CHAPITRE 2

MÉTHODES DE PÉNALITÉ

Nous considérons le problème d'optimisation avec contraintes suivant :

$$\begin{cases} \min f(x), \\ x \in \Omega, \end{cases} \tag{2.1}$$

où $f : \mathbb{R}^n \to \mathbb{R}$ est une fonction et Ω est un sous-ensemble fermé non vide de $\mathbb{R}^n$.

Comme d'habitude, la plupart du temps Ω sera défini au moyen des contraintes d'égalité et d'inégalité, c'est-à-dire

$$\Omega = \{x \in \mathbb{R}^n : h(x) = 0 \text{ et } g(x) \leq 0\}, \tag{2.2}$$

où $h(x) = (h_1(x), h_2(x), \ldots, h_m(x))^T$ et $g(x) = \left(g_1(x), g_2(x), \ldots, g_p(x)\right)^T$.

Définition 2.1. (Méthode de pénalité). L'idée d'une méthode de fonction de pénalité est de remplacer le problème d'optimisation avec contraintes (2.1) par un problème d'optimisation sans contraintes suivant :

$$\begin{cases} \min f(x) + \gamma P(x), \\ x \in \mathbb{R}^n, \end{cases} \tag{2.3}$$

où $\gamma > 0$ et P est une fonction continue telle que

1. $P(x) = 0$ pour tout $x \in \Omega$,
2. $P(x) > 0$ pour tout $x \notin \Omega$.

Une fonction P avec les propriétés ci-dessus est appelée fonction de pénalité pour Ω. Par exemple, si Ω est de la forme (2.2), alors une fonction de pénalité est donnée par :

$$P(x) = \sum_{i=1}^{m} |h_i(x)|^\alpha + \sum_{i=1}^{p} \left(g_i^+(x)\right)^\alpha, \tag{2.4}$$

où $\alpha \geq 1$ est une constante et $g^+ : \mathbb{R}^n \to \mathbb{R}_+$ est une fonction donnée par :

$$g_i^+(x) = \max(0, g_i(x)) = \begin{cases} g_i(x) & \text{si } g_i(x) > 0, \\ 0 & \text{si } g_i(x) \leq 0. \end{cases}$$

- ☞ Pour $\alpha = 1$, la fonction (2.4) est appelée fonction de pénalité linéaire.
- ☞ Pour $\alpha = 2$, la fonction (2.4) est appelée fonction de pénalité quadratique.

Exemple 2.1. Soient $g_1, g_2 : \mathbb{R} \to \mathbb{R}$ deux fonctions définies par :

$$g_1(x) = x - 2 \quad \text{et } g_2(x) = -(x+1)^3.$$

et soit L'ensemble des points réalisable est définie par :

$$\Omega = \{x \in \mathbb{R} : g_1(x) \leq 0,\ g_2(x) \leq 0\} = [-1, 2].$$

Nous avons :

$$g_1^+(x) = \max(0, g_1(x)) = \begin{cases} x-2 & \text{si } x > 2, \\ 0 & \text{si } x \leq 2. \end{cases}$$

et

$$g_2^+(x) = \max(0, g_2(x)) = \begin{cases} -(x+1)^3 & \text{si } x < -1, \\ 0 & \text{si } x \geq -1. \end{cases}$$

Pour $\alpha = 1$, On obtient :

$$P(x) = g_1^+(x) + g_2^+(x) = \begin{cases} x-2 & \text{si } x > 2, \\ 0 & \text{si } -1 \leq x \leq 2, \\ -(x+1)^3 & \text{si } x < -1. \end{cases}$$

Exemple 2.2. On considère le problème d'optimisation avec contraintes d'égalité suivant :

$$\begin{cases} \min x^T Q x \\ \|x\|^2 = 1, \end{cases} \tag{2.5}$$

où Q est une matrice carrée d'ordre n, symétrique et définie positive.

1. Soient $P(x) = \left(\|x\|^2 - 1\right)^2$ la fonction de pénalité quadratique et $\gamma > 0$ le paramètre de pénalité. Écrire un problème d'optimisation sans contraintes dont la solution x_γ rapproche la solution optimale du problème (2.5).
2. Monter que pour tout $\gamma > 0$, x_γ est un vecteur propre de Q.
3. Monter que $\|x_\gamma\|^2 - 1 = O(1/\gamma)$ quand $\gamma \to +\infty$.

Solution :

1. Le problème d'optimisation sans contraintes associe est donné par :

$$\begin{cases} \min q(x) = x^T Q x + \gamma \left[\|x\|^2 - 1\right]^2 \\ x \in \mathbb{R}^n \end{cases}$$

2. Nous avons :

$$q(x) \geq \lambda_{min} \|x\|^2 + \gamma \left[\|x\|^2 - 1\right]^2,$$

où λ_{min} la petite valeur propre de Q. Donc, $q(x) \xrightarrow[\|x\| \to +\infty]{} +\infty$ c'est à dire q est coercive et continue.
Alors, la fonction q admet au moins un point de minimum global noté x_γ.
D'après la condition nécessaire d'optimalité du premier ordre, on obtient :

$$\begin{aligned} \nabla q(x_\gamma) = 0 &\Leftrightarrow 2Qx_\gamma + 4\gamma \left[\|x_\gamma\|^2 - 1\right] x_\gamma = 0 \\ &\Leftrightarrow Qx_\gamma = 2\gamma \left[1 - \|x_\gamma\|^2\right] x_\gamma = \lambda_\gamma x_\gamma, \end{aligned}$$

où λ_γ est un scalaire. Donc, x_γ est un vecteur propre de Q.

3. Nous avons :

$$\lambda_\gamma = 2\gamma \left[1 - \|x_\gamma\|^2\right] \leq \lambda_{max},$$

où λ_{max} est plus grande valeur propre de Q. Pour $\gamma \to +\infty$, nous avons :

$$\|x_\gamma\|^2 - 1 = -\frac{\lambda_{max}}{2\gamma} = O\left(1/\gamma\right).$$

Nous analysons maintenant la méthode des pénalité dans un cadre plus général. On désigne par x^* une solution globale du problème (2.1). Soient P une fonction de pénalité et (γ_k) est une suite positive et croissante. On considère une fonction $q\left(\gamma_k, \cdot\right) : \mathbb{R}^n \to \mathbb{R}$ définie par :

$$q\left(\gamma_k, x\right) = f\left(x\right) + \gamma_k P\left(x\right).$$

Pour chaque $k \in \mathbb{N}$, nous pouvons écrire le problème d'optimisation sans contrainte associé suivant :

$$\begin{cases} \min q\left(\gamma_k, x\right) \\ x \in \mathbb{R}^n \end{cases} \tag{2.6}$$

On note par x^k le point de minimum de la fonction $q\left(\gamma_k, x\right)$ sur $\mathbb{R}^n$.

Le lemme suivant décrit certaines relations utiles entre le problème d'optimisation avec contraintes et les problèmes d'optimisation sans contraints associés.

Lemme 2.1. *Pour tout $k \in \mathbb{N}$, Nous avons les inégalités suivantes :*

1. $q\left(\gamma_k, x^k\right) \leq q\left(\gamma_{k+1}, x^{k+1}\right)$,
2. $P\left(x^{k+1}\right) \leq P\left(x^k\right)$,
3. $f\left(x^k\right) \leq f\left(x^{k+1}\right)$,
4. $f\left(x^k\right) \leq q\left(\gamma_k, x^k\right) \leq f\left(x^*\right)$.

Démonstration. 1. (γ_k) est croissante, alors nous avons :

$$\gamma_{k+1} \geq \gamma_k \Rightarrow q\left(\gamma_{k+1}, x^{k+1}\right) = f\left(x^{k+1}\right) + \gamma_{k+1} P\left(x^{k+1}\right) \geq f\left(x^{k+1}\right) + \gamma_k P\left(x^{k+1}\right) = q\left(\gamma_k, x^{k+1}\right). \tag{2.7}$$

D'autre part, x^k est un point de minimum global de $q\left(\gamma_k, x\right)$, donc on a :

$$q\left(\gamma_k, x^k\right) = f\left(x^k\right) + \gamma_k P\left(x^k\right) \leq f\left(x^{k+1}\right) + \gamma_k P\left(x^{k+1}\right) = q\left(\gamma_k, x^{k+1}\right). \tag{2.8}$$

En combinant (2.7) et (2.8), on obtient 1.

2. Les points x^k et x^{k+1} sont minimise respectivement les fonctions $q\left(\gamma_k, x\right)$ et $q\left(\gamma_{k+1}, x\right)$ sur $\mathbb{R}^n$. Alors, nous avons :

$$\begin{aligned} q\left(\gamma_k, x^k\right) &= f\left(x^k\right) + \gamma_k P\left(x^k\right) \leq f\left(x^{k+1}\right) + \gamma_k P\left(x^{k+1}\right), \\ q\left(\gamma_{k+1}, x^{k+1}\right) &= f\left(x^{k+1}\right) + \gamma_{k+1} P\left(x^{k+1}\right) \leq f\left(x^k\right) + \gamma_{k+1} P\left(x^k\right). \end{aligned}$$

En somment les deux inégalités précédentes, on obtient :

$$\gamma_k P\left(x^k\right) + \gamma_{k+1} P\left(x^{k+1}\right) \leq \gamma_{k+1} P\left(x^k\right) + \gamma_k P\left(x^{k+1}\right).$$

Réarrangement, nous obtenons :

$$\left(\gamma_{k+1} - \gamma_k\right) P\left(x^{k+1}\right) \leq \left(\gamma_{k+1} - \gamma_k\right) P\left(x^k\right) \Rightarrow P\left(x^{k+1}\right) \leq P\left(x^k\right).$$

3. Pour chaque $k \in \mathbb{N}$, x^k est un point de minimum global de la fonction $q(\gamma_k, x)$ sur $\mathbb{R}^n$, on obtient :
$$q\left(\gamma_k, x^k\right) = f\left(x^k\right) + \gamma_k P\left(x^k\right) \leq f\left(x^{k+1}\right) + \gamma_k P\left(x^{k+1}\right)$$
Donc, on a :
$$f\left(x^{k+1}\right) - f\left(x^k\right) \geq \gamma_k \left[P\left(x^k\right) - P\left(x^{k+1}\right)\right]$$
D'après la question 2., nous avons $P\left(x^k\right) - P\left(x^{k+1}\right) \geq 0$ et $\gamma_k > 0$, on obtient 3.
4. Nous avons x^k est un point de minimum global de la fonction $q(\gamma_k, x)$ sur $\mathbb{R}^n$, donc on a :
$$f(x^*) + \gamma_k P(x^*) \geq q\left(\gamma_k, x^k\right) = f\left(x^k\right) + \gamma_k P\left(x^k\right).$$
D'autre part, x^* est un point de minimum global du problème d'optimisation avec contraintes (2.1), donc $P(x^*) = 0$, c'est à dire
$$f(x^*) \geq f\left(x^k\right) + \gamma_k P\left(x^k\right).$$
Puisque $P\left(x^k\right) \geq 0$ et $\gamma_k > 0$, alors on a :
$$f(x^*) \geq f\left(x^k\right).$$

□

Avec Lemme 2.1, nous sommes prêts à prouver le théorème de convergence de la méthode de pénalité suivant :

Théorème 2.1. *Si* $\left(x^k\right)$ *converge vers* $\bar{x}$*, alors* $\bar{x}$ *est une solution optimale du problème d'optimisation avec contraintes* (2.1).

Démonstration. D'après Lemme 2.1, la suite $\left(q\left(\gamma_k, x^k\right)\right)$ est croissante et Majorée par $f(x^*)$. Par conséquent, la suite $\left(q\left(\gamma_k, x^k\right)\right)$ est convergente vers une limite suivante :
$$q^* = \lim_{k \to +\infty} q\left(\gamma_k, x^k\right) \quad \text{telle que} \quad q^* \leq f(x^*).$$
Puisque la fonction f est continue et d'après Lemme 2.1, $f\left(x^k\right) \leq f(x^*)$. Alors, nous avons :
$$\lim_{k \to +\infty} f\left(x^k\right) = f\left[\lim_{k \to +\infty} x^k\right] = f(\bar{x}) \leq f(x^*).$$
Puisque les suites $\left(f\left(x^k\right)\right)$ et $\left(q\left(\gamma_k, x^k\right)\right)$ sont convergentes, alors la suite suivante :
$$\left(\gamma_k P\left(x^k\right)\right) = \left(q\left(\gamma_k, x^k\right) - f\left(x^k\right)\right)$$
est convergente vers la limite suivante :
$$\lim_{k \to +\infty} q\left(\gamma_k, x^k\right) = q^* - f(\bar{x}).$$
Avec Lemme 2.1, la suite $\left(P\left(x^k\right)\right)$ est décroissante et minorée par 0, donc converge et par conséquent $\left(P\left(x^k\right)\right)$ est convergente. Puisque $\gamma_k \xrightarrow[k \to +\infty]{} +\infty$, nous concluons que
$$\lim_{k \to +\infty} P\left(x^k\right) = 0.$$
Par continuité de P, on a :
$$0 = \lim_{k \to +\infty} P\left(x^k\right) = P\left[\lim_{k \to +\infty} x^k\right] = P(\bar{x}).$$
D'où $\bar{x}$ est un point réalisable c'est à dire $\bar{x} \in \Omega$. Puisque $f(x^*) \geq f(\bar{x})$, nous concluons que $\bar{x}$ une solution optimale de problème d'optimisation avec contraintes (2.1). □

2.1 Méthode de pénalité quadratique (Beltrami)

Considérons le problème d'optimisation avec contraintes suivant :

$$\begin{cases} \min f(x), \\ h(x) = 0, \\ g(x) \leq 0, \end{cases} \tag{2.9}$$

où $f : \mathbb{R}^n \to \mathbb{R}$, $h = (h_1, \ldots, h_m)$, et $g = (g_1, \ldots, g_p)$ sont des fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}^n$.

Définition 2.2. La fonction de pénalité quadratique $q(\gamma, x)$ pour le problème (2.9) est

$$q(\gamma, x) := f(x) + \frac{\gamma}{2} \sum_{i=1}^{m} h_i^2(x) + \frac{\gamma}{2} \sum_{i=1}^{p} \left(g_i^+(x)\right)^2, \tag{2.10}$$

où γ est le paramètre de pénalité et $g_i^+(x) = \max\{0, g_i(x)\}$ avec $i = 1, \ldots, p$.

Remarque 2.1. Remarquons que si γ est tend vers $+\infty$, les contraintes sont pénalisées. L'idée est donc de considérer une suite de (γ_k) avec $\gamma_k \xrightarrow[k \to +\infty]{} +\infty$ et de chercher le point de minimum x_k de la fonction de pénalité quadratique $q(\gamma_k, x)$ pour chaque k.

Comme les deux termes de pénalité dans (2.10) sont différentiable, toute méthode d'optimisation sans contrainte peut être utilisée pour le calcul de x_k. Par conséquent, une suite de sous-problèmes de minimisation sans contrainte $\min\limits_{x \in \mathbb{R}^n} q(\gamma_k, x)$ est résolue pour chaque valeur du paramètre de pénalité γ_k.

La méthode de pénalité quadratique est basée sur la minimisation de la fonction de pénalité quadratique dans le cadre suivant :

Algorithme 1 (Méthode de pénalité quadratique)

1: Poser $k = 0$, choisi un point de départ initial x_0^s, une valeur initiale $\gamma_0 > 0$ pour le paramètre de pénalité et une suite de tolérances ε_k avec $\varepsilon_k \xrightarrow[k \to +\infty]{} 0$.
2: Trouver un point de minimum x_k du sous-problème $\min\limits_{x \in \mathbb{R}^n} q(\gamma_k, x)$, à partir de x_k^s, et terminer les itérations lorsque $\|\nabla_x q(\gamma_k, x)\| \leq \varepsilon_k$.
3: Si un test de convergence de la méthode est satisfait, alors s'arrêter à x_k comme solution approchée du problème sinon :
4: Choisi un nouveau paramètre de pénalité $\gamma_{k+1} > \gamma_k$.
5: Choisi un nouveau point de départ x_{k+1}^s, poser $k = k + 1$ et passer à l'étape 2.

Quelques remarques sont les suivantes :

Remarque 2.2.
1. La suite des paramètres de pénalité (γ_k) peut être sélectionnée de manière adaptative en fonction de la difficulté à minimiser la fonction de pénalité à chaque itération k. Lorsque la minimisation de $q(\gamma_k, x)$ est trop coûteuse, alors γ_{k+1} peut être choisi proche de γ_k, par exemple $\gamma_{k+1} = 1.5\gamma_k$.
   D'autre part, si la minimisation de $q(\gamma_k, x)$ est relativement facile à faire, alors γ_{k+1} peut être sélectionné plus grand que γ_k, par exemple $\gamma_{k+1} = 10\gamma_k$.
2. Dans le cas où le problème (2.9) n'a que des contraintes d'égalité, alors $q(\gamma_k, x)$ est différentiable. Par conséquent, les algorithmes d'optimisation sans contrainte (à l'étape 2) peuvent être utilisés pour trouver une solution approchée x_k. Cependant, la minimisation de $q(\gamma_k, x)$ devient plus difficile à mesure que γ_k devient plus grand.
   Au voisinage de point de minimum, la matrice Hessienne $\nabla^2 q(\gamma_k, x)$ devient mal conditionnée, et dans ce cas, les méthodes de quasi-Newton ou de gradient conjugué fonctionnent très mal.

Théorème 2.2. *Soit le problème d'optimisation avec contraintes d'égalité suivant :*

$$\begin{cases} \min f(x), \\ h(x) = 0, \end{cases} \tag{2.11}$$

où $f : \mathbb{R}^n \to \mathbb{R}$ et $h = (h_1, \ldots, h_m)$ sont des fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}^n$. On suppose que x_k, pour chaque $k \in \mathbb{N}$, point de minimum global de la fonction de pénalité quadratique suivante :

$$q(\gamma_k, x) := f(x) + \frac{\gamma_k}{2} \sum_{i=1}^{m} h_i^2(x), \tag{2.12}$$

où $\gamma_k \xrightarrow[k \to +\infty]{} +\infty$. Soit (x_k) une suite converge vers x^*.

Si x^* est un point régulier, alors x^* vérifiant la condition de Lagrange du problème (2.11)*. De plus, nous avons : $\gamma_k h(x_k) \xrightarrow[k \to +\infty]{} \lambda^*$ où $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)^T$ est l'unique vecteur de multiplicateur de Lagrange.*

Démonstration. D'après (2.12), nous avons :

$$\nabla q(\gamma_k, x_k) = \nabla f(x_k) + \sum_{i=1}^{m} \gamma_k h_i(x_k) \nabla h_i(x_k). \tag{2.13}$$

En utilisant (2.13) et le critère de test d'arrêt pour Algorithme 1, on obtient :

$$\left\| \nabla f(x_k) + \sum_{i=1}^{m} \gamma_k h_i(x_k) \nabla h_i(x_k) \right\| \leq \varepsilon_k \tag{2.14}$$

Avec l'inégalité suivante $\|a\| - \|b\| \leq \|a + b\|$, (2.14) devient :

$$\left\| \sum_{i=1}^{m} h_i(x_k) \nabla h_i(x_k) \right\| \leq \frac{1}{\gamma_k} \left[\varepsilon_k + \|\nabla f(x_k)\| \right] \tag{2.15}$$

Par passage à la limite $k \to +\infty$ pour l'inégalité (2.15), on obtient :

$$\sum_{i=1}^{m} h_i(x^*) \nabla h_i(x^*) = 0.$$

Mais x^* est un point régulier, alors $\{\nabla h_1(x^*), \ldots, \nabla h_m(x^*)\}$ sont linéairement indépendant, donc $h_i(x^*) = 0$ pour tout $i = 1, \ldots, m$.

On désigne par $A^T(x) = \nabla h(x)$ et $\lambda_k = \gamma_k h(x_k)$. Alors, nous avons :

$$A^T(x_k) \lambda_k = \nabla q(\gamma_k, x_k) - \nabla f(x_k).$$

Donc, on a :

$$\lambda_k = \left[A(x_k) A^T(x_k) \right]^{-1} A(x_k) \left[\nabla q(\gamma_k, x_k) - \nabla f(x_k) \right]. \tag{2.16}$$

Par passage à la limite $k \to +\infty$ pour (2.16), on trouve :

$$\lambda_k \xrightarrow[k \to +\infty]{} \lambda^* = -\left[A(x^*) A^T(x^*) \right]^{-1} A(x^*) \nabla f(x^*). \tag{2.17}$$

D'autre part, par passage à la limite pour (2.14), on obtient :

$$\nabla f(x^*) + A^T(x^*) \lambda^* = 0, \tag{2.18}$$

d'où, x^* vérifiant la condition de Lagrange du problème (2.11) avec λ^* est l'unique vecteur de multiplicateur de Lagrange. □

Exercice 2.1. Soit le problème d'optimisation suivant :

$$\begin{cases} \min f(x,y,z) = x^2 + y^2 + xy - 2y \\ x + y = 2. \end{cases}$$

Résoudre ce problème par la méthode de pénalité de Beltrami.

Solution : On pose $h(x,y) = x + y - 2$ et $P(x,y) = h^2(x,y)$. Soit $q(\gamma,x,y)$ telle que

$$q(\gamma,x,y) = f(x,y) + \gamma P(x,y).$$

1. On cherche les points critiques de la fonction q. Nous avons :

$$\nabla q(\gamma,x,y) = \nabla f(x,y) + \gamma \nabla P(x,y) = \nabla f(x,y) + 2\gamma h(x,y) \nabla h(x,y) = 0,$$

d'où, le système suivant :

$$\begin{cases} 2x + y + 2\gamma(x + y - 2) = 0, \\ 2y + x - 2 + 2\gamma(x + y - 2) = 0. \end{cases} \Rightarrow 2x + y = 2y + x - 2 \Rightarrow y = x + 2.$$

D'autre part, nous avons :

$$2(1+\gamma)x + (1+2\gamma)(x+2) = 4\gamma \Rightarrow (3+4\gamma)x = -2 \Rightarrow x_\gamma = -\frac{2}{3+4\gamma} \text{ et } y_\gamma = 2 + x_\gamma.$$

2. On pose $x = r\cos\theta$ et $y = r\sin\theta$. Nous avons :

$$q(\gamma,x,y) = r^2(1 + \cos\theta\sin\theta) - 2r\sin\theta + r\gamma\left(\cos\theta + \sin\theta - \frac{2}{r}\right) \xrightarrow[r\to+\infty]{} +\infty,$$

d'où, q est coercive donc (x_γ, y_γ) est un point de minimum global de q sur $\mathbb{R}^2$ pour tout $\gamma > 0$. Finalement, le point $(x_\gamma, y_\gamma) \to (0,2)$ est solution globale du problème considéré.

2.2 Méthode de pénalité exacte

Définition 2.3. La fonction de pénalité exacte $q(\gamma,x)$ du problème (2.9) est donnée par :

$$q(\gamma,x) = f(x) + \sum_{i=1}^{m} |h_i(x)| + \sum_{i=1}^{p} g_i^+(x), \tag{2.19}$$

où $\gamma > 0$ est le paramètre de pénalité et $g_i^+(x) = \max\{0, g_i(x)\}$ pour tout $i = 1,\ldots,p$.

Le théorème suivant est prouvé par Han et Mangasarian en 1979.

Théorème 2.3. *On suppose que le problème d'optimisation avec contraintes* (2.9) *admet une solution locale stricte* x^* *pour laquelle les conditions de Karush-Kuhn-Tucker sont vérifiées avec les multiplicateurs* $\lambda^* = (\lambda_1^*,\ldots,\lambda_m^*) \in \mathbb{R}^m$ *et* $\mu^* = (\mu_1^*,\ldots,\mu_p^*) \in \mathbb{R}_+^p$. *Alors,* x^* *est un point de minimum local de la fonction* $q(\gamma,x)$ *définie par* (2.19) *pour tout* $\gamma > \gamma^*$, *où*

$$\gamma^* = \max\left\{\max_{i=1,\ldots,m} |\lambda_i^*|, \max_{i=1,\ldots,p} \mu_i^*\right\}. \tag{2.20}$$

De plus, si les conditions suffisantes d'optimalités du seconde ordre du problème (2.9) *sont vérifiées au point* x^* *pour* $\gamma > \gamma^*$, *alors* x^* *est un point de minimum local strict de la fonction* $q(\gamma,x)$.

La définition suivante caractérise les points stationnaires de $q(\gamma, x)$.

Définition 2.4. Un point $\bar{x}$ est un point stationnaire pour la fonction de pénalité $q(\gamma, x)$ si

$$D(q(\gamma, \bar{x}); v) := \nabla f(\bar{x})^T v + \gamma \sum_{i=1}^{m} \left| \nabla h_i(\bar{x})^T v \right| + \gamma \sum_{i=1}^{p} \left[\nabla g_i(\bar{x})^T v \right]^+ \geq 0, \text{ pour tout } v \in \mathbb{R}^n. \quad (2.21)$$

De même, $\bar{x}$ est un point stationnaire pour la mesure suivante :

$$M(x) := \sum_{i=1}^{m} |h_i(x)| + \sum_{i=1}^{p} [g_i(x)]^+, \quad (2.22)$$

si $D(M(\bar{x}); v) \geq 0$, pour tout $v \in \mathbb{R}^n$.

Si un point est non réalisable pour (2.9) mais stationnaire par rapport à la mesure M, alors c'est un point stationnaire non réalisable.

Le théorème suivant montre que sous certaines hypothèses, les points stationnaires de $q(\gamma, x)$ sont des points vérifiant les conditions de K.K.T. du problème d'optimisation avec contraintes (2.9).

Théorème 2.4. *Soit $\bar{x}$ est un point stationnaire de la fonction de pénalité $q(\gamma, x)$ pour tout $\gamma > \bar{\gamma} > 0$, où $\bar{\gamma}$ est un réel donné. Alors, nous avons :*

1. *Si $\bar{x}$ est un réalisable pour le problème d'optimisation avec contraintes* (2.9), *alors il satisfait les conditions de K.K.T. pour ce problème.*
2. *Si $\bar{x}$ n'est pas réalisable pour le problème* (2.9), *alors c'est un point stationnaire non réalisable.*

2.3 La méthode Lagrangienne augmentée

On considère le problème d'optimisation avec contraintes d'égalités suivant :

$$\begin{cases} \min f(x), \\ h(x) = 0, \end{cases} \quad (2.23)$$

où $f : \mathbb{R}^n \to \mathbb{R}$ et $h_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, sont des fonctions de classe $\mathcal{C}^1$.

Définition 2.5. La fonction Lagrangienne augmentée est définie par :

$$\mathcal{L}(x, \lambda, \gamma) := f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \frac{\gamma}{2} \sum_{i=1}^{m} h_i^2(x), \quad (2.24)$$

où $\lambda \in \mathbb{R}^m$ et $\gamma > 0$.

Soit x_k est un point de minimum de la fonction lagrangienne augmentée $\mathcal{L}(x, \lambda^k, \gamma_k)$ avec (λ^k) et (γ_k) deux suites de $\mathbb{R}^m$ et $\mathbb{R}_+^*$.

Alors, avec la condition d'optimalité de premier ordre pour d'optimisation sans contraintes nous avons :

$$\nabla \mathcal{L}\left(x_k, \lambda^k, \gamma_k\right) = \nabla f(x_k) + \sum_{i=1}^{m} \left[\lambda_i^k + \gamma_k h_i(x_k) \right] \nabla h_i(x_k).$$

En comparant ce résultat avec la condition d'optimalité (2.18), on trouve

$$\lambda_i^* \approx \lambda_i^k + \gamma_k h_i(x_k), \ i = 1, \ldots, m. \quad (2.25)$$

D'après (2.25), on a :

$$h_i\left(x_k\right) \approx \frac{\lambda_i^* - \lambda_i^k}{\gamma_k},\ i = 1, \ldots, m.$$

On remarque que si x_k est point réalisable du problème (2.23), alors $h\left(x_k\right) = 0$. Dans ce cas, $\lambda^k = \lambda^*$ est le multiplicateur de Lagrange.

On considère la suite de récurrence suivante :

$$\lambda_i^{k+1} = \lambda_i^k + \gamma_k h_i\left(x_k\right),\ i = 1, \ldots, m. \tag{2.26}$$

L'algorithme du Lagrangien augmenté peut être présenté comme suit.

Algorithme 2 (Algorithme du Lagrangien augmenté)

---

1 : Poser $k = 0$, choisi des points de départ initiaux x_0^d et λ^0. Choisi une valeur initial $\gamma_0 > 0$ et tolérance $\varepsilon_0 > 0$.
2 : Trouver un point de minimum approché x_k de $\S, \lambda^{\|}, \gamma_{\|}$. Si $\left\|\nabla \mathcal{L}\left(x_k, \lambda^k, \gamma_k\right)\right\| \leq \varepsilon_k$.
3 : Si un test de convergence de la méthode est satisfait, alors s'arrêter à x_k comme solution approchée du problème sinon :
4 : Choisi un nouveau multiplicateur de Lagrange $\lambda^{k+1} = \lambda^k + \gamma_k h\left(x_k\right)$, choisi un paramètre de pénalité $\gamma_{k+1} > \gamma_k$ et une tolérance $\varepsilon_{k+1} > 0$.
5 : Choisi un nouveau point de départ $x_{k+1}^d = x_k$, poser $k = k + 1$ et passer à l'étape 2.

Théorème 2.5. *Soit x^* est point régulier du problème* (2.23). *Si x^* est une solution local strict du problème* (2.23) *(vérifiant la condition suffisante d'optimalité du seconde ordre), alors il existe $\gamma \bar{>} 0$ tel que pour tout $\gamma \geq \bar{\gamma}$, x^* est un point de minimum local strict de la fonction de Lagrange augmentée $\mathcal{L}\left(x, \lambda^*, \gamma\right)$ où λ^* est le vecteur de multiplicateur de Lagrange.*