Chapitre 5

Conditions d'optimalité en optimisation sous contraintes

5.1 Introduction

On s'intéresse maintenant à des problèmes d'optimisation de la forme

$$(\mathcal{PC}) \quad \begin{cases} \inf\limits_{x \in C} & f(x) \\ s.c. & g(x) \leqslant 0 \\ & h(x) = 0 \end{cases}$$

où $C \subsetneq \mathbb{R}^n$, les fonctions f, g et h sont différentiables au moins une fois, et f est typiquement non linéaire. On suppose qu'il existe p contraintes d'inégalité $(g_i)_{i \in [\![1,p]\!]}$ et q contraintes d'égalité $(h_j)_{j \in [\![1,q]\!]}$. Cependant nous étudierons le cas où g et h sont linéaires avec un intérêt tout particulier. Dans ce chapitre nous allons nous efforcer d'obtenir les conditions d'optimalité associées au problème sous contraintes. Les chapitres suivants mettront ensuite l'accent sur les méthodes numériques permettant de le résoudre. Nous nous intéresserons précisément dans ce chapitre aux problèmes sous contraintes d'égalité (Section 5.3) ou d'inégalité (Section 5.4).

5.2 Résultats d'existence et d'unicité

Dans le cas des problèmes d'optimisation sous contraintes, le théorème suivant est utile

Théorème 5.1 – Théorème d'existence (sous contraintes)

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction continue.

(1) Si C est un sous-ensemble compact (fermé borné) de $\mathbb{R}^n$, alors il existe $x^\star \in C$ tel que $f(x^\star) = \inf_{x \in C} f(x)$.

(2) Si C est un sous-ensemble fermé de $\mathbb{R}^n$, et si f est coercive (infinie à l'infini : $\lim_{\|x\| \to +\infty} f(x) = +\infty$), alors il existe $x^\star \in C$ tel que $f(x^\star) = \inf_{x \in C} f(x)$.

Démonstration 11

(1) Si C est un sous-ensemble compact de $\mathbb{R}^n$, comme f est continue, elle atteint ses bornes sur C, d'où l'existence de $x^\star$.

(2) Si f est coercive, alors il existe $R > 0$ tel que si $\|x\| > R$ alors $f(x) > f(0)$, donc $\inf_{x \in C} f(x) = \inf_{x \in C \bigcap \mathrm{BF}(0,R)} f(x)$. L'ensemble $C \bigcap \mathrm{BF}(0,R)$ est compact, comme intersection de deux compacts. Donc, par ce qui précède, il existe $x^\star \in C$ tel que $f(x^\star) = \inf_{x \in C \bigcap \mathrm{BF}(0,R)} f(x) = \inf_{x \in C} f(x)$.

Théorème 5.2 – Théorème d'unicité (sous contraintes)

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction continue. On suppose que f est strictement convexe et que C est convexe. Alors il existe au plus un élément $x^\star$ de C tel que $f(x^\star) = \inf_{x\in C} f(x)$.

Démonstration 12

Soit f strictement convexe, supposons qu'il existe $x^\star$, $\overline{x}$ dans C deux minimums de la fonction f sur $\mathbb{R}^n$ tels que

$$x^\star \neq \overline{x} \text{ et } f(x^\star) = f(\overline{x}) = \inf_{x\in\mathbb{R}^n} f(x)$$

Alors, si $\lambda \in]0,1[$, on a $\tilde{x} = \lambda.\overline{x} + (1-\lambda).x^\star \in C$ car C convexe, et comme f est strictement convexe

$$f(\tilde{x}) < \lambda f(\overline{x}) + (1-\lambda) f(x^\star) = f(x^\star) = \inf_{x\in\mathbb{R}^n} f(x)$$

Ceci fournit une contradiction, donc $x^\star = \overline{x}$.

5.3 Les conditions de Lagrange

On va tout d'abord s'intéresser au problème suivant, dit problème d'optimisation sous contraintes d'égalité seulement $(\mathcal{PCE})$

$$(\mathcal{PCE}) \quad \begin{cases} \inf\limits_{x\in\mathbb{R}^n} & f(x) \\ s.c. & h(x) = 0 \end{cases}$$

La raison par laquelle on s'intéresse en premier au problème sous contraintes d'égalité est qu'ils sont plus faciles à traiter. Nous allons dans un premier temps nous intéresser au cas où les contraintes sont linéaires.

5.3.1 Contraintes d'égalité linéaires

Un problème d'optimisation sous contraintes d'égalité linéaires est sous la forme

$$(\mathcal{PCE}) \quad \begin{cases} \inf\limits_{x\in C} & f(x) \\ s.c. & A\cdot x - b = 0 \end{cases}$$

où A est une matrice $q \times n$ avec $q \leqslant n$ et $b \in \mathbb{R}^q$. On notera $C = \{x \in \mathbb{R}^n,\ A\cdot x - b = 0\}$. Nous allons maintenant définir le concept de direction admissible dans C.

Définition 5.1 – Direction admissible (problème sous contraintes d'égalité linéaires)

On dit que $d \in \mathbb{R}^n$ est une *direction admissible en $x \in C$* s'il existe $\alpha > 0$ tel que

$$x + t.d \in C, \quad \forall t \in]0, \alpha]$$

On a $A\cdot(x + t.d) - b = t.A\cdot d$ puisque $x \in C$, et donc les directions admissibles d sont caractérisées par

$$A \cdot d = 0$$

Théorème 5.3 – Existence des multiplicateurs de Lagrange (problème sous contraintes d'égalité linéaires)

Soit $x^\star \in C = \{x \in \mathbb{R}^n,\ h(x) = 0\}$ un point régulier solution du problème sous contraintes d'égalité linéaires, vérifiant donc

$$f(x^\star) \leqslant f(x),\ \forall x \in C$$

Alors il existe nécessairement un vecteur des multiplicateurs de Lagrange $\lambda \in \mathbb{R}^q$ vérifiant

$$\nabla f(x^\star) + A^\top \cdot \lambda = 0$$

Si de plus A est de rang q alors λ est unique.

Démonstration 13 unicité du multiplicateur

Soit d une direction admissible, vérifiant donc $d \in \mathrm{Ker}(A)$. Pour tout $t \in \mathbb{R}$ on a

$$f(x^\star) \leqslant f(x^\star + t.d)$$

alors on en déduit que pour $\forall d \in \mathrm{Ker}(A)$,

$$\langle \nabla f(x^\star), d \rangle = \lim_{t \to 0} \frac{f(x^\star + t.d) - f(x^\star)}{t} = 0$$

Soit $\nabla f(x^\star) \in (\mathrm{Ker}(A))^\perp$, donc $\nabla f(x^\star) \in \mathrm{Im}(A^\top)$. Il existe donc un vecteur λ tel que

$$\nabla f(x^\star) = -A^\top \cdot \lambda$$

ce qui démontre le résultat. Pour l'unicité, supposons qu'il existe deux vecteurs λ_1 et λ_2 vérifiant

$$\nabla f(x^\star) = -A^\top \cdot \lambda_1 = -A^\top \cdot \lambda_2$$

On a donc

$$A^\top \cdot (\lambda_1 - \lambda_2) = 0$$

et donc $\lambda_1 - \lambda_2 = 0$ si A est de rang q.

Exemple 5.1 – Distance d'un point à un plan

On cherche à calculer la distance d'un point $x_0 \in \mathbb{R}^n$ au plan défini par l'équation $A \cdot x = b$, où $A \in \mathcal{M}_{q \times n}$ avec $\mathrm{rang}(A) = q$. Ce problème se pose sous la forme

$$(\mathcal{PCE}) \quad \begin{cases} \min\limits_{x \in \mathbb{R}^n} & \frac{1}{2}\|x_0 - x\|^2 \\ s.c. & A \cdot x = b \end{cases}$$

On pose donc $f(x) = \frac{1}{2}\|x_0 - x\|^2$. On a

$$\nabla f(x) = -(x_0 - x)$$

et donc le système d'optimalité est donné par

$$(x^\star - x_0) + A^\top \cdot \lambda^\star = 0 \quad (1)$$
$$A \cdot x^\star = b \quad (2)$$

En multipliant l'équation (1) par A on peut exprimer $\lambda^\star$ par

$$\lambda^\star = (A \cdot A^\top)^{-1} \cdot (A \cdot x_0 - b)$$

et on obtient en substituant $\lambda^\star$ dans l'équation (1)

$$x = (I - A^\top \cdot (A \cdot A^\top)^{-1} \cdot A) \cdot x_0 + A^\top \cdot (A \cdot A^\top)^{-1} \cdot b$$

Un problème voisin est celui de la projection d'une direction d sur le plan $A \cdot x = 0$. Le résultat précédent donne donc

$$d^\star = P \cdot d$$

avec $P = I - A^\top \cdot (A \cdot A^\top)^{-1} \cdot A$.

5.3.2 Contraintes d'égalité non linéaires

Nous étudions maintenant le problème sous la forme

$$(\mathcal{PCE}) \quad \begin{cases} \inf\limits_{x \in \mathbb{R}^n} & f(x) \\ s.c. & h(x) = 0 \end{cases}$$

où $h : \mathbb{R}^n \to \mathbb{R}^q$ est différentiable. On note comme précédemment $C = \{x \in \mathbb{R}^n,\ h(x) = 0\}$. Le concept de direction admissible dans C ne peut pas se définir comme pour les contraintes linéaires, car pour $x^\star \in C$ il peut ne pas exister $\alpha > 0$ et $d \in \mathbb{R}^n$ tels que $x^\star + t.d \in C$. On doit donc définir le concept de courbe admissible.

Considérons une courbe $x(t)$ définie pour $t \geqslant 0$ vérifiant

$$\begin{cases} x(t) \in C, \quad \forall t \in [-\alpha, \alpha],\ \alpha > 0 \\ x(0) = x^\star \end{cases}$$

Puisque $x(t) \in C$ on a $h_j(x(t)) = 0$ pour $1 \leqslant j \leqslant q$ et on peut écrire que

$$\frac{d}{dt} h_j(x(t)) = \langle \nabla h_j(x(t)), x'(t) \rangle = 0,\ 1 \leqslant j \leqslant q$$

Si on note $y = x'(0)$ le vecteur tangent à la courbe $x(t)$ en $t = 0$, on a donc

$$\langle \nabla h_j(x^\star), y \rangle = 0,\ 1 \leqslant j \leqslant q$$

Cela conduit à la définition suivante

Définition 5.2 – Direction admissible (problème sous contraintes d'égalité non linéaires)

On dit que $y \in \mathbb{R}^n$ est une direction admissible en $x^\star \in C$ s'il existe $\alpha > 0$ et une courbe $x(t)$ vérifiant

$$\begin{cases} x(t) \in C,\ \forall t \in [-\alpha, \alpha] \\ x(0) = x^\star \\ x'(0) = y \end{cases}$$

On notera alors $y \in T(x^\star)$.

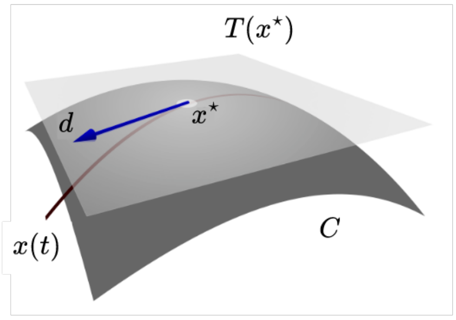

FIGURE 5.1 – Le plan $T(x^\star)$ est tangent à la surface C en $x^\star$, selon la courbe admissible $x(t)$, passant par $x(0) = x^\star$ où le vecteur d est tangent à la courbe.

L'ensemble $T(x^\star)$ définit le plan tangent (voir Figure 5.1) à C en $x^\star$. L'analyse faite précédemment montre que l'on a l'implication.

$$y \in T(x^\star) \Rightarrow \langle \nabla h_j(x^\star), y \rangle = 0,\ 1 \leqslant j \leqslant q$$

qui sera insuffisante pour montrer la condition nécéssaire d'optimalité. Nous allons donc maintenant nous attacher à montrer sous quelles conditions la relation $\langle \nabla h_j(x^\star), y \rangle = 0$ est une condition suffisante d'appartenance à $T(x^\star)$.

Définition 5.3 – Point régulier (problème sous contraintes d'égalité non linéaires)

On dit que $x^\star$ est un point régulier pour la contrainte $h(x) = 0$ si

(1) $h(x^\star) = 0$;

(2) les vecteurs $\nabla h_j(x^\star)$ sont linéairement indépendants.

Si on note $J_h(x^\star) \in \mathcal{M}_{n,q}$ la jacobienne de h, on a $J_h(x^\star) = \begin{bmatrix} \nabla h_1(x^\star) & \cdots & \nabla h_q(x^\star) \end{bmatrix}$, la condition d'indépendance linéaire des $\nabla h_j(x^\star)$ peut s'écrire

$$\mathrm{rang}(J_h(x^\star)) = q$$

et on a donc $J_h(x^\star)^\top \cdot x'(0) = 0$ pour toute courbe admissible $x(t)$.

Proposition 5.1

Si $x^\star$ est un point régulier pour la contrainte $h(x) = 0$, alors

$$J_h(x^\star)^\top \cdot y = 0 \Rightarrow y \in T(x^\star)$$

Démonstration 14

Soit $y \in \mathbb{R}^n$ vérifiant $\nabla h(x^\star)^\top \cdot y = 0$. On considère la courbe $x(t)$ donnée par

$$x(t) = x^\star + t.y + \nabla h(x^\star) \cdot u(t)$$

La fonction $u(t) \in \mathbb{R}^q$, pour l'instant inconnue, va être déterminée de telle façon que $h(x(t)) = 0$. On va pour cela poser le problème de la détermination de $u(t)$ sous la forme d'une équation implicite. On définit la fonction $F : \mathbb{R} \times \mathbb{R}^q \to \mathbb{R}^q$ par

$$F(t, u) = h\Big(x^\star + t.y + \nabla h(x^\star) \cdot u(t)\Big)$$

Le problème de la détermination de $u(t)$ se ramène donc à la résolution de l'équation

$$F(t, u) = 0$$

au voisinage du point $(0, 0)$. On a d'une part $F(0, 0) = h(x^\star) = 0$ et

$$\frac{\partial}{\partial u} F(t, u) = \nabla h(x^\star)^\top \cdot \nabla h\Big(x^\star + t.y + \nabla h(x^\star) \cdot u(t)\Big)$$

soit

$$\frac{\partial}{\partial u} F(0, 0) = \nabla h(x^\star)^\top \cdot \nabla h(x^\star)$$

La matrice $\frac{\partial}{\partial u} F(0, 0)$ est inversible puisque par hypothèse $\nabla h(x^\star)$ est de rang p. On peut alors appliquer le théorème des fonctions implicites : il existe un voisinage du point $(0, 0)$ et une fonction $u(t)$ tels que

$$F(t, u) = 0 \iff u = u(t)$$

Notons que l'on a donc nécéssairement $u(0) = 0$ puisque $F(0, 0) = 0$.
On a donc maintenant

$$x'(t) = y + \nabla h(x^\star) \cdot u'(t)$$

soit en $t = 0$ soit

$$x'(0) = y + \nabla h(x^\star) \cdot u'(0)$$

Montrons que $u'(0) = 0$. Pour cela on écrit que l'on a

$$\frac{d}{dt} h(x(t)) = \nabla h(x(t))^\top \cdot \Big(y + \nabla h(x^\star) \cdot u'(t)\Big) = 0$$

puisque $h(x(t)) = 0$, et donc en $t = 0$ la relation précédente prend la forme

$$\frac{d}{dt} h(x(t))\Big|_{t=0} = \nabla h(x^\star)^\top \cdot y + \nabla h(x^\star)^\top \cdot \nabla h(x^\star) \cdot u'(0) = 0$$

Le premier terme du second membre est nul par hypothèse, et donc $u'(0) = 0$ puisque $\nabla h(x^\star)^\top \cdot \nabla h(x^\star)$ est inversible. Donc $x'(0) = y$, soit $y \in T(x^\star)$, ce qui démontre le résultat annoncé.

Théorème 5.4 – Théorème de Lagrange (problème sous contraintes d'égalité non linéaires)

Soit $x^\star \in C = \{x \in \mathbb{R}^n,\ h(x) = 0\}$ un point régulier solution du problème sous contraintes d'égalité non linéaires, vérifiant donc

$$f(x^\star) \leqslant f(x),\ \forall x \in C$$

Alors il existe nécessairement un vecteur $\lambda \in \mathbb{R}^q$ unique vérifiant

$$\nabla f(x^\star) + J_h(x^\star) \cdot \lambda = 0$$

soit encore

$$\nabla f(x^\star) + \sum_{j=1}^{q} \lambda_j . \nabla h_j(x^\star) = 0$$

Les composantes du vecteur λ sont appelées multiplicateurs de Lagrange.

Démonstration 15

Considérons une courbe $x(t)$ définie pour $t \in [-\alpha, \alpha]$ vérifiant

$$\begin{cases} x(t) \in C, & \forall t \in [-\alpha, \alpha],\ \alpha > 0 \\ x(0) = x^\star \end{cases}$$

On a $f(x(0)) \leqslant f(x(t))$ pour tout $t \in [-\alpha, \alpha]$, donc nécessairement

$$\frac{d}{dt} f(x(t))\Big|_{t=0} = \langle \nabla f(x^\star), x'(0) \rangle = 0$$

ce qui signifie que $\nabla f(x^\star)$ se trouve dans l'orthogonal du plan tangent $T(x^\star)$ à C en $x^\star$. Si l'on utilise l'équivalence

$$T(x^\star) = \left(\mathrm{Ker}(\nabla h(x^\star))\right)^\top \iff T(x^\star)^\perp = \mathrm{Im}(\nabla h(x^\star))$$

il existe donc un vecteur $\lambda \in \mathbb{R}^q$ tel que

$$\nabla f(x^\star) = -\nabla h(x^\star) \cdot \lambda$$

L'unicité résulte du fait que $\nabla h(x^\star)$ est de rang q et se montre comme dans le cas linéaire.

5.4 Les conditions de Kuhn et Tucker

5.4.1 Contraintes d'inégalité

On s'intéresse maintenant au problème suivant, dit problème d'optimisation avec contraintes d'inégalité seulement

$$(\mathcal{PCI}) \quad \begin{cases} \inf\limits_{x \in C} & f(x) \\ s.c. & g(x) \leqslant 0 \end{cases}$$

où $g : \mathbb{R}^n \to \mathbb{R}^p$, est différentiable (il n'y a ici aucune condition sur p). On notera C l'ensemble des points admissibles, c'est-à-dire $C = \{x \in \mathbb{R}^n,\ g_i(x) \leqslant 0,\ i \in [\![1, p]\!]\}$.

Au point solution de ce problème d'optimisation avec contraintes d'inégalité il va de soi que les contraintes effectivement actives vérifieront $g_i(x^\star) = 0$. Cependant, puisque l'on ne sait pas a priori quelles sont ces contraintes, le passage de problème d'optimisation avec contraintes d'inégalité a un problème du type problème d'optimisation avec contraintes d'égalité n'est pas direct.

Définition 5.4 – Contrainte active(saturée)/*Active constraint*/积极约束

Soit une contrainte d'inégalité $g_i(x) \leqslant 0$ et $x^\star$ un point de $\mathbb{R}^n$. Si $x^\star$ satisfait $g_i(x^\star) < 0$, on dit que la contrainte est inactive en $x^\star$, tandis que si $x^\star$ satisfait $g_i(x^\star) = 0$, on dit que la contrainte est *active* ou *saturée en* $x^\star$. [a]
Notons $A(x^\star) = \{i \in [\![1, p]\!] \mid g_i(x^\star) = 0\} \subseteq \{1, \ldots, p\}$ l'ensemble des indices des contraintes actives en $x^\star$.

[a] Voir Figure 5.2

Définition 5.5 – Direction admissible (problème sous contraintes d'inégalité)

On dit que $y \in \mathbb{R}^n$ est une direction admissible en $x^\star \in C$ s'il existe $\alpha > 0$ et une courbe $x(t)$ vérifiant

$$\begin{cases} x(t) \in C, & \forall t \in [-\alpha, \alpha] \\ x(0) = x^\star \\ x'(0) = y \end{cases}$$

On notera alors $y \in C(x^\star)$.

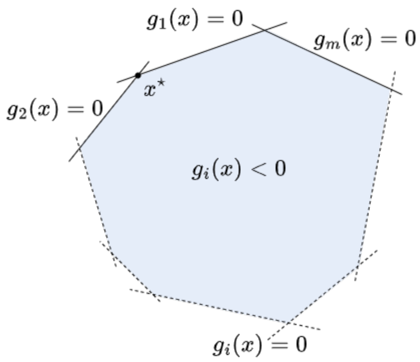

FIGURE 5.2 – g_1 et g_2 sont actives en $x^\star$ tandis que g_i ($i \neq 1$, $i \neq 2$) sont inactives en $x^\star$.

Proposition 5.2

Soit $y \in \mathbb{R}^n$ une direction admissible en $x^\star \in C$, alors on a nécessairement

$$\langle \nabla g_i(x^\star), y \rangle \leqslant 0,\ i \in A(x^\star)$$

Démonstration 16

Considérons une courbe $x(t)$ définie pour $t \in [-\alpha, \alpha]$ vérifiant

$$\begin{cases} x(t) \in C, \quad \forall t \in [-\alpha, \alpha],\ \alpha > 0 \\ x(0) = x^\star \\ x'(0) = y \end{cases}$$

Comme $g_i(x^\star) < 0$ pour $i \notin A(x^\star)$, on aura toujours $g_i(x(t)) < 0$ pour t suffisamment petit. Par contre, pour $i \in A(x^\star)$ on doit avoir $g_i(x(t)) \leqslant 0$ pour t suffisamment petit. Si on utilise le développement de Taylor de $g_i(x(t))$ en $t = 0$ on doit donc avoir

$$g_i(x^\star) + t\langle \nabla g_i(x^\star), y \rangle + t\varepsilon(t) \leqslant 0$$

Puisque $g_i(x^\star) = 0$ il faut donc nécessairement que l'on ait

$$\langle \nabla g_i(x^\star), y \rangle \leqslant 0$$

Définition 5.6 – Cône des directions admissibles

On dit qu'un arc paramétré $x : [0, \varepsilon] \to \mathbb{R}^n$ de classe $\mathcal{C}^n$ est une courbe admissible si $x(0) = x^\star$ et $x(t) \in C$, pour $t > 0$ au voisinage de 0.
On appelle direction admissible en $x^\star$ les vecteurs tangents à des courbes admissibles au point $x^\star$.
On note $C_{ad}(x^\star)$ le cône des directions admissibles en $x^\star$.

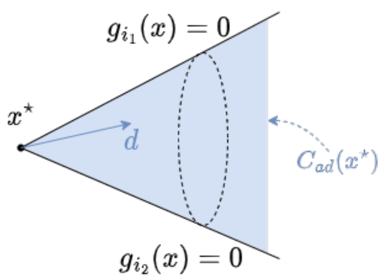

FIGURE 5.3 – Cône admissible : toutes les directions admissibles d en $x^\star$ forment le cône admissible en $x^\star$.

Comme dans le cas des contraintes d'égalité, on doit définir la notion de point régulier, qui est nécessaire pour que la condition précédente soit suffisante.

Définition 5.7 – Point régulier (problème sous contraintes d'inégalité)

On dit que $x^\star$ est un point régulier pour la contrainte $g(x) \leqslant 0$ si
(1) $g(x^\star) \leqslant 0$;
(2) les vecteurs $\nabla g_i(x^\star)$ sont linéairement indépendants.

Sous l'hypothèse de régularité de $x^\star$ on aura, comme dans le cas des contraintes d'égalité

$$\langle \nabla g_i(x^\star), y\rangle \leqslant 0,\ i \in A(x^\star) \ \Rightarrow\ y \in C(x^\star)$$

La proposition suivante permet d'effectuer le premier pas vers les conditions d'optimalité de Kuhn et Tucker.

Proposition 5.3

Soit $x^\star$ la solution du problème sous contraintes d'inégalité. Il existe $\eta > 0$ tel que

$$\forall x \in BO(x^\star, \eta),\ g_i(x) < 0,\ i \notin A(x^\star)$$

Alors $x^\star$ est la solution du problème

$$(\mathcal{PCI}) \quad \begin{cases} \inf\limits_{x \in BO(x^\star,\eta)} & f(x) \\ s.c. & g_i(x) = 0,\ i \in A(x^\star) \end{cases}$$

Ce résultat est uniquement dû à la continuité de g, et montre que l'on est localement ramené à un problème avec contraintes d'égalité.

5.4.2 Conditions de Kuhn et Tucker

Théorème 5.5 – Conditions de Kuhn et Tucker

Soit $x^\star \in C$ un point régulier solution du problème sous contraintes d'inégalité. Alors il existe un unique vecteur $\mu \in \mathbb{R}^p$ tel que

$$\nabla f(x^\star) + \sum_{i=1}^{p} \mu_i.\nabla g_i(x^\star) = 0$$
$$\mu_i \geqslant 0,\ i = 1, \ldots, p$$
$$\mu_i g_i(x^\star) = 0,\ i = 1, \ldots, p$$

Démonstration 17

La relation $\nabla f(x^\star)+\sum_{i=1}^p \mu_i.\nabla g_i(x^\star) = 0$, $\mu_i g_i(x^\star) = 0$ sont une conséquence directe du théorème de Lagrange, car il suffit de prendre $\mu_i = 0$ pour $i \notin A(x^\star)$. On peut ensuite montrer que $\mu_i \geqslant 0$ par l'absurde : supposons qu'il existe $k \in A(x^\star)$ tel que $\mu_k < 0$. On définit l'ensemble

$$S_k = \{x \in \mathbb{R}^n,\ g_i(x) = 0,\ i \in A(x^\star),\ i \neq k\}$$

On définit $y \in \mathbb{R}^n$ tel que

$$\langle \nabla g_i(x^\star), y\rangle = 0,\ i \in A(x^\star),\ i \neq k$$
$$\langle \nabla g_k(x^\star), y\rangle = -1$$

Alors y est une direction admissible en $x^\star$ puisque

$$\langle \nabla g_i(x^\star), y\rangle \leqslant 0,\ i \in A(x^\star)$$

et que $x^\star$ est un point régulier. Il existe donc une courbe $x(t) \in S_k$ et vérifiant de plus $x(t) \in C$, pour $t \in [-\alpha, \alpha]$, telle que $x'(0) = y$. On a donc

$$\begin{aligned}\frac{d}{dt}f(x(t))\Big|_{t=0} &= \langle \nabla f(x^\star), y\rangle \\ &= -\sum_{i\in A(x^\star)} \mu_i \langle \nabla g_i(x^\star), y\rangle \\ &= -\mu_k \langle \nabla g_k(x^\star), y\rangle = \mu_k < 0\end{aligned}$$

ce qui est impossible car f est minimum en $x^\star$.

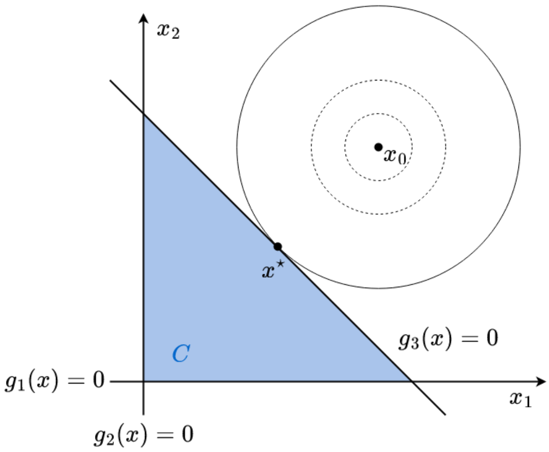

FIGURE 5.4 – Exemple de programme quadratique.

Exemple 5.2 – Interprétation géométrique des conditions de Kuhn et Tucker

On cherche à résoudre le problème

$$(\mathcal{PCI}) \quad \begin{cases} \min\limits_{x\in\mathbb{R}^2} & \frac{1}{2}\|x_0 - x\|^2 \\ s.c. & x_1 \geqslant 0 \\ & x_2 \geqslant 0 \\ & x_1 + x_2 \leqslant 1 \end{cases}$$

où $x_0 = (1, \frac{1}{2})$. Il s'agit d'un problème avec contraintes d'inégalité se mettant sous la forme $g(x) \leqslant 0$ avec

$$g(x) = (-x_1, -x_2, x_1 + x_2 - 1)$$

Sur le dessin 5.4, on peut s'assurer que très probablement seule la contrainte numéro 3 est active. On peut s'en persuader par le calcul de la façon suivante : on peut tenter de résoudre le système

$$\begin{aligned}\nabla f(x) + \mu_3.\nabla g_3(x) &= 0 \\ g_3(x) &= 0\end{aligned}$$

soit ici

$$\begin{aligned}x - x_0 + \mu_3\,(1, 1) &= 0 \\ x_1 + x_2 &= 1\end{aligned}$$

ou bien encore

$$\begin{aligned}x_1 + \mu_3 &= 1 \\ x_2 + \mu_3 &= \frac{1}{2} \\ x_1 + x_2 &= 1\end{aligned}$$

dont la solution est donnée par

$$x_1 = \frac{3}{4},\ x_2 = \frac{1}{4},\ \mu_3 = \frac{1}{4}$$

On a bien $\mu_3 \geqslant 0$ ce qui justifie *a posteriori* le choix de saturer la contrainte numéro 3.

5.5 Problème avec contraintes d'égalité et d'inégalité

On veut minimiser f sur l'ensemble C des contraintes données par les inégalités et les égalités

$$C = \{x \in \mathbb{R}^n,\ g_i(x) \leqslant 0,\ i \in [\![1, p]\!],\ h_j(x) = 0,\ j \in [\![1, q]\!]\}$$

où les g_i et h_j sont des fonctions de classe $\mathcal{C}^1$ de $\mathbb{R}^n$ dans $\mathbb{R}$. Donc, les problèmes d'optimisation peuvent s'écrire sous la forme

$$(\mathcal{PC}) \qquad \inf\{f(x),\ x \in C\}$$

5.5.1 Qualifications des contraintes

La compréhension des qualifications des contraintes est essentielle pour étudier les conditions de Karush-Kuhn-Tucker (KKT) dans le contexte des problèmes d'optimisation avec contraintes.

La qualification des contraintes est une condition qui garantit que les contraintes du problème ne sont pas «trop restrictives» ou «trop lâches», ce qui permet aux conditions KKT de s'appliquer. En d'autres termes, si les contraintes ne sont pas qualifiées, les conditions KKT ne peuvent pas garantir qu'une solution est optimale pour le problème d'optimisation avec contraintes.

Définition 5.8 – Qualification des contraintes/*Constraint qualification*/约束规范

Les contraintes sont dites qualifiées en $\overline{x}$ si les deux conditions suivantes sont satisfaites

(a) les vecteurs $\nabla h_1(\overline{x}), \ldots, \nabla h_q(\overline{x})$ sont linéairement indépendants ;

(b) il existe une direction $d \in \mathbb{R}^n$ telle que l'on ait, pour tout $i \in A(\overline{x})$ et $j \in [\![1, q]\!]$,

$$\langle \nabla g_i(\overline{x}), d\rangle \leqslant 0 \text{ et } \langle \nabla h_j(\overline{x}), d\rangle = 0$$

Remarque 5.1

On définit une ensemble

$$L(\overline{x}) = \left\{d \in \mathbb{R}^n,\ \langle \nabla g_i(\overline{x}), d\rangle \leqslant 0,\ \forall i \in A(\overline{x}),\ \langle \nabla h_j(\overline{x}), d\rangle = 0,\ \forall j \in [\![1, q]\!]\right\}$$

Alors, l'ensemble $L(\overline{x})$ est un cône polyédrique (convexe fermé non-vide). C'est en effet une approximation linéaire de l'ensemble C au point $\overline{x}$.

Propriété 5.1

Si $\overline{Z(\overline{x})} = L(\overline{x})$, alors les contraintes sont qualifiées.

Proposition 5.4

Pour tout $x^\star \in C$, *on a* $\overline{Z(x^\star)} \subset L(x^\star)$.

Démonstration 18

Soit $x^\star \in C$, alors pour tout $d \in Z(x^\star)$, il existe $\lambda^* > 0$, on a

$$\forall \lambda \in]0, \lambda^*],\ x^\star + \lambda.d \in C$$

C'est-à-dire,

$$\forall \lambda \in]0, \lambda^*],\ g(x^\star + \lambda.d) \leqslant 0,\ h(x^\star + \lambda.d) = 0$$

où

$$g_i(x^\star + \lambda.d) \leqslant g_i(x^\star) = 0, \forall i \in A(x^\star)$$
$$h_j(x^\star + \lambda.d) = 0, \forall j \in [\![1, q]\!]$$

Les développements limités à l'ordre 1 de g_i et h_j donnent

$$0 \geqslant g_i(x^\star + \lambda.d) - g_i(x^\star) = \underbrace{\lambda}_{>0} \langle \nabla g_i(x^\star), d\rangle + o(\|\lambda.d\|)$$
$$0 = h_j(x^\star + \lambda.d) = \underbrace{h_j(x^\star)}_{=0} + \underbrace{\lambda}_{>0} \langle \nabla h_j(x^\star), d\rangle + o(\|\lambda.d\|)$$

On a donc

$$\langle \nabla g_i(x^\star), d\rangle \leqslant 0, \ \forall i \in A(x^\star)$$
$$\langle \nabla h_j(x^\star), d\rangle = 0, \ \forall j \in [\![1, q]\!]$$

On obtient que

$$\forall x^\star \in C, \ Z(x^\star) \subset L(x^\star)$$

Comme l'ensemble $L(x^\star)$ est fermé et $Z(x^\star) \subset L(x^\star)$, donc

$$\forall x^\star \in C, \ \overline{Z(x^\star)} \subset L(x^\star)$$

Remarque 5.2

L'inclusion inverse de Proposition 5.4 n'est pas toujours vraie. Voici un contre-exemple.

Exemple 5.3 – Contre-exemple

Soit $D = \{x \in \mathbb{R}, \ -x^2 \geqslant 0\}$. Montrer que 0 n'est pas un point régulier de D.
Réponse : L'ensemble D est le singleton $\{0\}$, donc

$$Z(0) = \overline{Z(0)} = D = \{0\}$$

Mais $L(0) = \{d \in \mathbb{R}, \ -d \times 2 \times 0 \geqslant 0\} = \mathbb{R}$. On a donc l'inclusion stricte $\overline{Z(0)} \subset L(0)$. Donc 0 n'est pas un point régulier de D.

1. **Affinité (QC-A).** Si g et h sont des applications linéaires (ou affines), on a $\overline{Z(\overline{x})} = L(\overline{x})$.

Démonstration 19

D'après la Proposition 5.4, il suffit de montrer que $L(\overline{x}) \subset \overline{Z(\overline{x})}$.

Soit $d \in L(\overline{x})$, on a

$$\langle \nabla g_i(\overline{x}), d\rangle \leqslant 0, \ \forall i \in A(\overline{x})$$
$$\langle \nabla h_j(\overline{x}), d\rangle = 0, \ \forall j \in [\![1, q]\!]$$

Par affinité de h et de g, on a pour tout $\alpha > 0$,

$$h_j(\overline{x} + \alpha.d) = \underbrace{h_j(\overline{x})}_{=0} + \alpha \underbrace{\langle \nabla h_j(\overline{x}), d\rangle}_{=0} = 0$$
$$\forall i \in A(\overline{x}), \ g_i(\overline{x} + \alpha.d) = \underbrace{g_i(\overline{x})}_{=0} + \alpha \underbrace{\langle \nabla g_i(\overline{x}), d\rangle}_{\leqslant 0} \leqslant 0$$

Pour tout $i \in [\![1, p]\!] \setminus A(\overline{x})$, comme $g_i(\overline{x}) < 0$ et par la continuité de g_i, on obtient alors

$$\forall i \in [\![1, p]\!] \setminus A(\overline{x}), \ \exists \alpha^* > 0, \ \forall \alpha \in]0, \alpha^*], \ g_i(\overline{x} + \alpha.d) \leqslant 0$$

Donc

$$\exists \alpha^* > 0, \ \forall \alpha \in]0, \alpha^*], \ \overline{x} + \alpha.d \in C$$

Ceci implique

$$d \in \overline{Z(\overline{x})}$$

Donc

$$L(\overline{x}) \subset \overline{Z(\overline{x})}$$

2. **Slater (QC-S).** Si $(\mathcal{P})$ est un problème d'optimisation convexe (f convexe, g convexe, h affine), et qu'il existe au moins un point absolument réalisable.
   Autrement dit,
$$\exists \overline{x} \in C \text{ tel que } \exists i \in \llbracket 1,p \rrbracket, \, g_i(\overline{x}) < 0, \, \forall j \in \llbracket 1,q \rrbracket, \, h_j(\overline{x}) = 0$$

Démonstration 20

Par analogue comme dans la démonstration de (QC-A), on va montrer que $L(\overline{x}) \subset \overline{Z(\overline{x})}$.
Soit $d \in L(\overline{x})$, on a
$$\langle \nabla g_i(\overline{x}), d \rangle \leqslant 0, \, \forall i \in A(\overline{x})$$
$$\langle \nabla h_j(\overline{x}), d \rangle = 0, \, \forall j \in \llbracket 1,q \rrbracket$$
Par affinité de h_j, on a pour tout $\alpha > 0$,
$$h_j(\overline{x} + \alpha.d) = \underbrace{h_j(\overline{x})}_{=0} + \alpha \underbrace{\langle \nabla h_j(\overline{x}), d \rangle}_{=0} = 0$$
Pour tout $i \in A(\overline{x})$, par convexité de g_i, $g_i(x^\star) < 0$ et $g_i(\overline{x}) = 0$, on a
$$0 > g_i(x^\star) \geqslant \underbrace{g_i(\overline{x})}_{=0} + \langle \nabla g_i(\overline{x}), x^\star - \overline{x} \rangle = \langle \nabla g_i(\overline{x}), x^\star - \overline{x} \rangle$$
Pour tout $\lambda \in]0,1[$, notons $\hat{d}(\lambda) = (1-\lambda).d + \lambda.(x^\star - \overline{x})$, comme
$$\langle \nabla g_i(\overline{x}), d \rangle \leqslant 0, \text{ et} \langle \nabla g_i(\overline{x}), x^\star - \overline{x} \rangle < 0$$
donc
$$\langle \nabla g_i(\overline{x}), \hat{d}(\lambda) \rangle = \underbrace{(1-\lambda)}_{>0} \underbrace{\langle \nabla g_i(\overline{x}), d \rangle}_{\leqslant 0} + \underbrace{\lambda}_{>0} \underbrace{\langle \nabla g_i(\overline{x}), x^\star - \overline{x} \rangle}_{<0} < 0, \forall \lambda \in]0,1[$$
On obtient donc
$$\forall i \in A(\overline{x}), \, \forall \lambda \in]0,1[, \, g_i(\overline{x} + \lambda.\hat{d}(\lambda)) = \underbrace{g_i(\overline{x})}_{=0} + \underbrace{\alpha}_{>0} \underbrace{\langle \nabla g_i(\overline{x}), \hat{d}(\lambda) \rangle}_{<0} + o(\|\alpha.\hat{d}(\lambda)\|)$$
Alors, pour $\alpha \to 0^+$, on a
$$\forall \lambda \in]0,1[, \, \lim_{\alpha \to 0^+} \frac{g_i(\overline{x} + \alpha.\hat{d}(\lambda))}{\|\alpha.\hat{d}(\lambda)\|} = \frac{\langle \nabla g_i(\overline{x}), \hat{d}(\lambda) \rangle}{\|\alpha.\hat{d}(\lambda)\|} < 0$$
Donc
$$\forall \lambda \in]0,1[, \, \exists \alpha^* > 0, \, \forall \alpha \in]0, \alpha^*], \, g_i(\overline{x} + \alpha.\hat{d}(\lambda)) < 0$$
Par continuité de $\hat{d}$ et de g_i, il suffit de prendre $\lambda \to 0^+$ pour déduire que
$$\forall i \in A(\overline{x}), \, \exists \alpha^* > 0, \, \forall \alpha \in]0, \alpha^*], \, g_i(\overline{x} + \alpha.d) \leqslant 0$$
Pour tout $i \in \llbracket 1,p \rrbracket \backslash A(\overline{x})$, comme $g_i(\overline{x}) < 0$ et par continuité de g_i, on obtient alors
$$\forall i \in \llbracket 1,p \rrbracket \backslash A(\overline{x}), \, \exists \alpha^* > 0, \, \forall \alpha \in]0, \alpha^*], \, g_i(\overline{x} + \alpha.d) \leqslant 0$$
Donc
$$\exists \alpha^* > 0, \, \forall \alpha \in]0, \alpha^*], \, \overline{x} + \alpha.d \in C$$
Ceci implique
$$d \in \overline{Z(\overline{x})}$$
Donc
$$L(\overline{x}) \subset \overline{Z(\overline{x})}$$

3. **Indépendance linéaire (QC-IL).** Si les contraintes actives (c'est-à-dire les contraintes qui sont satisfaites avec égalité) sont linéairement indépendantes en tout point réalisable.
$$\sum_{i=1}^{p} \lambda_i \nabla g_i(x^\star) + \sum_{j=1}^{q} \mu_j \nabla h_j(x^\star) = 0$$
   où λ_i et μ_j sont des multiplicateurs de Lagrange associés aux contraintes actives d'inégalité et d'égalité respectivement, admet une unique solution $(\lambda, \mu) = (0,0)$.

4. **Mangasarian-Fromovitz (QC-MF).** S'il existe un point $x^\star \in \mathbb{R}^n$ tel que les conditions suivantes sont satisfaites
   — Les contraintes sont actives, c'est-à-dire $g_i(x^\star) = 0$ pour tout i tel que $g_i(x^\star) \leqslant 0$.

— Le gradient des contraintes actives est linéairement indépendant, c'est-à-dire que les vecteurs $\nabla g_i(x^\star)$ pour les i tels que $g_i(x^\star) = 0$ sont linéairement indépendants.

Alors, il existe un nombre $\epsilon > 0$ tel que, pour tout vecteur $d \in \mathbb{R}^n$ vérifiant $\|d\| < \epsilon$, toutes les contraintes actives restent satisfaites, c'est-à-dire que $g_i(x^\star + t.d) \leqslant 0$ pour tout t suffisamment petit.

5.5.2 Conditions de Karush, Kuhn et Tucker

Définition 5.9 – Lagrangien du problème $(\mathcal{P})$

On appelle Lagrangien du problème $(\mathcal{P})$ la fonction définie sur $\mathbb{R}^n \times \mathbb{R}_+^p \times \mathbb{R}^q$ par

$$\mathcal{L}(x, \mu, \lambda) = f(x) + \sum_{i=1}^{p} \mu_i g_i(x) + \sum_{j=1}^{q} \lambda_j h_j(x)$$

Les conditions de Karush-Kuhn-Tucker (ou Kuhn-Tucker) sont une généralisation des multiplicateurs de Lagrange qui permettent de résoudre des problèmes d'optimisation sous contraintes non linéaires.

Théorème 5.6 – Conditions de KKT (Karush-Kuhn-Tucker)

On introduit l'ensemble des contraintes

$$C = \{x \in \mathbb{R}^n,\ g_i(x) \leqslant 0,\ \forall i \in [\![1, p]\!]\,,\ h_j(x) = 0,\ \forall j \in [\![1, q]\!]\}$$

On suppose les contraintes qualifiées en $x^\star$. Alors, une condition nécessaire pour que $x^\star$ soit un minimum de f sur C est que

— ***les conditions nécessaires d'optimalité d'ordre 1***
Soit f, g_i, $i \in [\![1, p]\!]$, h_j, $j \in [\![1, q]\!]$, des fonctions différentiables en $x^\star$.

(a) (Réalisabilité primale)

$$x^\star \in C$$

(b) (Réalisabilité duale et stationnarité)

$$\exists \mu \in (\mathbb{R}_+)^p,\ \exists \lambda \in \mathbb{R}^q,\ \nabla_x \mathcal{L}(x^\star, \mu, \lambda) = \nabla f(x^\star) + \sum_{i=1}^{p} \mu_i.\nabla g_i(x^\star) + \sum_{j=1}^{q} \lambda_j.\nabla h_j(x^\star) = 0$$

où $\nabla_x \mathcal{L}(x^\star, \mu, \lambda)$ désigne le gradient de la fonctions $x \mapsto \mathcal{L}(x, \mu, \lambda)$ au point $x^\star$.

(c) (Complémentarité)

$$\mu_i g_i(x^\star) = 0,\ i \in [\![1, p]\!]$$

— ***les conditions nécessaires d'optimalité d'ordre 2***
Si de plus f, g_i, $i \in [\![1, p]\!]$, h_j, $j \in [\![1, q]\!]$ sont des fonctions deux fois différentiables en $x^\star$, alors $\exists \mu \in (\mathbb{R}_+)^p$, $\exists \lambda \in \mathbb{R}^q$,

$$\forall d \in \mathcal{T}(x^\star),\ \langle H_{\mathcal{L}}(x^\star, \mu, \lambda) \cdot d, d \rangle \geqslant 0$$

où $\mathcal{T}(x^\star)$ est un espace tangent défini par

$$\mathcal{T}(x^\star) = \left\{d \in \mathbb{R}^n,\ \langle \nabla g_i(x^\star), d \rangle = 0,\ \forall i \in A(x^\star),\ \langle \nabla h_j(x^\star), d \rangle = 0,\ \forall j \in [\![1, q]\!]\right\}$$

et $H_{\mathcal{L}}(x^\star, \mu, \lambda)$ est la matrice hessienne de la fonction $x \mapsto \mathcal{L}(x, \mu, \lambda)$ au point $x^\star$.

Remarque 5.3

- Un point vérifiant les conditions (KKT) est stationnaire (ou critique) pour le problème $(\mathcal{P})$. On dit que ce point est un point KKT.
Un point KKT n'est pas forcément un minimum local. (par exemple : $f(x) = x^3$ et $g_1(x) = -x - 1$ avec $p = 1$, $q = 0$.)

- Il est possible qu'un point non régulier soit un minimum local. Dans ce cas, on ne peut pas vérifier les conditions KKT.
- Si le saut de dualité s'annule, alors un point régulier est un minimum si, et seulement s'il vérifie les conditions KKT.

Démonstration 21

On peut retrouver le système d'optimalité en introduisant le lagrangien du problème qui est la fonction

$$\begin{aligned}\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q &\to \mathbb{R} \\ (x, \mu, \lambda) &\mapsto f(x) + \sum_{i=1}^{p} \mu_i g_i(x) + \sum_{j=1}^{q} \lambda_j h_j(x)\end{aligned}$$

Le nom des multiplicateurs de Lagrange vient du fait que les μ_i et λ_j multiplient les contraintes dans le lagrangien. Les conditions nécessaires d'optimalité d'écrivent alors

$$\nabla_x \mathcal{L}(x^\star, \mu, \lambda) = 0,\ \mu \in (\mathbb{R}_+)^p,\ \langle \mu, g(x^\star) \rangle = 0$$

Théorème 5.7 – Condition suffisante (cas convexe)

On dit que le problème de minimisation $\inf_{x \in C} f(x)$ ***est convexe si*** f ***est convexe et de classe*** $\mathscr{C}^1$***, si les contraintes d'inégalité sont aussi convexes et de classe*** $\mathscr{C}^1$ ***et si les contraintes d'égalité sont affines.***
Soit $(x^\star, \mu^\star, \lambda^\star)$ *un point vérifiant les conditions KKT, avec* f*,* g*,* h *deux fois différentiables en* $x^\star$*. On suppose que*

$$\forall d \in \mathcal{T}(x^\star) \backslash \{0\}, \langle H_{\mathcal{L}}(x^\star, \mu^\star, \lambda^\star) \cdot d, d \rangle > 0$$

Autrement dit, la matrice hessienne $H_{\mathcal{L}}(x^\star, \mu^\star, \lambda^\star)$ *est définie positive. Alors* $x^\star$ *est un minimum local du problème* $(\mathcal{P})$*.*

Démonstration 22

Soit $x^\star \in C$ un point KKT. On a donc

$$\nabla f(x^\star) = -\sum_{i=1}^{p} \mu_i . \nabla g_i(x^\star) - \sum_{j=1}^{q} \lambda_j . \nabla h_j(x^\star)$$

La fonction f étant convexe sur C,

$$\begin{aligned} f(x) - f(x^\star) &\geqslant \langle \nabla f(x^\star), x - x^\star \rangle \\ &= \sum_{i=1}^{p} \mu_i \langle \nabla g_i(x^\star), x^\star - x \rangle + \sum_{j=1}^{q} \lambda_j \langle \nabla h_j(x^\star), x^\star - x \rangle \end{aligned} \tag{5.1}$$

Par qualification des contraintes, les fonctions h_j étant affines, elles peuvent s'exprimer comme

$$h_j(x) = h_j(x^\star) + \langle \nabla h_j(x^\star), x - x^\star \rangle$$

Or, comme x et $x^\star$ sont dans C, on a $h_j(x) = h_j(x^\star) = 0$ car ce sont des contraintes d'égalité. Donc $\langle \nabla h_j(x^\star), x - x^\star \rangle = 0$, alors

$$f(x) - f(x^\star) \geqslant \sum_{i=1}^{p} \mu_i \langle \nabla g_i(x^\star), x^\star - x \rangle \geqslant \sum_{i=1}^{p} \mu_i (g_i(x^\star) - g_i(x))$$

où nous avons utilisé le fait que les contraintes g_i sont convexes. Par ailleurs, comme chaque $\mu_i \geqslant 0$ et $g_i(x) \leqslant 0$ pour tout $x \in C$, on a donc $\sum_{i=1}^{p} \mu_i g_i(x) \leqslant 0$ et donc

$$f(x) - f(x^\star) \geqslant \sum_{i=1}^{p} \mu_i g_i(x^\star)$$

Finalement, par la condition de complémentarité, si $g_i(x^\star) < 0$ alors $\mu_i = 0$, et si $g_i(x^\star) = 0$ alors $\mu_i \geqslant 0$, d'où la conclusion

$$f(x) \geqslant f(x^\star)$$

Proposition 5.5 – Condition nécessaire et suffisante (cas convexe)

Soit $x^\star$ un point KKT. Alors, $x^\star$ est un minimum local (donc global) de f si et seulement s'il existe des nombres positifs $\mu_1, \dots, \mu_p$ et des nombres réels $\lambda_1, \dots, \lambda_q$ tels que $\nabla_x \mathcal{L}(x^\star, \mu^\star, \lambda^\star) = 0$.

$$(KKT) \begin{cases} \nabla f(x^\star) + \sum_{i=1}^{p} \mu_i . \nabla g_i(x^\star) + \sum_{j=1}^{q} \lambda_j . \nabla h_j(x^\star) = 0 \\ \mu_i \geqslant 0, \quad \mu_i g_i(x^\star) = 0, \quad \forall i \in \{1, \dots, p\} \end{cases} \Longleftrightarrow x^\star \text{ solution}$$

Résumé : Démarche pour résoudre un problème de minimisation.

La démarche pour résoudre un problème de minimisation sous contraintes se décompose en quatre étapes.

(1) On montre a priori que le problème admet une solution. Pour cela, on pourra utiliser les conditions suffisantes d'existence (Théorème 5.7). Il sera parfois nécessaire d'utiliser des raisonnements spécifiques au problème donné.

(2) On cherche ensuite les points satisfaisant les conditions nécessaires de KKT. On note cet ensemble E_1.

(3) On cherche les points où la contrainte n'est pas qualifiée. On appelle cet ensemble E_2.

(4) Si le problème a une solution, le minimum appartient à $E_1 \bigcap E_2$. Un minimum est donc un point minimum dans $E_1 \bigcap E_2$.

Remarque 5.4

Dans le cas convexe, il suffit de trouver des points vérifiant les conditions nécessaires d'optimalité pour pouvoir conclure à l'existence d'un minimum.

Exemple 5.4

On considère le problème de minimisation sous contrainte

$$\inf_{x^2+y^2 \geqslant 1} f(x, y) = x^4 + 3y^4$$

(1) **Existence.** On sait que pour tous $(x, y) \in \mathbb{R}^2$,

$$f(x, y) \geqslant 2x^2 + 6y^2 - 4 \geqslant 2\|(x, y)\|^2 - 4 \xrightarrow[\|(x,y)\| \to +\infty]{} +\infty$$

On en déduit que f est infinie à l'infini, et l'ensemble des contraintes $C = \{(x, y) \in \mathbb{R}^2, g(x, y) \leqslant 0\}$, avec $g(x, y) = 1 - x^2 - y^2$ est fermé, ce qui garantit l'existence de solution(s) pour le problème d'optimisation ci-dessus.

(2) **Conditions d'optimalité au premier ordre.** Le théorème de Kuhn-Tucker assure l'existence de $\mu \in \mathbb{R}_+$ tel que $\nabla f(x, y) + \mu . \nabla g(x, y) = 0$.

(a) (Réalisabilité primale) $(x, y) \in C = \{(x, y) \in \mathbb{R}^2,\ g(x, y) = 1 - x^2 - y^2 \leqslant 0\}$.

(b) (Réalisabilité duale et stationnarité) On en a besoin que $\nabla_x \mathcal{L}(x^\star, \mu^\star, \lambda^\star) = 0$, alors

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x}(x^\star, \mu^\star, \lambda^\star) = 4x^3 - 2\mu x = 0 \\ \frac{\partial \mathcal{L}}{\partial y}(x^\star, \mu^\star, \lambda^\star) = 12y^3 - 2\mu y = 0 \end{cases}$$

(c) (Complémentarité) Avec la contrainte $g(x, y) = 1 - x^2 - y^2$,

$$\mu g(x, y) = \mu(x^2 + y^2 - 1) = 0$$

(3) Supposons que l'on ait $\mu = 0$. Alors, $x = y = 0$. Impossible car $(0, 0)$ n'appartient pas à l'ensemble des contraintes. Donc, on a nécessairement $\mu > 0$.

(4) Des deux premières équations, on tire que les minimums sont à choisir parmi

$$(x_1, y_1) = \left(0, \pm\sqrt{\frac{\mu}{6}}\right), (x_2, y_2) = \left(\pm\sqrt{\frac{\mu}{2}}, 0\right), (x_3, y_3) = \left(\pm\sqrt{\frac{\mu}{2}}, \pm\sqrt{\frac{\mu}{6}}\right)$$

— **Étude de** (x_1, y_1). Puisque $x^2 + y^2 = 1$, on obtient $\mu = 6$ dans ce cas, et $(x_1, y_1) = (0, \pm 1)$, et $f(x_1, y_1) = 3$.

— **Étude de** (x_2, y_2). Puisque $x^2 + y^2 = 1$, on obtient $\mu = 2$ dans ce cas, et $(x_2, y_2) = (\pm 1, 0)$, et $f(x_2, y_2) = 1$.

— **Étude de** (x_3, y_3). Puisque $x^2 + y^2 = 1$, on obtient $\mu = \frac{3}{2}$ dans ce cas, et $(x_3, y_3) = (\pm\frac{\sqrt{3}}{2}, \pm\frac{1}{2})$, et $f(x_3, y_3) = \frac{3}{4}$.

Finalement, on en déduit que $\inf_{(x,y)\in C} f(x, y) = f(x_3, y_3) = \frac{3}{4}$.