Devoir Maison d'Optimisation Numérique

Corrigé

essayez de le faire en deux heures max

Vrai ou faux 1 (4 points). Pour chacune des affirmations suivantes, dire si elle est vraie ou fausse. Pour chaque réponse, vous pouvez choisir de la justifier (et dans ce cas elle vous rapportera un score compris entre 0 et 1) ou pas (et alors une réponse erronée donnera $-0,5$ points, une réponse correcte $+0,5$). Pour justifier une réponse il faut soit donner un contre-exemple, soit s'appuyer sur des résultats vus en cours.

1. Toute fonction $f : \mathbb{R}^n \to \mathbb{R}$ strictement convexe admet un minimum.
2. Si $f : \mathbb{R} \to \mathbb{R}$ est une fonction C^1 qui s'annule en un point $\bar{x}$ et avec dérivée f' strictement positive partout, alors l'algorithme de Newton $x_{k+1} = x_k - f(x_k)/f'(x_k)$ converge vers $\bar{x}$ quelle que soit son initialisation.
3. Le sous-différentiel $\partial f(x)$ d'une fonction convexe $f : \mathbb{R} \to \mathbb{R}$ en un point x n'est jamais vide.
4. Le sous-différentiel $\partial f(x)$ d'une fonction $f \in C^1(\mathbb{R}^n)$ en un point x n'est jamais vide.

Solution

1. FAUX : pensez par exemple à $f(x) = e^x$.
2. FAUX : il y a plein d'exemples, comme on a vu en classe (par exemple pour $f(x) = \frac{x}{1+|x|}$, si on prend $x_0 = 1$ on a $f(x_0) = \frac{1}{2}$ et $f'(x_0) = \frac{1}{4}$, d'où $x_1 = -1$ mais $f(x_1) = -\frac{1}{2}$ et $f'(x_1) = \frac{1}{4}$, donc $x_2 = 1$ et on a un loop infini qui ne converge pas).
3. VRAI : il contient toujours l'intervalle $[f'_g(x), f'_d(x)]$
4. FAUX : si f est C^1 le sous-différentiel ne peut contenir que le gradient $\nabla f(x)$ mais, si par exemple f est concave plutôt que convexe, le graphe de f reste *au dessous* du graphe de la tangente, et pas au dessus.

Exercice 2 (5 points). Considérer la fonctionnelle

$$J(f) := \int_0^1 \left[\frac{1}{2}f'(t)^2 + \frac{1}{2}f(t)^2 + tf(t)\right] dt.$$

Écrire l'équation d'Euler-Lagrange, complétée par ses conditions aux bords, associée à la minimisation de J sur l'ensemble des fonctions régulières s'annulant en $t = 0$.

Trouver la solution de cette équation en la cherchant parmi les fonctions du type $f(t) = Ae^t + Be^{-t} + Ct + D$. Nous allons appeler $\bar{f}$ cette solution.

En calculant $J(\bar{f} + g) - J(\bar{f})$ (en développant les carrés et/ou en intégrant par partie), montrer que $\bar{f}$ minimise J dans l'ensemble

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

Solution On a dans ce cas $L(t,x,v) = \frac{1}{2}v^2 + \frac{1}{2}x^2 + tx$, d'où $\partial L/\partial x = x + t$ et $\partial L/\partial v = v$. L'équation d'Euler-Lagrange est alors

$$\frac{d}{dt}f'(t) = f(t) + t, \quad \text{c'est-à-dire} \quad f''(t) = f(t) + t.$$

C'est une équation d'ordre deux, elle a besoin de deux conditions au bord. Une est $f(0) = 0$. Par contre $f(1)$ n'est pas fixé, et on n'a pas le droit de le fixer. Il signifie qu'on optimise en fait une quantité du type

$$\int L(t, f(t), f'(t))dt + G(f(1)),$$

et dans ce cas on sait que la condition manquante est obtenue en imposant la condition de transversalité $\partial L/\partial v(1, f(1), f'(1)) = -G'(f(1))$. Dans notre cas on ne voit pas G, ce qui signifie qu'elle est nulle. On doit donc imposer $f'(1) = 0$ (condition de Neumann, typique).

On cherche donc à résoudre

$$\begin{cases} f''(t) = f(t) + t \\ f(0) = 0 \\ f'(1) = 0 \end{cases} .$$

On nous dit de chercher une solution du type $f(t) = Ae^t + Be^{-t} + Ct + D$ (puisque les exponentielles e^t et e^{-t} sont une base des solutions de l'équation homogène associée). On calcule

$$f'(t) = Ae^t - Be^{-t} + C, \quad f''(t) = Ae^t + Be^{-t}, \quad f''(t) - f(t) - t = -Ct - D - t,$$

d'où on doit imposer $C = -1$ et $D = 0$ (pour l'équation), $A+B = 0$ (pour $f(0) = 0$) et $Ae-Be^{-1}+C = 0$ (pour l'autre condition), ce qui donne

$$A = \frac{1}{e + e^{-1}}, \; B = -\frac{1}{e + e^{-1}}, \; C = -1, \; D = 0$$

et donc

$$\bar{f}(t) = \frac{1}{e + e^{-1}} e^t - \frac{1}{e + e^{-1}} e^{-t} - t.$$

Enfin, nous pouvons prendre un g arbitraire tel que $\bar{f} + g \in \mathcal{A}$, développer et obtenir

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

L'avant dernière intégrale est positive puisque elle contient des carrés, et la dernière est nulle. En effet, en intégrant par partie, on trouve bien

$$\int_0^1 \left[\bar{f}'(t) g'(t) + \bar{f}(t) g(t) + tg(t) \right] dt = \int_0^1 \left[-\bar{f}''(t) g(t) + \bar{f}(t) g(t) + tg(t) \right] dt + \left[f'(t) g(t) \right]_{t=0}^{t=1},$$

et l'intégrale est nulle à cause de $\bar{f}''(t) = \bar{f}(t) + t$ et les termes aux bornes sont nulles aussi parce que $g(0) = 0$ et $f'(1) = 0$. Ceci démontre que $J(\bar{f} + g) \geq J(\bar{f})$ et que $\bar{f}$ est minimal.

Exercice 3 (5 points)**.** Soit $C \subset \mathbb{R}^2$ le triangle $C := \{(x, y) \in \mathbb{R}^2 \,:\, x \geq 0, y \geq 0, x + y \leq 1\}$. Soit f la fonction donnée par $f(x, y) = -x - 2y - 2xy + \frac{1}{2}x^2 + \frac{1}{2}y^2$.

1. f est-elle convexe ? concave ?
2. Démontrer que tout minimum o maximum local de f sur C se trouve sur la frontière de C.
3. Démontrer que f admet bien un minimum et un maximum sur C.
4. Trouver le minimum de f sur C.
5. Trouver le maximum de f sur C.

Solution

1. f ,nest ni convexe ni concave, puisque elle est C^2 et sa matrice Hessienne est

$$\begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix},$$

qui a déterminant égal à $-3 < 0$, ce qui implique que ses valeurs propres sont de signe opposé.

2. Si il y avait un minimum ou maximum local à l'intérieur il devrait annuler ∇f ; or, $\nabla f(x,y) = (-1-2y+x, -2-2x+y) = 0$ seulement pour $(x,y) = (-\frac{5}{3}, -\frac{4}{3}) \notin C$ donc il n'est pas possible que f ait un minimum ou un maximum local à l'intérieur.
3. f est une fonction continue et C est fermé borné, donc compact.
4. Faisons une liste de candidats à l'optimisation : il n'y a pas de points à l'intérieur, mais il pourrait y avoir sur la frontière. Cette frontière est composée de 6 parties : les trois sommets $(0,0), (1,0)$ et $(0,1)$, et les trois côtés $(0,y)$ avec $y \in]0,1[$, $(x,0)$ avec $\in]0,1[$ et (x,y) avec $x, y \in]0,1[$ et $x+y=1$. Sur ce dernier côté nous pouvons appliquer le multiplicateur de Lagrange, en imposant $\nabla f = \lambda(1,1)$, où $(1,1)$ est le gradient de la fonction $x+y$. On trouve alors $-1-2y+x = \lambda = -2-2x+y$, d'où
$$\begin{cases} -1-2y+x = -2-2x+y \\ x+y=1, \end{cases},$$
ce qui donne $(x,y) = (\frac{1}{3}, \frac{2}{3})$. Pour les autres côtés il n'est pas nécessaire de faire de multiplicateurs de Lagrange, il suffit d'étudier $f(0,y) = -2y + \frac{1}{2}y^2$, dont la dérivée s'annulle pour $y=2$, c'est-à-dire hors de l'intervalle qui nous intéresse, et $f(x,0) = -x + \frac{1}{2}x^2$, dont la dérivée s'annulle pour $x=1$, ce qui est aussi hors de l'intervalle d'intérêt (c'est sur son bord). Les points intéressants sont alors $(0,0), (1,0), (0,1)$ et $(\frac{1}{3}, \frac{2}{3})$ et on a
$$f(0,0) = 0, \quad f(1,0) = -\frac{1}{2}, \quad f(0,1) = -\frac{3}{2}, \quad f(\frac{2}{3}, \frac{1}{3}) = -\frac{11}{6}.$$
Le minimum est alors réalisé en $(\frac{1}{3}, \frac{2}{3})$ et vaut $-\frac{11}{6}$.
5. La même analyse permet de dire que le maximum est réalisé en $(0,0)$ et vaut 0.

Exercice 4 (4 points)**.** Soit $f : \mathbb{R} \to \mathbb{R}$ la fonction définie par $f(x) = |x| + x^2$. Dire si f est convexe et calculer f^* et f^{**}.

Répondre aux mêmes questions concernant la fonction $\tilde{f} : \mathbb{R}^n \to$ définie par $\tilde{f}(x) = f(|x|)$ (ici $|x|$ désigne la norme du vecteur $x \in \mathbb{R}^n$).

Solution f est convexe puisque somme de fonctions convexes. En particulier cela implique $f^{**} = f$. Pour f^*, on peut écrire
$$f^*(y) := \sup_x \; xy - f(x) = \sup_x \; xy - |x| - x^2.$$
Remarquons d'abord que pour tout y fixé ce sup est atteint (puisque $xy - |x| - x^2 \to -\infty$ quand $x \to \pm\infty$) et qu'il est atteint en un point du même signe que y (sinon $-x$ donnerait un résultat meilleure). Considérons d'abord $y \geq 0$. Dans ce cas là on cherche un max sur les positifs, et on cherche donc à maximiser $x(y-1) - x^2$. La dérivée s'annule pour $y - 1 = 2x$, ce qui n'est possible que pour $y \geq 1$ (sinon on n'a pas $x \geq 0$). Si $0 \leq y < 1$ on a donc une dérivée strictement négatif, le maximum est donc réalisé en $x = 0$, et vaut 0. Pour $y \geq 1$ le maximum est réalisé en $x = (y-1)/2$, ce qui donne une valeur de $\frac{(y-1)^2}{4}$. On peut voir que pour $y < 0$ on a une valeur nulle pour $-1 < y < 0$ et égale à $\frac{(-y-1)^2}{4}$ pour $y \leq -1$. On peut résumer le tout en disant que
$$f^*(y) = \frac{(|y|-1)_+^2}{4}.$$
Quant à la fonction définie sur tout $\mathbb{R}^n$, elle est encore convexe comme somme de fonctions convexes. Pour calculer f^* on remarque que pour maximiser $x \cdot y - |x| - x^2$, parmi tous les vecteurs de même norme, il vaut mieux prendre celui qui est parallèle à y et dans la même direction. Disons $x = \lambda \frac{y}{|y|}$ avec $\lambda \geq 0$. Et le problème revient ensuite au même que tout à l'heure, puisqu'on doit maximiser $\lambda(|y|-1) - \lambda^2$, qui est la même fonction d'une variable qu'avant. La réponse est alors encore
$$f^*(y) = \frac{(|y|-1)_+^2}{4}.$$

Exercice 5 (6 points)**.** Considérons l'ensemble $K = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \ : \ |x_1|^3 + \cdots + |x_n|^3 \le 1\}$ et $p = (p_1, \ldots, p_n) \notin K$ un point à l'extérieur de K. Décrire de manière détaillée et explicite au moins une méthode numérique pour calculer la projection de p sur K et justifier sa convergence.

Cette projection pourrait se configurer comme un problème d'optimisation d'une fonction convexe (la distance à p, ou le carré de cette distance) sous contrainte convexe (K étant convexe). Pourquoi ne serait-il pas raisonnable *du tout* de considérer un algorithme de gradient projeté pour répondre à la question précédente ?

Solution S'agissant d'une optimisation sous contrainte, un algorithme adapté est, par exemple, celui d'Uzawa, dont la convergence a été prouvée en classe. L'algorithme du gradient projeté par contre est complètement inadapté parce qu'il demanderait à utiliser la projection sur K, ce qui est exactement ce qu'on cherche à calculer ! !

Pour minimiser $\frac{1}{2}|x-p|^2 \ : \ x \in K = \{x \ : \ g(x) \le 1\}$, l'algorithme d'Uzawa utilise la dualité, et cherche à résoudre

$$\max_{\lambda \ge 0} \ \min_x \frac{1}{2}|x-p|^2 + \lambda(g(x) - 1).$$

Si on appelle $F(\lambda) = \min_x \frac{1}{2}|x-p|^2 + \lambda(g(x)-1)$, l'algorithme d'Uzawa est un algorithme de gradient projeté pour la fonction concave F (qu'on cherche à maximiser), soumise à la contrainte $\lambda \ge 0$ (qui est simple à gérer). On sait que $F'(\lambda) = g(x_\lambda) - 1$ où x_λ est le x optimal correspondant à λ, celui qui minimise $\frac{1}{2}|x-p|^2 + \lambda(g(x) - 1)$. Ce x_λ peut être trouvé en imposant $x - p + \lambda \nabla g(x) = 0$. La suite des λ produite par l'algorithme doit satisfaire $\lambda_{k+1} = (\lambda_k + \rho(g(x_{\lambda_k}) - 1))_+$, où $\rho > 0$ est le pas de l'algorithme de gradient projeté. Ce ρ doit être suffisamment petit, et éventuellement variable.

Explicitons le calcul de x_λ : pour chaque composante i, on doit imposer $x_i - p_i + \lambda \partial_{x_i} g(x) = 0$. Comme $\partial_{x_i} g(x) = \pm 3x_i^2$ (le signe étant celui de x_i, il est nécessaire que x_i ait le même signe de p_i. Donc pour $p_i > 0$ on cherche $x_i > 0$ tel que $x_i - p_i + \lambda x_i^2 = 0$ et pour $p_i < 0$ on cherche $x_i < 0$ tel que $x_i - p_i - \lambda x_i^2 = 0$, ce qui donne

$$x_i = \frac{-1 + \sqrt{1 + 12\lambda p_i}}{6\lambda} \ \text{ si } p_i \ge 0, \quad x_i = \frac{1 - \sqrt{1 - 12\lambda p_i}}{6\lambda} \ \text{ si } p_i < 0.$$

L'algorithme itératif est donc obtenu en partant d'un couple $(x^{(0)}, \lambda_0)$ quelconque et prenant ensuite

$$x^{(k+1)} := \text{ donné par la formule précédente avec } \lambda = \lambda_k, \quad \lambda_{k+1} = \Big(\lambda_k + \rho(g(x^{(k+1)}) - 1)\Big)_+.$$

La suite $x^{(k)}$ si obtenue converge alors vers la projection de p sur K.