Chapitre 4 : Optimisation non-linéaire avec contraintes

4.1 Introduction

Nous avons traité des techniques d'optimisation d'une fonction non linaire d'une ou de plusieurs variables sans contraintes. Cependant, dans des situations plus réalistes, il faut également considérer des bornes communes sur les variables données par des **contraintes** du modèle étudié. Nous distinguons deux types de contraintes, des contraintes d'égalité et des contraintes d'inégalité

4.2 Optimisation avec contraintes d'égalité

De façon générale, les problèmes non linéaires avec contraintes d'égalité ont généralement la modelisation suivante :

$$\begin{aligned} &\max\ f(X) \\ &\text{s.c.} \\ &\left.\begin{aligned} h_1(X) &= 0 \\ &\vdots \\ h_m(X) &= 0 \end{aligned}\right\} \ m \text{ contraintes d'égalité.} \end{aligned}$$

Conditions d'optimalité

Les concepts d'optimalité (optimum local et absolu) restent les mêmes, à l'exception maintenant que tout point X_0 considéré doit être **admissible**, c'est-à-dire qu'il doit satisfaire toutes les contraintes $h_i(X_0) = 0$, $\forall i = 1,\dots m$.

Nous verrons, dans les sections suivantes, deux méthodes analytiques utilisée pour résoudre un problème non linéaire avec contrainte d'égalité.

4.2.1 Méthode de résolution par substitution

Cette technique de résolution est utilisée lorsqu'il y a une variable de plus que le nombre de contraintes et que ces contraintes sont simples (par exemple, linéaires). En fait, le problème est transformé par substitutions successives à un problème *d'une fonction d'une seule variable*. On peut ainsi utiliser les techniques d'optimisation non linéaire sans contraintes vues dans le chapitre précédents.

Exemple 4.1

Résoudre le problème $\max f(x,y) = -x^2 - y^2$ s.c. $3x + 4y - 25 = 0$.

Par la contrainte, on isole $y = \dfrac{25-3x}{4}$ que l'on substitue dans $f(x,y)$:

$f(x,y) = -x^2 - \left(\dfrac{25-3x}{4}\right)^2 = g(x)$. $g(x)$ est donc une fonction à une seule variable dont on trouve le point stationnaire : $g'(x^*) = 0 \Rightarrow -2x_1^* - 2\left(\dfrac{-3}{4}\right)\left(\dfrac{25-3x^*}{4}\right) = 0 \qquad \Rightarrow x^* = 3$.

La dérivée seconde étant toujours négative ($g^{(2)}(x) = -25/8, \forall x$), la fonction $g(x)$ est strictement concave et le point stationnaire trouvé est un maximum absolu.
L'évaluation de y est immédiate : $y^* = \dfrac{25-3(3)}{4} = 4$, d'où $(x^*, y^*) = (3, 4)$ est le point maximisant l'objectif du problème original, sous la contrainte imposée.

3.2.2 Conditions d'optimum du premier ordre

Considérons le problème suivant à deux variables : $\max \quad f(x,y)$ s.c. $h(x,y) = 0$.

Théorème 3.1 : Conditions (nécessaires) d'optimum du premier ordre

Soit $f(x,y)$ et $h(x,y)$ deux fonctions dérivables; si $f(x,y)$ admet un optimum local en $X^* = (x^*, y^*)$ sous l'hypothèse que $\nabla h(X^*) \neq (0,0)$, alors :

1. **il existe un $\lambda^* \in \Re$ tel que**

$$\left.\begin{array}{l} f'_x(X^*) - \lambda^*\, h'_x(X^*) = 0 \\ f'_y(X^*) - \lambda^*\, h'_y(X^*) = 0 \end{array}\right\} \Leftrightarrow \nabla f(X^*) - \lambda^* \nabla h(X^*) = (0,0).$$

2. $h(X^*) = 0$ **;**

λ^* est le ***multiplicateur de Lagrange***

La première condition est une généralisation du gradient nul, en tenant compte de cette même contrainte. La deuxième condition indique la contrainte est satisfaite par le point considéré X^*.
Par conséquence Un point X^* est un **point stationnaire** sil il satisfait ces conditions d'optimum du premier ordre . Il faut noter que rien n'est sûre que ce point pourra être un optimum.

Si l'une des dérivées partielles $h'_x(X^*)$ ou $h'_y(X^*)$ est nulle, on utilisera l'autre ratio pour calculer λ^*. Ceci est possible car, par hypothèse, on a $\nabla h(X^*) = (h'_x(X^*), h'_y(X^*)) \neq (0,0)$.

Retour sur l'exemple 4.1

Considérons à nouveau $\max f(x,y) = -x^2 - y^2$ s.c. $3x + 4y = 25$.

Vérifions le calcul de λ^* au point $(x^*, y^*) = (3, 4)$:

$$\lambda^* = \frac{f'_x(X^*)}{h'_x(X^*)} = -\frac{2x}{3} = -\frac{2(3)}{3} = -2 \text{ et } \lambda^* = \frac{f'_y(X^*)}{h'_y(X^*)} = -\frac{y}{2} = -\frac{4}{2} = -2.$$

Ce résultat vérifie aussi **que** $\nabla f(x^*, y^*) - \lambda^* \nabla h(x^*, y^*) = (-6, -8) - (-2)(3, 4) = (0, 0)$.

4.2.3 La méthode du Lagrangien

La méthode du lagrangien utilise la règle du multiplicateur de Lagrange pour résoudre des problèmes d'optimisation avec des contraintes d'égalité. Nous examinerons le cas d'une seule contrainte à deux variables $X = (x, y)$.

Les étapes de la méthode du Lagrangien

Soit le problème : $\max f(X)$ s.c. $h(X) = 0$.

1- Calculer le **Lagrangien** associé à ce problème la fonction :

$$L(x, y, \lambda) = L(x, y, \lambda) = f(x, y) - \lambda h(x, y) .$$

Le Lagrangien transforme le problème avec contrainte à un problème sans contrainte aprés l'introduction de la contrainte dans la fonction-objectif avec une certaine pénalité λ. On se retrouve ainsi à maximiser une *fonction à trois variables sans contrainte.*

2. Trouver les points stationnaires (x^*, y^*, λ^*) du Lagrangien $L(x, y, \lambda)$, cela veut dire que nous cherchons (x^*, y^*, λ^*) tel que $\nabla L(x^*, y^*, \lambda^*) = (0, 0, 0)$.

Nous déduisons alors le système suivant comportant trois équations :

$$\begin{aligned}
L'_x(x^*, y^*, \lambda^*) &= \quad f'_x(x^*, y^*) - \lambda^* h'_x(x^*, y^*) = 0 \\
L'_y(x^*, y^*, \lambda^*) &= \quad f'_y(x^*, y^*) - \lambda^* h'_y(x^*, y^*) = 0 \\
L'_\lambda(x^*, y^*, \lambda^*) &= \quad -h(x^*, y^*) = 0
\end{aligned}$$

Dans les deux premières équations, on reconnaît les conditions d'optimum du premier ordre : $\nabla f(x^*, y^*) - \lambda^* \nabla h(x,^* y^*) = (0,0)$**, alors que la troisième est équivalente à la contrainte de départ imposée par le modèle :** $h(x, y) = 0$**, au point** (x^*, y^*) **(point admissible).**

Remarque : pour un problème de minimisation remplacer – par + dans le lagrangien.

Remarque sur la nature du point stationnaire du Lagrangien

Les points stationnaires trouvés ne constituent pas forcément un optimum. Pour vérifier ça, il faut vérifier la convexité/concavité de la fonction-objectif.
Dans les autres cas, il faut calculer le **Hessien bordé** pour déterminer la nature de ces points.

Retour sur l'exemple 4.1

Résoudre par la méthode du Lagrangien :
$\max f(X) = -x^2 - y^2$ s.c. $h(X) = 3x + 4y = 25$.
On forme d'abord **le Lagrangien** : $L(x,y,\lambda) = -x^2 - y^2 - \lambda(3x + 4y - 25)$.
On optimise ensuite $L(x,y,\lambda)$, une fonction à **trois** variables pour laquelle on cherche les **points stationnaires** (x^*, y^*, λ^*) tel que $\nabla L(x^*, y^*, \lambda^*) = (0,0,0)$:

$$L'_x(x^*, y^*, \lambda^*) = -2x^* - 3\lambda^* = 0 \qquad \Rightarrow x^* = -\frac{3}{2}\lambda^*.$$

$$L'_y(x^*, y^*, \lambda^*) = -2y^* - 4\lambda^* = 0 \qquad \Rightarrow y^* = -\frac{4}{2}\lambda^*.$$

$$L'_\lambda(x^*, y^*, \lambda^*) = -\left(3x^* + 4y^* - 25\right) = 0.$$

Avec x^* et y^* dans la troisième équation, on a : $-\left(3\left(-\frac{3}{2}\lambda^*\right) + 4\left(-\frac{4}{2}\lambda^*\right) - 25\right) = 0$,

d'où $\frac{9}{2}\lambda^* + \frac{16}{2}\lambda^* = -25 \qquad \Rightarrow \qquad \lambda^* = -2$.

Reportant à son tour la valeur de λ^* dans les deux premières expressions, on retrouve :
$x^* = -\frac{3}{2}(-2) = 3$ et $y^* = -\frac{4}{2}(-2) = 4$.
Donc $(x^*, y^*, \lambda^*) = (3, 4, -2)$ est le seul point stationnaire du Lagrangien, et par conséquent, $(x^*, y^*) = (3, 4)$ est le seul point extrémal du problème original avec contrainte (trouvé sans avoir à énoncer les conditions d'optimum du premier ordre!).

Le fait que le point $(x^*, y^*, \lambda^*) = (3, 4, -2)$ soit un point stationnaire du Lagrangien ne garantit pas qu'il s'agisse d'un optimum. Il faut vérifier sa **nature**. Puisque la contrainte $h(x,y)$ est linéaire, le domaine est convexe. Il suffit alors de vérifier si la fonction-objectif est convexe/concave par l'analyse du signe de son Hessien. Dans le cas présent, le hessien est le suivant :

$$H(x,y) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

Le hessien est donc toujours défini négatif et la fonction-objectif $f(x,y)$ est strictement concave.
Par conséquent, le point stationnaire du Lagrangien correspond bien au maximum absolu du modèle sous contrainte.

4.2.4 Interprétation de λ^*

Soit le problème d'optimisation suivant: $\max \quad f(x,y) \quad$ s.c. $\quad h(x,y) = b$.

On voit clairement que la partie droite n'
Le Lagrangien associé à ce problème est donné par $L(x, y, \lambda) = f(x, y) - \lambda\ (h(x, y) - b)$. À l'optimalité, on obtient

$$L(x^*, y^*, \lambda^*) = f(x^*, y^*) - \lambda^*\ h(x^*, y^*) + \lambda^*\ b.$$

Ainsi, ce multiplicateur de Lagrange λ^* représente l'effet sur la valeur optimale de la fonction objectif $f(x^*, y^*)$ si l'on modifie d'une unité la valeur de *b*. Il faut bien s'entendre que c'est un calcul approximatif.

λ^* **donne une approximation de la variation de** $f(x^*, y^*)$ **lorsqu'on augmente d'une unité le côté droit** *b* **de la contrainte** $h(x, y) = b$ **.**

Retour sur l'exemple 4.1
max $f(x, y) = -x^2 - y^2$ s.c. $h(x, y) = 3x + 4y = 25$.
Pour ce problème, on a trouvé : $\lambda^* = -2$. Si on augmente b de une unité (donc $b = 26$), alors la valeur optimale de $f(X)$ devrait diminuer de 2 unités (**environ**).

Si on résolvait le nouveau problème avec la contrainte $h(x, y) = 3x + 4y = 26$ au lieu de 25, on obtiendrait la solution $x^* = 3{,}12$ $y^* = 4{,}16$ et $f(x^*, y^*) = -(3{,}12)^2 - (4{,}16)^2 = -27{,}04$. Ainsi, la diminution exacte de la fonction-objectif est de 2,04.

4.3 Optimisation avec contraintes d'inégalité (cas général)

Maintenant nous allons voir comment optimiser une fonction non linéaire avec des **contraintes d'inégalité**. Supposer le problème d'optimisation avec des contraintes d'égalité et d'inégalité suivant:

$$\begin{cases} \min_x \ f(X), \text{ Tel que,} \\ \vec{h}_i(X) = 0,\ i = 1, \ldots, l, \\ \vec{g}_i(X) \le 0,\ i = 1, \ldots, m. \end{cases}$$

De plus, supposer que f, h_i et g_i sont continument différentiables à un point x^*.

Le lagrangien pour ce problème est :

$$F(\underline{x}, \underline{\lambda}, \underline{\mu}) = f(\underline{x}) - \underline{\lambda}^T \left[\underline{h}(\underline{x})\ \right] - \underline{\mu}^T \left[\underline{g}(\underline{x})\right]$$

Où les éléments $\underline{\lambda}=\left(\lambda_1,\lambda_2 \ldots,\lambda_l\right)$ et $\underline{\mu}=\left(\mu_1,\mu_2 \ldots,\mu_m\right)$ sont appelés les multiplicateurs de Lagrange.

Si x^* est un minimum local qui remplit quelques conditions de régularité, alors il existe des constantes $\mu_i \; (i = 1, \ldots, m)$ et $\lambda_j \; (j = 1, \ldots, l)$ qui vérifient les conditions suivantes :

- **Stationnarité**

$$\nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{j=1}^{l} \lambda_j \nabla h_j(x^*) = 0,$$

- **Faisabilité principale**

$$h_j(x^*) = 0, \text{ for all } j = 1, \ldots, l$$
$$g_i(x^*) \leq 0, \text{ for all } i = 1, \ldots, m$$

- **Faisabilité Duelle**

$$\mu_i \geq 0 \; (i = 1, \ldots, m)$$

- **Relâchement complémentaire**

$$\mu_i g_i(x^*) = 0 \text{ for all } i = 1, \ldots, m.$$

Les conditions précédentes sont les **conditions nécessaires d'optimalité de karush Kuhn et Tucker (KKT)**, qui constitue le résultat fondamental de la programmation non linéaire (PNL) :

Les conditions suffisantes qu'un x* de point est un minimum local strict du problème simple classique de PNL, où f, g_j, et le h_i sont des fonctions deux fois différentiables sont :

- **Les conditions nécessaires de KKT sont vérifiées.**
- **La matrice Hessienne est définie positive :**

$$\square \quad \nabla^2 L(x^*) = \nabla^2 f(x^*) + \Sigma \mu_i \nabla^2 g_i(x^*) + \Sigma \lambda_i \nabla^2 h_i(x^*)$$

Remarque :

1. Pour un problème de maximisation remplacer le signe « + » par « – »
2. Dans certaines circonstances, les conditions de Kuhn-Tucker peuvent également être prises en tant que conditions suffisantes **si**
   - *F est une fonction concave (max problème) ou convexe (min problème)*
   - *hi,gi* sont *des fonctions convexes(concaves).*

Alors toute solution *x** qui satisfait les conditions de KKT est la solution optimale.

Un exemple des condictions de Karush-Kuhn-Tucker

Ce n'est que pour des problèmes très simples que nous pouvons utiliser les conditions Karush-Kuhn-Tucker pour résoudre un problème de programmation non linéaire. Considérez le problème suivant:

maximiser $f(x, y) = xy$
sous réserve de $x + y^2 \leq 2$
$x, y \geq 0$

Notez que la région des possibles est bornée, donc un maximum global doit exister:
Une fonction continue sur un ensemble fermé et borné contient un maximum.

Nous écrivons d'abord les contraintes comme
$g1\ (x, y) = x + y\ 2 \leq 2$, $g\ 2\ (x, y) = -x \leq 0$, $g\ 3\ (x, y) = -y \leq 0$.

Ainsi, les conditions KKT peuvent être écrites comme
$y - \lambda_1 + \lambda_2 = 0$
$x - 2y\lambda_1 + \lambda_3 = 0$
$\lambda_1(2 - x - y^2) = 0$
$\lambda_2\, x = 0$
$\lambda_3\, y = 0$
$x + y^2 \leq 2$
$x, y, \lambda_1, \lambda_2, \lambda_3 \geq 0$

Dans chacune des équations de «relâchement complémentaire» $\lambda_i\,(bi - gi\,(x1, ..., xn)) = 0$, au moins l'un des deux facteurs doit être égal à 0. Avec *n* telles conditions, il y aurait potentiellement être 2^n cas possibles à considérer. Cependant, avec un peu de réflexion, nous pourrions être en mesure de réduire cela considérablement.

Cas 1: Supposons $\lambda_1 = 0$. Alors la première condition KKT $y + \lambda_2 = 0$ et la seconde $x + \lambda_3 = 0$. Puisque chaque terme est non négatif, la seule façon qui peut arriver est si $x = y = \lambda_2 = \lambda_3 = 0$. En effet, les conditions KKT sont satisfaites lorsque $x = y = \lambda_1 = \lambda_2 = \lambda_3 = 0$ (bien que ce ne soit clairement pas un maximum local puisque $f(0, 0) = 0$ tandis que $f(x, y) > 0$ aux points à l'intérieur de la région des possibles).

Cas 2: Supposons $x + y\ 2 = 2$. Maintenant, au moins $x = 2 - y^2$ et y doit être positif.
Cas 2a: Supposons $x > 0$. Alors $\lambda_2 = 0$. La première condition KKT donne $\lambda_1 = y$. La deuxième condition KKT donne alors $x - 2y\lambda_1 + \lambda_3 = 2 - 3y^2 + \lambda_3 = 0$, donc $3y^2 = 2 + \lambda_3 > 0$, et $\lambda\ 3 = 0$. Ainsi $y = \sqrt{2} / 3$, et $x = 2 - 2/3 = 4/3$. Encore une fois toutes les conditions KKT sont remplies.
Cas 2b: Supposons $x = 0$, c'est-à-dire $y = \sqrt{2}$. Puisque $y > 0$ nous avons $\lambda_3 = 0$. A partir de la deuxième condition KKT, nous devons avoir $\lambda_1 = 0$. Mais cela nous ramène au cas 1.
Nous concluons qu'il n'y a que deux candidats pour un max local : $(0, 0)$ et $(4/3, \sqrt{2} / 3)$.
Le maximum global est à $(4/3, \sqrt{2} / 3)$.

4.4 Les méthodes de résolution numériques

Toutes les techniques de résolution des problèmes non linéaires vues sont des techniques analytiques praticables pour des équations et fonctions simples. Malheureusement, dans la réalité les fonctions et les contraintes sont plus compliquées. En effet, face à un problème d'optimisation compliqué, soit en raison du grand nombre d'inconnues et/ou soit en raison de la formulation mathématique complexe, il est difficile de pouvoir trouver une solution analytique exacte. Le recours la résolution « numérique » du problème est recommandées. La plupart des algorithmes numériques d'optimisation sont itératifs et permettent de s'approcher pas à pas vers la solution. Il existe plusieurs algorithmes numériques dont l'efficacité dépend de la nature du problème comme la descente du gradient, la méthode Newton-Raphson

Référence

1. Sylvain Perron, *cours de modélisation et optimisation*, école de gestion canada,2008.
2. Avriel, Mordecai *Nonlinear Programming: Analysis and Methods.* Dover Publishing. 2003.
3. Bertsekas, Dimitri P. *Nonlinear Programming: 2nd Edition.* Athena Scientific. 1999.