$  be a matrix  $m \times n$ . Exactly one of the two following alternatives holds:

I The system  $Ax = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $A^T y \geq 0, b^T y < 0$  has a solution  $y \in R^m$ .

I The system  $[a_{\bullet 1} \ a_{\bullet 2} \ \dots \ a_{\bullet n}]x = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n, b^T y < 0$  has a solution  $y \in R^m$ .

**Corollary 7: (Theorem of alternatives)** Let  $A$  be a matrix  $m \times n$ . Exactly one of the two following alternatives holds:

I The system  $Ax = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $A^T y \geq 0, b^T y < 0$  has a solution  $y \in R^m$ .

**Proof.** It is easy to verify that the two alternatives cannot hold simultaneously, since otherwise the following relation would be verified:

$$0 > b^T y = x^T A^T y \geq 0$$

a contradiction.

I The system  $Ax = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $A^T y \geq 0, b^T y < 0$  has a solution  $y \in R^m$ .

I The system  $[a_{\bullet 1} \ a_{\bullet 2} \ \dots \ a_{\bullet n}]x = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n, b^T y < 0$  has a solution  $y \in R^m$ .

Let  $a_{\bullet j}, j = 1, \dots, n$ , be the  $j^{\text{ième}}$  column of  $A$ .

- I The system  $Ax = b, x \geq 0$  has a solution  $x \in R^n$
- II The system  $A^T y \geq 0, b^T y < 0$  has a solution  $y \in R^m$ .

- I The system  $[a_{\bullet 1} \ a_{\bullet 2} \ \dots \ a_{\bullet n}]x = b, x \geq 0$  has a solution  $x \in R^n$
- II The system  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n, b^T y < 0$  has a solution  $y \in R^m$ .

Let  $a_{\bullet j}, j = 1, \dots, n$ , be the  $j^{\text{ième}}$  column of  $A$ .

Consider the alternative II. It is verified or not.

I The system  $Ax = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $A^T y \geq 0, b^T y < 0$  has a solution  $y \in R^m$ .

I The system  $[a_{\bullet 1} \ a_{\bullet 2} \ \dots \ a_{\bullet n}]x = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n, b^T y < 0$  has a solution  $y \in R^m$ .

Let  $a_{\bullet j}, j = 1, \dots, n$ , be the  $j^{\text{ième}}$  column of  $A$ .

Consider the alternative II. It is verified or not.

In the case where it is not holding, it follows that the system

$A^T y \geq 0, b^T y < 0$  has no solution  $y \in R^m$ ;



I The system  $Ax = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $A^T y \geq 0, b^T y < 0$  has a solution  $y \in R^m$ .

I The system  $[a_{\bullet 1} \ a_{\bullet 2} \ \dots \ a_{\bullet n}]x = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n, b^T y < 0$  has a solution  $y \in R^m$ .

Let  $a_{\bullet j}, j = 1, \dots, n$ , be the  $j^{\text{ième}}$  column of  $A$ .

Consider the alternative II. It is verified or not.

In the case where it is not holding, it follows that the system

$A^T y \geq 0, b^T y < 0$  has no solution  $y \in R^m$ ;

i.e., the system  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n, b^T y < 0$  has no solution  $y \in R^m$ .

I The system  $Ax = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $A^T y \geq 0, b^T y < 0$  has a solution  $y \in R^m$ .

I The system  $[a_{\bullet 1} \ a_{\bullet 2} \ \dots \ a_{\bullet n}]x = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n, b^T y < 0$  has a solution  $y \in R^m$ .

Let  $a_{\bullet j}, j = 1, \dots, n$ , be the  $j^{\text{ième}}$  column of  $A$ .

Consider the alternative II. It is verified or not.

In the case where it is not holding, it follows that the system

$A^T y \geq 0, b^T y < 0$  has no solution  $y \in R^m$ ;

i.e., the system  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n, b^T y < 0$  has no

solution  $y \in R^m$ . Then for all  $y \in R^m$ ,

if  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n$ , then necessarily  $b^T y \geq 0$ .

I The system  $Ax = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $A^T y \geq 0, b^T y < 0$  has a solution  $y \in R^m$ .

I The system  $[a_{\bullet 1} \ a_{\bullet 2} \ \dots \ a_{\bullet n}]x = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n, b^T y < 0$  has a solution  $y \in R^m$ .

Let  $a_{\bullet j}, j = 1, \dots, n$ , be the  $j^{\text{ième}}$  column of  $A$ .

Consider the alternative II. It is verified or not.

In the case where it is not holding, it follows that the system

$A^T y \geq 0, b^T y < 0$  has no solution  $y \in R^m$ ;

i.e., the system  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n, b^T y < 0$  has no

solution  $y \in R^m$ . Then for all  $y \in R^m$ ,

if  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n$ , then necessarily  $b^T y \geq 0$ .

Then by theorem 6, there exists a vector  $x \in R^n, x \geq 0$  such

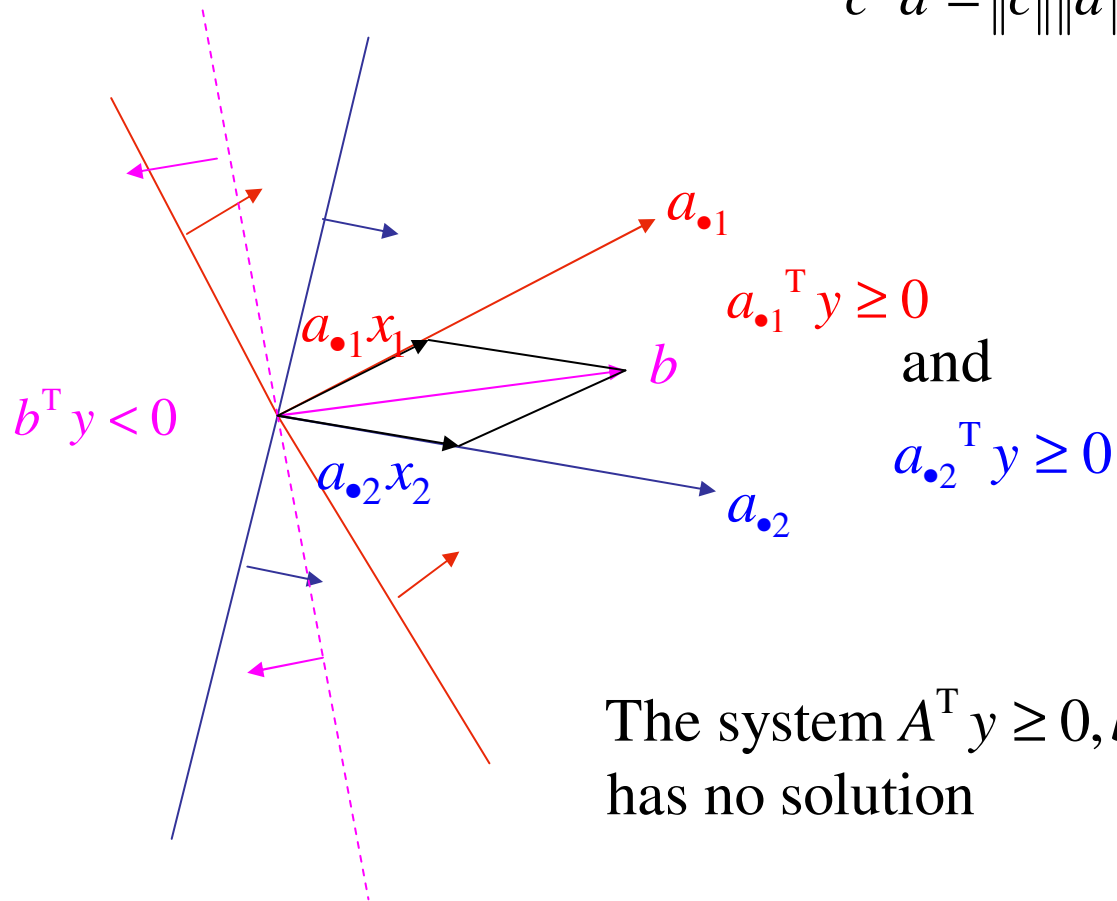
that  $Ax = a_{\bullet 1}x_1 + \dots + a_{\bullet n}x_n = b$ , and alternative I holds. □

- I The system  $Ax = b, x \geq 0$  has a solution  $x \in R^n$
- II The system  $A^T y \geq 0, b^T y < 0$  has a solution  $y \in R^m$

- I The system  $[a_{\bullet 1} \ a_{\bullet 2} \ \dots \ a_{\bullet n}]x = b, x \geq 0$  has a solution  $x \in R^n$
- II The system  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n, b^T y < 0$  has a solution  $y \in R^m$ .

The case where the system  $Ax = b, x \geq 0$  has a solution.

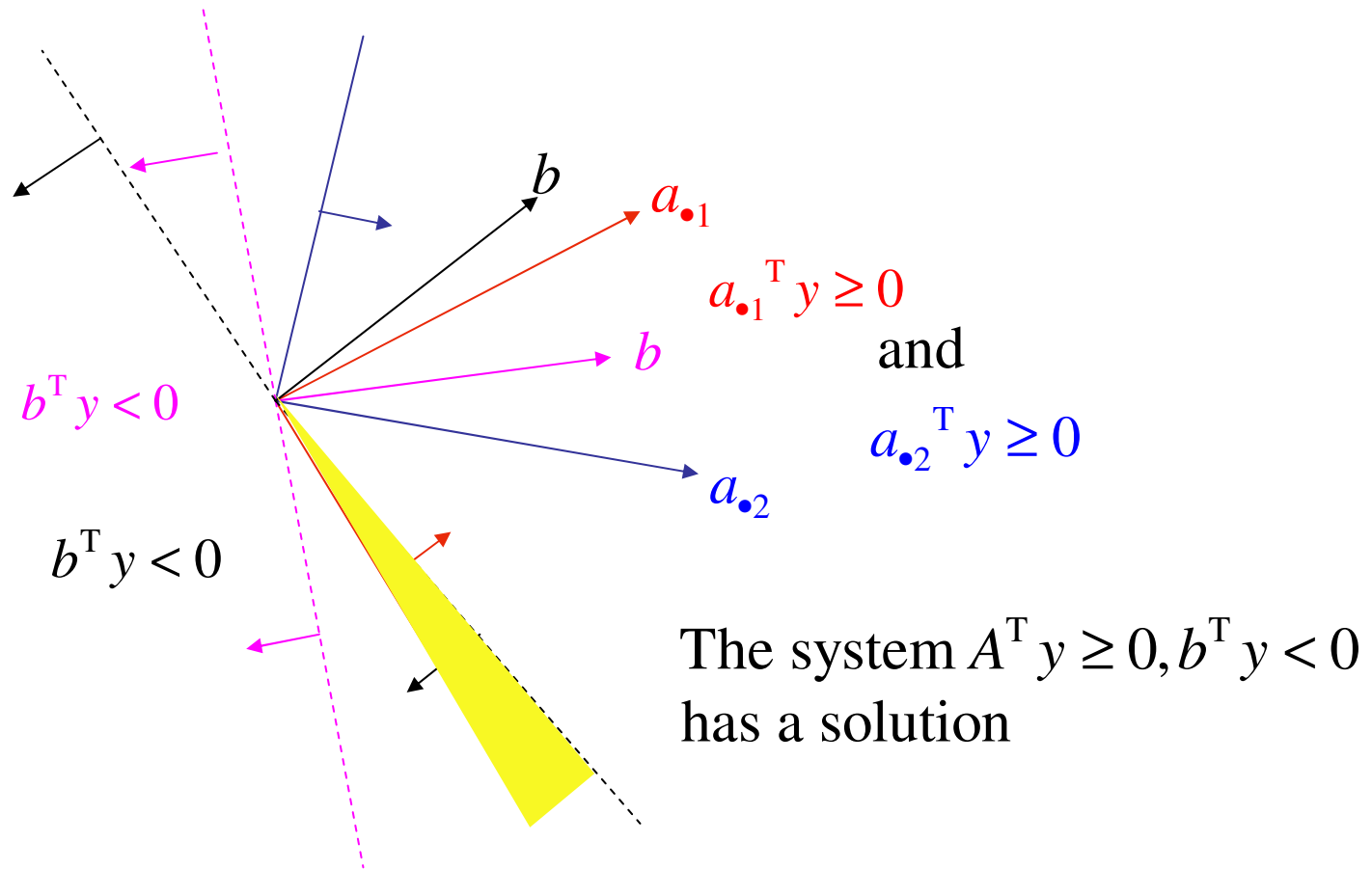
$$c^T d = \|c\| \|d\| \cos \theta$$



- I The system  $Ax = b, x \geq 0$  has a solution  $x \in R^n$
- II The system  $A^T y \geq 0, b^T y < 0$  has a solution  $y \in R^m$

- I The system  $[a_{\bullet 1} \ a_{\bullet 2} \ \dots \ a_{\bullet n}]x = b, x \geq 0$  has a solution  $x \in R^n$
- II The system  $a_{\bullet j}^T y \geq 0, j = 1, \dots, n, b^T y < 0$  has a solution  $y \in R^m$ .

The case where the system  $Ax = b, x \geq 0$  has no solution.



Now return to analyse the necessity of the K-K-T conditions using Corollary 7.

Now return to analyse the necessity of the K-K-T conditions using Corollary 7.

Suppose that  $x^* \in X$ ,  $f_i(x^*) \leq 0$ ,  $i = 1, \dots, m$ , is a local optimal solution of problem 2.

$$\begin{array}{ll} \text{Min} & f(x) \\ \text{Sujet à} & f_i(x) \leq 0 \quad i = 1, \dots, m \\ & x \in X \end{array} \quad (2)$$

Now return to analyse the necessity of the K-K-T conditions using Corollary 7.

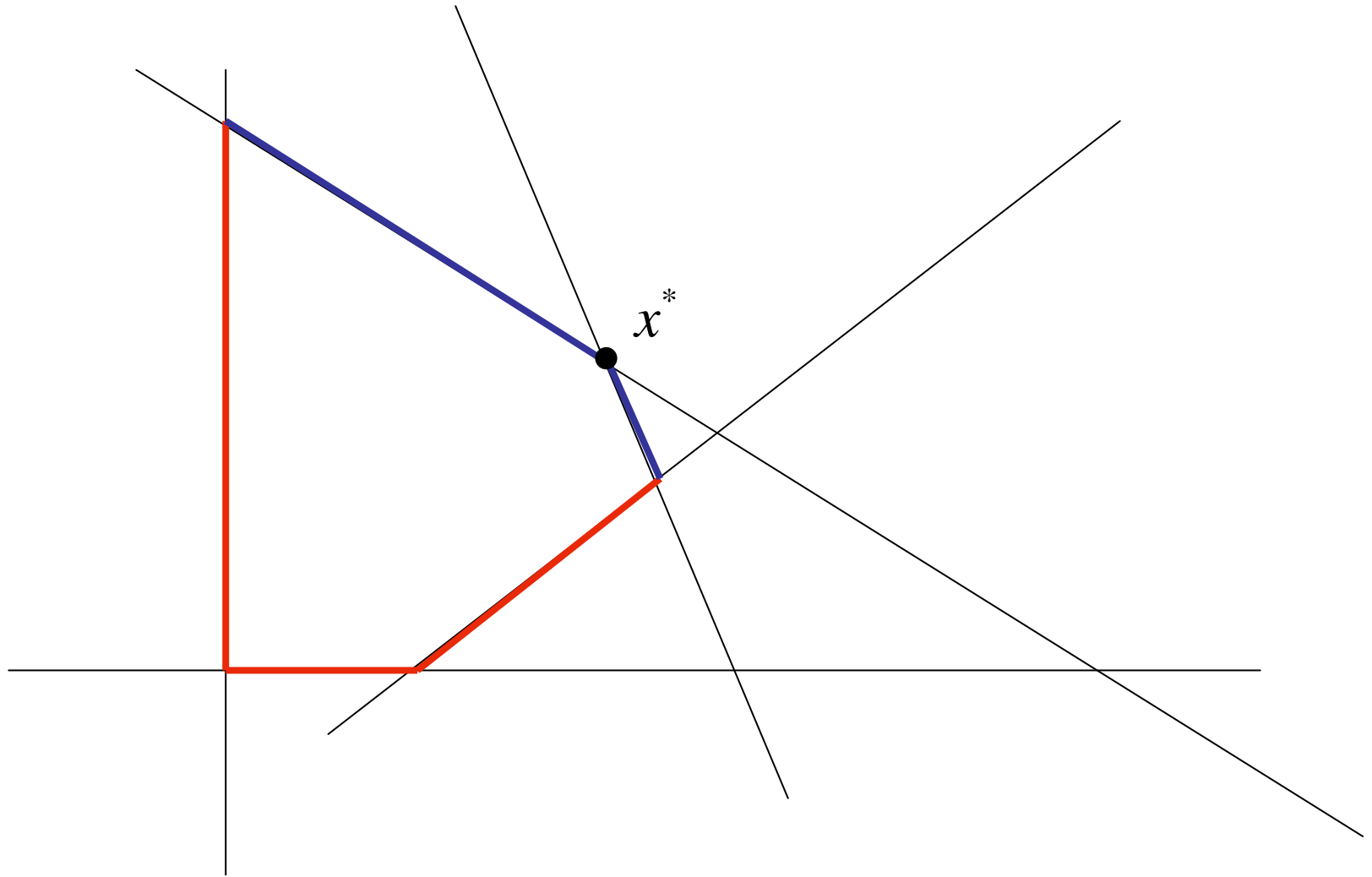
Suppose that  $x^* \in X$ ,  $f_i(x^*) \leq 0$ ,  $i = 1, \dots, m$ , is a local optimal solution of problem 2.

$$\begin{array}{ll} \text{Min} & f(x) \\ \text{Sujet à} & f_i(x) \leq 0 \quad i = 1, \dots, m \\ & x \in X \end{array} \quad (2)$$

**Notation:** Denote the set of **active constraints**

$$A(x^*) = \{i : f_i(x^*) = 0\} = \{i_1, \dots, i_k\} \subset \{1, \dots, m\}.$$





Now return to analyse the necessity of the K-K-T conditions using Corollary 7.

Suppose that  $x^* \in X$ ,  $f_i(x^*) \leq 0$ ,  $i = 1, \dots, m$ , is a local optimal solution of problem 2.

$$\begin{array}{ll} \text{Min} & f(x) \\ \text{Sujet à} & f_i(x) \leq 0 \quad i = 1, \dots, m \\ & x \in X \end{array} \quad (2)$$

**Notation:** Denote the set of **active constraints**

$$A(x^*) = \{i : f_i(x^*) = 0\} = \{i_1, \dots, i_k\} \subset \{1, \dots, m\}.$$

**Assumption to be verified :** Suppose that we can show that there is no vector  $d \in R^n$  such that

$$\begin{array}{ll} \nabla f_i(x^*)^T d \leq 0 & i \in A(x^*) \\ \nabla f(x^*)^T d < 0 \end{array} \quad (5.6)$$

**Assumption to be verified** : Suppose that we can show that

there is no vector  $d \in R^n$  such that

$$\nabla f_i(x^*)^T d \leq 0$$

$$\nabla f(x^*)^T d < 0$$

$$i \in A(x^*)$$

$$\begin{bmatrix} \nabla f_{i_1}(x^*)^T d \leq 0 \\ \vdots \\ \nabla f_{i_k}(x^*)^T d \leq 0 \end{bmatrix}$$

$$\begin{bmatrix} \nabla f_{i_1}(x^*)^T \\ \vdots \\ \nabla f_{i_k}(x^*)^T \end{bmatrix} d \leq 0$$

**Assumption to be verified** : Suppose that we can show that

there is no vector  $d \in R^n$  such that

$$\nabla f_i(x^*)^T d \leq 0$$

$$\nabla f(x^*)^T d < 0$$

$$i \in A(x^*)$$

$$\begin{bmatrix} \nabla f_{i_1}(x^*)^T d \leq 0 \\ \vdots \\ \nabla f_{i_k}(x^*)^T d \leq 0 \end{bmatrix}$$

$$\begin{bmatrix} \nabla f_{i_1}(x^*)^T \\ \vdots \\ \nabla f_{i_k}(x^*)^T \end{bmatrix} d \leq 0$$

Then the system

$$\left[ \nabla f_{i_1}(x^*), \dots, \nabla f_{i_k}(x^*) \right]^T d \leq 0,$$

$$\nabla f(x^*)^T d < 0$$

has no solution.

**Assumption to be verified** : Suppose that we can show that

there is no vector  $d \in R^n$  such that

$$\nabla f_i (x^*)^T d \leq 0 \quad i \in A(x^*)$$

$$\nabla f (x^*)^T d < 0$$

$$\begin{bmatrix} \nabla f_{i_1} (x^*)^T d \leq 0 \\ \vdots \\ \nabla f_{i_k} (x^*)^T d \leq 0 \end{bmatrix}$$

$$\begin{bmatrix} \nabla f_{i_1} (x^*)^T \\ \vdots \\ \nabla f_{i_k} (x^*)^T \end{bmatrix} d \leq 0$$

Then the system

$$\left[ \nabla f_{i_1} (x^*), \dots, \nabla f_{i_k} (x^*) \right]^T d \leq 0,$$

$$\nabla f (x^*)^T d < 0$$

has no solution.

Now apply Corollary 7:

$$\left\{ \text{II} : - \left[ \nabla f_{i_1} (x^*), \dots, \nabla f_{i_k} (x^*) \right]^T d \geq 0, \nabla f (x^*)^T d < 0 \text{ has no solution } d \in R^n \right\}$$

implies that

$$\left\{ \text{I} : - \left[ \nabla f_{i_1} (x^*), \dots, \nabla f_{i_k} (x^*) \right] \tilde{\lambda}^* = \nabla f (x^*), \tilde{\lambda}^* = \left[ \lambda_{i_1}^*, \dots, \lambda_{i_k}^* \right] \geq 0 \text{ has a solution} \right\}$$

**Assumption to be verified** : Suppose that we can show that

there is no vector  $d \in R^n$  such that

$$\begin{aligned} \nabla f_i(x^*)^T d &\leq 0 & i \in A(x^*) \\ \nabla f(x^*)^T d &< 0 \end{aligned}$$

$$\begin{bmatrix} \nabla f_{i_1}(x^*)^T d \leq 0 \\ \vdots \\ \nabla f_{i_k}(x^*)^T d \leq 0 \end{bmatrix}$$

$$\begin{bmatrix} \nabla f_{i_1}(x^*)^T \\ \vdots \\ \nabla f_{i_k}(x^*)^T \end{bmatrix} d \leq 0$$

Then the system

$$\begin{bmatrix} \nabla f_{i_1}(x^*), \dots, \nabla f_{i_k}(x^*) \end{bmatrix}^T d \leq 0, \quad \nabla f(x^*)^T d < 0$$

has no solution.

Now apply Corollary 7:

$$\left\{ \text{II} : - \begin{bmatrix} \nabla f_{i_1}(x^*), \dots, \nabla f_{i_k}(x^*) \end{bmatrix}^T d \geq 0, \nabla f(x^*)^T d < 0 \text{ has no solution } d \in R^n \right\}$$

implies that

$$\left\{ \text{I} : - \begin{bmatrix} \nabla f_{i_1}(x^*), \dots, \nabla f_{i_k}(x^*) \end{bmatrix} \tilde{\lambda}^* = \nabla f(x^*), \tilde{\lambda}^* = [\lambda_{i_1}^*, \dots, \lambda_{i_k}^*] \geq 0 \text{ has a solution} \right\}$$

**Corollary 7: (Theorem of alternatives)** Let  $A$  be a matrix  $m \times n$ . Exactly one of the two following alternatives holds:

I The system  $Ax = b, x \geq 0$  has a solution  $x \in R^n$

II The system  $A^T y \geq 0, b^T y < 0$  has a solution  $y \in R^m$ .

Now apply Corollary 7:

$$\left\{ \text{II} : -\left[ \nabla f_{i_1}(x^*), \dots, \nabla f_{i_k}(x^*) \right]^T d \geq 0, \nabla f(x^*)^T d < 0 \text{ has no solution } d \in R^n \right\}$$

implies that

$$\left\{ \text{I} : -\left[ \nabla f_{i_1}(x^*), \dots, \nabla f_{i_k}(x^*) \right] \tilde{\lambda}^* = \nabla f(x^*), \tilde{\lambda}^* = \left[ \lambda_{i_1}^*, \dots, \lambda_{i_k}^* \right] \geq 0 \text{ has a solution} \right\}.$$

Now apply Corollary 7:

$$\left\{ \text{II: } -\left[ \nabla f_{i_1}(x^*), \dots, \nabla f_{i_k}(x^*) \right]^T d \geq 0, \nabla f(x^*)^T d < 0 \text{ has no solution } d \in R^n \right\}$$

implies that

$$\left\{ \text{I: } -\left[ \nabla f_{i_1}(x^*), \dots, \nabla f_{i_k}(x^*) \right] \tilde{\lambda}^* = \nabla f(x^*), \tilde{\lambda}^* = \left[ \lambda_{i_1}^*, \dots, \lambda_{i_k}^* \right] \geq 0 \text{ has a solution} \right\}.$$

This can be written as

$$\left\{ \text{I: } - \sum_{i \in A(x^*)} \lambda_i^* \nabla f_i(x^*) = \nabla f(x^*), \tilde{\lambda}^* = \left[ \lambda_{i_1}^*, \dots, \lambda_{i_k}^* \right] \geq 0 \text{ has a solution} \right\}$$



Now apply Corollary 7:

$$\left\{ \text{II} : - \left[ \nabla f_{i_1} (x^*), \dots, \nabla f_{i_k} (x^*) \right]^T d \geq 0, \nabla f (x^*)^T d < 0 \text{ has no solution } d \in R^n \right\}$$

implies that

$$\left\{ \text{I} : - \left[ \nabla f_{i_1} (x^*), \dots, \nabla f_{i_k} (x^*) \right] \tilde{\lambda}^* = \nabla f (x^*), \tilde{\lambda}^* = \left[ \lambda_{i_1}^*, \dots, \lambda_{i_k}^* \right] \geq 0 \text{ has a solution} \right\}.$$

This can be written as

$$\left\{ \text{I} : - \sum_{i \in A(x^*)} \lambda_i^* \nabla f_i (x^*) = \nabla f (x^*), \tilde{\lambda}^* = \left[ \lambda_{i_1}^*, \dots, \lambda_{i_k}^* \right] \geq 0 \text{ has a solution} \right\}$$

Let  $\lambda_i^* = 0$  for all  $i \notin A(x^*)$ . Then

$$\sum_{i \in A(x^*)} \lambda_i^* \nabla f_i (x^*) = 0$$

$$\lambda_i^* f_i (x^*) = 0 \quad \text{for all } i \notin A(x^*).$$

This can be written as

$$\left\{ \text{I: } - \sum_{i \in A(x^*)} \lambda_i^* \nabla f_i(x^*) = \nabla f(x^*), \tilde{\lambda}^* = [\lambda_{i_1}^*, \dots, \lambda_{i_k}^*] \geq 0 \text{ has a solution} \right\}$$

Let  $\lambda_i^* = 0$  for all  $i \notin A(x^*)$ . Then

$$\sum_{i \notin A(x^*)} \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad \text{for all } i \notin A(x^*).$$

This can be written as

$$\left\{ \text{I: } - \sum_{i \in A(x^*)} \lambda_i^* \nabla f_i(x^*) = \nabla f(x^*), \tilde{\lambda}^* = [\lambda_{i_1}^*, \dots, \lambda_{i_k}^*] \geq 0 \text{ has a solution} \right\}$$

Let  $\lambda_i^* = 0$  for all  $i \notin A(x^*)$ . Then

$$\sum_{i \notin A(x^*)} \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad \text{for all } i \notin A(x^*).$$

Hence we obtain the K-K-T conditions

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, n$$

$$f_i(x^*) \leq 0 \quad i = 1, \dots, n$$

$$\lambda_i^* \geq 0 \quad i = 1, \dots, n$$

This can be written as

$$\left\{ \text{I: } - \sum_{i \in A(x^*)} \lambda_i^* \nabla f_i(x^*) = \nabla f(x^*), \tilde{\lambda}^* = [\lambda_{i_1}^*, \dots, \lambda_{i_k}^*] \geq 0 \text{ has a solution} \right\}$$

Let  $\lambda_i^* = 0$  for all  $i \notin A(x^*)$ . Then

$$\sum_{i \notin A(x^*)} \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \text{ for all } i \notin A(x^*).$$

Hence we obtain the K-K-T conditions

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, n$$

$$f_i(x^*) \leq 0 \quad i = 1, \dots, n$$

$$\lambda_i^* \geq 0 \quad i = 1, \dots, n$$

This can be written as

$$\left\{ \text{I: } - \sum_{i \in A(x^*)} \lambda_i^* \nabla f_i(x^*) = \nabla f(x^*), \tilde{\lambda}^* = [\lambda_{i_1}^*, \dots, \lambda_{i_k}^*] \geq 0 \text{ has a solution} \right\}$$

Let  $\lambda_i^* = 0$  for all  $i \notin A(x^*)$ . Then

$$\sum_{i \notin A(x^*)} \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad \text{for all } i \notin A(x^*).$$

Hence we obtain the K-K-T conditions

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, n$$

$$f_i(x^*) \leq 0 \quad i = 1, \dots, n$$

$$\lambda_i^* \geq 0 \quad i = 1, \dots, n$$

BUT

**Assumption to be verified** : Suppose that we can show that there is no vector  $d \in R^n$  such that

$$\begin{aligned} \nabla f_i(x^*)^T d &\leq 0 \quad i \in A(x^*) & (5.6) \\ \nabla f(x^*)^T d &< 0 \end{aligned}$$

is not necessarily verified for all local solution  $x^*$  for all problems as illustrated in the following example.

$$\text{Min } f(x_1, x_2) = -x_1$$

$$\text{s.t. } f_1(x_1, x_2) = (x_1 - 1)^3 + x_2 \leq 0$$

$$f_2(x_1, x_2) = -x_1 \leq 0$$

$$f_3(x_1, x_2) = -x_2 \leq 0.$$

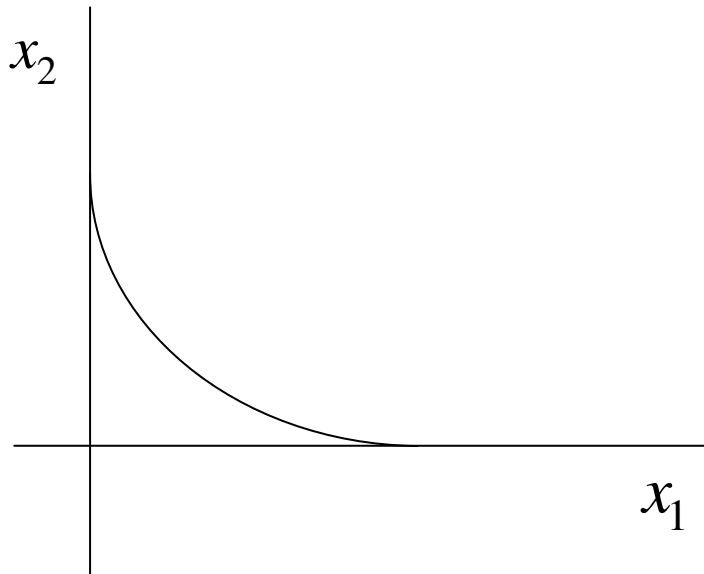
$$\text{Min } f(x_1, x_2) = -x_1$$

$$\text{s.t. } f_1(x_1, x_2) = (x_1 - 1)^3 + x_2 \leq 0$$

$$f_2(x_1, x_2) = -x_1 \leq 0$$

$$f_3(x_1, x_2) = -x_2 \leq 0.$$

The feasible domain of this problem is illustrated in the figure below under the curve  $f_1(x_1, x_2)$ , above the  $x_1$  axis, and on the right of the  $x_2$  axis.





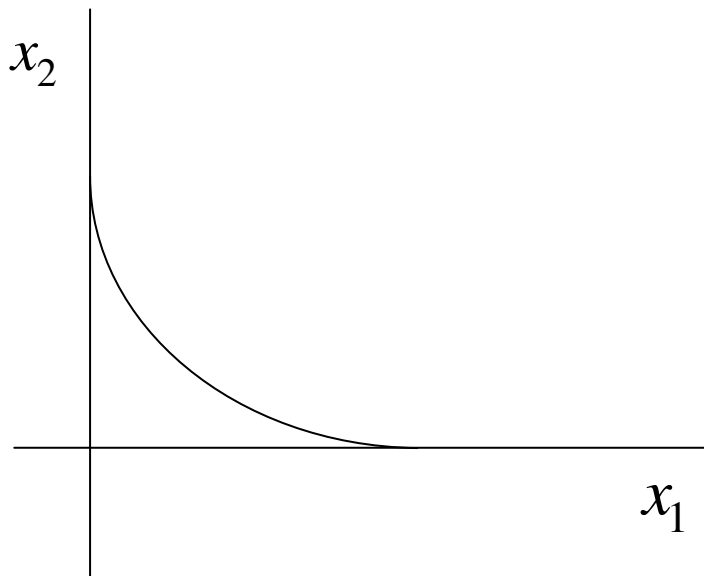
$$\text{Min } f(x_1, x_2) = -x_1$$

$$\text{s.t. } f_1(x_1, x_2) = (x_1 - 1)^3 + x_2 \leq 0$$

$$f_2(x_1, x_2) = -x_1 \leq 0$$

$$f_3(x_1, x_2) = -x_2 \leq 0.$$

The feasible domain of this problem is illustrated in the figure below under the curve  $f_1(x_1, x_2)$ , above the  $x_1$  axis, and on the right of the  $x_2$  axis.



It is easy to verify that

$x^* = [1, 0]^T$  is a global optimal solution of this problem.

Moreover,  $A(x^*) = \{1, 3\}$ .

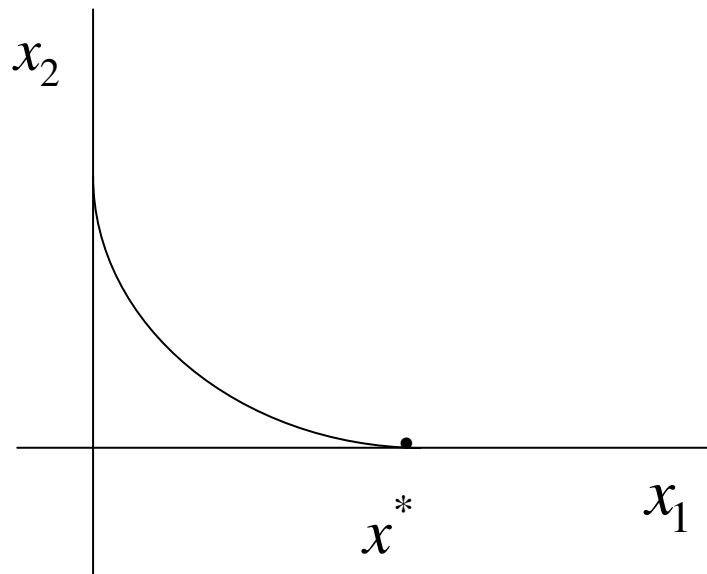
$$\text{Min } f(x_1, x_2) = -x_1$$

$$\text{s.t. } f_1(x_1, x_2) = (x_1 - 1)^3 + x_2 \leq 0$$

$$f_2(x_1, x_2) = -x_1 \leq 0$$

$$f_3(x_1, x_2) = -x_2 \leq 0.$$

The feasible domain of this problem is illustrated in the figure below under the curve  $f_1(x_1, x_2)$ , above the  $x_1$  axis, and on the right of the  $x_2$  axis.



It is easy to verify that

$x^* = [1, 0]^T$  is a global optimal solution of this problem.

Moreover,  $A(x^*) = \{1, 3\}$ .

$$\text{Min } f(x_1, x_2) = -x_1$$

$$\text{s.t. } f_1(x_1, x_2) = (x_1 - 1)^3 + x_2 \leq 0$$

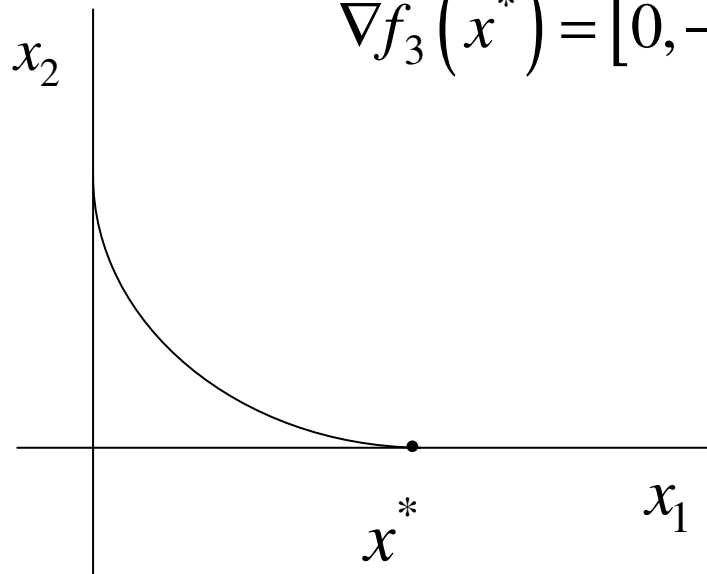
$$f_2(x_1, x_2) = -x_1 \leq 0$$

$$f_3(x_1, x_2) = -x_2 \leq 0.$$

$$\nabla f(x^*) = [-1, 0]^T,$$

$$\nabla f_1(x) = [3(x_1 - 1)^2, 1]^T \text{ et } \nabla f_1(x^*) = [0, 1]^T$$

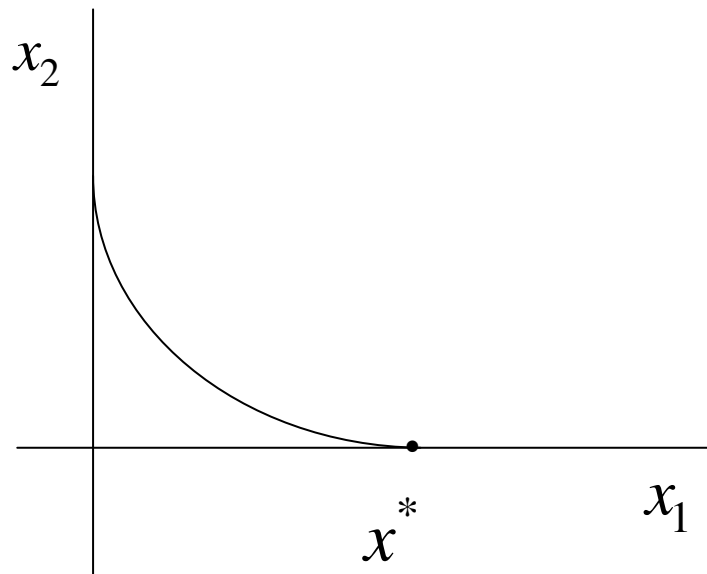
$$\nabla f_3(x^*) = [0, -1]^T$$



$$\nabla f(x^*) = [-1, 0]^T,$$

$$\nabla f_1(x) = [3(x_1 - 1)^2, 1] \text{ et } \nabla f_1(x^*) = [0, 1]^T$$

$$\nabla f_3(x^*) = [0, -1]^T$$



$$\nabla f(x^*) = [-1, 0]^T,$$

$$\nabla f_1(x) = [3(x_1 - 1)^2, 1] \text{ et } \nabla f_1(x^*) = [0, 1]^T$$

$$\nabla f_3(x^*) = [0, -1]^T$$

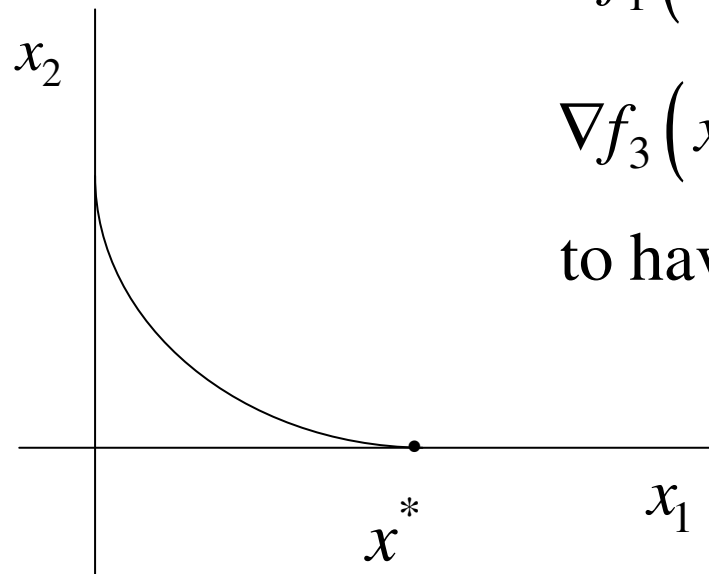
We would like the system

$$\nabla f(x^*)^T d < 0$$

$$\nabla f_1(x^*)^T d \leq 0$$

$$\nabla f_3(x^*)^T d \leq 0$$

to have no solution  $d$



$$\nabla f(x^*) = [-1, 0]^T,$$

$$\nabla f_1(x) = [3(x_1 - 1)^2, 1] \text{ et } \nabla f_1(x^*) = [0, 1]^T$$

$$\nabla f_3(x^*) = [0, -1]^T$$

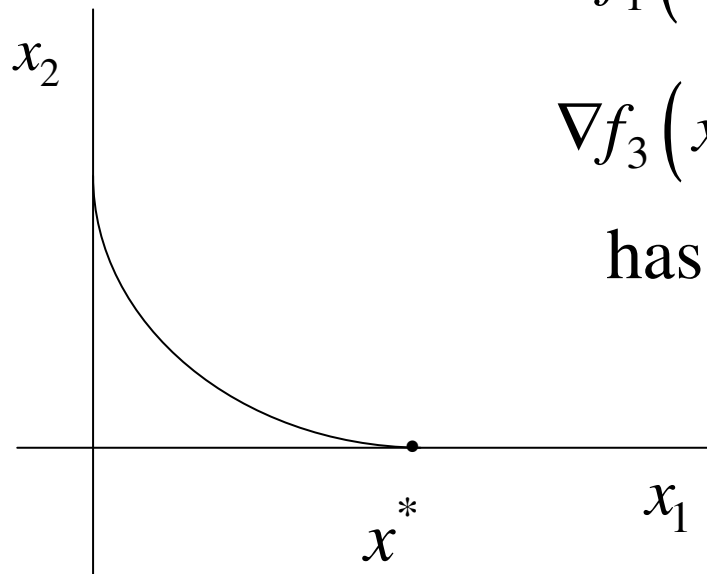
But the system

$$\nabla f(x^*)^T d = -d_1 < 0$$

$$\nabla f_1(x^*)^T d = d_2 \leq 0$$

$$\nabla f_3(x^*)^T d = -d_2 \leq 0$$

has solution  $d = [1, 0]$ .



But the system

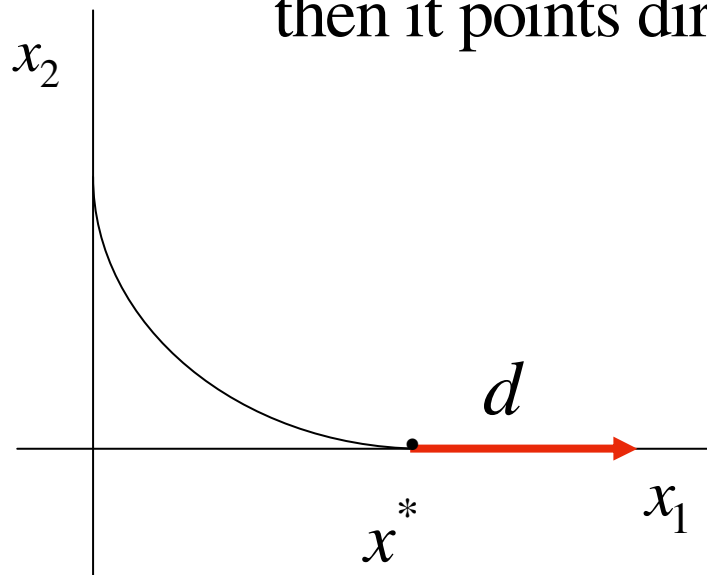
$$\nabla f(x^*)^T d = -d_1 < 0$$

$$\nabla f_1(x^*)^T d = d_2 \leq 0$$

$$\nabla f_3(x^*)^T d = -d_2 \leq 0$$

has a solution  $d = [1, 0]$ .

If we look at the 1a direction  $d = [1, 0]$  at the point  $x^*$ , then it points directly out of the feasible domain.



We will restrict the constraints of the problems to eliminate those where such a situation exists.

# Kuhn-Tucker constraints qualification

**Notation:**  $\tilde{R}$  denotes the feasible domain of problem 2

$$\delta\tilde{R} = \left\{ x \in \tilde{R} : \exists i, 1 \leq i \leq m, \text{ such that } f_i(x) = 0 \right\}.$$

**Definition.** Let the point  $\bar{x} \in \delta\tilde{R}$ ,  $f_1, \dots, f_m$  satisfy the **constraints qualification** at the point  $\bar{x}$  if for any vector  $\hat{d}$  where the system

$\nabla f_i(\bar{x})^T \hat{d} \leq 0, i \in A(\bar{x})$ , is verified, there exists a differentiable function  $\alpha: [0, 1] \rightarrow \tilde{R}$  such that  $\alpha(0) = \bar{x}$  and  $\alpha'(0) = \sigma \hat{d}, \sigma > 0$ .



**Notation:**  $\tilde{R}$  denotes the feasible domain of problem (5.2)

$$\delta\tilde{R} = \{x \in \tilde{R} : \exists i, 1 \leq i \leq m, \text{ such that } f_i(x) = 0\}.$$

**Definition.** Let the point  $\bar{x} \in \delta\tilde{R}$ ,  $f_1, \dots, f_m$  satisfy the **constraints qualification** at the point  $\bar{x}$  if for any vector  $\hat{d}$  where the system  $\nabla f_i(\bar{x})^T \hat{d} \leq 0$ ,  $i \in A(\bar{x})$ , is verified, there exists a differentiable function  $\alpha: [0,1] \rightarrow \tilde{R}$  such that  $\alpha(0) = \bar{x}$  and  $\alpha'(0) = \sigma \hat{d}$ ,  $\sigma > 0$ .

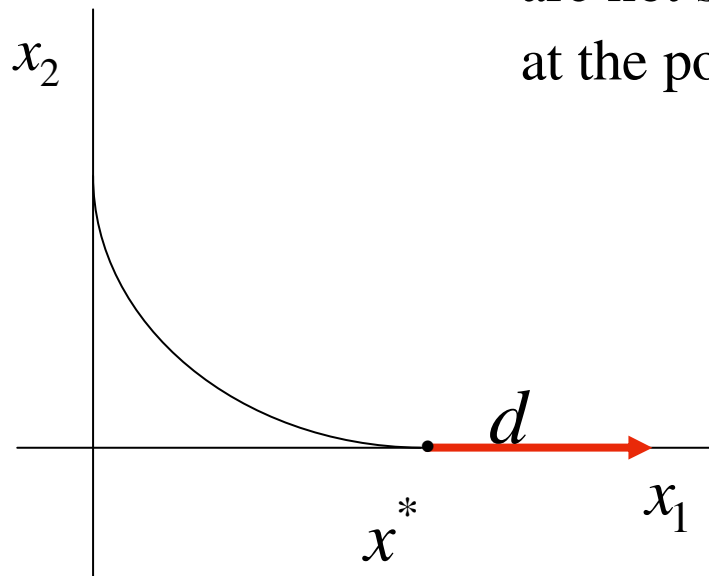
In the preceding example, the constraints  $f_1, f_2, f_3$  are not satisfying the constraints qualification at the point  $x^* \in \delta\tilde{R}$ .

**Notation:**  $\tilde{R}$  denotes the feasible domain of problem (5.2)

$$\delta\tilde{R} = \{x \in \tilde{R} : \exists i, 1 \leq i \leq m, \text{ such that } f_i(x) = 0\}.$$

**Definition.** Let the point  $\bar{x} \in \delta\tilde{R}$ ,  $f_1, \dots, f_m$  satisfy the **constraints qualification** at the point  $\bar{x}$  if for any vector  $\hat{d}$  where the system  $\nabla f_i(\bar{x})^T \hat{d} \leq 0$ ,  $i \in A(\bar{x})$ , is verified, there exists a differentiable function  $\alpha: [0, 1] \rightarrow \tilde{R}$  such that  $\alpha(0) = \bar{x}$  and  $\alpha'(0) = \sigma \hat{d}$ ,  $\sigma > 0$ .

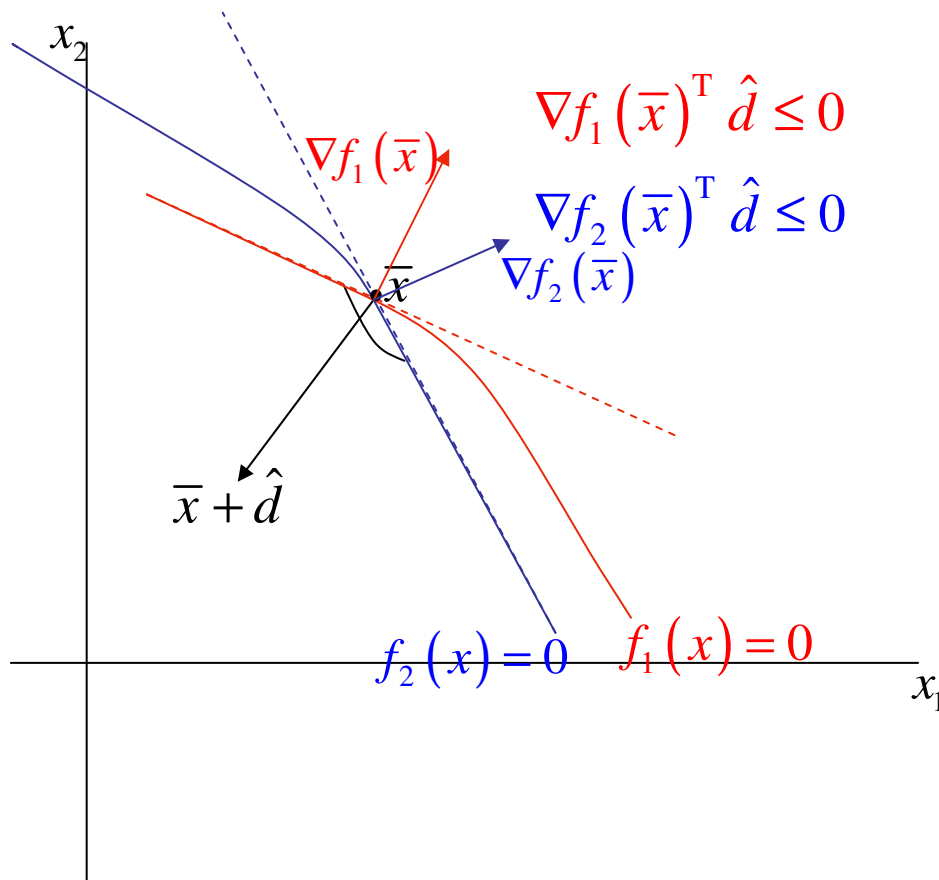
In the preceding example, the constraints  $f_1, f_2, f_3$  are not satisfying the constraints qualification at the point  $x^* \in \delta\tilde{R}$ .



Indeed, there is no differentiable function  $\alpha$  taking its values in  $\tilde{R}$  and having its slope at 0 equal to a positive multiple of  $d$ , since  $d$  points out directly outside  $\tilde{R}$ .

**Definition.** Let the point  $\bar{x} \in \delta\tilde{R}$ ,  $f_1, \dots, f_m$  satisfy the **constraints qualification** at the point  $\bar{x}$  if for any vector  $\hat{d}$  where the system  $\nabla f_i(\bar{x})^T \hat{d} \leq 0$ ,  $i \in A(\bar{x})$ , is verified, there exists a differentiable function  $\alpha: [0,1] \rightarrow \tilde{R}$  such that  $\alpha(0) = \bar{x}$  and  $\alpha'(0) = \sigma \hat{d}$ ,  $\sigma > 0$ .

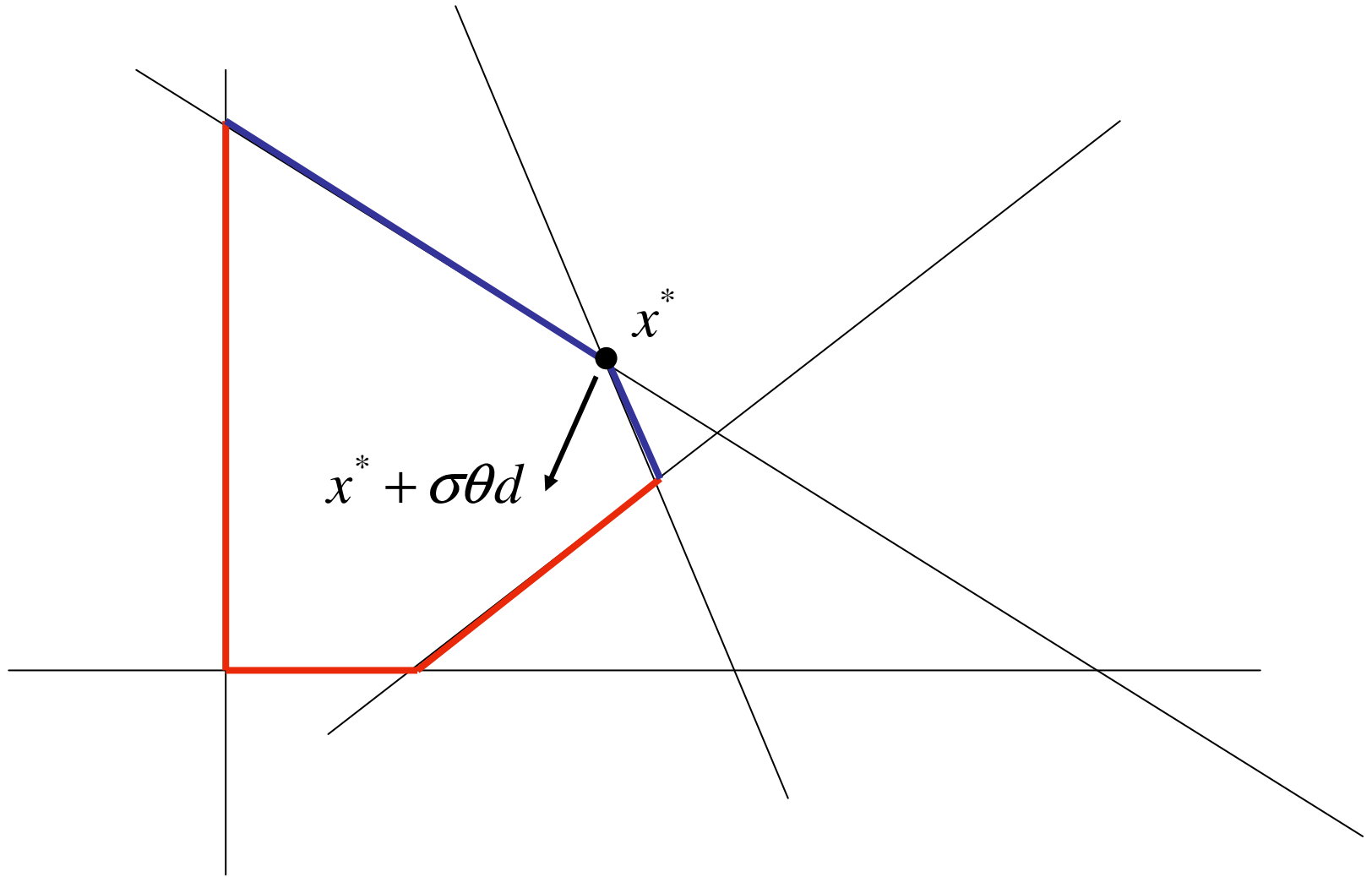
**Graphic interpretation.**



$$\alpha(\theta) = \bar{x} + \theta \hat{d}$$

$$\alpha(0) = \bar{x}$$

$$\alpha'(0) = \hat{d}$$



**Theorem 8: (Necessity of K-K-T optimality conditions)** Let  $x^* \in X$  be a local optimal solution of problem 2, and let  $X$  be open. Furthermore, if  $x^* \in \delta\tilde{R}$ , then  $f_1, \dots, f_m$  satisfy the constraints qualification at point  $x^*$ . Then there exists a vector of multipliers  $\lambda^* = [\lambda_1^*, \dots, \lambda_m^*] \geq 0$  such that

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0$$

$\begin{aligned} &\text{Min } f(x) \\ &\text{s.t. } f_i(x) \leq 0 \quad i = 1, \dots, m \\ &\quad x \in X \end{aligned}$
--

$$i = 1, \dots, m.$$

**Theorem 8: (Necessity of K-K-T optimality conditions)** Let  $x^* \in X$  be a local optimal solution of problem 2, and let  $X$  be open. Furthermore, if  $x^* \in \delta\tilde{R}$ , then  $f_1, \dots, f_m$  satisfy the constraints qualification at point  $x^*$ . Then there exists a vector of multipliers  $\lambda^* = [\lambda_1^*, \dots, \lambda_m^*] \geq 0$  such that

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, m.$$

$\begin{aligned} \text{Min } & f(x) \\ \text{s.t. } & f_i(x) \leq 0 \quad i = 1, \dots, m \\ & x \in X \end{aligned}$
---

**Proof.** If  $x^*$  is a point inside the feasible domain  $\tilde{R}$  (i.e.,  $f_i(x^*) < 0$  for all  $i$ ), then let  $\lambda_i^* = 0$  for all  $i = 1, \dots, m$ .

**Theorem 8: (Necessity of K-K-T optimality conditions)** Let  $x^* \in X$  be a local optimal solution of problem 2, and let  $X$  be open. Furthermore, if  $x^* \in \delta\tilde{R}$ , then  $f_1, \dots, f_m$  satisfy the constraints qualification at point  $x^*$ . Then there exists a vector of multipliers  $\lambda^* = [\lambda_1^*, \dots, \lambda_m^*] \geq 0$  such that

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, m.$$

$\begin{aligned} &\text{Min } f(x) \\ &\text{s.t. } f_i(x) \leq 0 \quad i = 1, \dots, m \\ &\quad x \in X \end{aligned}$
--

**Proof.** If  $x^*$  is a point inside the feasible domain  $\tilde{R}$  (i.e.,  $f_i(x^*) < 0$  for all  $i$ ), then let  $\lambda_i^* = 0$  for all  $i = 1, \dots, m$ . Indeed, in this case, if  $\nabla f(x^*)$  would take a value different from 0, then the direction  $d = -\nabla f(x^*)$  would be a descent direction for  $f$  at  $x^*$ .

## Recall

**Lemma:** Let  $X \subset R^n$ ,  $f \in C^1 / X$ , and  $x \in X$ . If  $d \in R^n$  is a feasible direction at  $x$  and  $\nabla f(x)^T d < 0$ , then there exists a scalar  $\xi > 0$  such that for all  $0 < \tau \leq \xi$   
 $f(x + \tau d) < f(x)$ .  
(i.e.,  $d$  is a **descent direction** at  $x$ .)

**Proof :** Since  $\lim_{\theta \rightarrow 0} \frac{f(x + \theta d) - f(x)}{\theta} = \nabla f(x)^T d < 0$ , then there exists a scalar  $\xi > 0$  such that for all  $\tau \neq 0$ ,  $-\xi \leq \tau \leq \xi$ ,  
$$\frac{f(x + \tau d) - f(x)}{\tau} < 0.$$

Then restrict  $\tau$  to be positive in order to have

$$f(x + \tau d) - f(x) < 0 \quad \text{ou} \quad f(x + \tau d) < f(x). \quad \square$$



**Theorem 8: (Necessity of K-K-T optimality conditions)** Let  $x^* \in X$  be a local optimal solution of problem 2, and let  $X$  be open. Furthermore, if  $x^* \in \delta\tilde{R}$ , then  $f_1, \dots, f_m$  satisfy the constraints qualification at point  $x^*$ . Then there exists a vector of multipliers  $\lambda^* = [\lambda_1^*, \dots, \lambda_m^*] \geq 0$  such that

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \dots, m.$$

$\begin{aligned} &\text{Min } f(x) \\ &\text{s.t. } f_i(x) \leq 0 \quad i = 1, \dots, m \\ &\quad x \in X \end{aligned}$
--

**Proof.** If  $x^*$  is a point inside the feasible domain  $\tilde{R}$  (i.e.,  $f_i(x^*) < 0$  for all  $i$ ), then let  $\lambda_i^* = 0$  for all  $i = 1, \dots, m$ . Indeed, in this case, if  $\nabla f(x^*)$  would take a value different from 0, then the direction  $d = -\nabla f(x^*)$  would be a descent direction for  $f$  at  $x^*$ . Hence, it would exist a scalar  $\tau > 0$  sufficiently small to have  $(x^* + \tau d) \in B_\varepsilon(x^*) \cap \tilde{R}$  and  $f(x^* + \tau d) < f(x^*)$ , a contradiction.

Let  $x^* \in \delta\tilde{R}$ . Let us show that

**Assumption to be verified** : Suppose that we can show that there is no vector  $d \in R^n$  such that

$$\nabla f_i(x^*)^T d \leq 0 \quad i \in A(x^*) \quad (5.6)$$

$$\nabla f(x^*)^T d < 0$$

is verified under the assumptions of the theorem.

Let  $x^* \in \delta\tilde{R}$ . Let us show that

**Assumption to be verified** : Suppose that we can show that

there is no vector  $d \in R^n$  such that

$$\nabla f_i(x^*)^T d \leq 0 \quad i \in A(x^*) \quad (5.6)$$

$$\nabla f(x^*)^T d < 0$$

is verified under the assumptions of the theorem. Indeed, for contradiction

assume that such a vector  $\hat{d}$  would exist. Since  $f_1, \dots, f_m$  satisfy the constraints qualification at point  $x^*$ , then there exists a differentiable function  $\alpha: [0, 1] \rightarrow \tilde{R}$  such that  $\alpha(0) = x^*$  and  $\alpha'(0) = \sigma \hat{d}, \sigma > 0$ .

Let  $x^* \in \delta \tilde{R}$ . Let us show that

**Assumption to be verified** : Suppose that we can show that

there is no vector  $d \in R^n$  such that

$$\nabla f_i (x^*)^T d \leq 0 \quad i \in A(x^*) \quad (5.6)$$

$$\nabla f (x^*)^T d < 0$$

is verified under the assumptions of the theorem. Indeed, for contradiction

assume that such a vector  $\hat{d}$  would exist. Since  $f_1, \dots, f_m$  satisfy the constraints qualification at point  $x^*$ , then there exists a differentiable function  $\alpha : [0, 1] \rightarrow \tilde{R}$

such that  $\alpha(0) = x^*$  and  $\alpha'(0) = \sigma \hat{d}$ ,  $\sigma > 0$ . Thus,

$$\lim_{\theta \rightarrow 0} \frac{f(x + \theta d) - f(x)}{\theta} = \nabla f(x)^T d$$

$$\lim_{\theta \rightarrow 0} \frac{f(\alpha(\theta)) - f(x^*)}{\theta} = \nabla f(x^*)^T \alpha'(0) = \sigma \nabla f(x^*)^T \hat{d} < 0,$$

implying the existence of  $\hat{\theta} \in [0, 1]$  sufficiently small to have  $\alpha(\hat{\theta}) \in B_\epsilon(x^*)$

such that  $f(\alpha(\hat{\theta})) < f(x^*)$ , a contradiction since  $\alpha(\hat{\theta}) \in \tilde{R}$ .

Let  $x^* \in \delta\tilde{R}$ . Let us show that

**Assumption to be verified** : Suppose that we can show that

there is no vector  $d \in R^n$  such that

$$\nabla f_i(x^*)^T d \leq 0 \quad i \in A(x^*) \quad (5.6)$$

$$\nabla f(x^*)^T d < 0$$

is verified under the assumptions of the theorem. Indeed, for contradiction

assume that such a vector  $\hat{d}$  would exist. Since  $f_1, \dots, f_m$  satisfy the constraints qualification at point  $x^*$ , then there exists a differentiable function  $\alpha: [0, 1] \rightarrow \tilde{R}$

such that  $\alpha(0) = x^*$  and  $\alpha'(0) = \sigma \hat{d}$ ,  $\sigma > 0$ . Thus,

$$\lim_{\theta \rightarrow 0} \frac{f(x + \theta d) - f(x)}{\theta} = \nabla f(x)^T d$$

$$\lim_{\theta \rightarrow 0} \frac{f(\alpha(\theta)) - f(x^*)}{\theta} = \nabla f(x^*)^T \alpha'(0) = \sigma \nabla f(x^*)^T \hat{d} < 0,$$

implying the existence of  $\hat{\theta} \in [0, 1]$  sufficiently small to have  $\alpha(\hat{\theta}) \in B_\varepsilon(x^*)$

such that  $f(\alpha(\hat{\theta})) < f(x^*)$ , a contradiction since  $\alpha(\hat{\theta}) \in \tilde{R}$ .

The rest of the proof is completed as before when we were assuming that the hypothesis was verified.

# **K-K-T condition for linear programming**

K-K-T conditions  $\leftrightarrow$  duality and complementary slackness results.

# **K-K-T condition for linear programming**

- K-K-T conditions  $\leftrightarrow$  duality and complementary slackness results.  
For linear programming,
- K-K-T conditions are sufficient since linear functions are convex.

# K-K-T condition for linear programming

K-K-T conditions  $\leftrightarrow$  duality and complementary slackness results.

For linear programming,

- K-K-T conditions are sufficient since linear functions are convex.
- K-K-T conditions are necessary since linear functions always satisfy the constraints qualification.



# K-K-T condition for linear programming

K-K-T conditions  $\leftrightarrow$  duality and complementary slackness results.

For linear programming,

- K-K-T conditions are sufficient since linear functions are convex.
- K-K-T conditions are necessary since linear functions always satisfy the constraints qualification.

Consider the following linear programming problem:

$$\begin{aligned} & \text{Min} \quad \sum_{j=1}^n c_j x_j \\ & \text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, \dots, m \\ & \quad \quad \quad x_j \geq 0 \quad j = 1, \dots, n. \end{aligned}$$

Consider the following linear programming problem:

$$\begin{aligned} & \text{Min } \sum_{j=1}^n c_j x_j \\ & \text{s.t. } \sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, \dots, m \\ & \quad \quad \quad x_j \geq 0 \quad j = 1, \dots, n. \end{aligned}$$

Consider the following linear programming problem:

$$\begin{aligned} \text{Min} \quad & \sum_{j=1}^n c_j x_j \\ \text{s.t.} \quad & \sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, \dots, m \\ & x_j \geq 0 \quad j = 1, \dots, n. \end{aligned}$$

This problem is equivalent to

$$\begin{aligned} \text{Min} \quad & \sum_{j=1}^n c_j x_j \\ \text{s.t.} \quad & - \sum_{j=1}^n a_{ij} x_j + b_i \leq 0 \quad i = 1, \dots, m \\ & - x_j \leq 0 \quad j = 1, \dots, n. \end{aligned}$$

Consider the following linear programming problem:

$$\begin{aligned} & \text{Min} \quad \sum_{j=1}^n c_j x_j \\ & \text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, \dots, m \\ & \quad \quad \quad x_j \geq 0 \quad j = 1, \dots, n. \end{aligned}$$

This problem is equivalent to

$$\begin{aligned} & \text{Min} \quad \sum_{j=1}^n c_j x_j \\ & \text{s.t.} \quad - \sum_{j=1}^n a_{ij} x_j + b_i \leq 0 \quad i = 1, \dots, m \quad \lambda_i \\ & \quad \quad \quad - x_j \leq 0 \quad j = 1, \dots, n. \quad \lambda_{m+j} \end{aligned}$$

Associate a multiplier  $\lambda_i$  with each of the  $m$  first constraints and a multiplier  $\lambda_{m+j}$  with each of the  $n$  last constraints.

$$\begin{aligned}
& \text{Min} \quad \sum_{j=1}^n c_j x_j \\
& \text{s.t.} \quad - \sum_{j=1}^n a_{ij} x_j + b_i \leq 0 \quad i = 1, \dots, m \quad \lambda_i \\
& \quad \quad \quad - x_j \leq 0 \quad j = 1, \dots, n. \quad \lambda_{m+j}
\end{aligned}$$

Associate a multiplier  $\lambda_i$  with each of the  $m$  first constraints and a multiplier  $\lambda_{m+j}$  with each of the  $n$  last constraints.

$$\begin{aligned} \text{Min} \quad & \sum_{j=1}^n c_j x_j \\ \text{s.t.} \quad & - \sum_{j=1}^n a_{ij} x_j + b_i \leq 0 \quad i = 1, \dots, m \quad \lambda_i \\ & - x_j \leq 0 \quad j = 1, \dots, n. \quad \lambda_{m+j} \end{aligned}$$

Associate a multiplier  $\lambda_i$  with each of the  $m$  first constraints and a multiplier  $\lambda_{m+j}$  with each of the  $n$  last constraints.

The K-K-T conditions are the following:

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

$$\begin{aligned} \frac{\partial f(x)}{\partial x_j} + \sum_{i=1}^m \lambda_i \frac{\partial f_i(x)}{\partial x_j} + \sum_{i=m+1}^{m+n} \lambda_i \frac{\partial f_i(x)}{\partial x_j} &= c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} = 0 \quad j = 1, \dots, n \\ \lambda_i \left( - \sum_{j=1}^n a_{ij} x_j + b_i \right) &= 0 \quad i = 1, \dots, m \\ \lambda_{m+j} (-x_j) &= 0 \quad j = 1, \dots, n \\ - \sum_{j=1}^n a_{ij} x_j + b_i &\leq 0 \quad i = 1, \dots, m \\ - x_j &\leq 0 \quad j = 1, \dots, n \\ \lambda_j &\geq 0 \quad j = 1, \dots, m+n. \end{aligned}$$

**Primal** Min  $\sum_{j=1}^n c_j x_j$

s.t.  $\sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, \dots, m$

$x_j \geq 0 \quad j = 1, \dots, n$

K-K-T conditions

$$c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} = 0 \quad j = 1, \dots, n$$

$$\lambda_i \left( -\sum_{j=1}^n a_{ij} x_j + b_i \right) = 0 \quad i = 1, \dots, m$$

$$\lambda_{m+j} (-x_j) = 0 \quad j = 1, \dots, n$$

$$-\sum_{j=1}^n a_{ij} x_j + b_i \leq 0 \quad i = 1, \dots, m$$

$$-x_j \leq 0 \quad j = 1, \dots, n$$

$$\lambda_i \geq 0 \quad i = 1, \dots, m+n$$

Consider the dual problem:

$$\text{Primal} \quad \text{Min} \quad \sum_{j=1}^n c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, \dots, m$$

$$x_j \geq 0 \quad j = 1, \dots, n$$

$$\text{Dual} \quad \text{Max} \quad \sum_{i=1}^m b_i y_i$$

$$\text{s.t.} \quad \sum_{i=1}^m a_{ij} y_i \leq c_j \quad j = 1, \dots, n$$

$$y_i \geq 0 \quad i = 1, \dots, m.$$

K-K-T conditions

$$c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} = 0 \quad j = 1, \dots, n$$

$$\lambda_i \left( - \sum_{j=1}^n a_{ij} x_j + b_i \right) = 0 \quad i = 1, \dots, m$$

$$\lambda_{m+j} (-x_j) = 0 \quad j = 1, \dots, n$$

$$- \sum_{j=1}^n a_{ij} x_j + b_i \leq 0 \quad i = 1, \dots, m$$

$$-x_j \leq 0 \quad j = 1, \dots, n$$

$$\lambda_i \geq 0 \quad i = 1, \dots, m+n$$



Consider the dual problem:

$$\text{Primal} \quad \text{Min} \quad \sum_{j=1}^n c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, \dots, m$$

$$x_j \geq 0 \quad j = 1, \dots, n$$

$$\text{Dual} \quad \text{Max} \quad \sum_{i=1}^m b_i y_i$$

$$\text{s.t.} \quad \sum_{i=1}^m a_{ij} y_i \leq c_j \quad j = 1, \dots, n$$

$$y_i \geq 0 \quad i = 1, \dots, m.$$

K-K-T conditions

$$c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} = 0 \quad j = 1, \dots, n$$

$$\lambda_i \left( -\sum_{j=1}^n a_{ij} x_j + b_i \right) = 0 \quad i = 1, \dots, m$$

$$\lambda_{m+j} (-x_j) = 0 \quad j = 1, \dots, n$$

$$-\sum_{j=1}^n a_{ij} x_j + b_i \leq 0 \quad i = 1, \dots, m$$

$$-x_j \leq 0 \quad j = 1, \dots, n$$

$$\lambda_i \geq 0 \quad i = 1, \dots, m+n$$

The vector  $[\lambda_1, \dots, \lambda_m]$  is a feasible solution for the dual: for  $j = 1, \dots, n$

$$c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} = 0$$

$$c_j - \sum_{i=1}^m \lambda_i a_{ij} = \lambda_{m+j} \geq 0$$

$$\sum_{i=1}^m \lambda_i a_{ij} \leq c_j$$

Furthermore,

$$\lambda_i \geq 0 \quad i = 1, \dots, m.$$

Consider the dual problem:

$$\text{Primal} \quad \text{Min} \quad \sum_{j=1}^n c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, \dots, m$$

$$x_j \geq 0 \quad j = 1, \dots, n$$

$$\text{Dual} \quad \text{Max} \quad \sum_{i=1}^m b_i y_i$$

$$\text{s.t.} \quad \sum_{i=1}^m a_{ij} y_i \leq c_j \quad j = 1, \dots, n$$

$$y_i \geq 0 \quad i = 1, \dots, m.$$

K-K-T conditions

$$c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} = 0 \quad j = 1, \dots, n$$

$$\lambda_i \left( -\sum_{j=1}^n a_{ij} x_j + b_i \right) = 0 \quad i = 1, \dots, m$$

$$\lambda_{m+j} (-x_j) = 0 \quad j = 1, \dots, n$$

$$-\sum_{j=1}^n a_{ij} x_j + b_i \leq 0 \quad i = 1, \dots, m$$

$$-x_j \leq 0 \quad j = 1, \dots, n$$

$$\lambda_i \geq 0 \quad i = 1, \dots, m+n$$

Consider the dual problem:

$$\text{Primal} \quad \text{Min} \quad \sum_{j=1}^n c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, \dots, m$$

$$x_j \geq 0 \quad j = 1, \dots, n$$

$$\text{Dual} \quad \text{Max} \quad \sum_{i=1}^m b_i y_i$$

$$\text{s.t.} \quad \sum_{i=1}^m a_{ij} y_i \leq c_j \quad j = 1, \dots, n$$

$$y_i \geq 0 \quad i = 1, \dots, m.$$

K-K-T conditions

$$c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} = 0 \quad j = 1, \dots, n$$

$$\lambda_i \left( -\sum_{j=1}^n a_{ij} x_j + b_i \right) = 0 \quad i = 1, \dots, m$$

$$\lambda_{m+j} (-x_j) = 0 \quad j = 1, \dots, n$$

$$-\sum_{j=1}^n a_{ij} x_j + b_i \leq 0 \quad i = 1, \dots, m$$

$$-x_j \leq 0 \quad j = 1, \dots, n$$

$$\lambda_j \geq 0 \quad j = 1, \dots, m+n$$

The vector  $[\lambda_1, \dots, \lambda_m]$  is an optimal solution for the dual: for  $j = 1, \dots, n$

$$x_j \left( c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} \right) = 0$$

$$c_j x_j - \sum_{i=1}^m \lambda_i a_{ij} x_j = \lambda_{m+j} x_j = 0$$

$$\sum_{j=1}^n c_j x_j = \sum_{j=1}^n \sum_{i=1}^m \lambda_i a_{ij} x_j$$

Consider the dual problem:

$$\text{Primal} \quad \text{Min} \quad \sum_{j=1}^n c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, \dots, m$$

$$x_j \geq 0 \quad j = 1, \dots, n$$

$$\text{Dual} \quad \text{Max} \quad \sum_{i=1}^m b_i y_i$$

$$\text{s.t.} \quad \sum_{i=1}^m a_{ij} y_i \leq c_j \quad j = 1, \dots, n$$

$$y_i \geq 0 \quad i = 1, \dots, m.$$

K-K-T conditions

$$c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} = 0 \quad j = 1, \dots, n$$

$$\lambda_i \left( -\sum_{j=1}^n a_{ij} x_j + b_i \right) = 0 \quad i = 1, \dots, m$$

$$\lambda_{m+j} (-x_j) = 0 \quad j = 1, \dots, n$$

$$-\sum_{j=1}^n a_{ij} x_j + b_i \leq 0 \quad i = 1, \dots, m$$

$$-x_j \leq 0 \quad j = 1, \dots, n$$

$$\lambda_j \geq 0 \quad j = 1, \dots, m+n$$

The vector  $[\lambda_1, \dots, \lambda_m]$  is an optimal solution for the dual:

$$\sum_{j=1}^n c_j x_j = \sum_{j=1}^n \sum_{i=1}^m \lambda_i a_{ij} x_j$$

For  $i = 1, \dots, m$

$$\lambda_i \left( -\sum_{j=1}^n a_{ij} x_j + b_i \right) = 0$$

$$\lambda_i b_i = \lambda_i \sum_{j=1}^n a_{ij} x_j$$

$$\sum_{i=1}^m \lambda_i b_i = \sum_{i=1}^m \sum_{j=1}^n \lambda_i a_{ij} x_j$$

Consider the dual problem:

$$\text{Primal} \quad \text{Min} \quad \sum_{j=1}^n c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, \dots, m$$

$$x_j \geq 0 \quad j = 1, \dots, n$$

$$\text{Dual} \quad \text{Max} \quad \sum_{i=1}^m b_i y_i$$

$$\text{s.t.} \quad \sum_{i=1}^m a_{ij} y_i \leq c_j \quad j = 1, \dots, n$$

$$y_i \geq 0 \quad i = 1, \dots, m.$$

K-K-T conditions

$$c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} = 0 \quad j = 1, \dots, n$$

$$\lambda_i \left( - \sum_{j=1}^n a_{ij} x_j + b_i \right) = 0 \quad i = 1, \dots, m$$

$$\lambda_{m+j} (-x_j) = 0 \quad j = 1, \dots, n$$

$$- \sum_{j=1}^n a_{ij} x_j + b_i \leq 0 \quad i = 1, \dots, m$$

$$-x_j \leq 0 \quad j = 1, \dots, n$$

$$\lambda_j \geq 0 \quad j = 1, \dots, m+n$$

The vector  $[\lambda_1, \dots, \lambda_m]$  is an optimal solution for the dual:

$$\sum_{j=1}^n c_j x_j = \sum_{j=1}^n \sum_{i=1}^m \lambda_i a_{ij} x_j$$

$$\sum_{i=1}^m \lambda_i b_i = \sum_{i=1}^m \sum_{j=1}^n \lambda_i a_{ij} x_j$$

Consequently

$$\sum_{j=1}^n c_j x_j = \sum_{i=1}^m \lambda_i b_i$$

and the result follows from the weak duality theorem.

Consider the dual problem:

$$\text{Primal} \quad \text{Min} \quad \sum_{j=1}^n c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, \dots, m$$

$$x_j \geq 0 \quad j = 1, \dots, n$$

$$\text{Dual} \quad \text{Max} \quad \sum_{i=1}^m b_i y_i$$

$$\text{s.t.} \quad \sum_{i=1}^m a_{ij} y_i \leq c_j \quad j = 1, \dots, n$$

$$y_i \geq 0 \quad i = 1, \dots, m.$$

K-K-T conditions

$$c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} = 0 \quad j = 1, \dots, n$$

$$\lambda_i \left( - \sum_{j=1}^n a_{ij} x_j + b_i \right) = 0 \quad i = 1, \dots, m$$

$$\lambda_{m+j} (-x_j) = 0 \quad j = 1, \dots, n$$

$$- \sum_{j=1}^n a_{ij} x_j + b_i \leq 0 \quad i = 1, \dots, m$$

$$-x_j \leq 0 \quad j = 1, \dots, n$$

$$\lambda_i \geq 0 \quad i = 1, \dots, m+n$$

Consider the dual problem:

$$\text{Primal} \quad \text{Min} \quad \sum_{j=1}^n c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, \dots, m$$

$$x_j \geq 0 \quad j = 1, \dots, n$$

$$\text{Dual} \quad \text{Max} \quad \sum_{i=1}^m b_i y_i$$

$$\text{s.t.} \quad \sum_{i=1}^m a_{ij} y_i \leq c_j \quad j = 1, \dots, n$$

$$y_i \geq 0 \quad i = 1, \dots, m.$$

K-K-T conditions

$$c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} = 0 \quad j = 1, \dots, n$$

$$\lambda_i \left( -\sum_{j=1}^n a_{ij} x_j + b_i \right) = 0 \quad i = 1, \dots, m$$

$$\lambda_{m+j} (-x_j) = 0 \quad j = 1, \dots, n$$

$$-\sum_{j=1}^n a_{ij} x_j + b_i \leq 0 \quad i = 1, \dots, m$$

$$-x_j \leq 0 \quad j = 1, \dots, n$$

$$\lambda_j \geq 0 \quad j = 1, \dots, m+n$$

For  $j = 1, \dots, n$

$$c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} = 0$$

$$x_j \left( c_j - \sum_{i=1}^m \lambda_i a_{ij} - \lambda_{m+j} \right) = 0$$

$$x_j \left( c_j - \sum_{i=1}^m \lambda_i a_{ij} \right) = x_j \lambda_{m+j} = 0.$$

For  $i = 1, \dots, m$

$$\lambda_i \left( -\sum_{j=1}^n a_{ij} x_j + b_i \right) = 0.$$