Chapitre 2

Minimisation avec contraintes

On s'intéresse dans ce chapitre au problème d'optimisation avec contraintes, dont on va présenter sa forme générale, théorèmes d'existence et d'unicité et les conditions d'optimalité.

2.1 Formulation du problème

[3] Soit $f : \mathbb{R}^n \to \mathbb{R}$. On appelle problème de minimisation avec contraintes le problème suivant :

$$(P)\begin{cases} \min f(x), \\ x \in C, C \subset \mathbb{R}^n \end{cases} .$$

où C est l'ensemble des contraintes, C domaine non vide et fermé de $\mathbb{R}^n$.

Le problème (P) s'écrit aussi sous la forme générale suivante :

$$(P)\begin{cases} \min f(x), \\ h_i(x) = 0, \; i = 1, \ldots, p, \\ g_j(x) \le 0, \; j = 1, \ldots, q, \end{cases}$$

- *Les fonctions $f, h_i, i = \overline{1,p}$ et $g_j, j = \overline{1,q}$ définies de C dans $\mathbb{R}$ sont de classe C^1;*

- *Les conditions* $h_i(x) = 0$, $i = \overline{1,p}, g_j(x) \leq 0$, $j = \overline{1,q}$ *et* $x \in \mathbb{R}^n$ *sont appelées contraintes du problème (P).*

Définitions

[3]

- *On appelle solution réalisable ou admissible tout vecteur x vérifiant les contraintes du probléme (P).*

  L'ensemble $C = \{x \in : h_i(x) = 0, \forall i = \overline{1,p}$ *et* $g_j(x) \leq 0, \forall j = \overline{1,q}\}$ *est appelé ensemble des solutions réalisables ou admissibles.*

- *On appelle solution optimale du problème (P) (ou minimum global) une solution réalisable* x^* *qui minimise f(x) sur l'ensemble de toutes les solutions réalisables, c.à.d :*

$$\forall x \in C, \ f(x^*) \leq f(x).$$

- *On dit qu'un point* x^* *est un optimum (minimum) local de f si, et seulement si, il existe un voisinage* $V(x^*)$ *de* x^* *tel que* x^* *soit un min global de f sur* $V(x^*)$ *:*

$$\forall x \in V(x^*) \cap C, \ f(x^*) \leq f(x).$$

- *Un minimum est dit strict si les inégalités dans les définitions précédentes sont strictes.*

- *Si pour* $x \in C$ *et pour* $j \in \{1, 2, \ldots, q\}$ *on a* $g_j(x) = 0$, *on dit que la contrainte* g_j *est saturée ou active en x.*

  Une contrainte qui n'est pas active est dite inactive. On note I(x) l'ensemble des indices j correspondants aux contraintes actives en x :

$$I(x) = \{j \in \{1, 2, \ldots, q\} : g_j(x) = 0\}.$$

- *Un vecteur d est une direction de descente s'il existe* τ *tel que :*

$$f(x + td) < f(x), \ \ t \in [0, \tau].$$

- *Soit* $f : \mathbb{R}^n \to \mathbb{R},\ C \subseteq \mathbb{R}^n,\ x \in C$

  d est une direction admissible si $\exists \lambda' > 0$ *tel que :* $x + \lambda d \in C,\ \forall\, 0 \leq \lambda \leq \lambda'$

Exemple 2.1.1. *Considérons le problème ci-dessous :*

$$(P_0)\begin{cases} \min\ f(x,y), \\ g_1(x,y) = (x+1)^2 + (y-2)^2 - 4 \leq 0, \\ g_2(x,y) = -y + \frac{3}{5}x + \frac{7}{5} \leq 0, \\ g_3(x,y) = y + x - 3 \leq 0, \\ (x,y) \in \mathbb{R}^2. \end{cases}$$

1. *considérons les directions* $d_0 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ *et* $d_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ *et le point* $x^0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

   (voir Figure 2.1)

2. *Quelles sont les contraintes actives en* x^0 *?*

   $g_1(x^0) = g_1(1,2) = (1+1)^2 + (2-2)^2 - 4 = 0,$ g_1 *est active*

   $g_2(x^0) = g_2(1,2) = -2 + \frac{3}{5} \times 1 + \frac{7}{5} = 0,$ g_2 *est active*

   $g_3(x^0) = g_3(1,2) = 2 + 1 - 3 = 0,$ g_3 *est active*

   $I(x^0) = I(1,2) = \{1,2,3\}$

3. d_0 *est-elle une direction admissible pour* f *en* x^0 *?*

   Il faut avoir $\langle \nabla g_j(x^0), d_0 \rangle \leq 0, \forall j = \overline{1,3}$ *car* g_1, g_2 *et* g_3 *sont actives.*

   $$\nabla g_1(x,y) = \begin{pmatrix} 2(x+1) \\ 2(y-2) \end{pmatrix} \Rightarrow \nabla g_1(1,2) = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$$

   $$\nabla g_2(x,y) = \begin{pmatrix} \frac{3}{5} \\ -1 \end{pmatrix} \Rightarrow \nabla g_2(1,2) = \begin{pmatrix} \frac{3}{5} \\ -1 \end{pmatrix}$$

   $$\nabla g_3(x,y) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Rightarrow \nabla g_3(1,2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

   $$\langle \nabla g_1(1,2), d_0 \rangle = \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} 4 \\ 0 \end{pmatrix} = -4 \leq 0$$

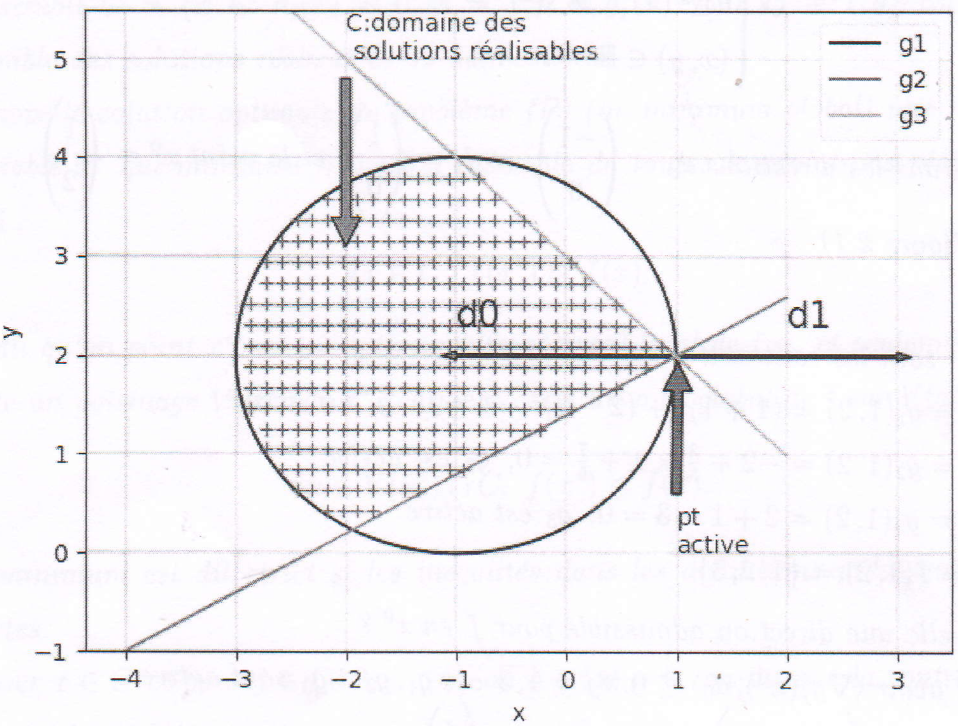

FIGURE 2.1 – Domaine des solutions réalisables

$$\langle \nabla g_2(1,2), d_0 \rangle = \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} \frac{3}{5} \\ -1 \end{pmatrix} = \frac{-3}{5} \leq 0$$

$$\langle \nabla g_3(1,2), d_0 \rangle = \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = -1 \leq 0$$

Donc d_0 est une direction admissible pour f en x^0.

4. *d_1 est-elle une direction admissible pour f en x^0 ?*

$$\langle \nabla g_1(1,2), d_1 \rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 4 \\ 0 \end{pmatrix} = 4 \geq 0$$

Donc d_1 n'est pas une direction admissible pour f en x^0

2.2 Résultats d'existence et d'unicité

Théorème 2.2.1. *[3](Existence)*

Supposons que f est continue, que C est un sous-ensemble fermé non vide de $\mathbb{R}^n$ et que l'une des conditions suivantes est satisfaite :

1. *Soit C est borné ;*
2. *Soit f est coercive.*

Alors le problème (P) admet au moins une solution.

Preuve 2.2.1. *[3] Soit $(x_n)_{n\in\mathbb{N}}$ une suite minimisante dans $f(C)$ i.e d'élément de C telle que $lim_{n\to+\infty}f(x_n) = inf f(C)$ comme C est fermé et borné, il existe une sous suite etraite $(x_\sigma(n))$ qui converge vers un $x^* \in C$ cette suite exraite vérifie : $x_{\sigma(n)} \to x^*$ et $f(x_{\sigma(n)}) \to inf_{y\in C}f(y)$ or f est continue, d'où par Unicité de la limite, il suit $f(x^*) = inf_{y\in C}f(y)$ avec $x^* \in C$ et f réalise son minimun sur C.$\square$*

Théorème 2.2.2. *(Existence et unicité)[3]*

Soit f une fonction continue et strictement convexe et soit C est un sous-ensemble non vide, convexe et fermé de $\mathbb{R}^n$.

Si C est borné ou si f est coercive, alors il existe un unique $x^* \in C$ solution de (P).

Preuve 2.2.2. *[3] supposons que f admette au moins un minimum m et soit $x_1 \neq x_2$ dans C réalisant ce minimum : $f(x_1) = f(x_2) = m$ par strict convexité de la fonction f on a alors :*

$$f(\frac{x_1 + x_2}{2}) < \frac{1}{2}(f(x_1) + f(x_2)) = m$$

ceci contredit le fait que m et le minimum donc $x_1 = x_2 = x^*$. □

2.3 Conditions d'optimalité du 1^er ordre

2.3.1 Conditions d'optimalité du 1^er ordre générales

Théorème 2.3.1. *(Condition nécessaire du 1^er ordre)[1]*
Si f est une fonction gâteaux différentiable et si C est un convexe et fermé de $\mathbb{R}^n$, alors toute solution x^* du problème (P) vérifie la propriété suivante :

$$\forall x \in C, \langle \nabla f(x^*), x - x^* \rangle \geq 0. \tag{2.1}$$

Preuve 2.3.1. *Soient x^* une solution du problème (P) et $x \in C$, alors :*

$$\forall x \in C, \ f(x^*) \leq f(x).$$

C est convexe, alors $f(tx+(1-t)x^*) \leq t f(x) + (1-t) f(x^*)$

$$x^* + t(x - x^*) \in C, \ \forall t \in [0, 1].$$ $\Rightarrow$ ~~...~~ $t(f(x) - f(x^*)) + f(x^*)$

Donc

$$\checkmark \ \forall x \in C, \ f(x^* + t(x - x^*)) - f(x^*) \geq 0.$$

D'où

$$\forall x \in C, \ \forall t \in [0, 1], \ \frac{f(x^* + t(x - x^*)) - f(x^*)}{t} \geq 0.$$

Alors

$$\forall x \in C, \ \forall t \in [0, 1], \ \lim_{t \to 0^+} \frac{f(x^* + t(x - x^*)) - f(x^*)}{t} \geq 0.$$

Par conséquent,

$$\forall x \in C, \ \langle \nabla f(x^*), \ x - x^* \rangle \geq 0.$$

Théorème 2.3.2. *[3](C.N.S. du 1er ordre dans le cas convexe)*
Soient f une fonction convexe et gâteaux différentiable et C est un convexe et fermé de $\mathbb{R}^n$. Soit x^* un élément quelconque de C.
La conditions (2.1) est nécessaire et suffisante pour que x^* soit solution du problème (P).

Preuve 2.3.2. *[3] La condition est nécessaire (voir Théorème 2.3.1). Il reste à montrer qu'elle est suffisante.*
Soit x^* un élément de C, comme f est convexe, alors :

$$\forall x \in C,\ f(x) \geq f(x^*) + \langle \nabla f(x^*), x - x^* \rangle.$$

Puisque $\langle \nabla f(x^*), x - x^* \rangle \geq 0$, alors $\forall x \in C, f(x) \geq f(x^*)$. D'où x^* est une solution du problème (P).

2.3.2 Conditions d'optimalité du 1er ordre, cas en égalité

[8] Lorsqu'on a que des contraintes d'égalité, le problème (P) se réduit à la forme suivante :

$$(P_e)\begin{cases} \min\ f(x), \\ h_i(x) = 0,\ i = 1, \dots, p, \\ x \in \mathbb{R}^n. \end{cases}$$

Théorème 2.3.3. *[3](C.N. du 1er ordre - Contraintes en égalité)*
On suppose que :

- *f, h_i pour $i = \overline{1,p}$, sont de classe C^1 dans $\mathbb{R}^n$;*
- *Le problème (P_e) admet une solution x^* ;*
- *Les p vecteurs de $\mathbb{R}^n$: $\nabla h_1(x^*), \nabla h_2(x^*), \dots, \nabla h_p(x^*)$ sont linéairement indépendants (et donc $p \leq n$).*

Alors il existe p réels $\lambda_1^*, \lambda_2^*, \dots, \lambda_p^*$ tels que

$$\nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla h_i(x^*) = 0. \tag{2.2}$$

Preuve 2.3.3. *[3] Analogue à celle du théorème de Karush-Kuhn-Tucker qui sera démontré par la suite.*

Définition 2.3.1. *[3](Multiplicateurs de Lagrange)*
Les réels $\lambda_1^*,\ \lambda_2^*, \dots,\ \lambda_p^*$ du théorème 2.2.3 sont appelés multiplicateurs de Lagrange.

Exemple 2.3.1. *cas $\mathbb{R}^2$*

On veut résourdre : $$\begin{cases} min\ f(x,y) = xy \\ h(x,y) = 4x^2 + y^2 - 4 = 0 \\ (x,y) \in \mathbb{R}^2 \end{cases}$$ $\longrightarrow (1,0) \in F;\ F \neq \emptyset$

Le Lagrangien pour ce problèm est

$$\mathcal{L}(x,y,\lambda) = f(x,y) + \lambda h(x,y), \quad \forall \lambda \in \mathbb{R}$$

$$\mathcal{L}(x,y,\lambda) = xy + \lambda(4x^2 + y^2 - 4)$$

Le système à résoudre pour trouver les points critiques est donc

$$\nabla \mathcal{L}(x,y,\lambda) = 0 \Rightarrow \begin{cases} \dfrac{\partial \mathcal{L}(x,y,\lambda)}{\partial x} = 0 \\ \dfrac{\partial \mathcal{L}(x,y,\lambda)}{\partial y} = 0 \\ \dfrac{\partial \mathcal{L}(x,y,\lambda)}{\partial \lambda} = 0 \end{cases} \Rightarrow \begin{cases} y + 8\lambda x = 0 \\ x + 2\lambda y = 0 \\ 4x^2 + y^2 - 4 = 0 \end{cases}$$

$$\Rightarrow \begin{cases} y = -8\lambda x \\ x = -2\lambda y \\ 4x^2 + y^2 - 4 = 0 \end{cases} \Rightarrow \begin{cases} y = -8\lambda(-2\lambda y) = 16\lambda^2 y \\ x = -2\lambda y \\ 4x^2 + y^2 - 4 = 0 \end{cases} \Rightarrow \begin{cases} y(1 - 16\lambda^2) = 0 \\ x = -2\lambda y \\ 4x^2 + y^2 - 4 = 0 \end{cases}$$

- *Si* $y = 0$:

$$\begin{cases} x = 0 \\ 4(0)^2 + (0)^2 = 0 \neq 4 \end{cases} \qquad \textit{alors pas de solution dans ce cas}$$

- *Si* $\lambda = \pm\frac{1}{4}$ *(car* $1 - 16\lambda^2 = 0$*)* :

$$\begin{cases} x = -2(\pm\frac{1}{4}y) = \pm\frac{1}{2}y \\ 4x^2 + y^2 = 4 \end{cases} \Rightarrow \begin{cases} x = \pm\frac{1}{2}y \\ 4(\pm\frac{1}{2}y)^2 + y^2 = 4 \end{cases}$$

$$\Rightarrow \begin{cases} x = \pm\frac{1}{2}y \\ y^2 = 2 \end{cases} \Rightarrow \begin{cases} x = \pm\frac{1}{2}y \\ y = \pm\sqrt{2} \end{cases}$$

dans ce cas on obtient 4 condidats $(\sqrt{2}/2, \sqrt{2})$; $(\sqrt{2}/2, -\sqrt{2})$; $(-\sqrt{2}/2, \sqrt{2})$; $(-\sqrt{2}/2, -\sqrt{2})$

Le tableau suivant permet de déterminer la nature des points critiques :

(x_0, y_0)	$(\frac{\sqrt{2}}{2}, \sqrt{2})$	$(\frac{\sqrt{2}}{2}, -\sqrt{2})$	$(\frac{-\sqrt{2}}{2}, \sqrt{2})$	$(\frac{-\sqrt{2}}{2}, -\sqrt{2})$
$f(x_0, y_0)$	*1*	*-1*	*-1*	*1*

les deux points critiques $(\frac{\sqrt{2}}{2}, -\sqrt{2})$, $(-\frac{\sqrt{2}}{2}, \sqrt{2})$ *sont des points minimaux.*

Exemple 2.3.2. *cas* $\mathbb{R}^3$

On veut résourdre :
$$\begin{cases} Min\ f(x_1, x_2, x_3) = 2x_1^2 + x_2^2 + 4x_3^2 \\ h_1(x_1, x_2, x_3) = x_1 + 2x_2 - x_3 = 6 \\ h_2(x_1, x_2, x_3) = 2x_1 - 2x_2 + 3x_3 = 12 \\ (x_1, x_2, x_3) \in \mathbb{R}^3 \end{cases}$$

les conditions de Lagrange :

$$\begin{cases} \nabla f(x_1, x_2, x_3) + \sum_{i=1}^{2} \lambda_i \nabla h_i(x_1, x_2, x_3) = 0. \\ \lambda_1, \lambda_2 \in \mathbb{R}. \end{cases}$$

$$\begin{cases} \begin{pmatrix} 4x_1 \\ 2x_2 \\ 8x_3 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ (x_1 + 2x_2 - x_3 - 6) = 0. \\ (2x_1 - 2x_2 + 3x_3 - 12) = 0. \\ \lambda_1, \lambda_2 \in \mathbb{R}. \end{cases}$$

— *Si* $\lambda_1 = 0$ *et* $\lambda_2 = 0$ *alors :* $\begin{cases} 4x_1 = 0 \\ 2x_2 = 0 \\ 8x_3 = 0 \end{cases} \Longrightarrow \begin{cases} x_1 = 0 \\ x_2 = 0 \\ x_3 = 0 \end{cases}$

$h_1(0,0,0) = 0 + 2 \times 0 - 0 = 0 \neq 6$ *non réalisable.*

— *Si* $\lambda_1 \neq 0$ *et* $\lambda_2 = 0$ *alors :* $\begin{cases} 4x_1 + \lambda_1 = 0 \\ 2x_2 + 2\lambda_1 = 0 \\ 8x_3 - \lambda_1 = 0 \\ x_1 + 2x_2 - x_3 - 6 = 0 \end{cases} \Longrightarrow \begin{cases} x_1 = \frac{6}{9.5} \\ x_2 = \frac{24}{9.5} \\ x_3 = \frac{-3}{9.5} \\ \lambda_1 = \frac{-24}{9.5} \end{cases}$

$h_1(\frac{6}{9.5}, \frac{24}{9.5}, \frac{-3}{9.5}) = \frac{6}{9.5} + 2 \times \frac{24}{9.5} - \frac{-3}{9.5} = \frac{57}{9.5} \neq 6$ *non réalisable.*

— *Si* $\lambda_1 = 0$ *et* $\lambda_2 \neq 0$ *alors :* $\begin{cases} 4x_1 + 2\lambda_2 = 0 \\ 2x_2 - 2\lambda_2 = 0 \\ 8x_3 - 3\lambda_2 = 0 \\ 2x_1 - 2x_2 + 3x_3 - 12 = 0 \end{cases} \Longrightarrow \begin{cases} x_1 = \frac{48}{25} \\ x_2 = \frac{-96}{25} \\ x_3 = \frac{36}{25} \\ \lambda_2 = \frac{-96}{25} \end{cases}$

$h_1(\frac{48}{25}, \frac{-96}{25}, \frac{36}{25}) = \frac{48}{25} + 2 \times \frac{-96}{25} - \frac{36}{25} = \frac{-180}{25} \neq 6$ *non réalisable.*

— $\lambda_1 \neq 0$ *et* $\lambda_2 \neq 0$ *alors On pose la matrice suivante :*

$$\begin{pmatrix} 1 & 2 & -1 & 0 & 0 \\ 2 & -2 & 3 & 0 & 0 \\ 4 & 0 & 0 & 1 & 2 \\ 0 & 2 & 0 & 2 & -2 \\ 0 & 0 & 8 & -1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 12 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{cases} x_1 = 5.044 \\ x_2 = 1.19 \\ x_3 = 1.43 \\ \lambda_1 = -7.52 \\ \lambda_2 = -6.32 \end{cases}$$

$h_1(5.044, 1.19, 1.43) = 5.044 + 2 \times 1.19 - 1.43 = 6$ *réalisable alors* h_1 *est active.*

$h_2(5.044, 1.19, 1.43) = 2 \times 5.044 - 2 \times 1.19 + 3 \times 1.43 = 12$ *réalisable alors* h_2 *est active.*

la valeur de la fonction f *au point* $(5.044, 1.19, 1.43)$ *est :*

$f(5.044, 1.19, 1.43) = 2 \times 5.044^2 + 1.19^2 + 4 \times 1.43^2 = 60.48$

2.3.3 Conditions d'optimalité du 1^{er} ordre, cas en égalité et en inégalité

[3] Considérons un problème d'optimisation sous forme générale :

$$(P) \begin{cases} \min \ f(x), \\ h_i(x) = 0, \ i = 1, \dots, \ p \\ g_j(x) \leq 0, \ j = 1, \dots, \ q \\ x \in \mathbb{R}^n. \end{cases}$$

Théorème 2.3.4. *[3](Conditions d'optimalité non qualifiées)*
On suppose que f, h_i *pour* $i = \overline{1,p}$ *et* g_j *pour* $j = \overline{1,q}$ *sont de classe* C^1. *Soit* x^* *une solution de* (P).
Alors il existe $\lambda^* = (\lambda_1^*, \ \lambda_2^*, \dots, \ \lambda_p^*) \in \mathbb{R}^p$, $\mu^* = (\mu_1^*, \ \mu_2^*, \dots, \ \mu_q^*) \in \mathbb{R}_+^q$ *et* $\mu_0^* \in \mathbb{R}_+$

tels que :

- $\forall j \in \{0, 1, \ldots, q\}, : \quad \mu_j^* \geq 0;$ (2.3a)
- $h_i(x^*) = 0$ *pour* $i = \overline{1,p}$ *et* $g_j(x^*) \leq 0$ *pour* $j = \overline{1,q};$ (2.3b)
- $\forall j \in \{1, \ldots, q\} \ : \ \mu_j^* \ g_j(x^*) = 0;$ (2.3c)
- $\mu_0^* \nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla h_i(x^*) + \sum_{i=1}^{q} \mu_j^* \nabla g_j(x^*) = 0.$ (2.3d)

Preuve 2.3.4. *La preuve consiste à remplacer le problème* (P) *avec contraintes par la suite de problèmes sans contraintes :*

$$(P_k) \begin{cases} \min \ f_k(x) = f(x) + k\alpha(x), \\ x \in \mathbb{R}^n \end{cases}$$

où $\alpha : \mathbb{R}^n \to \mathbb{R}$ *est une fonction de pénalisation des contraintes et* $k > 0$. α *est choisie de façon à avoir* (P) *et* (P_k) *équivalents (c.à.d. ayant les mêmes solutions).*

Pour tout entier k*, considérons le problème :*

$$(P_k) \begin{cases} \min \ f_k(x), \\ x \in B(x^*, \rho). \end{cases}$$

où

$$f_k(x) = f(x) + \frac{k}{2}\sum_{i=1}^{p}[h_i(x)]^2 + \frac{k}{2}\sum_{j=1}^{q}[g_j^+(x)]^2 + \|x - x^*\|^2, \tag{2.4}$$

avec $g_j^+(x) = max(0, g_j(x))$ *et* $B(x^*, \rho)$ *est une boule fermée (compact) centrée en* x^* *et de rayon* $\rho > 0$.

Remarque 2.3.1. *Les fonction* g_j *sont de classe* C^1*, alors* $(g_j^+)^2$ *est différentiable et*

$$\frac{d(g_j^+)^2}{dx}(x) = 2\frac{dg_i}{dx}(x)g_j^+(x).$$

• Le problème (P_k) a au moins une solution x_k

[3] En effet, f_k est continue, elle atteint donc son minimum sur le compact $B(x^*, \rho)$.

• La suite (x_k) converge vers x^*

[3] La suite (x_k) est dans le compact $B(x^*, \rho)$ et on peut en extraire une sous-suite, noté (x_k), qui converge vers $\tilde{x} \in B(x^*, \rho)$.
Comme $f_k(x_k) \leq f_k(x^*) = f(x^*) < +\infty$ car x_k est le mimnimum de f_k sur $B(x^*, \rho)$, nous avons alors

$$\sum_{i=1}^{p}[h_i(x_k)]^2 + \sum_{i=1}^{q}[g_j^+(x_k)]^2 \leq \frac{2}{k}[f(x^*) - f(x_k) - \|x_k - x^*\|^2]. \tag{2.5}$$

$[f(x^*) - f(x_k) - \|x_k - x^*\|^2]$ étant borné, donc

$$\forall i = \overline{1,p}, \ \lim_{k\to+\infty} h_i(x_k) = h_i(\tilde{x}) = 0,$$

et

$$\forall j = \overline{1,q}, \ \lim_{k\to+\infty} g_j^+(x_k) = g_j^+(\tilde{x}) = 0.$$

Alors $\tilde{x}$ est une solution réalisable.
D'autre part :

$$f(x_k) + \|x_k - x^*\|^2 \leq f_k(x_k) \leq f(x^*)$$

car $\frac{k}{2}[h_i(x_k)]^2 + \frac{k}{2}\sum\limits_{j=1}^{q}[g_j^+(x_k)]^2 \geq 0$ (c'est une somme de carrés).
Comme x^* est une solution du problème (P), alors

$$f(\tilde{x}) + \|\tilde{x} - x^*\|^2 \leq f(x^*) \leq f(\tilde{x}).$$

Donc $\|\tilde{x} - x^*\|^2 = 0$. Par conséquent, $\tilde{x} = x^*$.
Ce raisonnement peut être fait pour toute valeur d'adhérence de la suite (x_k).
Donc la suite (x_k) converge vers x^*.

- **Condition d'optimalité pour (P_k)**

[3] Comme (x_k) converge vers x^*, elle est dans $B(x^*, \rho)$ à partir d'un certain rang et donc x_k est un minimum local (sans contraintes) de f_k. Par conséquent, $\nabla f_k(x_k) = 0$ pour k assez grand, c.à.d.

$$\nabla f(x_k) + k \sum_{i=1}^{p} [h_i(x_k) \nabla h_i(x_k)] + k \sum_{j=1}^{q} [g_j^+(x_k) \nabla g_j(x_k)] + 2(x_k - x^*) = 0. \quad (2.6)$$

Posons

$$S_k = \left(1 + k^2 \sum_{i=1}^{p} [h_i(x_k)]^2 + k^2 \sum_{j=1}^{q} [g_j^+(x_k)]^2\right)^{\frac{1}{2}}$$

$\mu_0^k = \frac{1}{S_k}$, $\lambda_i^k = \frac{k h_i(x_k)}{S_k}$, *pour* $i = \overline{1,p}$ *et* $\mu_j^k = \frac{k g_j^+(x_k)}{Sk}$, *pour* $j = \overline{1,q}$.
La relation (2.6) devient :

$$\mu_0^k \nabla f(x_k) + \sum_{i=1}^{p} [\lambda_i^k \nabla h_i(x_k)] + \sum_{j=1}^{q} [\mu_j^k \nabla g_j(x_k)] + \frac{2}{S_k}(x_k - x^*) = 0. \quad (2.7)$$

Comme le vecteur $(\lambda_1^k, \ldots, \lambda_p^k, \mu_0^k, \mu_1^k, \ldots, \mu_q^k) \in \mathbb{R}_+^p \times \mathbb{R}_+ \times \mathbb{R}^q$ est de norme 1, on peut en extraire une sous-suite convergente vers $(\lambda_1^k, \ldots, \lambda_p^k, \mu_0^k, \mu_1^k, \ldots, \mu_q^k) \neq 0$ et passer à la limite dans (2.7). Alors, on aura :

$$\mu_0^* \nabla f(x^*) + \sum_{i=1}^{p} [\lambda_i^* \nabla h_i(x^*)] + \sum_{j=1}^{q} [\mu_j^* \nabla g_j(x^*)] = 0 \quad (2.8)$$

car toutes les fonctions considérées sont continues et $S_k \geq 1$. Notons que $\mu_j^* \geq 0, \forall j = \overline{0,q}$.

- **Relation (2.3c)**

[3] Si $g_j(x^*) < 0$, alors $g_j(x_k) < 0$ à partir d'un certain rang et $\mu_j^k = 0$. Par conséquent, en passant à la limite $\mu_j^* = 0$.

Remarque 2.3.2. *[3] Quelques remarques importantes :*

1. *Les réels λ_i^* et μ_j^* sont les multiplicateurs de Lagrange.*

2. *La relation (2.3b) est une relation de réalisabilité, c.à.d. que tout point x véerifiant (2.3b) est une solution réalisable.*
3. *La relation (2.3c) est une relation de complémentarité.*
4. *Les conditions du théorème (2.3d) sont dites non qualifiées car le réel μ_0^* peut être nul et on n'a pas de renseignements sur le minimum de f puisqu'elle n'apparaît nulle part dans la relation d'optimalité. Donc il est important de donner des conditions qui permettent d'assurer que μ_0^* soit non nul.*

   De telles conditions sont dites conditions de qualification ou de régularité. Lorsqu'elles sont vérifiées le problème est dit qualifié.

Définition 2.3.2. *[3](Point régulier ou condition de qualification 1)*
On dit qu'un élément $x^* \in \mathbb{R}^n$ est régulier pour les contraintes h_i pour $i = \overline{1,p}$ et g_j pour $j = \overline{1,q}$:

- *s'il est réalisable : $h_i(x^*) = 0$ pour $i = \overline{1,p}$ et $g_j(x^*) \leq 0$ pour $j = \overline{1,q}$*
- *si les vecteurs $\nabla h_i(x^*)$ pour $i = \overline{1,p}$ sont linéairement indépendants*
- *et si on peut trouver une direction $d \in \mathbb{R} \setminus \{0\}$ tel que*

$$\langle \nabla h_i(x^*), d \rangle = 0, \ i = \overline{1,p} \text{ et } \langle \nabla g_j(x^*), d \rangle < 0, \ \forall j = \overline{1,q}.$$

On dit aussi que x^* vérifie la condition de qualification de Mangasarian-Fromowitz que nous noterons (2.3.2).
La condition (2.3.3) suivante est plus forte que la condition (2.3.2).

Définition 2.3.3. *[3](Condition de qualification 2)*
On dit qu'un élément $x^* \in \mathbb{R}^n$ est régulier pour les contraintes h_i pour $i = \overline{1,p}$ et g_j pour $j = \overline{1,q}$:

- *s'il est réalisable,*
- *et si les vecteurs $\nabla h_i(x^*), \nabla g_j(x^*)$ pour $i = \overline{1,p}$ et $j = \overline{1,q}$ sont linéairement indépendants*

Exemple 2.3.3. *considérons le problème :*
$$\begin{cases} min\ f(x,y) = 4x^2 + 5y^2 + 6xy - 6x - 6y + \\ y - x + \frac{3}{11} = 0 \\ x + y - 1 \leq 0 \\ x \geq 0 \\ y \geq 0 \\ (x,y) \in \mathbb{R}^2 \end{cases}$$

le point (0,0) n'est pas active

f est un fonction continue et strictement convexe.

$$\nabla h(x,y) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \ \nabla g_1(x,y) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ \nabla g_2(x,y) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \ \nabla g_3(x,y) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

la famille $\{\nabla h(x,y),\ \nabla g_1(x,y),\ \nabla g_2(x,y),\ \nabla g_3(x,y)\}$ sont libres.

Théorème 2.3.5. *[3](Conditions de Karush-Kuhn-Tucker)*

On suppose que f, h_i pour $i = \overline{1,p}$ et g_j pour $j = \overline{1,q}$ sont de classe C^1. Soit x^* une solution de (P).

On suppose que x^* est un point régulier pour les contraintes $h_i,\ i = \overline{1,p}$ et $g_j,\ j = \overline{1,q}$.

Alors il existe $\lambda^* = (\lambda_1^*,\ \lambda_2^*, \ldots,\ \lambda_p^*) \in \mathbb{R}^p$ et $\mu^* = (\mu_1^*,\ \mu_2^*, \ldots,\ \mu_q^*) \in \mathbb{R}_+^q$ tels que :

- $\forall j \in \{1, \ldots, q\},\ \mu_j^* \geq 0;$ (2.9a)
- $h_i(x^*) = 0$ *pour* $i = \overline{1,p}$ *et* $g_j(x^*) \leq 0$ *pour* $j = \overline{1,q};$ (2.9b)
- $\forall j \in \{1, \ldots, q\} : \mu_j^* g_j(x^*) = 0;$ (2.9c)
- $\nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^{q} \mu_j^* \nabla g_j(x^*) = 0.$ (2.9d)

Preuve 2.3.5. *On montre que sous les hypothèses de régularité (2.3.2), le réel μ_0^* donné dans le théorème (2.3d) est non nul.*

On a $(\lambda_1^*,\ \lambda_2^*, \ldots,\ \lambda_p^*,\ \mu_0^*,\ \mu_1^*,\ \mu_2^*, \ldots,\ \mu_q^*) \neq 0$. On suppose que $\mu_0^* = 0$ et supposons aussi que $\mu_j^* = 0,\ \forall j \in I(x^*)$.

Alors le vecteur $\lambda^* = (\lambda_1^*, \ \lambda_2^*, \ldots, \ \lambda_p^*) \neq 0$, *on a alors* $\sum_{i=1}^{p} \lambda_i^* \nabla h_i(x^*) = 0$. *Ce résultat contredit l'indépendances linéaire des* $\nabla h_i(x^*), i = \overline{1,p}$.

Donc il existe $j_0 \in I(x^*)$ *tel que* $\mu_{j_0}^* \neq 0$.

Nous avons, en prenant la direction d donnée par (2.3.2) :

$$0 = \sum_{i=1}^{p} \lambda_i^* \langle \nabla h_i(x^*), d\rangle + \sum_{j=1}^{q} \mu_j^* \langle \nabla g_j(x^*), d\rangle < \mu_{j_0}^* \langle \nabla g_{j_0}(x^*), d\rangle < 0,$$

d'où la contradiction.

L'ensemble des contraintes (2.9a)-(2.9d) sont appelées conditions de Karush-Kuhn-Tucker, notées K.K.T.

Définition 2.3.4. *(Lagrangien)*

On appelle lagrangien du problème (P) la fonction définie de $\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q$ *dans* $\mathbb{R}$ *par :*

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^{p} \lambda_i h_i(x) + \sum_{j=1}^{q} \mu_j g_j(x).$$

Remarque 2.3.3. *La relation (2.9d) s'écrit*

$$\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) = 0,$$

où ∇_x *désigne le gradient du lagrangien par rapport à la première variable.*

Dans le cas convexe, nous avons le résultat suivant :

Théorème 2.3.6. *(C.N.S. dans le cas convexe)*

On suppose que f, h_i *pour* $i = \overline{1,p}$ *et* g_j *pour* $j = \overline{1,q}$ *sont de classe* C^1 *et que* f *et* $g_j, j = \overline{1,q}$ *sont convexes et* $h_i, i = \overline{1,p}$ *sont affines. On suppose aussi que* x^* *est un point régulier. Alors,*

$$(x^* \text{ est solution de } (P)\ ) \Leftrightarrow (\text{les conditions } (2.9a)\text{-}(2.9d) \text{ sont satisfaites}).$$

Preuve 2.3.6. *[3] On montre que si les conditions (2.9a)-(2.9d) sont satisfaites alors* x^* *est une solution du problème (P).*

De (2.9b), on déduit que x^* est réalisable. Comme toutes les fonctions sont convexes, alors le lagrangien est convexe par rapport à la variable x et la condition (2.9d) est équivalente à dire que x^* est un minimum de $\mathcal{L}(x, \lambda^*, \mu^*)$. On obtient donc :

$$\forall x \in \mathbb{R}^n : \mathcal{L}(x^*, \lambda^*, \mu^*) \leq \mathcal{L}(x, \lambda^*, \mu^*)$$

Si $x \in C$, alors $h_i(x) = 0$ et $g_j(x) \leq 0$ de sorte que

$$\sum_{i=1}^{p} \lambda_i^* h_i(x) + \sum_{j=1}^{q} \mu_j^* g_j(x) = \sum_{j=1}^{q} \mu_j^* g_j(x) \leq 0,$$

puisque $\mu_j^* \geq 0$ d'après (2.9a).
De plus avec la relation de complémentarité (2.9c), on voit que

$$\mathcal{L}(x^*, \lambda^*, \mu^*) = f(x^*).$$

Finalement, on obtient

$$f(x^*) = \mathcal{L}(x^*, \lambda^*, \mu^*) \leq \mathcal{L}(x, \lambda^*, \mu^*) = f(x^*) \leq f(x).$$

Ce qui preuve que x^* est solution du problème (P).

Exemple 2.3.4. *On veut résourdre :* $\begin{cases} \min f(x,y) = 4x^2 + 5y^2 + 6xy - 6x - 6y + 3 \\ x + y \leq 1 \\ x \geq 0 \\ y \geq 0 \\ (x,y) \in \mathbb{R}^2 \end{cases}$

On écrit les contraintes sous la forme générale : $\begin{cases} \min f(x,y) = 4x^2 + 5y^2 + 6xy - 6x - 6y + 3 \\ x + y - 1 \leq 0 \\ x \geq 0 \\ y \geq 0 \\ (x,y) \in \mathbb{R}^2 \end{cases}$

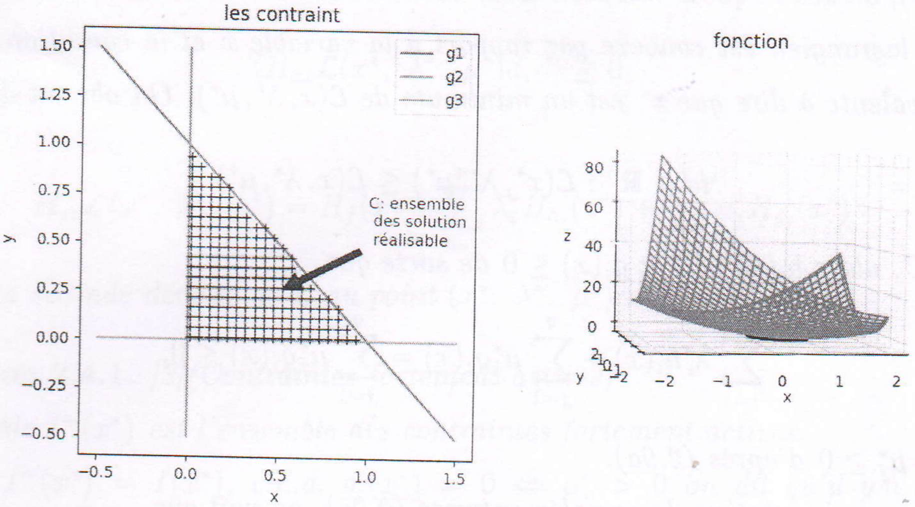

FIGURE 2.2 – la fonction f et les contraintes g_i

les condition de KKT

on a

$$\begin{cases} \nabla f(x,y) + \sum_{i=1}^{3} u_i \nabla g_1(x,y) = 0 \\ u_1, u_2, u_3 \geq 0 \\ u_i g_i(x,y) = 0, \quad i = \overline{1,3} \end{cases}$$

$$\Rightarrow \begin{cases} \begin{pmatrix} 8x + 6y - 6 \\ 10y + 6x - 6 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + u_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 0 \\ u_1, u_2, u_3 \geq 0 \\ u_1(x + y - 1) = 0 \\ u_2(-x) = 0 \\ u_3(-y) = 0 \end{cases}$$

- *si* $u_1 = 0$ *et* $u_2 = 0$ *et* $u_3 = 0$

$$\begin{cases} 8x + 6y = 0 \\ 10y + 6x - 6 = 0 \end{cases} \Rightarrow \begin{cases} x = \frac{6}{11} \\ y = \frac{3}{11} \end{cases} \textit{réalisable}$$

d'après le théorème de Weierstrass le pt$(\frac{6}{11}, \frac{3}{11})$ *est minimum global.*

- $u_1 = u_3 = 0$ *et* $u_2 > 0 \Rightarrow x = 0$ *et* $y = \frac{6}{11}$ *et* $u_2 = \frac{-24}{10} < 0$ *refuse.*
- *si* $u_1 = u_2 = 0$ *et* $u_3 > 0 \Rightarrow y = 0$ *et* $x = \frac{6}{8}$ *et* $u_3 = \frac{-3}{2} < 0$ *refuse.*
- $u_2 = u_3 = 0$ *et* $u_1 > 0 \Rightarrow x = \frac{2}{3}$ *et* $y = \frac{1}{3}$ *et*$u_1 = \frac{-4}{3} < 0$ *refuse.*
- *si* $u_1 > 0$ *et* $u_2 > 0$*et* $u_3 = 0 \Rightarrow y = 1$ *et* $x = 0$ *et* $u_1 = -4 < 0$ *et* $u_2 = u_1 = -4 < 0$ *refuse.*
- *si* $u_1 > 0$ *et* $u_3 > 0$*et* $u_2 = 0 \Rightarrow x = 1$ *et* $y = 0$ *et* $u_1 = u_3 = -2 < 0$ *refuse.*
- *si* $u_1 = 0$ *et* $u_2 > 0$*et* $u_3 > 0 \Rightarrow y = 0$ *et* $x = 0$ *et* $u_2 = u_3 = -6 < 0$ *et* $u_2 = u_1 = -4 < 0$ *refuse.*
- *si* $u_1 > 0$ *et* $u_2 > 0$*et* $u_3 > 0 \Rightarrow u_2 = -u_3$ *et* $u_1 = 6$ *et* $x = y = 0$ *et refuse.*

2.4 Conditions d'optimalité du second ordre

Les conditions d'optimalité du premier ordre permettent de déterminer les bons condidats à la solution de (P). Les Conditions d'optimalité du second ordre vont permettre dans un premiers temps de restreindre encore le nombre de candidats.

Théorème 2.4.1. *[9](Condition nécessaire du second ordre)*
On suppose que f, h_i, $i = \overline{1,p}$ *et* g_j, $j = \overline{1,q}$ *sont de classe* C^2, *que* x^* *est un minimum de* f *sur* C *et que la condition (CQ1) est vérifiée.*
Alors il existe $\lambda^* = (\lambda_1^*, \lambda_2^*, \ldots, \lambda_p^*) \in \mathbb{R}^p$ *et* $\mu^* = (\mu_1^*, \mu_2^*, \ldots, \mu_q^*) \in \mathbb{R}^q$ *tels que :*

- *Les relations (2.9a)-(2.9d) de K.K.T. sont satisfaites*
- *Pour toute direction* $d \in \mathbb{R}^n \setminus \{0\}$ *vérifiant :*

$$\begin{cases} \langle \nabla h_i(x^*), d\rangle = 0, & \textit{pour } i = \overline{1,p}; \\ \langle \nabla g_j(x^*), d\rangle = 0, & \textit{pour } j \in I^+(x^*); \\ \langle \nabla g_j(x^*), d\rangle \leq 0, & \textit{pour } j \in I(x^*) \setminus I^+(x^*); \end{cases} \tag{2.10}$$

où $I^+(x^*) = \{j \in \{1, 2, \ldots, q\} : g_j(x^*) = 0$ *et* $\mu_j^* > 0\}$, *on a*

$$\langle H_{xx}\mathcal{L}(x^*,\ \lambda^*,\ \mu^*)d, d\rangle \geq 0, \tag{2.11}$$

avec

$$H_{xx}\mathcal{L}(x^*,\ \lambda^*,\ \mu^*) = H_f(x) + \sum_{i=1}^{p} \lambda_i^* H_{h_i}(x^*) + \sum_{j=1}^{q} \mu_j^* H_{g_j}(x^*)$$

désigne la seconde dérivée de $\mathcal{L}$ *au point* $(x^*,\ \lambda^*,\ \mu^*)$.

Définition 2.4.1. *[3](Contraintes fortement actives)*
L'ensemble $I^+(x^*)$ *est l'ensemble des contraintes fortement actives.*
Lorsque $I^+(x^*) = I(x^*)$, *c.à.d.* $g_j(x^*) = 0 \Leftrightarrow \mu_j^* > 0$ *on dit qu'il y'a stricte complémentarité.*
Le résultat suivant donne une condition suffisante du second ordre.

Théorème 2.4.2. *[3](Condition suffisante du second ordre)*
On suppose que $f, h_i,\ i = \overline{1,p}$ *et* $g_j,\ j = \overline{1,q}$ *sont de classe* $\mathcal{C}^2$. *Soit* $x^* \in \mathbb{R}^n$ *vérifiant les conditions de K.K.T. avec les multiplicateurs de Lagrange* $\lambda^*,\ \mu^*$.
Si la matrice hessienne du lagrangien au point $(x^*,\ \lambda^*,\ \mu^*)$, $\nabla^2_{xx}\mathcal{L}(x^*,\ \lambda^*,\ \mu^*)$, *est définie positive sur le sous-espace*

$$\tau = \{d \in \mathbb{R}^n \setminus \{0\} : <\nabla h_i(x^*), d\rangle = 0,\ i = \overline{1,p} \text{ et } <\nabla g_j(x^*), d\rangle = 0, j \in I^+(x^*)\}$$

alors x^* *est un minimum strict de* f *sur* C.