Université Paris-Sud
M1 Ingénierie Mathématique
24 mars 2011

Examen d'Optimisation Numérique – CORRIGE

Exercice 1 (4 points)**.** Soit $E \subset \mathbb{R}^n$ l'ellipsoïde $E = \left\{ x = (x_1, \ldots, x_n) \in \mathbb{R}^n \,:\, \sum_{i=1}^n \left(\frac{x_i}{a_i}\right)^2 \leq 1 \right\}$, où les nombres $a_i > 0$ sont fixés. Soit $x \notin E$.

Démontrer qu'il existe une unique valeur de $\mu > 0$ telle que

$$\sum_{i=1}^n \frac{a_i^2 x_i^2}{(a_i^2 + \mu)^2} = 1.$$

Cette valeur sera notée $\bar{\mu}$.

Soit maintenant $\bar{x} = P_E(x)$ la projection sur E de x. Justifier que cette projection existe et est unique, et prouver qu'elle est donnée par

$$\bar{x}_i = \frac{a_i x_i}{a_i^2 + \bar{\mu}}.$$

Solution

La fonction

$$\phi(\mu) := \sum_{i=1}^n \frac{a_i^2 x_i^2}{(a_i^2 + \mu)^2}$$

est une fonction continue et strictement décroissante de la variable $\mu \in \mathbb{R}_+$. On a $\lim_{\mu \to +\infty} \phi(\mu) = 0$ et $\phi(0) = \sum_{i=1}^n x_i^2/a_i^2$. Or, $\sum_{i=1}^n x_i^2/a_i^2 > 1$ puisque $x \notin E$. Donc, par le théorème des valeurs intermédiaires il existe une valeur $\bar{\mu}$ qui donne $\phi(\bar{\mu}) = 1$ et cette valeur est unique par monotonie. *remarque : on aurait pu également expliciter l'équation en μ qui est une équation d'ordre deux, et vérifier qu'elle n'a qu'une seule solution positive, les calculs étant cependant un peu moins évidents.*

Chercher la projection de x sur E revient à minimiser $f(y) := |x - y|^2$, parmi les $y \in E$. Ce minimum existe parce que f est continue et E compact et il est unique parce que f est strictement convexe et E convexe (*on aurait pu également évoquer tout court le théorème vu en classe, comme quoi la projection sur les convexes fermés est bien définie*).

Or, $\nabla f(y) = 2(y - x)$, et c'est impossible d'avoir $\nabla f(y) = 0$ avec $y \in E$ puisque ∇f ne s'annule que en x, qui n'appartient pas à E. Donc le minimum est sur la frontière et on veut appliquer le multiplicateur de Lagrange.

La fonction $g(y) = \sum_{i=1}^n \left(\frac{y_i}{a_i}\right)^2$ a pour gradient

$$\nabla g(y) = \left(\frac{2y_1}{a_1^2}, \frac{2y_2}{a_2^2}, \ldots, \frac{2y_n}{a_n^2} \right),$$

qui ne s'annule qu'en $y = 0$. On peut alors appliquer la méthode des multiplicateurs de Lagrange puisque nous sommes intéressés à la surface $\{g = 1\}$ et sur cette surface on a $\nabla g \neq 0$.

Si y est un point de minimum on a alors $\nabla f(y) + \lambda \nabla g(y)$, ce qui signifie

$$\forall i = 1, \ldots, n \quad y_i - x_i + \lambda \frac{y_i}{a_i^2} \Rightarrow x_i a_i^2 = y_i(a_i^2 + \lambda).$$

Ceci permet de dire $y_i = \frac{x_i a_i^2}{a_i^2+\lambda}$. De plus, il faut $g(y) = 1$, ce qui signifie

$$\sum_{i=1}^{n} \frac{1}{a_i^2} \left(\frac{x_i a_i^2}{a_i^2 + \lambda} \right)^2 = 1.$$

Ceci impose $\lambda = \bar{\mu}$ si on peut prouver $\lambda \geq 0$. Pour cela, on peut soit utiliser les conditions de Kuhn-Tucket, qui nous donnent le signe du multiplicateur, ou remarquer qu'il faut $\nabla f(y) \cdot \nabla g(y) \leq 0$ puisque, pour h petit, on a $y - h\nabla g(y) \in E$ (puisque $E = \{g \leq 1\}$ et $g(y - h\nabla g(y)) = 1 - h|\nabla g(y)|^2 + o(h) < 1$) et $f(y - h\nabla g(y)) = f(y) - h\nabla f(y) \cdot \nabla g(y) + o(h) \geq f(y)$, ce qui implique $\nabla f(y) \cdot \nabla g(y) \leq 0$ et donc $\lambda \geq 0$.

Exercice 2 (6 points)**.** Résoudre le problème

$$\min\{J(f) := \int_0^{\pi} \left[\frac{1}{2} f'(t)^2 - \sin(t) f(t) \right] dt\,;\ f \in \mathcal{A}\},$$

où

$$\mathcal{A} := \{f \in C^1([0, \pi])\ :\ f(0) = f(\pi) = 0\}.$$

Indiquer la valeur minimale de J sur $\mathcal{A}$ ainsi que la ou les fonctions f la réalisant, en prouvant qu'il s'agit bien d'un minimiseur.

Solution

Posons $L(t, x, v) = \frac{1}{2}|v|^2 - \sin(t)x$. L'équation d'Euler-Lagrange de ce problème est

$$\frac{d}{dt} \frac{\partial L}{\partial v}(t, f(t), f'(t)) = \frac{\partial L}{\partial x}(t, f(t), f'(t)) \Rightarrow f''(t) = -\sin(t).$$

Le conditions au bord viennent de $f \in \mathcal{A}$, qui impose $f(0) = f(\pi) = 0$.

Les solutions de $f''(t) = -\sin(t)$ sont $f(t) = \sin(t) + At + B$ (il faut remarquer qu'il ne s'agit pas vraiment ici d'une équadiff, mais juste d'un double calcul de primitives). La condition $f(0) = 0$ impose $B = 0$ et ensuite $f(\pi) = \sin(\pi) + B\pi = 0$ impose $B = 0$.

La seule solution est donc $f(t) = \sin(t)$. C'est donc la seule fonction de $\mathcal{A}$ qui puisse minimiser J. De plus, la fonction L étant convexe par rapport à (x, v), satisfaire l'équation d'Euler-Lagrange avec les conditions au bord est aussi suffisant pour minimiser. Donc f minimise, ce que l'on pourrait vérifier également comme suit.

Soit $h \in \mathcal{A}$ arbitraire et définissons $g = h - f$. Remarquons que $g \in \mathcal{A}$ aussi (ce n'est pas vrai en général, ce qui se passe en général est que g s'annule sur les point du bord où les valeurs des fonctions de $\mathcal{A}$ sont fixées ; en général, $\mathcal{A}$ n'est pas un espace vectoriel mais seulement affine).

Calculons

$$\begin{aligned} J(f+g) &= \int_0^{\pi} \left[\frac{1}{2} f'(t)^2 + \frac{1}{2} g'(t)^2 + f'(t)g'(t) - \sin(t)f(t) - \sin(t)g(t) \right] dt \\ &= J(f) + \frac{1}{2} \int_0^{\pi} g'(t)^2 dt + \int_0^{\pi} \left[f'(t)g'(t) - \sin(t)g(t) \right] dt. \end{aligned}$$

Or, l'intégrale de $g'(t)^2$ est positive et pour l'autre on a

$$\int_0^{\pi} \left[f'(t)g'(t) - \sin(t)g(t) \right] dt = \int_0^{\pi} \left[-f''(t)g(t) - \sin(t)g(t) \right] dt = 0$$

car $f''(t) = -sint(t)$ (les termes aux bornes disparaissent puisque g s'annule en 0 et π). Donc $J(f+g) \geq J(f)$, c'est-à-dire $J(h) \geq J(f)$ et f minimise.

L'exercice demander à calculer la valeur minimale, il faut donc calculer maintenant $J(f)$. On a

$$\int_0^\pi \left[\frac{1}{2}f'(t)^2 - \sin(t)f(t)\right] dt = \int_0^\pi \left[\frac{1}{2}\cos(t)^2 - \sin(t)^2\right] dt = \frac{\pi}{4}.$$

Exercice 3 (6 points)**.** Étant donné $a \in \mathbb{R}^n$ on considère $f_a : \Omega \to \mathbb{R}$ définie par

$$f_a(x) = -\ln(1-|x|^2) + < a, x >,$$

où $\Omega = \{x \in \mathbb{R}^n \,:\, |x| < 1\}$

1. Montrer que f_a est strictement convexe sur Ω.
2. Considérer le problème de minimisation

$$(\mathcal{P}_a) \quad \min\left\{ f_a(x) \,:\, |x| \leq \frac{1}{2}, < a, x > \leq 0\right\}.$$

   (a) Résoudre $(\mathcal{P}_a)$ pour $a = 0$.
   (b) Justifier que le minimum de $(\mathcal{P}_a)$ existe également pour $a \neq 0$.
   (c) En supposant $a \neq 0$ et en appelant $\bar{x}$ le minimiseur, montrer que $\bar{x}$ et a sont colinéaires.
   (d) En déduire que $< a, \bar{x} > < 0$, et qu'on peut donc ignorer cette contrainte dans les conditions d'optimalité données par les multiplicateurs de Lagrange. Expliciter alors ces conditions sous la forme d'un système d'équations.
   (e) Trouver le minimiseur $\bar{x}$.

Solution

La fonction $x \mapsto < a, x >$ est linéaire et donc convexe (mais non strictement). La fonction $x \mapsto \ln(1-|x|^2)$ est strictement concave sur son domaine de définition, qui inclut Ω, en tant que composition de deux fonctions strictement concaves, dont celle à l'extérieur croissante (*la seule concavité des deux fonctions ne suffit pas, imaginez par exemple de composer une fonction concave avec la fonction $x \mapsto -x$, qui est concave mais change le signe de la concavité du résultat*). Pour vérifier cela, prenez $f : \mathbb{R}^n \to \mathbb{R}$ concave et $g : \mathbb{R} \to \mathbb{R}$ concave croissante. On a alors

$$g(f((1-t)x + ty)) > g((1-t)f(x) + tf(y)) > (1-t)g(f(x)) + tg(f(y)),$$

la première inégalité étant justifiée par la concavité de f et la croissance de g, la deuxième par la concavité de g.

Globalement, f_a résulte strictement convexe.

Pour le cas $a = 0$ il suffit de remarquer qu'on veut minimiser $-\ln(1-|x|^2)$. Donc maximiser $1-|x|^2$. Donc minimiser $|x|^2$. La seule solution est donc $x = 0$.

Pour $a \neq 0$ le minimum existe néanmoins puisque l'ensemble $\{x \,:\, |x| \leq \frac{1}{2}, < a, x > \leq 0\}$ est compact et f est continue.

Pour démontrer que a et $\bar{x}$ sont colinéaires il suffit de remarquer que, parmi tous les vecteurs qui ont la même norme que $\bar{x}$, il vaut mieux choisir celui qui a le même direction de $-a$. En effet, changer la direction ne change rien à la partie avec le logarithme, qui ne dépend que de $|x|$, mais peut changer le produit scalaire. Et le produit scalaire est minimale quand les directions sont opposées (et maximales quand elles sont égales).

On vent de dire que $\bar{x} = -\lambda a$ avec $\lambda \geq 0$. Il faut juste démontrer $\lambda \neq 0$. Si c'était le cas, on aurait $\bar{x} = 0$. Or, si on prend le vecteur $y = -ta$ ce vecteur, pour $t > 0$ suffisamment petit, est admissible (il satisfait $|y| \leq 1/2$ et $< a, y > = -t|a|^2 \leq 0$) et $f_a(y) = t^2|a|^2 - t|a|^2 + o(t^2)$, ce qui donne $f(y) < 0 = f(0)$ pour t petit. Donc 0 ne peut pas être un minimum pour $a \neq 0$. La contrainte $< a, x > \leq 0$ n'étant pas saturée par l'optimum $\bar{x}$, on peut alors dire que

– soit $|\bar{x}| < 1/2$ et $\bar{x}$ est à l'intérieur, et alors $\nabla f_a(\bar{x}) = 0$, c'est-à-dire

$$\frac{2x}{1-|x|^2} + a = 0,$$

– soit $|\bar{x}| = 1/2$ et $\bar{x}$ est sur le bord, et alors $\nabla f_a(\bar{x}) + \nabla g(\bar{x}) = 0$, où g est une fonction qui exprime la contrainte, qu'on peut écrire $|x|^2 = \frac{1}{4}$ et prendre donc $g(x) = |x|^2$; on a donc

$$\begin{cases} \frac{2x}{1-|x|^2} + a + 2\lambda x = 0, \\ |x|^2 = \frac{1}{4}. \end{cases}$$

éventuellement on peut remplacer la valeur de $|x|^2$ dans le dénominateur de ∇f_a.

Pour trouver le minimiseur on peut se contenter de le chercher parmi les vecteurs du type $t\frac{a}{|a|}$ avec $t \in [0, 1/2]$, en minimisant donc la fonction $h(t) = f_a(t\frac{a}{|a|}) = -\ln(1-t^2) - t|a|$. Cette fonction est une fonction convexe dont la dérivée vaut $h'(t) = \frac{2t}{1-t^2} - |a|$. On sait déjà que cette fonction n'est pas minimisée en $t = 0$, elle est donc minimisée soit en $t = 1/2$, soit en un point $t \in]0, 1/2[$ tel que $\frac{2t}{1-t^2} = |a|$. Ce deuxième cas est vérifié quand

$$|a|t^2 + 2t - |a| = 0 \Rightarrow t = \frac{-1 \pm \sqrt{1+|a|^2}}{|a|},$$

de ces deux solutions seule $\frac{-1+\sqrt{1+|a|^2}}{|a|}$ est positive. Il y a donc deux cas

– soit $\frac{-1+\sqrt{1+|a|^2}}{|a|} < 1/2$, ce qui correspond à $|a| < \frac{4}{3}$, et alors le minimiseur $\bar{x}$ est $\frac{-1+\sqrt{1+|a|^2}}{|a|^2}a$,

– soit $\frac{-1+\sqrt{1+|a|^2}}{|a|} \geq 1/2$, ce qui correspond à $|a| \geq \frac{4}{3}$, et alors $\bar{x} = \frac{1}{2}\frac{a}{|a|}$

Exercice 4 (4 points)**.** Soient B_+ et B_- les boules fermées dans $\mathbb{R}^2$ de rayon 1 centrées en $(1, 0)$ et $(-1, 0)$, respectivement, et Q le carré $[-1, 1] \times [-1, 1]$. Soit $K = B_+ \cup B_- \cup Q$. Dessiner cet ensemble K, qui est un convexe ; écrire une formule pour la projection sur K, qu'on appellera P_K. La formule doit être de la forme $P_K(x, y) = \dots$ avec des calculs explicites et, éventuellement, des cas à distinguer.

Solution

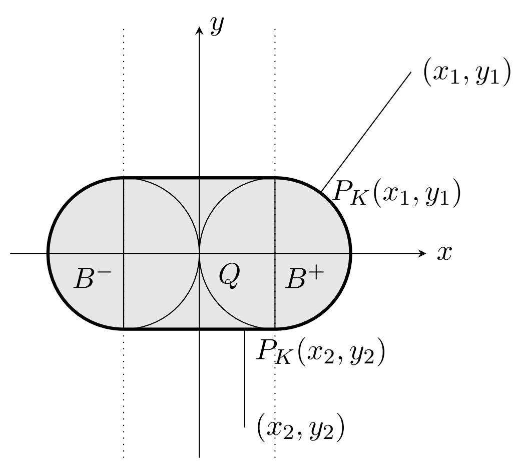

Figure 1 – Le convexe K

Le dessin du convexe K est donné en figure. Il est facile de de rendre compte (éventuellement) en distinguant plusieurs cas) que la projection d'un point (x, y) sur K est soit la projection sur B^+ (si $x \geq 1$), soit sur B^- (si $x \leq -1$), soit sur Q (si $-1 \leq x \leq 1$).

En connaissant les projections sur des rectangles et des disques on peut conclure

$$P_K(x,y) = \begin{cases} \left(1+\frac{x-1}{\sqrt{(x-1)^2+y^2}}, \frac{y}{\sqrt{(x-1)^2+y^2}}\right) & \text{si } x \geq 1 \text{ et } \sqrt{(x-1)^2+y^2} \geq 1, \\ (x,y) & \text{si } x \geq 1 \text{ et } \sqrt{(x-1)^2+y^2} < 1, \\ (x,1) & \text{si } -1 < x < 1 \text{ et } y \geq 1, \\ (x,y) & \text{si } -1 < x < 1 \text{ et } -1 < y < 1, \\ (x,-1) & \text{si } -1 < x < 1 \text{ et } y < -1, \\ (x,y) & \text{si } x \leq -1 \text{ et } \sqrt{(x+1)^2+y^2} < 1, \\ \left(-1+\frac{x+1}{\sqrt{(x+1)^2+y^2}}, \frac{y}{\sqrt{(x+1)^2+y^2}}\right) & \text{si } x \leq -1 \text{ et } \sqrt{(x+1)^2+y^2} \geq 1. \end{cases}$$

Cette même projection peut également s'écrire à l'aide d'autres formules équivalentes, notamment en utilisant des expressions avec min et max. Par exemple, définissons

$$R(x,y) = \min\{\sqrt{(x-1)^2+y^2}, \sqrt{(x+1)^2+y^2}\},$$

qui correspond à la plus petite des deux distances aux points $(0,1)$ et $(0,-1)$. On a alors

$$P_K(x,y) = \begin{cases} (1,0) + \frac{(x-1,y)}{\max\{R(x,y),1\}} & \text{si } x \geq 1 \\ (x, \min\{1, \max\{y,-1\}\})) & \text{si } -1 < x < 1 \\ (-1,0) + \frac{(x+1,y)}{\max\{R(x,y),1\}} & \text{si } x \leq -1, \end{cases}$$

ce qui est exactement la même chose.

Exercice 5 (4 points)**.** Une quantité variable est observée aux instants t_i et donne lieu à des observations k_i. Ces observations devraient découler d'une fonction f, qui est supposé être de type sinusoïdal (la fréquence étant connue aussi) et on veut trouver la fonction qui approche mieux les observations au sens des moindres carrés. On considère toutes les fonctions de la forme $f(t) = A + B\sin(\omega t - \phi)$, ω étant une fréquence fixée. Le but est de trouver A, B et ϕ qui minimisent

$$J(A,B,\phi) := \sum_{i=1}^{n} (A + B\sin(\omega t_i - \phi) - k_i)^2.$$

S'agit-il d'un problème d'optimisation convexe (c'est-à-dire : J est-elle une fonction convexe de 3 variables) ? On le réécrit sous la forme suivante : on cherche maintenant une fonction g donnée par $g(t) = A + B\sin(\omega t) + C\cos(\omega t)$ qui minimise

$$H(A,B,C) := \sum_{i=1}^{n} (A + B\sin(\omega t_i) + C\cos(\omega t_i) - k_i)^2.$$

Justifier que ce problème est équivalent au précédent, en montrant que la classe des fonctions g de ce type coïncide avec la classe des fonctions f introduite précédemment. S'agit-il maintenant d'un problème convexe ?

Réécrire ce dernier problème comme un problème d'optimisation quadratique, en trouvant une matrice symétrique M, un vecteur v et une constante c qui permettent de l'écrire sous la forme

$$\min \quad \frac{1}{2} < Mx, x > - < v, x > + c.$$

Suggérer un algorithme qui trouve sa solution exactement en un petit nombre d'itérations, et estimer ce nombre.

Solution

La fonction J n'est pas convexe puisqu'elle est périodique mais non constante par rapport à sa troisième variable ϕ.

L'équivalence entre la minimisation de J et de H vient du fait que toute fonction s'exprimant comme $B\sin(\omega t+\phi)$ pour $(B,\phi)\in\mathbb{R}^2$ s'exprime également comme combinaison linéaire de $\sin(\omega t)$ et $\cos(\omega t)$, en utilisant l'égalité $B\sin(\omega t+\phi) = B\cos\phi\sin(\omega t) + B\sin\phi\cos(\omega t)$. Cette transformation est également inversible, puisque $B'\sin(\omega t)+C'\cos(\omega t) = \sqrt{(B')^2+(C')^2}\sin(\omega t+\phi)$ pour $\phi = \arctan(C'/B')$ (et $\phi = \pi/2$ si $B' = 0$).

La fonction H est maintenant convexe puisque quadratique et définie positive. Elle peut également s'expliciter comme

$$
\begin{aligned}
H(A,B,C) \;=\;& nA^2 + B^2\sum_{i=1}^n \sin(\omega t_i)^2 + C^2\sum_{i=1}^n \cos(\omega t_i)^2 \\
&+2AB\sum_{i=1}^n \sin(\omega t_i) + 2AC\sum_{i=1}^n \cos(\omega t_i) + 2BC\sum_{i=1}^n \sin(\omega t_i)\cos(\omega t_i) \\
&+2A\sum_{i=1}^n k_i + 2B\sum_{i=1}^n \sin(\omega t_i)k_i + 2C\sum_{i=1}^n \cos(\omega t_i)k_i + \sum_{i=1}^n k_i^2,
\end{aligned}
$$

ce qui donne une matrice M de taille 3×3 de la forme

$$
M = 2\begin{pmatrix}
n & \sum_{i=1}^n \sin(\omega t_i) & \sum_{i=1}^n \cos(\omega t_i) \\
\sum_{i=1}^n \sin(\omega t_i) & \sum_{i=1}^n \sin(\omega t_i)^2 & \sum_{i=1}^n \sin(\omega t_i)\cos(\omega t_i) \\
\sum_{i=1}^n \cos(\omega t_i) & \sum_{i=1}^n \sin(\omega t_i)\cos(\omega t_i) & \sum_{i=1}^n \cos(\omega t_i)^2
\end{pmatrix},
$$

un vecteur v de la forme

$$
v = -2\left(\sum_{i=1}^n k_i, \sum_{i=1}^n \sin(\omega t_i)k_i, \sum_{i=1}^n \cos(\omega t_i)k_i\right),
$$

et une constante

$$
c = \sum_{i=1}^n k_i^2.
$$

L'algorithme du gradient conjugué trouverait la solution *exacte* de ce problème en 3 itérations (la dimension de l'inconnue étant ici 3).