TRAVAUX PRATIQUES : OPTIMISATION AVEC CONTRAINTES

Problème 1

Projection orthogonale sur un hyperplan

Soient $d \in \mathbb{R}^n$ un vecteur non nul, $a \in \mathbb{R}$, et l'hyperplan H de $\mathbb{R}^n$:

$$H = \{x \in \mathbb{R}^n \mid \langle d, x \rangle = a\}$$

Donné un point $u_0 \in \mathbb{R}^n$ on cherche à déterminer son projeté orthogonal u sur H.

1) - Observer que u est solution du problème de minimisation convexe :

$$\min_{u \in H} \frac{1}{2}\|u - u_0\|^2$$

Justifier de l'existence d'une unique solution, que l'on déterminera. Quelle est la distance de u_0 à H ?

2) - Application : on considère dans $\mathbb{R}^3$, le plan H d'équation $x + 2y - z = 2$. Déterminer sous matlab les coordonnées du projeté othogonal de $u_0 = (2, 3, 0)$ sur H par 3 méthodes :

a) - En n'utilisant que les opérateurs `+,-,*,/,'`.
b) - En utilisant la fonction `inv()`.
c) - En utilisant la fonction `quadprog()`.

Quelle est la distance de u_0 à H ?

Problème 2

Minimum avec contraintes d'une fonction définie avec des valeurs absolues

La fonction f est définie sur $\mathbb{R}^2$ par :

$$f(x, y) = |x - y^2 + 1| + |2x^2 - 3y - 2| + |x^2 - y + 1|$$

1) - Justifier de l'existence d'un extremum global pour f.

2) - Tracer la surface d'équation $z = f(x, y)$ pour $-2 \leqslant x \leqslant 2$ et $-2 \leqslant y \leqslant 2$

```
[X,Y]=meshgrid(-2:0.05:2,-2:0.05:2);
f=abs(X-Y.^2+1)+abs(2*X.^2-3*Y-2)+abs(X.^2-Y+1);
mesh(X,Y,f)
```

Afin de situer plus précisement les points critiques de f, on pourra tracer les courbes équipotentielles de cette fonction :

```
figure,contour(X,Y,f,50)
```

3) - Chercher les minimums de f sur le carré $\mathcal{D} = [-2, 2] \times [-2, 2]$. On utilisera la fonction Matlab `fmincon`. On choisira différents points initiaux pour l'algorithme utilisé, en s'aidant du tracé des courbes équipotentielles. On pourra aussi modifier le domaine du tracé afin de visualiser plus finement certaines régions.

Problème 3

Méthode des trapèzes optimale

Les résultats de cet exercice sont effectivement utilisés en électronique. D'un point de vue plus mathématique, on pourrait intituler ce problème : "méthode des trapèzes optimale". La méthode des trapèzes permet le calcul approché d'une intégrale définie sur un intervalle fermé $[a, b]$. La subdivision de cet intervalle est constituée, en général, de points régulièrement espacés. Le problème posé est le suivant : pour un nombre de points donnés dans la subdivision de $[a, b]$ existe–t–il une répartition de ces points qui minimise l'erreur d'approximation de l'intégrale ?

1) On suppose la fonction à intégrer continue. Que répondre au problème posé ?
On suppose f convexe ou concave et 2 fois différentiable, quelles conditions nécessaires du 1^{er} et 2^{nd} ordre doit vérifier une solution du problème ?

Nous traiterons ce problème dans le cas particulier de la fonction f définie sur $[0, \frac{\pi}{2}]$ par $f(x) = \sin x$.

2) - *Cas de la subdivision comportant un seul point.* Ce cas est à traiter "à la main". Trouver la valeur de α, unique point de la subdivision qui minimise l'erreur $\mathcal{A}$ lorsqu'on calcule une valeur approchée de $\int_0^{\frac{\pi}{2}} \sin x \; dx$ par la méthode des trapèzes avec un seul point dans la subdivision de $[a, b]$.

3) - *Cas de la subdivision comportant deux points.* Résoudre le problème en utilisant la fonction `fmincon`. On vérifera que l'on obtient bien un minimum.

Problème 4

Un cascadeur de masse m saute dans un gouffre. Il est relié à deux élastiques E_1 et E_2, de raideurs respectives k_1 et k_2, de longueurs à vide respectives ℓ_1 et ℓ_2. Les points d'ancrage des élastiques sont distants de L. (voir dessin ci–dessous)

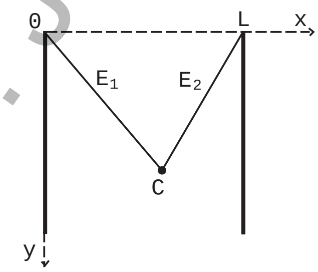

Figure 1: Position d'équilibre du cascadeur

On cherche la position d'équilibre du cascadeur.
1) - Formaliser ce problème par une recherche de minimum. Justifier de l'existence d'un minimum.
2) - Déterminer la position d'équilibre du cascadeur sous matlab à l'aide de la fonction `fmincon`.
Indication : à l'équilibre, l'énergie potentielle du cascadeur est minimale. De plus, le gouffre est suffisamment profond !
On prendra, comme valeurs numériques :

$k_1 = 15 \quad ; \quad l_1 = 25 \quad ; \quad k_2 = 30 \quad ; \quad l_2 = 40 \quad ; \quad L = 10 \quad ; \quad m = 70 \quad ; \quad g = 10.$

Correction

Problème 1

1) - Le projeté orthogonal u de u_0 sur H a la propriété (bien connue) d'être l'élément $u \in H$ minimisant la distance $\|u-u_0\|$ et donc aussi $\frac{1}{2}\|u-u_0\|^2$ (car $x \longrightarrow x^2$ est croissante sur $\mathbb{R}_+$). Posons $u_0 = (x_0, y_0, z_0)$. L'application $u = (x,y,z) \longrightarrow f(u) = \frac{1}{2}((x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2)$ est une fonction quadratique et la contrainte $\varphi(u) = \langle d, u\rangle - a$ est affine. La matrice Hessienne A de f est la matrice identité, et A est donc définie positive. On sait alors que f admet un unique minimum global u caractérisé par le système d'équations :

$$\begin{cases} \nabla_x \mathcal{L}(u,\lambda) = \overrightarrow{0} \\ \langle d, u\rangle = a \end{cases} \iff \begin{cases} u - u_0 + \lambda d = \overrightarrow{0} \\ \langle d, u\rangle = a \end{cases}$$

où λ désigne le multiplicateur de Lagrange.

Il découle de la première équation que :

$$\langle u-u_0, d\rangle + \lambda \|d\|^2 = 0$$

en appliquant la seconde équation on détermine λ :

$$\lambda = \frac{\langle d, u_0\rangle - a}{\|d\|^2}$$

et en remplaçant dans la première équation on détermine u :

$$u = u_0 - \frac{\langle d, u_0\rangle - a}{\|d\|^2} d$$

la distance de u_0 à H est :

$$dist(u_0, H) = \langle u-u_0, u-u_0\rangle^{\frac{1}{2}} = \sqrt{\left(\frac{\langle d, u_0\rangle - a}{\|d\|^2}\right)^2 \|d\|^2} = \frac{|\langle d, u_0\rangle - a|}{\|d\|}$$

2) a) Avec la formule obtenue en 1), il suffit de saisir le code :

```
>> a=2; d=[1 2 -1]'; u0=[2 3 0]';
>> u=u0-(d'*u0-a)/(d'*d)*d
```

2) b) f est la fonction quadratique strictement convexe $f(u) = \frac{1}{2}{}^t u.I.u - {}^t u_0.u + \frac{\|u_0\|^2}{2}$ dont on cherche le minimum sous la contrainte affine $x + 2y - z = 2$. Ainsi u est la solution du système de Cramer :

$$\begin{cases} x & & & +\lambda & = & 2 \\ & y & & +2\lambda & = & 3 \\ & & z & -\lambda & = & 0 \\ x & +2y & -z & & = & 2 \end{cases}$$

Que l'on résout sous matlab en saisissant le code :

```
>> M=[1 0 0 1;0 1 0 2;0 0 1 -1;1 2 -1 0];
>> c=[2 3 0 2]';
>> u=inv(M)*c
```

2) c) - En utilisant la fonction `quadprog` :

```
>> A=[1 0 0;0 1 0;0 0 1]; b=-u0;
>> Aeq=d'; beq=a;
>> u=quadprog(A,b,[],[],Aeq,beq)
```

Dans tous les cas matlab retourne le résultat :

```
u=[1;1;1]
```

Et l'on en déduit la distance de u_0 à H :

```
>> dist=sqrt((u-u0)'*(u-u0))
dist=2.4495
```

Problème 2

1) - La fonction f admet un minimum global et aucun maximum car elle est coercive.

Montrons que f est coercive, en procédant par l'absurde. Si f n'est pas coercive il existe $(x_n, y_n)_n$ une suite de points de $\mathbb{R}^2$ avec $\lim_{n\to\infty} \|(x_n, y_n)\| = +\infty$ et $f(x_n, y_n)$ qui ne converge pas vers $+\infty$. Nécessairement il existe une sous-suite $(f(x_{\sigma n}, y_{\sigma n}))_n$ convergeant vers une valeur finie ($= \liminf f(x_n, y_n) \geq 0$). Donc les 3 suites positives $|x_{\sigma n} - y_{\sigma n}^2 + 1|_n$, $|2x_{\sigma n}^2 - 3y_{\sigma n} - 2|_n$ et $|y_{\sigma n} - x_{\sigma n}^2 + 1|_n$ doivent avoir aussi une limite finie, et en particulier être bornées, soit trois conditions à vérifier.

La troisième condition montre que $y_{\sigma n} = x_{\sigma n}^2 + b(n)$ avec $b(n)$ bornée. Avec la deuxième condition $2x_{\sigma n}^2 - 3x_{\sigma n}^2 = -x_{\sigma n}^2$ est borné, et donc $x_{\sigma n}$ aussi. Avec la première condition $y_{\sigma n}^2$ est borné et donc $y_{\sigma n}$ aussi. Ainsi $x_{\sigma n}$ et $y_{\sigma n}$ sont tous deux bornés. Celà contredit le fait que $\lim_{n\to\infty} \|(x_{\sigma n}, y_{\sigma n})\| = +\infty$. Ainsi f est coercive.

2) - Tracé de la surface et des équipotentielles :

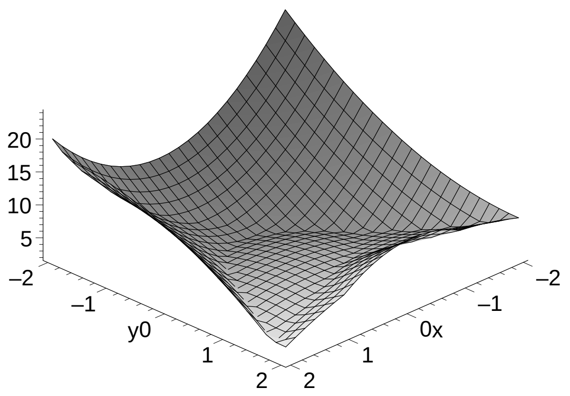

Figure 2: Surface $z = f(x, y)$

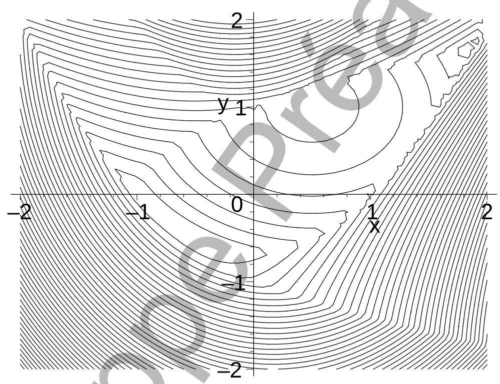

Figure 3: Equipotentielles

3) - Les 2 minimums sont bien visibles sur la figure 3. Voici quelques résultats obtenus en partant de points initiaux différents.

point initial	minimum obtenu	valeur de f
(0 , 0)	(-0.4963 , -0.5025)	2
(1 , 0)	(1 , 0)	4
(0.5 , 1)	(0.5 , 1)	5.25
(1.5 , 1.5)	(2 , 2)	4
(1.8 , 1.7)	(1.8834 , 1.6980)	2.849

Les cas 2, 3, 4 du tableau précédent illustrent les difficultés de l'optimisation.

Problème 3

1) - Soit $A(\alpha_1, \alpha_2, \ldots, \alpha_n)$ l'aire délimitée par les trapèzes. En posant $\alpha_0 = a, \alpha_{n+1} = b$,

$$A(\alpha_1, \alpha_2, \ldots, \alpha_n) = \sum_{i=0}^{n} (\alpha_{i+1} - \alpha_i) \frac{f(\alpha_i) + f(\alpha_{i+1})}{2}$$

Le problème consiste à minimiser :

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) \xrightarrow{E} |\int_a^b f(t)dt - A(\alpha_1, \alpha_2, \ldots, \alpha_n)|$$

sous les contraintes inégalitaires :

$$\alpha_0 \leq \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n \leq \alpha_{n+1}$$

Le fait que f soit continu entraîne que A et par suite E sont continus. Montrons que le domaine $\mathcal{D}$ est compact. Il est borné puisque $a \leq \alpha_i \leq b$ pour tout $i = 1, 2, \ldots, n$. Montrons qu'il est fermé. En posant pour tout $i = 0, 1, \ldots, n$, $\psi_i(\alpha_1, \alpha_2, \ldots, \alpha_n) = \alpha_i - \alpha_{i+1}$, le domaine $\mathcal{D}$ s'écrit

$$\mathcal{D} = \bigcap_{i=0}^{n} \psi_i^{-1}(]-\infty, 0])$$

et donc $\mathcal{D}$ est un fermé de $\mathbb{R}^n$ puisque les ψ_i sont continues. Ainsi $\mathcal{D}$ est compact, et puisque E est continue elle admet un minimum global sur $\mathcal{D}$.

Si f est convexe on peut supprimer la valeur absolue dans l'expression de E (positive) qui est alors 2 fois différentiable. Même chose si f est concave mais on cherche désormais un maximum de E (négative). Toutes les contraintes étant affines on peut appliquer la condition nécessaire de (KKT) :

$$\nabla_x \mathcal{L}(u,\mu) = \nabla E(u) + \sum_{i=0}^{n} \mu_i \nabla \psi_i(u) = \overrightarrow{0}$$

$$\mu_i \geq 0, \quad i = 0, 1, \ldots, n$$

$$\mu_i . \psi_i(u) = 0, \quad i = 0, 1, \ldots, n$$

Or remarquer que toutes les contraintes inégalitaires sont non saturées ! Donc tous les multiplicateurs de Lagrange-KKT sont nuls, et la condition nécessaire devient :

$$\nabla E(u) = 0 \quad \text{(équation d'Euler)}$$

L'extrémum se trouve dans l'intérieur du domaine, et la condition nécessaire du second ordre s'écrit :

$$\nabla^2 E(u) \text{ semi-définie positive ou négative}$$

selon si f est convexe ou concave. Une condition suffisante est alors : $\nabla E(u) = 0$ et $\nabla^2 E(u)$ est définie positive (respectivement négative).

Les contraintes inégalitaires étant inactives, en l'absence de contraintes égalitaires, le problème se restreint à un ouvert, l'intérieur du domaine, et les notions du chapitre II (optimisation sans contraintes) s'appliquent.

2) - *Cas de la subdivision comportant un seul point.* Il faut donc trouver $\alpha \in [0\ ,\ \frac{\pi}{2}]$ minimisant l'aire du domaine compris entre la sinusoïde et les 2 trapèzes. Cette aire $\mathcal{A}$ est égale à :

$$\mathcal{A} = 1 - \left(\frac{\pi}{2} - \alpha + \frac{\pi}{2}\sin\alpha\right)$$

La dérivée s'annule pour $\frac{\pi}{2}\cos\alpha = 1$. La solution de cette équation, dans l'intervalle $[0\ ,\ \frac{\pi}{2}]$, est :

$$\alpha^* = \arccos\frac{2}{\pi} \simeq 0.8807$$

La valeur minimale de l'erreur est $\mathcal{A}(\alpha^*) = 0.0493$

3) - *Cas de la subdivision comportant deux points.* En utilisant la fonction Matlab `fmincon`, on obtient les deux points cherchés de la subdivision. Il est préférable, pour cette recherche, de fournir le gradient de la fonction à minimiser. Voir pour cela le fichier `A2.m` :

```
function [f,gradf]=A2(x)
f=1-pi/4-1/2*(x(2)*sin(x(1))-x(1)*sin(x(2))+pi/2*sin(x(2))-x(2));
if nargout > 1
    gradf(1)=-1/2*(x(2)*cos(x(1))-sin(x(2)));
    gradf(2)=-1/2*(sin(x(1))-x(1)*cos(x(2))+pi/2*cos(x(2))-1);
end
```

Le minimum est obtenu par :

```
options=optimset('LargeScale','off','GradObj','on');
[xx,ffval,exitflag,output,lambda,grad,hessian]=fmincon(@A2,...
[0,0],[],[],[],[],[0,0],[pi/2,pi/2],[],options)
determinant2=det(hessian)
trace2=trace(hessian)
valeurspropres2=eig(hessian)
```

Résultats :

```
xx = 0.6389    1.1228

ffval = 0.0213
```

On vérifie que la matrice hessienne est définie positive, soit en remarquant que le déterminant et la trace sont strictement positifs, soit en calculant les 2 valeurs propres.

4) - Résultat du problème à 3 points :

```
xxx =  0.5109    0.8929    1.2385

fffval =  0.0118
```

Résultat du problème à 4 points :

```
xxxx = 0.4301    0.7498    1.0361    1.3066

ffffval = 0.0075
```

Problème 4

1) - L'énergie potentielle du cascadeur vaut :

$$f(x_1,x_2) = \frac{1}{2}k_1\left(\sqrt{x_1^2+x_2^2}-l_1\right)^2 + \frac{1}{2}k_2\left(\sqrt{(L-x_1)^2+x_2^2}-l_2\right)^2 - mgx_2$$

Il faut la minimiser avec les contraintes :

$$0 \leqslant x_1 \leqslant 10 \qquad \text{et} \qquad 0 \leqslant x_2.$$

L'application f est coercive sur le fermé $\mathcal{D} = [0,10] \times [0,+\infty]$, en effet : puisque x_1 est borné, $f(x_1,x_2)$ a pour équivalent lorsque x_2 est proche de l'infini :

$$f(x_1,x_2) \underset{+\infty}{\approx} \frac{1}{2}(k_1+k_2)x_2^2 - mgx_2$$

qui tend vers $+\infty$ quand x_2 tend vers l'infini. Puisque x_1 est borné, $\lim_{\|(x_1,x_2)\|\to\infty} = +\infty$. Ainsi f est coercive sur un fermé, il existe donc un minimum global.

2) - Il faut prendre les contraintes :

$$0 \leqslant x_1 \leqslant 10 \qquad \text{et} \qquad 0 \leqslant x_2 \leqslant 10000$$

La valeur 10000 est prise pour majorer (très largement) x_2. Il est préférable, pour cette recherche, de fournir le gradient de la fonction à minimiser. Voir pour cela le fichier `ep.m` :

```
function [f,gradf]=ep(x)
k1=15;
k2=30;
l1=25;
l2=40;
L=10;
m=70;
g=10;
f=0.5*k1*(sqrt(x(1)^2+x(2)^2)-l1)^2+0.5*k2*(sqrt((x(1)-L)^2+x(2)^2)-l2)^2-m*g*x(2);
if nargout > 1
    gradf(1)=k1*(sqrt(x(1)^2+x(2)^2)-l1)*x(1)/sqrt(x(1)^2+x(2)^2)...
+k2*(sqrt((x(1)-L)^2+x(2)^2)-l2)*(x(1)-L)/sqrt((x(1)-L)^2+x(2)^2);
    gradf(2)=k1*(sqrt(x(1)^2+x(2)^2)-l1)*x(2)/sqrt(x(1)^2+x(2)^2)...
+k2*(sqrt((x(1)-L)^2+x(2)^2)-l2)*x(2)/sqrt((x(1)-L)^2+x(2)^2)-m*g;
end
```

Le minimum est obtenu par :

```
clc
clear all
options=optimset('LargeScale','off','GradObj','on');
[xx,fval,exitflag,output,lambda,grad,hessian]=fmincon(@ep,[4.5,50],[],[],[],[],...
[0,0],[10,10000],[],options)
determinantH=det(hessian)
traceH=trace(hessian)
valeurs_propresH=eig(hessian)
```

On vérifie que la matrice hessienne est définie positive, soit en remarquant que le déterminant et la trace sont strictement positifs, soit en calculant les 2 valeurs propres.
Résultats :

```
xx = 4.5426   50.3667

fval = -2.8648e+004
```