3.4 Optimisation sous contraintes

3.4.1 Définitions

Soit $E = \mathbb{R}^n$, soit $f \in C(E, \mathbb{R})$, et soit K un sous ensemble de E. On s'intéresse à la recherche de $\bar{u} \in K$ tel que :

$$\begin{cases} \bar{u} \in K \\ f(\bar{u}) = \inf\limits_K f \end{cases} \tag{3.48}$$

Ce problème est un problème de minimisation avec contrainte (ou "sous contrainte") au sens où l'on cherche u qui minimise f en restreignant l'étude de f aux éléments de K. Voyons quelques exemples de ces contraintes (définies par l'ensemble K), qu'on va expliciter à l'aide des p fonctions continues, $g_i \in C(E, \mathbb{R})$ $i = 1 \ldots p$.

1. **Contraintes égalités.** On pose $K = \{x \in E, g_i(x) = 0 \; i = 1 \ldots p\}$. On verra plus loin que le problème de minimisation de f peut alors être résolu grâce au théorème des multiplicateurs de Lagrange (voir théorème 3.34).
2. **Contraintes inégalités.** On pose $K = \{x \in E, g_i(x) \leq 0 \; i = 1 \ldots, p\}$. On verra plus loin que le problème de minimisation de f peut alors être résolu grâce au théorème de Kuhn–Tucker (voir théorème 3.38).
   - *Programmation linéaire.* Avec un tel ensemble de contraintes K, si de plus f est linéaire, c'est-à-dire qu'il existe $b \in \mathbb{R}^n$ tel que $f(x) = b \cdot x$, et les fonctions g_i sont affines, c'est-à-dire qu'il existe $b_i \in \mathbb{R}^n$ et $c_i \in \mathbb{R}$ tels que $g_i(x) = b_i \cdot x + c_i$, alors on dit qu'on a affaire à un problème de "programmation linéaire". Ces problèmes sont souvent résolus numériquement à l'aide de l'algorithme de Dantzig, inventé vers 1950.
   - *Programmation quadratique.* Avec le même ensemble de contraintes K, si de plus f est quadratique, c'est-à-dire si f est de la forme $f(x) = \dfrac{1}{2} Ax \cdot x - b \cdot x$, et les fonctions g_i sont affines, alors on dit qu'on a affaire à un problème de "programmation quadratique".
3. **Programmation convexe.** Dans le cas où f est convexe et K est convexe, on dit qu'on a affaire à un problème de "programmation convexe".

3.4.2 Existence – Unicité – Conditions d'optimalité simple

Théorème 3.28 (Existence)**.** *Soit $E = \mathbb{R}^n$ et $f \in C(E, \mathbb{R})$.*

1. *Si K est un sous-ensemble fermé borné de E, alors il existe $\bar{\boldsymbol{x}} \in K$ tel que $f(\bar{\boldsymbol{x}}) = \inf\limits_K f$.*
2. *Si K est un sous-ensemble fermé de E, et si f est croissante à l'infini, c'est–à–dire que $f(x) \to +\infty$ quand $|x| \to +\infty$, alors $\exists \bar{\boldsymbol{x}} \in K$ tel que $f(\bar{\boldsymbol{x}}) = \inf\limits_K f$*

DÉMONSTRATION –

1. Si K est un sous-ensemble fermé borné de E, comme f est continue, elle atteint ses bornes sur K, d'où l'existence de $\bar{\boldsymbol{x}}$.
2. Si f est croissante à l'infini, alors il existe $R > 0$ tel que si $\|x\| > R$ alors $f(x) > f(0)$; donc $\inf\limits_K f = \inf\limits_{K \cap B_R} f$, où B_R désigne la boule de centre 0 et de rayon R. L'ensemble $K \cap B_R$ est compact, car intersection d'un fermé et d'un compact. Donc, par ce qui précède, il existe $\bar{\boldsymbol{x}} \in K$ tel que $f(\bar{\boldsymbol{x}}) = \inf\limits_{K \cap B_R} f = \inf\limits_{B_R} f$.

■

Théorème 3.29 (Unicité)**.** *Soit $E = \mathbb{R}^n$ et $f \in C(E, \mathbb{R})$. On suppose que f est strictement convexe et que K est convexe. Alors il existe au plus un élément $\bar{\boldsymbol{x}}$ de K tel que $f(\bar{\boldsymbol{x}}) = \inf_K f$.*

DÉMONSTRATION – Supposons que $\bar{\boldsymbol{x}}$ et $\bar{\bar{\boldsymbol{x}}}$ soient deux solutions du problème (3.48), avec $\bar{\boldsymbol{x}} \neq \bar{\bar{\boldsymbol{x}}}$
Alors $f(\frac{1}{2}\bar{\boldsymbol{x}} + \frac{1}{2}\bar{\bar{\boldsymbol{x}}}) < \dfrac{1}{2}f(\bar{\boldsymbol{x}}) + \dfrac{1}{2}f(\bar{\bar{\boldsymbol{x}}}) = \inf_K f$. On aboutit donc à une contradiction. ∎

Des théorèmes d'existence 3.28 et d'unicité 3.29 on déduit immédiatement le théorème d'existence et d'unicité suivant :

Théorème 3.30 (Existence et unicité)**.** *Soient $E = \mathbb{R}^n$, $f \in C(E, \mathbb{R}^n)$ une fonction strictement convexe et K un sous ensemble convexe fermé de E. Si K est borné ou si f est croissante à l'infini, c'est–à–dire si $f(x) \to +\infty$ quand $\|x\| \to +\infty$, alors il existe un unique élément $\bar{\boldsymbol{x}}$ de K solution du problème de minimisation* (3.48)*, i.e. tel que $f(\bar{\boldsymbol{x}}) = \inf_K f$*

Remarque 3.31. *On peut remplacer $E = \mathbb{R}^n$ par E espace de Hilbert de dimension infinie dans le dernier théorème, mais on a besoin dans ce cas de l'hypothèse de convexité de f pour assurer l'existence de la solution (voir cours de maîtrise).*

Proposition 3.32 (Condition simple d'optimalité)**.** *Soient $E = \mathbb{R}^n$, $f \in C(E, \mathbb{R})$ et $\bar{\boldsymbol{x}} \in K$ tel que $f(\bar{\boldsymbol{x}}) = \inf_K f$. On suppose que f est différentiable en $\bar{\boldsymbol{x}}$*

1. *Si $\bar{\boldsymbol{x}} \in \overset{\circ}{K}$ alors $\nabla f(\bar{\boldsymbol{x}}) = 0$.*
2. *Si K est convexe, alors $\nabla f(\bar{\boldsymbol{x}}) \cdot (x - \bar{\boldsymbol{x}}) \geq 0$ pour tout $x \in K$.*

DÉMONSTRATION –
1. Si $\bar{\boldsymbol{x}} \in \overset{\circ}{K}$, alors il existe $\varepsilon > 0$ tel que $B(\bar{\boldsymbol{x}}, \varepsilon) \subset K$ et $f(\bar{\boldsymbol{x}}) \leq f(x)\ \forall x \in B(\bar{\boldsymbol{x}}, \varepsilon)$. Alors on a déjà vu (voir preuve de la Proposition 3.7 page 214) que ceci implique $\nabla f(\bar{\boldsymbol{x}}) = 0$.
2. Soit $x \in K$. Comme $\bar{\boldsymbol{x}}$ réalise le minimum de f sur K, on a : $f(\bar{\boldsymbol{x}} + t(x - \bar{\boldsymbol{x}})) = f(tx + (1-t)\bar{\boldsymbol{x}}) \geq f(\bar{\boldsymbol{x}})$ pour tout $t \in]0, 1]$, par convexité de K. On en déduit que
$$\frac{f(\bar{\boldsymbol{x}} + t(x - \bar{\boldsymbol{x}})) - f(\bar{\boldsymbol{x}})}{t} \geq 0 \text{ pour tout } t \in]0, 1].$$
En passant à la limite lorsque t tend vers 0 dans cette dernière inégalité, on obtient : $\nabla f(\bar{\boldsymbol{x}}) \cdot (x - \bar{\boldsymbol{x}}) \geq 0$. ∎

3.4.3 Conditions d'optimalité dans le cas de contraintes égalité

Dans tout ce paragraphe, on considèrera les hypothèses et notations suivantes :

$$\begin{aligned} & f \in C(\mathbb{R}^n, \mathbb{R}), \quad g_i \in C^1(\mathbb{R}^n, \mathbb{R}), \;\; i = 1 \ldots p\,; \\ & K = \{u \in \mathbb{R}^n, \;\; g_i(u) = 0 \;\; \forall i = 1 \ldots p\}\,; \\ & g = (g_1, \ldots, g_p)^t \in C^1(\mathbb{R}^n, \mathbb{R}^p) \end{aligned} \tag{3.49}$$

Remarque 3.33 (Quelques rappels de calcul différentiel)**.**

Comme $g \in C^1(\mathbb{R}^n, \mathbb{R}^p)$, si $u \in \mathbb{R}^n$, alors $Dg(u) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$, ce qui revient à dire, en confondant l'application linéaire $Dg(u)$ avec sa matrice, que $Dg(u) \in \mathcal{M}_{p,n}(\mathbb{R})$. Par définition, $Im(Dg(u)) = \{Dg(u)z,\ z \in \mathbb{R}^n\} \subset \mathbb{R}^p$, et rang$(Dg(u)) = \dim(Im(Dg(u))) \leq p$. On rappelle de plus que

$$Dg(u) = \begin{pmatrix} \dfrac{\partial g_1}{\partial x_1}, & \dots, & \dfrac{\partial g_1}{\partial x_n} \\ \dots, & \ddots, & \dots \\ \dfrac{\partial g_p}{\partial x_1}, & \dots, & \dfrac{\partial g_p}{\partial x_n} \end{pmatrix},$$

et que rang $(Dg(u)) \leq \min(n,p)$. De plus, si rang $(Dg(u)) = p$, alors les vecteurs $(Dg_i(u))_{i=1\dots p}$ sont linéairement indépendants dans $\mathbb{R}^n$.

Théorème 3.34 (Multiplicateurs de Lagrange)**.** *Soit $\bar{u} \in K$ tel que $f(\bar{u}) = \inf_K f$. On suppose que f est différentiable en $\bar{u}$ et $\dim(Im(Dg(\bar{u})) = p$ (ou rang $(Dg(\bar{u})) = p$), alors :*

$$\textit{il existe } (\lambda_1, \dots, \lambda_p)^t \in \mathbb{R}^p \textit{ tels que } \nabla f(\bar{u}) + \sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{u}) = 0.$$

(Cette dernière égalité a lieu dans $\mathbb{R}^n$)

DÉMONSTRATION – Pour plus de clarté, donnons d'abord une idée "géométrique" de la démonstration dans le cas $n = 2$ et $p = 1$. On a dans ce cas $f \in C^1(\mathbb{R}^2, \mathbb{R})$ et $K = \{(x,y) \in \mathbb{R}^2 g(x,y) = 0\}$, et on cherche $u \in K$ tel que $f(u) = \inf_K f$. Traçons dans le repère (x,y) la courbe $g(x,y) = 0$, ainsi que les courbes de niveau de f. Si on se "promène" sur la courbe $g(x,y) = 0$, en partant du point P_0 vers la droite (voir figure 3.1), on rencontre les courbes de niveau successives de f et on se rend compte sur le dessin que la valeur minimale que prend f sur la courbe $g(x,y) = 0$ est atteinte lorsque cette courbe est tangente à la courbe de niveau de f : sur le dessin, ceci correspond au point P_1 où la courbe $g(x,y) = 0$ est tangente à la courbe $f(x,y) = 3$. Une fois qu'on a passé ce point de tangence, on peut remarquer que f augmente.
On utilise alors le fait que si φ est une fonction continûment différentiable de $\mathbb{R}^2$ dans $\mathbb{R}$, le gradient de φ est orthogonal à toute courbe de niveau de φ, c'est-à-dire toute courbe de la forme $\varphi(x,y) = c$, où $c \in \mathbb{R}$. (En effet, soit $(x(t), y(t))$, $t \in \mathbb{R}$ un paramétrage de la courbe $g(x,y) = c$, en dérivant par rapport à t, on obtient : $\nabla g(x(t), y(t)) \cdot (x'(t), y'(t))^t = 0$). En appliquant ceci à f et g, on en déduit qu'au point de tangence entre une courbe de niveau de f et la courbe $g(x,y) = 0$, les gradients de f et g sont colinéaires. Et donc si $\nabla g(u) \neq 0$, il existe $\lambda \neq 0$ tel que $\nabla f(u) = \lambda \nabla g(u)$.

Passons maintenant à la démonstration rigoureuse du théorème dans laquelle on utilise le théorème des fonctions implicites [5].
Par hypothèse, $Dg(\bar{u}) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ et $Im(Dg(\bar{u})) = \mathbb{R}^p$. Donc il existe un sous espace vectoriel F de $\mathbb{R}^n$ de dimension p, tel que $Dg(\bar{u})$ soit bijective de F dans $\mathbb{R}^p$. En effet, soit $(e_1 \dots e_p)$ la base canonique de $\mathbb{R}^p$, alors pour tout $i \in \{1, \dots p\}$, il existe $y_i \in \mathbb{R}^n$ tel que $Dg(\bar{x})y_i = e_i$. Soit F le sous espace engendré par la famille $\{y_1 \dots y_p\}$; on remarque que cette famille est libre, car si $\sum_{i=1}^p \lambda_i y_i = 0$, alors $\sum_{i=1}^p \lambda_i e_i = 0$, et donc $\lambda_i = 0$ pour tout $i = 1, \dots p$. On a ainsi montré l'existence d'un sous espace F de dimension p telle que $Dg(\bar{x})$ soit bijective (car surjective) de F dans $\mathbb{R}^p$.
Il existe un sous espace vectoriel G de $\mathbb{R}^n$, tel que $\mathbb{R}^n = F \bigoplus G$. Pour $v \in F$ et $w \in G$; on pose $\bar{g}(w,v) = g(v+w)$ et $\bar{f}(w,v) = f(v+w)$. On a donc $\bar{f} \in C(G \times F,\ \mathbb{R})$ et $\bar{g} \in C^1(G \times F,\ \mathbb{R})$. De plus, $D_2\bar{g}(w,v) \in \mathcal{L}(F,\ \mathbb{R}^p)$, et pour tout $z \in F$, on a $D_2\bar{g}(w,v)z = Dg(v+w)z$.
Soit $(\bar{v}, \bar{w}) \in F \times G$ tel que $\bar{u} = \bar{v} + \bar{w}$. Alors $D_2\bar{g}(\bar{w}, \bar{v})z = Dg(\bar{u})z$ pour tout $z \in F$. L'application $D_2\bar{g}(\bar{w}, \bar{v})$ est une bijection de F sur $\mathbb{R}^p$, car, par définition de F, $Dg(\bar{u})$ est bijective de F sur $\mathbb{R}^p$.

5. **Théorème des fonctions implicites** Soient p et q des entiers naturels, soit $h \in C^1(\mathbb{R}^q \times \mathbb{R}^p, \mathbb{R}^p)$, et soient $(\bar{x}, \bar{y}) \in \mathbb{R}^q \times \mathbb{R}^p$ et $c \in \mathbb{R}^p$ tels que $h(\bar{x}, \bar{y}) = c$. On suppose que la matrice de la différentielle $D_2h(\bar{x}, \bar{y}) (\in \mathcal{M}_p(\mathbb{R}))$ est inversible. Alors il existe $\varepsilon > 0$ et $\nu > 0$ tels que pour tout $x \in B(\bar{x}, \varepsilon)$, il existe un unique $y \in B(\bar{y}, \nu)$ tel que $h(x,y) = c$. on peut ainsi définir une application ϕ de $B(\bar{x}, \varepsilon)$ dans $B(\bar{y}, \nu)$ par $\phi(x) = y$. On a $\phi(\bar{x}) = \bar{y}$, $\phi \in C^1(\mathbb{R}^p, \mathbb{R}^p)$ et $D\phi(x) = -[D_2h(x, \phi(x))]^{-1} \cdot D_1h(x, \phi(x))$.

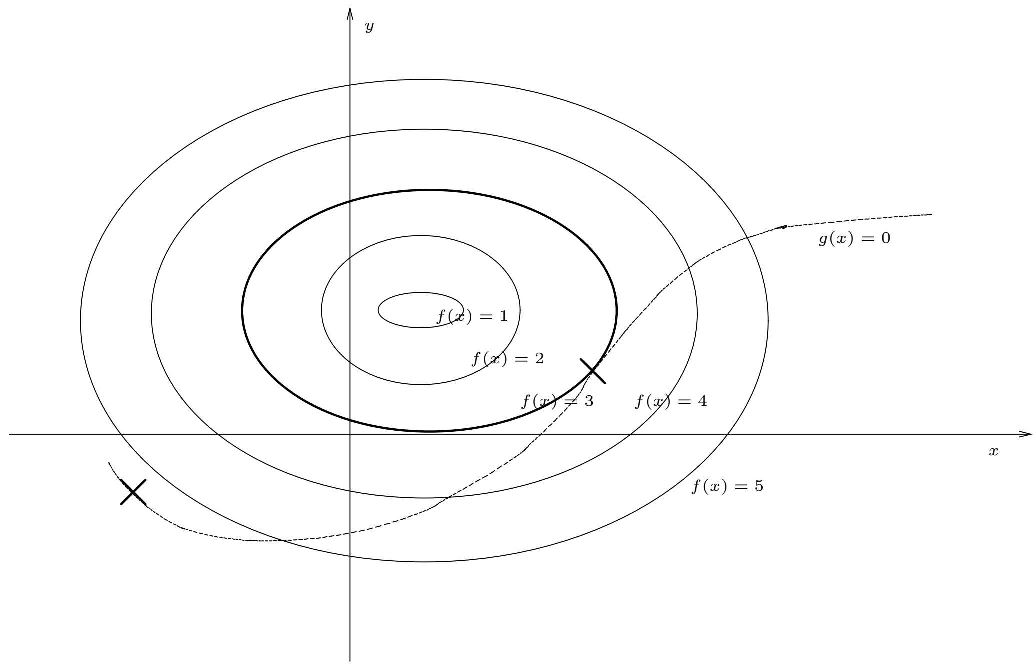

FIGURE 3.1: Interprétation géométrique des multiplicateurs de Lagrange

On rappelle que $K = \{u \in \mathbb{R}^n : g(u) = 0\}$ et on définit $\overline{K} = \{(w,v) \in G \times F,\ \bar{g}(w,v) = 0\}$. Par définition de $\bar{f}$ et de $\bar{g}$, on a

$$\begin{cases} (\bar{w},\bar{v}) \in \overline{K} \\ \bar{f}(\bar{w},\bar{v}) \le f(w,v)\ \ \forall (w,v) \in \overline{K} \end{cases} \tag{3.50}$$

D'autre part, le théorème des fonctions implicites (voir note de bas de page 262) entraîne l'existence de $\varepsilon > 0$ et $\nu > 0$ tels que pour tout $w \in B_G(\bar{w},\varepsilon)$ il existe un unique $v \in B_F(\bar{v},\nu)$ tel que $\bar{g}(w,v) = 0$. On note $v = \phi(w)$ et on définit ainsi une application $\phi \in C^1(B_G(\bar{w},\varepsilon), B_F(\bar{v},\nu))$.
On déduit alors de (3.50) que :

$$\bar{f}(\bar{w},\phi(\bar{w})) \le \bar{f}(w,\phi(w))),\ \forall w \in B_G(\bar{w},\varepsilon),$$

et donc

$$f(\bar{u}) = f(\bar{w} + \phi(\bar{w})) \le f(w + \phi(w)),\ \forall w \in B_G(\bar{w},\varepsilon).$$

En posant $\psi(w) = \bar{f}(w,\phi(w))$, on peut donc écrire

$$\psi(\bar{w}) = \bar{f}(\bar{w},\phi(\bar{w})) \le \psi(w),\ \forall w \in B_G(\bar{w},\varepsilon).$$

On a donc, grâce à la proposition 3.32,

$$D\psi(\bar{w}) = 0. \tag{3.51}$$

Par définition de ψ, de $\bar{f}$ et de $\bar{g}$, on a :

$$D\psi(\bar{w}) = D_1\bar{f}(\bar{w},\phi((\bar{w})) + D_2\bar{f}(\bar{w},\phi(\bar{w}))D\phi(\bar{w}).$$

D'après le théorème des fonctions implicites,

$$D\phi(\bar{w}) = -[D_2\bar{g}(\bar{w},\phi((\bar{w}))]^{-1}D_1\bar{g}(\bar{w},\phi((\bar{w})).$$

On déduit donc de (3.51) que

$$D_1\bar{f}(\bar{w},\phi((\bar{w}))w - [D_2\bar{g}(\bar{w},\phi((\bar{w}))]^{-1}D_1\bar{g}(\bar{w},\phi((\bar{w}))w = 0, \text{ pour tout } w \in G. \tag{3.52}$$

De plus, comme $D_2\bar{g}(\bar{w},\phi((\bar{w}))]^{-1}D_2\bar{g}(\bar{w},\phi((\bar{w})) = Id$, on a :

$$D_2\bar{f}(\bar{w},\phi((\bar{w}))z - D_2\bar{f}(\bar{w},\phi((\bar{w}))\,[D_2\bar{g}(\bar{w},\phi((\bar{w}))]^{-1}\,D_2\bar{g}(\bar{w},\phi((\bar{w}))z = 0,\ \forall z \in F. \tag{3.53}$$

Soit $x \in \mathbb{R}^n$, et $(z,w) \in F \times G$ tel que $x = z + w$. En additionant (3.52) et (3.53), et en notant

$$\Lambda = -D_2\bar{f}(\bar{w},\phi((\bar{w}))\,[D_2\bar{g}(\bar{w},\phi((\bar{w}))]^{-1}\,,$$

on obtient :

$$Df(\bar{u})x + \Lambda Dg(\bar{u})x = 0,$$

ce qui donne, en transposant : $\nabla f(\bar{u}) + \sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{u}) = 0$, avec $\Lambda = (\lambda_1, \ldots, \lambda_p)$.

■

Remarque 3.35 (Utilisation pratique du théorème de Lagrange)**.** Soit $f \in C^1(\mathbb{R}^n, \mathbb{R})$, $g = (g_1, \ldots, g_p)^t$ avec $g_i \in C(\mathbb{R}^n, \mathbb{R})$ pour $i = 1, \ldots, p.$, et soit $K = \{u \in \mathbb{R}^n, \quad g_i(u) = 0, \quad i = 1, \ldots, p\}$.
Le problème qu'on cherche à résoudre est le problème de minimisation (3.48) qu'on rappelle ici :

$$\begin{cases} \bar{u} \in K \\ f(\bar{u}) = \inf_K f \end{cases}$$

D'après le théorème des multiplicateurs de Lagrange, si $\bar{u}$ est solution de (3.48) et $Im(Dg(\bar{u})) = \mathbb{R}^p$, alors il existe $(\lambda_1, \ldots, \lambda_p) \in \mathbb{R}^p$ tel que $\bar{u}$ est solution du problème

$$\begin{cases} \dfrac{\partial f}{\partial x_j}(\bar{u}) + \displaystyle\sum_{i=1}^{p} \lambda_i \dfrac{\partial g_i}{\partial x_j} = 0, j = 1, \ldots n, \\ g_i(\bar{u}) = 0, \quad i = 1, \ldots, p. \end{cases} \tag{3.54}$$

Le système (3.54) est un système non linéaire de de $(n+p)$ équations et à $(n+p)$ inconnues $(\bar{x}, \ldots, \bar{x}_n, \lambda_i \ldots \lambda_p)$. Ce système sera résolu par une méthode de résolution de système non linéaire (Newton par exemple).

Remarque 3.36. *On vient de montrer que si $\bar{x}$ solution de* (3.48) *et $Im(Dg(\bar{x})) = \mathbb{R}^p$, alors $\bar{x}$ solution de* (3.54)*. Par contre, si $\bar{x}$ est solution de* (3.54)*, ceci n'entraîne pas que $\bar{x}$ est solution de* (3.48)*.*

Des exemples d'application du théorème des multiplicateurs de Lagrange sont donnés dans les exercices 126 page 266 et 127 page 266.

3.4.4 Contraintes inégalités

Soit $f \in C(\mathbb{R}^n, \mathbb{R})$ et $g_i \in C^1(\mathbb{R}^n, \mathbb{R}) \quad i = 1, \ldots, p$, on considère maintenant un ensemble K de la forme : $K = \{x \in \mathbb{R}^n, \; g_i(x) \le 0 \; \forall i = 1 \ldots p\}$, et on cherche à résoudre le problème de minimisation (3.48) qui sécrit :

$$\begin{cases} \bar{x} \in K \\ f(\bar{x}) \le f(x), \quad \forall x \in K. \end{cases}$$

Remarque 3.37. *Soit $\bar{x}$ une solution de* (3.48) *et supposons que $g_i(\bar{x}) < 0$, pour tout $i \in \{1, \ldots, p\}$. Il existe alors $\varepsilon > 0$ tel que si $x \in B(\bar{x}, \varepsilon)$ alors $g_i(x) < 0$ pour tout $i = 1, \ldots, p$.*
On a donc $f(\bar{x}) \le f(x) \quad \forall x \in B(\bar{x}, \varepsilon)$. On est alors ramené à un problème de minimisation sans contrainte, et si f est différentiable en $\bar{x}$, on a donc $\nabla f(\bar{x}) = 0$.

On donne maintenant sans démonstration le théorème de Kuhn-Tücker qui donne une caractérisation de la solution du problème (3.48).

Théorème 3.38 (Kuhn–Tucker)**.** *Soit $f \in C(\mathbb{R}^n, \mathbb{R})$, soit $g_i \in C^1(\mathbb{R}^n, \mathbb{R})$, pour $i = 1, \ldots, p$, et soit $K = \{\boldsymbol{x} \in \mathbb{R}^n, \; g_i(\boldsymbol{x}) \le 0 \; \forall i = 1 \ldots p\}$. On suppose qu'il existe $\bar{x}$ solution de* (3.48)*, et on pose $I(\bar{\boldsymbol{x}}) = \{i \in \{1, \ldots, p\}; |g_i(\bar{\boldsymbol{x}}) = 0\}$. On suppose que f est différentiable en $\bar{\boldsymbol{x}}$ et que la famille (de $\mathbb{R}^n$) $\{\nabla g_i(\bar{\boldsymbol{x}}), i \in I(\bar{\boldsymbol{x}})\}$ est libre. Alors il existe une famille $(\lambda_i)_{i \in I(\bar{\boldsymbol{x}})} \subset \mathbb{R}_+$ telle que*

$$\nabla f(\bar{\boldsymbol{x}}) + \sum_{i \in I(\bar{\boldsymbol{x}})} \lambda_i \nabla g_i(\bar{\boldsymbol{x}}) = 0.$$

Remarque 3.39.

1. Le théorème de Kuhn-Tucker s'applique pour des ensembles de contrainte de type inégalité. Si on a une contraite de type égalité, on peut évidemment se ramener à deux contraintes de type inégalité en remarquant que $\{h(x) = 0\} = \{h(x) \leq 0)\} \cap \{-h(x) \leq 0\}$. Cependant, si on pose $g_1 = h$ et $g_2 = -h$, on remarque que la famille $\{\nabla g_1(\bar{x}), \nabla g_2(\bar{x})\} = \{\nabla h(\bar{x}), -\nabla h(\bar{x})\}$ n'est pas libre. On ne peut donc pas appliquer le théorème de Kuhn-Tucker sous la forme donnée précédemment dans ce cas (mais on peut il existe des versions du théorème de Kuhn-Tucker permettant de traiter ce cas, voir Bonans-Saguez).
2. Dans la pratique, on a intérêt à écrire la conclusion du théorème de Kuhn-Tucker (i.e. l'existence de la famille $(\lambda_i)_{i \in I(\bar{x})}$) sous la forme du système de $n + p$ équations et $2p$ inéquations à résoudre suivant :

$$\left\{\begin{array}{l}
\nabla f(\bar{x}) + \displaystyle\sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x}) = 0, \\
\lambda_i g_i(\bar{x}) = 0, \quad \forall i = 1, \ldots, p, \\
g_i(\bar{x}) \leq 0, \quad \forall i = 1, \ldots, p, \\
\lambda_i \geq 0, \quad \forall i = 1, \ldots, p.
\end{array}\right.$$