**TD Extrema d'une fonction sans contrainte**

Exercice1

Soit la fonction $f(x,y) = x^4 + y^4 - 4xy$ définie sur $\mathbb{R}^2$.

1° Montrer que *f* est coercive à l'aide de la norme *sup* définie par $\|(x,y)\|_\infty = sup\{|x|,|y|\}$.

2° En déduire le minimum global de *f* sur $\mathbb{R}^2$.

3° Faire une étude locale des points critiques.

Exercice2

Soit la fonction $f(x,y,z) = e^{x-y} + e^{y-x} + z^2$ définie sur $\mathbb{R}^3$.

1° Calculer son hessien et étudier son signe. Montrer que *f* est convexe sur $\mathbb{R}^3$.

2° Déterminer les points critiques de *f*. En déduire le minimum global de *f*.

Exercice3. Minimiser une fonction quadratique convexe.

Soit la fonction quadratique $\frac{1}{2}x\bullet Ax + b\bullet x + c$ définie sur $\mathbb{R}^n$, avec *A* matrice carrée, symétrique, à entrées réelles et *b* vecteur colonne.

1° Calculer $f(x+\delta)$ et comparer au développement de Taylor au point $y = x + \delta$ autour du point *x* . En déduire l'expression du gradient et du hessien de *f* en fonction de *A* et *b*.

2° Vérifier le résultat sur $f(x) = x_1^2 + x_2^2 + x_3^2 - x_1x_3 - x_2x_3 - 4x_1$

En étudiant le signe du hessien de *f*, montrer que *f* est strictement convexe.

Déterminer le minimum global de *f*.

3° Soit $f(x) = x_1^2 + x_2^2 + \frac{x_3^2}{4} - x_1x_3 - 2x_1 + x_3$

En étudiant le signe du hessien de *f*, montrer que *f* est convexe.

Déterminer le minimum global de *f* .

4° Soit $f(x) = x_1^2 + x_2^2 + \frac{x_3^2}{4} - x_1x_3 - 2x_1 - x_3$

4-*a*° *f* admet-elle un point critique ?

4-*b*° En considérant le point $x_t = t\begin{pmatrix}1\\0\\2\end{pmatrix}$, montrer que *f* admet $-\infty$ pour minimum. On note *A* le hessien de *f*, vérifier que $Ax_t = 0$.

4-*c*° Généraliser la question à une fonction *f* quadratique quelconque $f(x) = \frac{1}{2}x\bullet Ax + b\bullet x + c$ et avec $b \neq 0$

Exercice4. Caractérisation des fonctions quadratiques coercives

Soit une matrice $A$ de *n* lignes et *n* colonnes, symétrique, à entrées réelles. On note $\lambda_{min}$ la plus petite valeur propre de $A$.

1° Montrer que $x \bullet Ax \geq \lambda_{min} \|x\|^2$ où la norme $\|x\|$ désigne la norme euclidienne du vecteur *x*.

2° Soit la forme quadratique $q(x) = x \bullet Ax$ définie sur $\mathbb{R}^n$, avec *A* matrice carrée, symétrique, à entrées réelles.

Montrer que *q* est coercive si et seulement si *A* est définie positive.

3° Soit la fonction quadratique $\frac{1}{2} x \bullet Ax + b \bullet x + c$ définie sur $\mathbb{R}^n$, avec *A* matrice carrée, symétrique, à entrées réelles et *b* vecteur colonne.

Montrer que f est coercive si et seulement si *A* est définie positive.

Exercice 5. Maximisation du log de vraisemblance .

On a un échantillon de *n* nombres réels $x_i$ *i*= 1 à *n*.

On fait l'hypothèse qu'ils suivent une loi normale de densité $f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ .

On cherche $\mu$ et $\sigma>0$ les paramètres de cette loi qui maximisent la fonction de vraisemblance de cet échantillon $V(\mu, \sigma) = \prod_{i=1}^{n} f(x_i, \mu, \sigma)$. Ce qui est équivalent à maximiser le logarithme de cette fonction : $LV(\mu, \sigma) = \log(\prod_{i=1}^{n} f(x_i, \mu, \sigma))$ car le logarithme est croissant.

1° Chercher le point critique de LV.

2° Calculer le hessien de LV en ce point critique. En déduire que le point critique est un maximum local strict de LV.

**Correction**

Exercice1.

1° On peut minorer la fonction $f(x,y) = x^4 + y^4 - 4xy$ par $\|(x,y)\|_\infty{}^4 - 4\|(x,y)\|_\infty{}^2$.

Elle apparaît alors clairement coercive.

2° On sait donc que *f* admet un minimum. Celui-ci se trouve nécessairement parmi les points critiques.

Il y a 3 points critiques : $P_1 = \binom{0}{0}, P_2 = \binom{1}{1}, P_3 = \binom{-1}{-1}$

*f*($P_1$)=0, *f*($P_2$)=*f*($P_3$)=-2. Donc $P_2$ et $P_3$ sont minimums globaux.

3° $P_1$ n'est pas un minimum local car c'est un point selle car le hessien de *f* en $P_1$ est indéfini.

Localement $P_2$ et $P_3$ sont minimums stricts car le hessien de *f* en $P_2$ et $P_3$ est défini positif.

Exercice2

Soit la fonction $f(x,y,z) = e^{x-y} + e^{y-x} + z^2$ définie sur $\mathbb{R}^3$.

1° On peut calculer son hessien et montrer qu'il est semi-défini positif sur $\mathbb{R}^3$. Il en résulte que *f* est convexe.

On peut aussi considérer *f* comme une somme de 3 fonctions de x,y,z qui sont convexes. Alors par théorème, son hessien est semi-définie positif sur $\mathbb{R}^3$.

2° Sachant que *f* est convexe ce qui est équivalent à dire que son hessien est semi-défini positif sur $\mathbb{R}^3$, pour minimiser *f* il suffit de chercher un point critique.

On cherche les points qui annulent le gradient de *f* et on trouve des points de la forme $\begin{pmatrix} x \\ x \\ 0 \end{pmatrix}$ c'est-à-dire une droite de vecteur directeur $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

Exercice3. Minimiser une fonction quadratique convexe.

Question 1°

Fonction quadratique $f(x) = \frac{1}{2}x \bullet Ax + b \bullet x + c$ avec A symétrique.

$f(x+\delta) = \frac{1}{2}(x+\delta) \bullet A(x+\delta) + b \bullet (x+\delta) + c$

On développe $f(x+\delta) = \frac{1}{2}x \bullet Ax + \delta \bullet Ax + \frac{1}{2}\delta \bullet A\delta + b \bullet (x+\delta) + c$ car A symétrique entraîne que $\delta \bullet Ax = x \bullet A\delta$.

Ensuite $f(x+\delta) = \frac{1}{2}x \bullet Ax + b \bullet x + c + \delta \bullet (Ax+b) + \frac{1}{2}\delta \bullet A\delta = f(x) + \delta \bullet (Ax+b) + \frac{1}{2}\delta \bullet A\delta$

En identifiant au développement de Taylor : $\nabla f(x) = Ax + b$ et $Hf = A$

Question 2° $f(x) = x_1^2 + x_2^2 + x_3^2 - x_1x_3 - x_2x_3 - 4x_1$ . Le hessien de $f$ est

$$A = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

Il est défini positif car les mineurs principaux sont $\Delta_1$=2>0 , $\Delta_2$=4>0 , $\Delta_3$=6+0-2=4>0. Donc $f$ est strictement convexe.

On cherche un point critique

$$Ax + b = 0 \text{ c'est-à-dire } \begin{cases} 2x_1 - x_3 = 4 & (1) \\ 2x_2 - x_3 = 0 & (2) \\ -x_1 - x_2 + 2x_3 = 0 & (3) \end{cases}$$

Unique solution : $x_1 = 3,\ x_2 = 1, x_3 = 2$

Question 3° Soit $f(x) = x_1^2 + x_2^2 + \frac{x_3^2}{4} - x_1x_3 - 2x_1 + x_3$. . Le hessien de $f$ est

$$A = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & \frac{1}{2} \end{pmatrix}$$

Il est semi-défini positif car les mineurs principaux sont $\Delta_1$=2>0 , $\Delta_2$=4>0 , $\Delta_3$=0. Donc $f$ est convexe.

Recherche de points critiques :

$$Ax + b = 0 \text{ c'est-à-dire } \begin{cases} 2x_1 - x_3 = 2 & (1) \\ 2x_2 = 0 & (2) \\ -x_1 + \frac{1}{2}x_3 = -1 & (3) \end{cases}$$

Les solutions sont les points de la forme : $x_t = t\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$,

Reportons dans $f$ :

$$f(x_t) = \tfrac{1}{2}x_t \bullet Ax_t + b \bullet x_t = \tfrac{1}{2}x_t \bullet Ax_t + \tfrac{1}{2}b \bullet x_t + \tfrac{1}{2}b \bullet x_t = \tfrac{1}{2}x_t \bullet (Ax_t + b) + \tfrac{1}{2}b \bullet x_t = \tfrac{1}{2}b \bullet x_t$$

Ce qui fait $f(x_t) = -1$. Notons que le vecteur $u_t = t\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ satisfait le système $Au_t = 0$ c'est-à-dire $u_t$ est orthogonal aux lignes de A et que –b est combinaison linéaire des colonnes de A. Comme A est symétrique, les colonnes de A sont les lignes de A c'est pourquoi le point $u_t$ est orthogonal à b. Ainsi $f(x_t)$ ne dépend pas de $t$ et le minimum est bien fini. Ceci se généralise pour toute fonction $f$ quadratique avec un hessien semi-défini positif (non inversible) et admettant un point critique.

Question 4-*a*° Soit $f(x) = x_1^2 + x_2^2 + \frac{x_3^2}{4} - x_1x_3 - 2x_1 - x_3$ . Le hessien de $f$ est

$$A = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & \frac{1}{2} \end{pmatrix}$$

Il est semi-défini positif car les mineurs principaux sont >0 sauf le dernier qui est nul. Donc $f$ est convexe.

Recherche de points critiques :

$Ax + b = 0$ c'est-à-dire $\begin{cases} 2x_1 - x_3 = 2 & (1) \\ 2x_2 = 0 & (2) \\ -x_1 + \frac{1}{2}x_3 = 1 & (3) \end{cases}$

Il n'y a pas de solution car les équations (1) et (3) sont incompatibles.

Question 4-*b*°

$x_t = t\begin{pmatrix}1\\0\\2\end{pmatrix}$, $f(x_t) = t^2 + \frac{1}{4}4t^2 - 2t^2 - 4t = -4t$ tends vers -∞ quand t tends vers +∞.

Question 4-*c*°

Supposons que $Ax = -b$ n'a pas de solution i.e. le gradient de *f* ne s'annule pas. Cela signifie que –*b* n'appartient pas à $\mathcal{L}(A)$ l'espace engendré par les colonnes de A. Donc –*b* a une composante $b_N$ dans l'orthogonal $\mathcal{L}(A)^{\perp}$ et une autre $b_R$ dans $\mathcal{L}(A)$ et –*b* = $b_R + b_N$ .

$b_N$ est orthogonal aux colonnes de A. Comme A est symétrique, $b_N$ est aussi orthogonal aux lignes de A. Et donc $b_N$ vérifie le système $Ab_N = 0$. On voit que $tb_N$ est aussi une solution du système quelque-soit le scalaire t.

Alors $f(tb_N) = \frac{1}{2}tb_N \bullet Atb_N + b \bullet tb_N + c = 0 + tb \bullet b_N + c$.

Or $b \bullet b_N = (-b_R - b_N) \bullet b_N = -\|b_N\|^2 \neq 0$ et on peut faire tendre $f(tb_N)$ vers moins l'infini lorsque t tends vers +∞.

Dans le cas de l'exemple : $b_N = \frac{4}{5}\begin{pmatrix}1\\0\\2\end{pmatrix}$ et $b_R = \frac{3}{5}\begin{pmatrix}2\\0\\-1\end{pmatrix}$. La somme fait bien $-b = \begin{pmatrix}2\\0\\1\end{pmatrix}$ et $b_N \bullet b_R = 0$ .

<u>Exercice 4</u>.

1° Montrer que $x \bullet Ax \geq \lambda_{min}\|x\|^2$ où la norme $\|x\|$ désigne la norme euclidienne du vecteur *x*.

A symétrique donc elle admet n vecteur propres unitaires et orthogonaux $u_i$ de valeurs propres $\lambda_i$ i=1,…,n

Un vecteur x quelconque peut se décomposer dans la base des $u_i$ : $x = \sum_{i=1}^{n} a_i u_i$

$$x \bullet Ax = \left(\sum_{i=1}^{n} a_i u_i\right) \bullet A\left(\sum_{i=1}^{n} a_i u_i\right) = \left(\sum_{i=1}^{n} a_i u_i\right) \bullet \left(\sum_{i=1}^{n} a_i \lambda_i u_i\right) =$$

$$\sum_{i=1}^{n} \lambda_i (a_i)^2 u_i \bullet u_i = \sum_{i=1}^{n} \lambda_i (a_i)^2 \geq \lambda_{min} \sum_{i=1}^{n} (a_i)^2$$

Or $\|x\|^2 = x \bullet x = (\sum_{i=1}^{n} a_i u_i) \bullet (\sum_{i=1}^{n} a_i u_i) = \sum_{i=1}^{n} (a_i)^2 u_i \bullet u_i = \sum_{i=1}^{n} (a_i)^2$

Finalement : $x \bullet Ax \geq \lambda_{min}\|x\|^2$

2° Soit la forme quadratique $q(x) = x \bullet Ax$ définie sur $\mathbb{R}^n$, avec *A* matrice carrée, symétrique, à entrées réelles.

Montrer que *q* est coercive si et seulement si *A* est définie positive.

A° Supposons A définie positive. Alors $\lambda_{min} > 0$.

$$q(x) = x \bullet Ax \geq \lambda_{min} \|x\|^2$$

Donc immédiatement, on voit que lorsque $\|x\| \to +\propto$, alors $q(x) \to +\propto$

B° Supposons q coercive.

Soit u vecteur propre unitaire (norme 1) associé à la plus petite valeur propre $\lambda_{min}$. Soit $\alpha$ scalaire.

$$q(\alpha u) = \alpha^2 u \bullet Au = \alpha^2 \lambda_{min}$$

Quand $\alpha \to \propto$ la norme $\|\alpha u\| \to \infty$ et donc $q(\alpha u) \to \propto$ et donc $\lambda_{min} > 0$ .

3° Soit la fonction quadratique $f(x) = \frac{1}{2} x \bullet Ax + b \bullet x + c$ définie sur $\mathbb{R}^n$, avec *A* matrice carrée, symétrique, à entrées réelles et *b* vecteur colonne.

Montrer que f est coercive si et seulement si *A* est définie positive.

A° Supposons A définie positive. Alors $\lambda_{min} > 0$. En appliquant Cauchy-Schwarz à $b \bullet x$ , on obtient :

$$f(x) = \frac{1}{2} x \bullet Ax + b \bullet x + c \geq \frac{1}{2} \lambda_{min} \|x\|^2 - \|b\| \|x\| + c$$

f(*x*) est minorée par un polynôme de la norme de *x* dont le coefficient du terme de plus haut degré est $\frac{1}{2} \lambda_{min} > 0$. Ce polynôme tend vers $+\propto$ quand la norme de *x* tend vers plus l'infini. Donc f est coercive.

B° Supposons f coercive. Soit u vecteur propre unitaire associé à $\lambda_{min}$ la plus petite valeur propre de A. Soit $\alpha$ scalaire. $f(\alpha u) = \frac{1}{2} \alpha^2 u \bullet Au + \alpha b \bullet u + c = \frac{1}{2} \alpha^2 \lambda_{min} + \alpha b \bullet u + c$

2 cas :

- $b \bullet u \leq 0$, on prend $\alpha > 0$
- $b \bullet u > 0$, on prend $\alpha < 0$.

Dans les 2 cas, on a $\alpha b \bullet u \leq 0$ et on a $f(\alpha u) = \frac{1}{2} \alpha^2 \lambda_{min} + \alpha b \bullet u + c \leq \frac{1}{2} \alpha^2 \lambda_{min} + c$

On fait $|\alpha| \to +\propto$ et alors $f(\alpha u) \to +\propto$ , ce qui entraine que $\lambda_{min} > 0$ .

<u>Exercice5. Maximisation du log de vraisemblance.</u>

$LV(\mu, \sigma) = \sum_{i=1}^{n} \log\big(f(x_i, \mu, \sigma)\big) = -n \log \sigma\sqrt{2\pi} - \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2}$

1° Recherche du point crtique

On dérive par rapport à μ :

$\frac{\partial LV(\mu,\sigma)}{\partial \mu} = \sum_{i=1}^{n} \frac{2(x_i-\mu)}{2\sigma^2} = 0 \rightarrow \frac{1}{\sigma^2}(\sum_{i=1}^{n} x_i - n\mu) = 0 \rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$

On dérive par rapport à σ :

$\frac{\partial LV(\mu,\sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \rightarrow n = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$ . Comme on connait déjà μ, on peut le remplacer par $\hat{\mu}$ .

2° Le hessien de LV

$\frac{\delta^2 LV(\mu,\sigma)}{(\delta\mu)^2} = -\frac{n}{\sigma^2} < 0, \frac{\delta^2 LV(\mu,\sigma)}{\delta\mu\delta\sigma} = \frac{\delta^2 LV(\mu,\sigma)}{\delta\sigma\delta\mu} = -\frac{2}{\sigma^3}(\sum_{i=1}^{n} x_i - n\mu), \frac{\delta^2 LV(\mu,\sigma)}{(\delta\sigma)^2} = \frac{n}{\sigma^2} - \frac{3}{\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2$

Le premier mineur principal du hessien est négatif.

Calculons le deuxième mineur principal (le déterminant du hessien) au point $\hat{\mu}, \hat{\sigma}$.

$-\frac{n}{\sigma^2}\left(-\frac{2n}{\sigma^2}\right) - \left(-\frac{2}{\sigma^3}\right)^2 (\sum_{i=1}^{n} x_i - n\hat{\mu})^2 = \frac{2n^2}{\sigma^4}$ car $\sum_{i=1}^{n} x_i - n\hat{\mu} = 0$. Le déterminant est donc >0.

Donc le hessien est défini négatif au point critique $\hat{\mu}, \hat{\sigma}$. Par théorème de cours (condition suffisante d'optimalité locale), le point critique trouvé est bien un maximum local strict.