Correction de l'examen d'Optimisation Non-Linéaire

IENAC07 - 28 janvier 2008

1 Question de cours : Optimisation sur un ensemble ouvert

Soient f une fonction multivariable de $\mathbb{R}^n$ dans $\mathbb{R}$, Ω un ouvert de $\mathbb{R}^n$ et (P_Ω) le problème de minimisation de f sur Ω, $\hat{x}$ est solution de (P_Ω) si et seulement si $\hat{x} \in \Omega$ et $\hat{x}$ solution du problème de minimisation de f sur $\mathbb{R}^n$.

Cela implique que lors de la résolution d'un problème d'optimisation sous contraintes, on peut laisser de côté les contraintes définissant des ouverts (typiquement les contraintes inégalité strictes), ces dernières ne pouvant jamais saturées. En revanche il faut vérifier *a posteriori* que les solutions trouvées respectent bien ces contraintes laissées à part initialement.

2 Exercice : Théorème de la dualité

1. Le problème d'optimisation sous contraintes correspondant s'écrit :

$$\left\{ \begin{array}{c} \min\limits_{x \in \mathbb{R}^n} \sum\limits_{i=1}^n \frac{1}{x_i} \\ \sum\limits_{i=1}^n x_i^2 - 1 \leq 0 \\ \forall i \in \{1, \ldots, n\}, -x_i < 0 \end{array} \right.$$

   D'après le théorème d'optimisation sur les ensembles ouverts on sait qu'on peut négliger les contraintes inégalité strictes tant qu'on les vérifie *a posteriori*, le problème s'écrit donc :

$$P_\Omega : \left\{ \begin{array}{l} \min\limits_{x \in \mathbb{R}^n} f(x) \\ h(x) \leq 0 \end{array} \right.$$

   avec $h(x) = \sum\limits_{i=1}^n x_i^2 - 1$.

2. Afin de montrer que f est convexe on se propose de calculer son Hessien. On a :

$$\frac{\partial f}{\partial x}(x) = \begin{bmatrix} -1/x_1^2 \\ \vdots \\ -1/x_n^2 \end{bmatrix} \text{ et } \frac{\partial^2 f}{\partial x^2}(x) = \begin{bmatrix} 2/x_1^3 & & \\ & \ddots & \\ & & 2/x_n^3 \end{bmatrix}$$

   $\frac{\partial^2 f}{\partial x^2}(x)$ est donc une matrice diagonale dont les termes sont positifs si $x_i > 0$ ce qui correspond à notre contrainte stricte. Le hessien de f est donc défini positif sur le domaine admisible et donc la fonction f est strictement convexe sur le domaine admissible.

   Par ailleurs $\frac{\partial^2 h}{\partial x^2}(x) = 2I$ où I est la matrice identité de taille n, donc la contrainte h est également convexe sur le domaine admissible.

3. Pour pouvoir appliquer le théorème de la dualité, il nous faut vérifier au préalable que f et h sont convexes (question 2) et que les contraintes sont qualifiées.

   Pour montrer que les contraintes (la contrainte en fait) sont qualifiées, on peut utiliser plusieurs arguments du théorème de qualification. Ici il est immédiat de voir que h est convexe (on vient de le montrer) et d'intérieur non vide par exemple.

   On peut donc appliquer le théorème de la dualité qui nous indique que résoudre le problème P_Ω est équivalent à résoudre le problème dual :

$$\max_{\mu > 0} \min_{x \in \mathbb{R}^n} \mathcal{L}(x, \mu)$$

   où $\mathcal{L}(x, \mu)$ est le Lagrangien associé au problème P_Ω et s'écrit :

$$\mathcal{L}(x, \mu) = f(x) + \mu h(x) = \sum_{i=1}^{n} \frac{1}{x_i} + \mu \left(\sum_{i=1}^{n} x_i^2 - 1 \right)$$

   On s'intéresse au problème de minimisation sans contraintes en premier et on cherche à annuler le gradient de $\mathcal{L}$ par rapport à x. $\frac{\partial \mathcal{L}}{\partial x}(x) = \begin{bmatrix} \vdots \\ -1/x_i^2 + 2\mu x_i \\ \vdots \end{bmatrix}$

$\frac{\partial \mathcal{L}}{\partial x}(x) = 0 \Rightarrow 2\mu x_i^3 = 1$
et donc $x_i = (2\mu)^{-1/3}$ et $\mu \neq 0$
On a donc le $x(\mu)$ qui minimise le Lagrangien par rapport à x. On introduit la fonction duale $\psi(\mu) = \mathcal{L}(x(\mu), \mu)$.

$$\begin{aligned}\psi(\mu) &= \sum_{i=1}^{n} (2\mu)^{1/3} + \mu \left(\sum_{i=1}^{n} (2\mu)^{-2/3} - 1 \right) \\ &= n\mu^{1/3} \left[2^{1/3} + 2^{-2/3} \right] - \mu = \frac{3n\mu^{1/3}}{2^{2/3}} - \mu\end{aligned}$$

   On cherche à maximiser $\psi(\mu)$ par rapport à μ, on a :

$$\psi'(\mu) = \frac{n}{(2\mu)^{2/3}} - 1$$

   Donc :

$$\psi'(\mu) = 0 \Rightarrow \mu = \frac{n^{3/2}}{2}$$

   On vérifie immédiatement que $\mu > 0$ et qu'il s'agit bien d'un maximum. L'optimum est donc atteint en :

$$x = \frac{1}{\sqrt{n}}$$

   Ce point vérifie bien les contraintes inégalité strictes il s'agit donc bien de la solution de notre problème.

   On aurait pu intuiter ce résultat en faisant un dessin en dimension 2 ; il apparait alors que la surface de la fonction objectif présente des parois escarpées le long des axes et que plus on s'éloigne de l'origine en ligne droite, plus on descend sur la surface. Il s'agit donc de s'éloigner au plus de l'origine en restant loin des axes (et en restant dans le domaine). Graphiquement, cela correspond à s'éloigner selon la direction de la première bissectrice. Or la contrainte nous impose de rester dans le cercle de rayon 1, le minimum sous contraintes est donc trouvé sur la contrainte, donc en $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ pour la dimension 2. Ce raisonnement graphique se généralise facilement à la dimension n et permet de s'attendre au résultat de l'application du théorème de la dualité.

3 Problème : le problème de tir

1. L'intégration des équations du principe fondamental de la dynamique donne successivement :
   $\frac{\partial x}{\partial t} = v \cos\theta$
   $\frac{\partial y}{\partial t} = -gt + v \sin\theta$

   $x(t) = v\cos\theta t$
   $y(t) = -\frac{1}{2}gt^2 + v\sin\theta t$
   A la date finale t_f, on a :
   $x_f = v\cos\theta t_f$
   $0 = -\frac{1}{2}gt_f^2 + v\sin\theta t_f$
   On en tire la valeur de t_f :
   $t_f = \frac{x_f}{v\cos\theta} = \frac{2v\sin\theta}{g}$

   Ainsi que la relation entre v et θ définissant l'ensemble des solutions au problème de tir :
   $v^2 \sin 2\theta = gx_f$

2. La dernière relation nous indique que $v^2\sin(2\theta) > 0$ donc $v > 0$ et $\theta \in]0; \pi/2[$.

3. On minimise le temps de vol sous la contrainte inégalité $v \leq v_m$, on a donc :

$$
\begin{aligned}
f_1(v,\theta) &= \frac{2v\sin\theta}{g} \\
g(v,\theta) &= v^2\sin 2\theta - gx_f \\
h(v,\theta) &= v - v_m \\
m_1(v) &= -v \\
m_2(\theta) &= -\theta \\
m_3(\theta) &= \theta - \pi/2
\end{aligned}
$$

4. D'après le théorème d'optimisation sur les ensembles ouverts on peut chercher une solution à (P_1) sans tenir compte des contraintes m_1, m_2 et m_3 tant qu'on vérifie *a posteriori* que les solutions trouvées vérifient bien ces contraintes. On cherche donc à montrer que les contraintes h et g sont qualifiées. En posant $X = (v, \theta)$ on a :

$$\frac{\partial g}{\partial X} = \begin{bmatrix} 2v\sin 2\theta \\ 2v^2\cos 2\theta \end{bmatrix} \text{ et } \frac{\partial h}{\partial X} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

   On veut montrer que la famille des gradients des contraintes égalité et des contraintes inégalité *saturées* forme une famille libre.

Si h n'est pas saturées, alors la famille $\left\{\frac{\partial g}{\partial X}\right\}$ est libre car $v \neq 0$.

Si h est saturée alors $v = v_m$. Les deux vecteurs forment une famille libre si et seulement si v est non nul (toujours le cas dans le domaine) et $\theta \neq \pi/4$ (les autres cas qui annulent $\cos 2\theta$ ne sont pas dans le domaine). Or $\theta \neq \pi/4$ implique que $v_m^2 = gx_f$ ce qui n'est pas possible losque $v_m^2 > gx_f$. La famille de ces gradients est donc bien libre, les contraintes sont qualifiées.

Le cas $v_m^2 = gx_f$ pose problème pour la qualification car alors on ne peut plus montrer que les gradients des contraintes saturées sont linéairement indépendants. C'est d'autant plus problématique que dans ce cas, il n'y a qu'un unique point dans le domaine. En effet, pour tous les $v < v_m$ on a $v < gx_f$ et donc $\frac{gx_f}{v} > 1$ et donc $\sin 2\theta > 1$; il n'existe donc pas de θ acceptable pour les $v < v_m$. On a donc alors nécessairement $v = v_m$ et $\theta = \pi/4$. Il est difficile de montrer la qualification mais il s'avère que c'est inutile : il n'y a qu'un point dans le domaine, c'est le minimum.

On sait que le tir à $\theta = \pi/4$ est le tir qui porte le plus loin sous es hypothèses de modélisation que l'on a faites (pas de frottements). Physiquement, ce cas correspond au cas où l'énergie fournie à l'obus initialement est tout juste suffisante pour atteindre la cible, il n'y a donc *que* le tir à $\theta = \pi/4$ qui permette de respecter les contraintes (d'atteindre la cible).

De même, le cas $v_m^2 < gx_f$ correspond au cas où le domaine est vide : aucun point de vérifie les contraintes. Il s'agit du cas où l'énergie communiquée initialement au projectile est de toute façon insuffisante pour atteindre la cible, quel que soit l'angle de tir.

5. D'après la question précédente les contraintes sont qualifiées. On peut donc appliquer le théorème de Kuhn et Tucker. Le Lagrangien du problème s'écrit :

$$\mathcal{L}(v, \theta, \lambda, \mu) = \frac{2v \sin 2\theta}{g} + \lambda \left(v^2 \sin 2\theta - gx_f\right) + \mu \left(v - v_m\right)$$

Sa dérivée première vaut :

$$\frac{\partial \mathcal{L}}{\partial X} = \begin{bmatrix} \frac{2\sin\theta}{g} + 2\lambda v \sin 2\theta + \mu \\ \frac{2v\cos\theta}{g} + 2\lambda v^2 \cos 2\theta \end{bmatrix}$$

La condition d'optimalité du premier ordre nous indique que pour tout (v, θ) solution, il existe $\lambda \in \mathbb{R}$ et $\mu \in \mathbb{R}^+$ tels que :

$$\frac{2\sin\theta}{g} + 2\lambda v \sin 2\theta + \mu = 0 \tag{1}$$

$$\frac{2v\cos\theta}{g} + 2\lambda v^2 \cos 2\theta = 0 \tag{2}$$

$$v^2 \sin 2\theta - gx_f = 0 \tag{3}$$

$$\mu(v - v_m) = 0 \tag{4}$$

La quatrième équation implique de distinguer deux cas :

<u>1er cas : $\mu = 0$</u>

On a :

$$\frac{2\sin\theta}{g} = -2\lambda v \sin 2\theta \tag{5}$$

$$\frac{2v\cos\theta}{g} = -2\lambda v^2 \cos 2\theta \tag{6}$$

On a $\sin\theta$ et $\sin 2\theta \neq 0$ car $\theta \neq 0$ ou $\pi/2$ et $v \neq 0$. On peut donc diviser 6 par 5. On obtient $\tan\theta = \tan 2\theta$, ce qui implique que $\theta = 0$ ce qui est hors-domaine. Il n'existe donc pas de solution admissible pour $\mu = 0$.

2nd cas : $v = v_m$

La contrainte 3 implique : $\sin 2\theta = \dfrac{gx_f}{v_m^2}$, d'où :

$$\theta = \frac{1}{2}\arcsin\frac{gx_f}{v_m^2} \text{ ou } \frac{\pi}{2} - \frac{1}{2}\arcsin\frac{gx_f}{v_m^2}$$

Le paramètre de Lagrange vaut :

$$\lambda = \frac{-2\cos\theta}{gv\cos 2\theta}$$

Le paramètre de Kuhn et Tucker vaut :

$$\begin{aligned}
\mu &= \frac{-2\sin\theta}{g} + 4\frac{\cos\theta}{gv\cos 2\theta}v\sin 2\theta \\
&= \frac{2}{g}\left[2\cos\theta\tan 2\theta - \sin\theta\right] \\
&= \frac{2}{g}\left[2\cos\theta\frac{2\tan\theta}{1-\tan^2\theta} - \sin\theta\right] \\
&= \frac{3\sin\theta\cos^2\theta + \sin\theta}{\cos 2\theta}
\end{aligned}$$

Cette dernière écriture de μ nous permet d'affirmer que pour le θ solution qui est plus petit que $\pi/4$ on a bien $\mu \geq 0$ tandis que pour le θ solution qui appartient à l'intervalle $]\pi/4;\pi/2[$, μ est négatif. On ne garde donc que la première solution :

$$\theta = \frac{1}{2}\arcsin\frac{gx_f}{v_m^2}, \quad v = v_m$$

On pouvait intuiter ce résultat puisqu'au final le résultat de l'optimisation nous indique que pour atteindre la cible le plus vite possible, il faut lancer le projectile le plus fort possible ($v = v_m$), deux angles de tir découlent du fait de fixer la vitesse initiale, on prend l'angle le plus faible.

6. Le nouveau problème d'optimisation est :

$$(P_2) : \begin{cases} \displaystyle\min_{(v,\theta)\in\mathbb{R}^2} \frac{2v\sin\theta}{g} + \frac{1}{2}Cv^2 \\ g(v,\theta) = 0,\ h(v,\theta) \leq 0 \\ m_1(v) < 0,\ m_2(\theta) < 0,\ m_3(\theta) < 0 \end{cases}$$

Le raisonnement précédent pour la qualification des contraintes tient toujours. Le nouveau Lagrangien s'écrit :

$$\mathcal{L}(v,\theta,\lambda,\mu) = \frac{2v\sin\theta}{g} + \frac{1}{2}Cv^2 + \lambda\left(v^2\sin 2\theta - gx_f\right) + \mu\left(v - v_m\right)$$

Et sa dérivée :

$$\frac{\partial\mathcal{L}}{\partial X} = \begin{bmatrix} \frac{2\sin\theta}{g} + Cv + 2\lambda v\sin 2\theta + \mu \\ \frac{2v\cos\theta}{g} + 2\lambda v^2\cos 2\theta \end{bmatrix}$$

La condition du premier ordre s'écrit donc :

$$\begin{aligned}
\frac{2\sin\theta}{g} + Cv + 2\lambda v \sin 2\theta + \mu = 0 \\
\frac{2v\cos\theta}{g} + 2\lambda v^2 \cos 2\theta = 0 \\
v^2 \sin 2\theta - g x_f = 0 \\
\mu(v - v_m) = 0
\end{aligned}$$

Paradoxalement, l'introduction du terme quadratique - qui implique le terme en Cv de la dérivée première - rend la résolution du système plus complexe. En effet on ne peut plus écarter le cas $\mu = 0$ aussi facilement que précédemment. Le compromis énergie-temps tire la solution vers l'intérieur du domaine : on peut s'attendre à ce que, lorsque C tend vers l'infini, θ tende vers $\pi/4$ et v vers $\sqrt{g x_f}$ car c'est la solution dans le cas où on cherche à optimser l'énergie mise en jeu (minimiser la vitesse transmise). Le mélange de termes linéaires et de termes trigonométriques rend les équations 6 à 6 complexes à résoudre.

7. Cf. cours sur les résolutions numériques.

8. Il s'agit simplement d'un adaptation de l'exercice 2 du TD1. Plutôt que de résoudre le problème sous contraintes égalité et inégalité, on se propose de tenter d'inverser la relation 3 afin d'obtenir directement v en fonction de $theta$. Cette inversion peut se faire de façon formelle mais elle est peu utile pour résoudre le reste du problème car il est difficile de résoudre des équations trigonométriques. A la place, étant donné la forme de la fonction $v(\theta)$ fournie par la résolution numérique, on se propose d'admettre une petite erreur sur la contrainte égalité afin de simplifier la résolution. On choisit d'approcher la fonction $v(\theta)$ par un polynôme d'ordre 2. On note $p(\theta) = a_2\theta^2 + a_1\theta + a_0$ ce polynôme.

   On dispose d'un ensemble de points de mesure θ_j et de la valeur exacte de v pour chaque θ_j. En utilisant les notations de l'énoncé (qui correspondent à celles du TD4), on a $v(\theta_j) = G_j$. On cherche à minimiser l'erreur commise lors de l'approximation, au sens des moindres carrés. On cherche donc à minimiser sans contraintes la quantité $\sum_{j=1}^{m} (p(\theta_j) - G_j)^2$ en fonction des paramètres a_0, a_1 et a_2.

   Come indiqué dans la correction du TD1, le problème se résout très simplement en posant :

$$X = \begin{bmatrix} 1 & \theta_1 & \theta_1^2 \\ \vdots & \vdots & \vdots \\ 1 & \theta_m & \theta_m^2 \end{bmatrix}, \; G = \begin{bmatrix} G_1 \\ \vdots \\ G_m \end{bmatrix} \text{ et } a = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

   La fonction à minimiser s'écrit alors $f(a) = (Xa - G)^2 = a^T X^T X a - 2G^T X a + G^T G$.

   On pose alors $Q = 2X^T X$ et $b = -2X^T G$. On cherche alors à minimiser sans contraintes la fonction quadratique :

$$f(a) = \frac{1}{2} a^T Q a + b^T a + G^T G$$

   On montre alors (Cf correction du TD1) que Q est définie positive et la solution s'écrit :

$$a = -Q^{-1} b$$