Conditions d'Optimalité II

Rappel

— x_* min local de $(P_X) = \inf_{x \in X} f(x)$

$\overset{\text{PK}}{\Longrightarrow}$ $$\boxed{\nabla f(x_*) \in (T_{x_*} X)^+}$$

— On considère cette fois le pb

$$\begin{cases} \min f(x) & f : \mathbb{E} \to \mathbb{R} \\ c_E(x) = 0 & \in \mathbb{R}^{m_E} \\ c_I(x) \leq 0 & \in \mathbb{R}^{m_I} \end{cases}$$

$c_i(x) \leq 0, \ \forall i \in I$

$$X = \{x \in \mathbb{E} : c_E(x) = 0, \ c_I(x) \leq 0\}$$

$(T_x X)^+$?

— Lemme de Farkas

$$\boxed{\overline{A(K)} = \{y \in F : A^* y \in K^+\}^+}$$

→ gênante, mais non présente si

K est polyédrique = polyèdre convexe

$= \bigcap_{\text{finie}}$ de $\frac{1}{2}$ espaces fermés

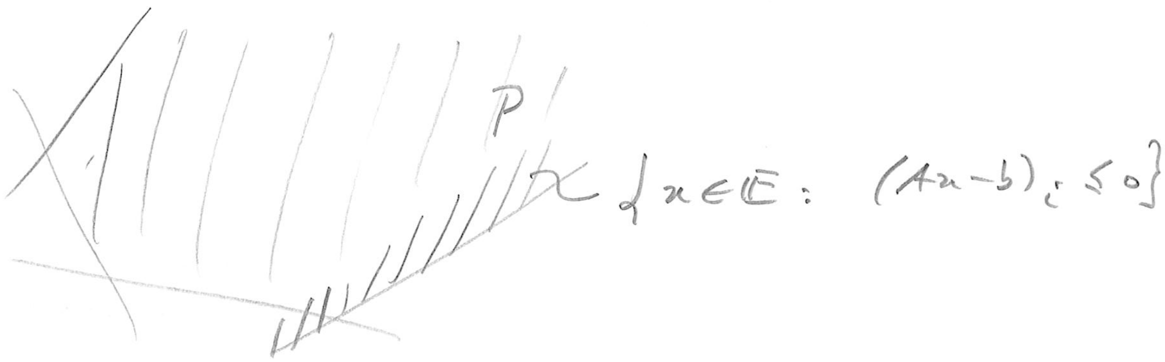
D Polyèdre convexe

Un polyèdre convexe d'un espace vectoriel $\mathbb{E}$ est un ensemble de la forme

$$P = \{x \in \mathbb{E} : Ax \leq b\}$$

$(Ax-b)_i \leq 0 \quad \forall i \in [1:m]$

où $A : \mathbb{E} \to \mathbb{R}^m$, $b \in \mathbb{R}^m$ et l'inégalité est comprise composante par composante. C'est donc une intersection d'un nombre fini m de demi-espaces fermés.



C'est donc un fermé.

Prop

Si
- $\mathbb{E}$ et $\mathbb{F}$ sont deux e.v. de dim finie
- P un polyèdre convexe de $\mathbb{E}$
- $T : \mathbb{E} \to \mathbb{F}$ linéaire

alors $T(P)$ est un polyèdre convexe de $\mathbb{F}$

donc $T(P)$ est fermé

<u>Dém</u> On peut supposer que $\mathbb{E} = \mathbb{R}^n$

1e étape <u>cas où $T = \Phi_n$ où</u>

<u>$\Phi_n : \mathbb{R}^n \to \mathbb{R}^{n-1} : (x_1, \dots, x_n) \longmapsto (x_1, \dots, x_{n-1})$</u>

(projection canonique)

Il s'agit de montrer que $\Phi_n(P)$ est un polyèdre convexe, c-à-d que $\Phi_n(P)$ est défini par un nombre <u>fini</u> d'inégalités affines. Soit $x \in P$. Donc

$$\sum_{j=1}^{n} a_{ij}\, x_j \le b_i, \qquad \forall i \in [1:m] \quad [1]$$

Pour montrer que, lorsque x vérifie [1], $\Phi_n(x) := (x_1, \dots, x_{n-1})$ est dans un polyèdre convexe, on élimine x_n des inégalités dans [1] (c'est <u>l'élimination de Fourier</u>). On a par [1] :

$$(*) \begin{cases} \forall i_1 \text{ t.q. } a_{i_1 n} > 0 : & x_n \le \dfrac{1}{a_{i_1 n}} \left(b_{i_1} - \displaystyle\sum_{j=1}^{n-1} a_{i_1 j}\, x_j \right) \\ \forall i_2 \text{ t.q. } a_{i_2 n} < 0 : & x_n \ge \dfrac{1}{a_{i_2 n}} \left(b_{i_2} - \displaystyle\sum_{j=1}^{n-1} a_{i_2 j}\, x_j \right) \\ \forall i_3 \text{ t.q. } a_{i_3 n} = 0 : & 0 \le b_{i_3} - \displaystyle\sum_{j=1}^{n-1} a_{i_3 j}\, x_j . \end{cases}$$

Montrons que $\Phi_n(P)$ est l'ensemble des $\bar{x} \in \mathbb{R}^{n-1}$ tels que $\forall (i_1, i_2, i_3) \in [1:m]^3$ tels que $a_{i_1 n} > 0$, $a_{i_2 n} < 0$, $a_{i_3 n} = 0$, on a

$$(**) \begin{cases} \dfrac{1}{a_{i_2 n}} \left(b_{i_2} - \displaystyle\sum_{j=1}^{n-1} a_{2j}\, \bar{x}_j \right) \le \dfrac{1}{a_{i_1 n}} \left(b_{i_1} - \displaystyle\sum_{j=1}^{n} a_{i_1 j}\, \bar{x}_j \right) \\ 0 \le b_{i_3} - \displaystyle\sum_{j=1}^{n-1} a_{i_3 j}\, \bar{x}_j \end{cases}$$

En effet

- $x \in P \Rightarrow \bar{x} = \Phi_n(x)$ vérifie $(**)$ grâce à $(*)$
- $\bar{x}$ vérifie $(**) \Rightarrow$ en prenant x_n tel que

$$\sup_{\substack{i_2 \text{ t.q.} \\ a_{i_2 n} < 0}} \frac{1}{a_{i_2 n}} \left(b_{i_2} - \sum_{j=1}^{n-1} a_{i_2 j} \bar{x}_j \right) \leq x_n \leq \inf_{\substack{i_1 \text{ t.q.} \\ a_{i_1 n} < 0}} \frac{1}{a_{i_1 n}} \left(b_{i_1} - \sum_{j=1}^{n} a_{i_1 j} \bar{x}_j \right)$$

alors le vecteur $x = (\bar{x}_1, \dots, \bar{x}_{n-1}, x_n)$ vérifie $(*)$ et donc $x \in P$

On a donc bien défini $\Phi_n(P)$ par un nombre fini (et grand !) d'inégalités affines. Donc $\Phi_n(P)$ est un polyèdre convexe

2ᵉ étape — cas général Soit $F = \mathbb{R}^p$

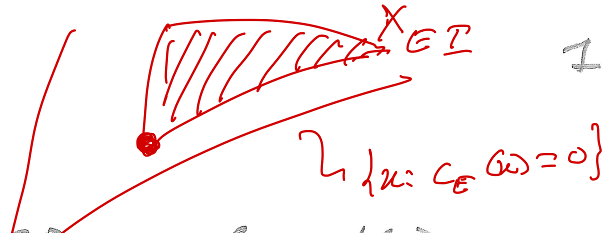
$$T(P) = \{ Tx : \underbrace{Ax \leq b}_{\text{pour un } x \in \mathbb{R}^n} \}$$

$$= \{ y \in \mathbb{R}^p : y = Tx \text{ et } \underbrace{Ax \leq b}_{\text{pour un } x \in \mathbb{R}^n} \}$$

$$= \underbrace{(\Phi_{p+1} \circ \dots \circ \Phi_{p+n}) \underbrace{\{ \overset{\in \mathbb{R}^p \times \mathbb{R}^n}{(y,x)} : \begin{array}{l} y \leq Tx \\ -y \leq -Tx \\ Ax \leq b \end{array} \}}_{\text{c'est un polyèdre convexe de } \mathbb{R}^m \times \mathbb{R}^n}}_{\text{par l'étape 1, c'est aussi 1 polyèdre convexe de } \mathbb{R}^p}$$

$\square$

I) Le problème (P_{EI})



On cherche à obtenir des CO pour le problème

$$(P_{EI})\quad\begin{cases}\min f(x) & \\ c_i(x) = 0, & i \in E \\ c_i(x) \leq 0, & i \in I\end{cases}$$

$|E| = m_E \in \mathbb{N}$
$|I| = m_I \in \mathbb{N}$

où f et $c_i : \mathbb{E} \longrightarrow \mathbb{R}$

- On suppose que (E, I) est une partition de $[1:m]$

  $E \cap I = \emptyset, \quad E \cup I = [1:m]$

  $m_E := |E|, \quad m_I := |I|, \quad m = m_E + m_I$

- On note $c : \mathbb{E} \longrightarrow \mathbb{R}^m$ la fonction définissant les contraintes

- Pour $v \in \mathbb{R}^m$, on note

  $v_E = (v_i)_{i \in E}$ et $v_I = (v_i)_{i \in I}$

  Donc les contraintes s'écrivent

  $c_E(x) = 0 \qquad c_I(x) \leq 0$

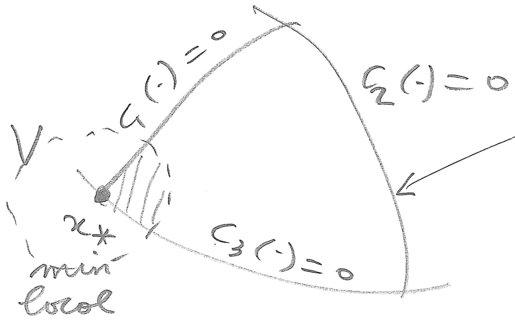
- L'ensemble admissible du problème se note

  $X_{EI} := \{x \in \mathbb{E} : c_E(x) = 0, \ c_I(x) \leq 0\}$

- On dit que (P_{EI}) est convexe si

  $\begin{cases}\bullet\ X_{EI} \text{ est convexe} \\ \bullet\ f \text{ est convexe sur } X_{EI}\end{cases}$

Notion de contrainte active



Cette contrainte ne joue pas de rôle dans l'optimalité locale en x_* ; donc elle n'apparaîtra pas dans les CO

- On dit que c_i est active en x si $c_i(x) = 0$

- On note

$$I^0(x) := \{ i \in I : c_i(x) = 0 \}$$

l'ensemble des indices des contraintes d'inégalité actives en x.

$$I^0_* := I^0(x_*).$$

(dans l'exemple ci-dessus : $I^0_* = \{1, 3\}$)

- On sait pas a priori quelles sont les contraintes d'inégalité qui seront actives en la solution ; il y a 2^{m_I} possibilités de réaliser cela

  $\Longrightarrow$ difficulté redoutable (à l'origine de la NP-ardité des problèmes d'optimisation avec contraintes d'inégalité)

Donc

si $I = \emptyset$ $\Longrightarrow$ problème relevant de l'analyse, "assez facile" à résoudre

si $I \neq \emptyset$ $\Longrightarrow$ problème relevant de l'analyse convexe ; toujours difficile à résoudre

2) CO du premier ordre

On cherche à exprimer la CN1 de Peano-Kantorovitch

$$\nabla f(x_*) \in \left(T_{x_*} X_{EI}\right)^+$$

Il faut donc expliciter $T_x X_{EI}$ et son dual.

$X_{EI} := \{x \in E : c_E(x) = 0,\ c_I(x) \le 0\}$

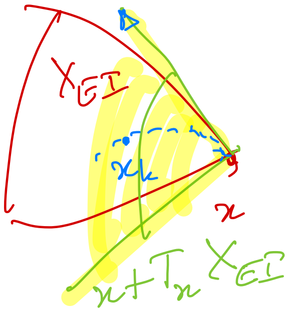
A Le cône tangent et son dual

Soit $x \in X_{EI}$. On a

$c_{E \cup I^0(x)}$ différentiable

$$\boxed{T_x X_{EI} \subset \underbrace{T_x' X_{EI}}_{\text{s'appelle le cône linéarisant}} := \{d \in E : c_E'(x) \cdot d = 0,\ c'_{I^0(x)}(x) \cdot d \le 0\}}$$

seules les contraintes d'inégalité actives en x interviennent



Dém $d \in T_x X_{EI} \setminus \{0\}$

$\Rightarrow \exists (x_k) \subset X_{EI},\ \exists (t_k) \downarrow 0 : \dfrac{x_k - x}{t_k} \to 0$

$\Rightarrow \forall i \in E \cup I^0(x)$ on a

$$\underbrace{c_i(x_k)}\ = \underbrace{c_i(x)} + c_i'(x)\,\frac{(x_k - x)}{t_k} + o\left(\frac{\|x_k - x\|}{t_k}\right)$$

	$c_i(x_k)$	$c_i(x)$	$c_i'(x)\frac{(x_k-x)}{t_k} + o\left(\frac{\|x_k-x\|}{t_k}\right)$
$i \in E$	$= 0$	$= 0$	$= 0$
$i \in I^0(x)$	≤ 0	$= 0$	≤ 0

à la limite on obtient $c_E'(x)\, d = 0$

$c'_{I^0(x)}(x)\, d \le 0$

□

En général $\underbrace{T_x X_{EI}}_{\text{pas néc. convexe}} \neq \underbrace{T_x' X_{EI}}_{\text{polyèdre convexe}}$ car

Dfn : On dit que les contraintes c de (P_{EI}) sont qualifiées pour représenter X_{EI} si
$$T_x X_{EI} = T_x' X_{EI}$$

Rmq

1) On aime bien avoir qualification, comme cela on a une formule explicite du cône tangent

2) On peut décrire X_{EI} par beaucoup de fonctions (contrainte) $c : \mathbb{E} \to \mathbb{R}^m$. Certaines sont adéquates, d'autres pas. La qualification sert à sélectionner les bonnes représentations.

Par exemple, il est tentant de remplacer
$$c(x) \leq 0 \in \mathbb{R}^m, \quad m \gg 1$$
$X = \{x : c(x) \leq 0\}$ $\quad c(x)^+ = \max(0, c(x))$

par $\tilde{c}(x) := \frac{1}{2} \| c(x)^+ \|_2^2 = 0 \in \mathbb{R}$

$X = \{x : \tilde{c}(x) = 0\}$ (1 seule contrainte d'égalité différentiable)

Mais cette contrainte n'est généralement pas qualifiée car
$$\nabla \tilde{c}(x) = c'(x)^T \underbrace{c(x)^+}_{\substack{\| \\ 0 \text{ en } x \text{ admissible}}} = 0$$
$$\Rightarrow T_x' X = \mathbb{E}^n \text{ probablement } \neq T_x X \;!!$$

3) On verra des conditions suffisantes de qualification de contraintes plus loin.

On peut calculer le dual de $T_x X_{EI}$ en utilisant le lemme de Farkas

- $\mathbb{E}$, $\mathbb{F}$ deux e. euclidiens
- $A : \mathbb{E} \to \mathbb{F}$ linéaire (adjoint A^*)
- K cône convexe non vide de $\mathbb{E}$

$$\boxed{\overline{A(K)} = \{ y \in \mathbb{F} : A^* y \in K^+ \}^+} \qquad (F)$$

Cas particulier on note $A_E := c_E'(x)$, $A_J := c'_{I^0(x)}(x)$

$$\{ d \in \mathbb{E} : A_E d = 0,\ A_J d \leq 0 \}^+ = ?$$

$A^* d \in K^+$ $A = ?$ $K = ?$

c'est le membre de droite de (F)

$\mathbb{F} := \mathbb{E}$ p.s. euclidien

$A^* := \begin{pmatrix} A_E \\ A_J \end{pmatrix} \Rightarrow A = (A_E^T \ \ A_J^T)$

$K^+ := \{0\}_{\mathbb{R}^{m_E}} \times \mathbb{R}^{|J|}_{-}$ $\Leftarrow$ $K = \mathbb{R}^{m_E} \times \mathbb{R}^{|J|}_{-}$

$\{ K^+ = \underbrace{(\mathbb{R}^{m_E})^+}_{\{0_{\mathbb{R}^{m_E}}\}} \times \underbrace{(\mathbb{R}^{|J|}_{-})^+}_{\mathbb{R}^{|J|}_{-}}$

$$\boxed{\begin{aligned} &\{ d \in \mathbb{E} : A_E d = 0,\ A_J d \leq 0 \}^+ \\ &= \{ A_E^T z_E + A_J^T z_J : z_J \leq 0 \} \end{aligned}}$$

$z_E \in \mathbb{R}^{m_E}$

[B] CN1 de Karush, Kuhn et Tucker

Deux notions importantes

(1) Le lagrangien de $(P_{E I})$ est la fonction

$$\ell : \mathbb{E} \times \mathbb{R}^m \longrightarrow \mathbb{R}$$
$$(x, \underset{=}{\lambda}) \longmapsto \boxed{\ell(x,\lambda) = f(x) + \overbrace{\lambda^T c(x)}^{\sum_{i=1}^{m} \lambda_i c_i(x)}}$$

(2) Conditions de complémentarité (multiplicateur) : soit F et $G : \mathbb{G} \to \mathbb{R}^p$. Alors l'écriture

$$\boxed{0 \leqslant F(z) \perp G(z) \geqslant 0}$$

veut dire

$$F(z) \geqslant 0, \quad G(z) \geqslant 0, \quad F(z)^T G(z) = 0 \quad \underbrace{\sum_{i=1}^{p} F_i(z)\, G_i(z)}_{\geqslant 0} = 0$$

Avec la positivité de $F(z)$ et $G(z)$

$$\boxed{\begin{aligned} & F(z)^T G(z) = 0 \\ \Longleftrightarrow \quad & \forall i \in [1:p] : \quad F_i(z)\, G_i(z) = 0 \end{aligned}}$$

($\Leftarrow$) clair

($\Rightarrow$) car $0 = F(z)^T G(z) = \sum_{i=1}^{p} \underbrace{F_i(z)\, G_i(z)}_{\geqslant 0}$

$$\Rightarrow F_i(z)\, G_i(z) = 0, \quad \forall i$$

□

CN1 (KKT)

Si • x_* min local de (P_{EI})

• f et $c_{E\cup I_*^0}$ sont différentiables en x_*

• contraintes qualifiées en x_*

Alors $\exists\, \lambda_* \in \mathbb{R}^m$ tel que

$$(KKT) \quad \begin{cases} \nabla_x \ell(x_*,\lambda_*) = 0 \\ c_E(x_*) = 0 \\ 0 \leq (\lambda_*)_I \perp c_I(x_*) \leq 0 \end{cases}$$

$\nabla f(x_*) \neq 0$ qén

Démo

$$\nabla f(x_*) \in (T_{x_*} X_{EI})^+ \qquad \text{[Peano-Kantorovitch]}$$

$$= (T'_{x_*} X_{EI})^+ \qquad \text{[Qualification]}$$

$$= \{ d \in E : c'_E(x_*)\, d = 0,\ c'_{I^0_*}(x_*)\, d \leq 0 \}^+$$

$$= \{ -c'_E(x_*)^T z_E \;\overline{\cancel{+}}\; c'_{I^0_*}(x_*)^T z_{I^0_*} : z_{I^0_*} \overset{\geq}{\cancel{\leq}} 0 \}$$

$z_E \in \mathbb{R}^{m_E}$

[Farkas]

c'est 1 résultat d'existence !!

$\Rightarrow \exists\, z_{E\cup I^0_*}$ t.q.

$$\nabla f(x_*) + c'_E(x_*)^T z_E + c'_{I^0_*}(x_*)^T z_{I^0_*} + c'_{I\setminus I^0_*}(x_*)^T 0 = 0$$

Si (λ_*) est déf par $(\lambda_*)_i = \begin{cases} z_i & \text{si } i \in E\cup I^0_* \\ 0 & \text{sinon} \end{cases}$

alors $\nabla f(x_*) + c'(x_*)^T \lambda_* = 0$ ou $\nabla_x \ell(x_*,\lambda_*) = 0$

si $c_i(x_*) < 0$ $(i\in I)$, on a $(\lambda_*)_i = 0$

□

Rmq

1) La condition de complémentarité

$$0 \leq (\lambda_*)_I \perp c_I(x_*) \leq 0$$

se lit

$$(\lambda_*)_I \geq 0, \quad c_I(x_*) \leq 0 \quad \text{et} \quad (\lambda_*)_I^T c_I(x_*) = 0$$

Lorsque $(\lambda_*)_I \geq 0$, $c_I(x_*) \leq 0$, on a

$$(\lambda_*)_I^T c_I(x_*) = 0$$

$$\Leftrightarrow \quad \forall i \in I : \quad (\lambda_*)_i \, c_i(x_*) = 0$$

$$\Leftrightarrow \quad \forall i \in I : \quad c_i(x_*) < 0 \Rightarrow (\lambda_*)_i = 0$$

(les $(\lambda_*)_i$ associés aux contraintes inactives sont nuls ; donc les contraintes inactives n'apparaissent pas dans (KKT))

2) On dit que l'on a complémentarité stricte si

$$\forall i \in I : \quad c_i(x_*) < 0 \Leftrightarrow (\lambda_*)_i = 0$$

3) Le vecteur λ_* est appelé le multiplicateur optimal associé à x_* (il multiplie la contrainte dans le lagrangien).

Il y a un multiplicateur par contrainte.

L'ensemble des multiplicateurs associés à x_* est le polyèdre convexe

$$\Lambda_* = \{ \lambda_* \in \mathbb{R}^m : \nabla f(x_*) + c'(x_*)^* \lambda_* = 0, \; (\lambda_*)_I \geq 0, \; (\lambda_I)_{I \setminus I_*^0} = 0 \}$$

c) CS1 pour un problème convexe

> **si** • (P_{EI}) est convexe [X_{EI} convexe, f convexe sur X_{EI}]
> • f et $c_{E\cup I_*^0}$ différentiables en $x_* \in X_{EI}$
> • (x_*, λ_*) vérifie (KKT)
>
> **alors** x_* est un min global de (P_{EI})

Dém Il suffit de montrer que

$$f'(x_*)(x-x_*) \geq 0, \quad \forall x \in X_{EI} \quad ?$$

On a $\nabla f(x_*) + c'(x_*)^T \lambda_* = 0$

$$\lambda_*^T \underbrace{c'(x_*)(x-x_*)}_{\gamma} \leq 0, \quad \forall x \in X_{EI} \; ?$$

$$\gamma_i = c'(x_*)(x-x_*) = \lim_{t \downarrow 0} \frac{1}{t}\left[\underline{c(x_* + t(x-x_*))} - c(x_*) \right]_i$$

$(1-t)x_* + tx \in X_{EI}$
si $t > 0$ petit
par convexité de X_{EI}
et $x_*, x \in X_{EI}$

$i \in E$: $c_i(x_* + t(x-x_*)) = 0$, $c_i(x_*) = 0$
$\Rightarrow \gamma_i = 0$

$i \in I_*^0$: $\left.\begin{array}{l} c_i(x_*) = 0 \\ c_i(x_* + t(x-x_*)) \leq 0 \end{array}\right\} \Rightarrow \gamma_i \leq 0$, $(\lambda_*)_i \geq 0$
$\Rightarrow (\lambda_*)_i \gamma_i \leq 0$

$i \in I \setminus I_*^0$: γ_i ? $(\lambda_*)_i = 0$

$\square$

D Qualification de contraintes en x. ($T_x X = T'_x X$)

Voici les 4 conditions *suffisantes* de qualification de contraintes les plus utilisées. (elles impliquent donc la qualification). Il faut en utiliser 1 seule pour toutes les contraintes (pas une par contrainte !!).

(QC-A) [affinité] $c_{E\cup I^0(x)}$ est affine dans un voisinage de x (optim linéaire)

(QC-S) [Slater] (optim convexe)

- c_E est affine
- les c_i $(i \in I^0(x))$ sont convexes
- $\exists\, \check{x} \in X_{EI}$: $c_{I^0(x)}(\check{x}) < 0$

(QC-IL) [indépendance linéaire] (optim non linéaire pour débutants)

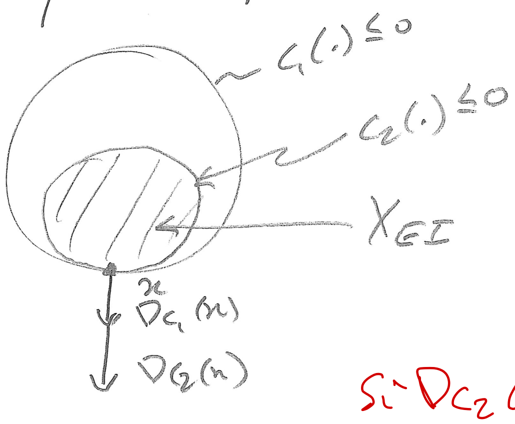
$\{\nabla c_i(x) : i \in E\cup I^0(x)\}$ sont linéairement indépendants

(QC-MF) [Mangasarian-Fromovitz] (optim non linéaire pour les connaisseurs)

$$\left.\begin{array}{l}\displaystyle\sum_{i\in E\cup I^0(x)} \alpha_i \nabla c_i(x) = 0 \\ \alpha_{I^0(x)} \geq 0\end{array}\right\} \Rightarrow \alpha_{E\cup I^0(x)} = 0$$

On a $\boxed{(QC\text{-}IL) \Rightarrow (QC\text{-}MF)}$

Mais pas d'équivalence. Contre-exemple



(QC-IL) non vérifiée

(QC-MF) vérifiée.

$$\alpha_1 \nabla c_1(x) + \alpha_2 \nabla c_2(x) = 0$$

$$\alpha_1 \geq 0,\quad \alpha_2 \geq 0$$

$$\Rightarrow \underbrace{\alpha_1 + \alpha_2 \beta}_{\geq 0} = 0$$

Si $\nabla c_2(x) = \overset{\geq 0}{\beta}\, \nabla c_1(x)$