J.F.C.

FONCTION CONVEXE D'UNE VARIABLE

Q1 a) Soit I un intervalle de $\mathbb{R}$.

Par définition : $\forall (x,y)\in I^2,\ \forall z\in\mathbb{R},\ x\leq z\leq y \Rightarrow z\in I$.

Soit $(a,b)\in I^2$.

ou $a\leq b$ et alors $\forall z\in[a,b],\ z\in I$; ou $a\geq b$ et alors $\forall z\in[b,a],\ z\in I$.

Ainsi si $a\leq b$: $[a,b]\subset I$ et si $a\geq b$: $[b,a]\subset I$. Donc I est convexe.

Remarque : Il est clair qu'un convexe de $\mathbb{R}$ est un intervalle. Ainsi les convexes de $\mathbb{R}$ sont les intervalles de $\mathbb{R}$.

b) Soit $(a,b)\in I^2$ tel que $a<b$.

$\forall\lambda\in]0,1[,\ \varphi(\lambda a+(1-\lambda)b)\leq\lambda\varphi(a)+(1-\lambda)\varphi(b)$.

$\forall\lambda\in]0,1[,\quad \varphi(\lambda a+(1-\lambda)b)-\varphi(a)\leq(1-\lambda)(\varphi(b)-\varphi(a))$ et $\lambda(\varphi(b)-\varphi(a))\leq\varphi(b)-\varphi(\lambda a+(1-\lambda)b)$

$\forall\lambda\in]0,1[,\ \dfrac{\varphi(\lambda a+(1-\lambda)b)-\varphi(a)}{(1-\lambda)(b-a)}\leq\dfrac{\varphi(b)-\varphi(a)}{b-a}\leq\dfrac{\varphi(b)-\varphi(\lambda a+(1-\lambda)b)}{\lambda(b-a)}=\dfrac{\varphi(\lambda a+(1-\lambda)b)-\varphi(b)}{\lambda(a-b)}$

$\forall\lambda\in]0,1[,\ \dfrac{\varphi(\lambda a+(1-\lambda)b)-\varphi(a)}{\underbrace{(\lambda a+(1-\lambda)b)-a}_{(1-\lambda)(b-a)\,!}}\underset{①}{\leq}\dfrac{\varphi(b)-\varphi(a)}{b-a}\underset{②}{\leq}\dfrac{\varphi(\lambda a+(1-\lambda)b)-\varphi(b)}{\underbrace{(\lambda a+(1-\lambda)b)-b}_{\lambda(a-b)\,!}}$

Remarquons que $\lim\limits_{\lambda\to1}(\lambda a+(1-\lambda)b)=a$; en faisant tendre λ vers 1 dans ① il vient alors $\varphi'(a)\leq\dfrac{\varphi(b)-\varphi(a)}{b-a}$.

Remarquons encore que $\lim\limits_{\lambda\to0}(\lambda a+(1-\lambda)b)=b$; en faisant tendre λ vers 0 dans ② il vient alors $\dfrac{\varphi(b)-\varphi(a)}{b-a}\leq\varphi'(b)$.

Ainsi $\forall(a,b)\in I^2,\ a<b\Rightarrow\varphi'(a)\leq\dfrac{\varphi(b)-\varphi(a)}{b-a}\leq\varphi'(b)$.

Ainsi φ' est croissante sur I.

b) Si φ' est croissante sur I. Montrons que φ est convexe sur I.

Soit $(a,b)\in I^2$ et soit $\lambda\in[0,1]$. Montrons que : $\varphi(\lambda a+(1-\lambda)b)\leq\lambda\varphi(a)+(1-\lambda)\varphi(b)$.

Si $a=b$ ou si $\lambda\in\{0,1\}$ c'est clair. Supposons alors $a\neq b$ et $\lambda\in]0,1[$.

Pour fixer les idées, supposons $a<b$. Alors $a<\lambda a+(1-\lambda)b<b$.

Les accroissements finis appliqués à φ sur $[a,\lambda a+(1-\lambda)b]$ et $[\lambda a+(1-\lambda)b,b]$ donnent :

$\exists u\in]a,\lambda a+(1-\lambda)b[,\quad \dfrac{\varphi(\lambda a+(1-\lambda)b)-\varphi(a)}{\lambda a+(1-\lambda)b-a}=\varphi'(u)$. Et

$\exists v\in]\lambda a+(1-\lambda)b,b[,\quad \dfrac{\varphi(b)-\varphi(\lambda a+(1-\lambda)b)}{b-(\lambda a+(1-\lambda)b)}=\varphi'(v)$.

$u<v$ donc $\varphi'(u)\leq\varphi'(v)$; ainsi $\dfrac{\varphi(\lambda a+(1-\lambda)b)-\varphi(a)}{(1-\lambda)(b-a)}\leq\dfrac{\varphi(b)-\varphi(\lambda a+(1-\lambda)b)}{\lambda(b-a)}$

donc $\lambda\varphi(\lambda a+(1-\lambda)b)-\lambda\varphi(a)\leq(1-\lambda)\varphi(b)-(1-\lambda)\varphi(\lambda a+(1-\lambda)b)$.

Ainsi $\varphi(\lambda a+(1-\lambda)b)\leq\lambda\varphi(a)+(1-\lambda)\varphi(b)$. Ceci achève de prouver que φ est convexe sur I.

Ainsi si φ' est croissante sur I, φ est convexe sur I.

c) Ce qui précède a montré que : φ convexe sur $I\Leftrightarrow\varphi'$ croissante sur I.

Comme : φ' croissante sur $I\Leftrightarrow\varphi''$ positive sur I.

Finalement : φ convexe sur $I\Leftrightarrow\varphi'$ croissante sur $I\Leftrightarrow\varphi''$ positive sur I.

Q2 Supposons φ convexe sur I. Soit $(a,x)\in I^2$.

- Si $a=x$: $\varphi'(a)(x-a)+\varphi(a)=\varphi(a)\leq\varphi(a)=\varphi(x)$
- Supposons $a<x$. D'après Q1 b) : $\varphi'(a)\leq\dfrac{\varphi(x)-\varphi(a)}{x-a}$ et comme

$x-a>0$: $\varphi'(a)(x-a)\leq\varphi(x)-\varphi(a)$ donc : $\varphi'(a)(x-a)+\varphi(a)\leq\varphi(x)$

- Supposons $a>x$. D'après Q1 b) : $\dfrac{\varphi(a)-\varphi(x)}{a-x}\leq\varphi'(a)$;

$a-x>0$ donne alors : $\varphi(a)-\varphi(x)\leq\varphi'(a)(a-x)$ donc $\varphi'(a)(x-a)+\varphi(a)\leq\varphi(x)$.

Ainsi si φ est convexe sur I : $\forall a\in I,\forall x\in I,\ \varphi(x)\geq\varphi'(a)(x-a)+\varphi(a)$.

Réciproquement supposons que $\forall a\in I$, $\forall x\in I$, $\varphi(x)\geq \varphi'(a)(x-a)+\varphi(a)$.
Montrons que φ est convexe sur I ; montrons, par exemple, que φ' est croissante sur I.
Soit $(a,b)\in I^2$ tel que : $a<b$.
Alors $\varphi(b)\geq \varphi'(a)(b-a)+\varphi(a)$ et $\varphi(a)\geq \varphi'(b)(a-b)+\varphi(b)$.
Ainsi $\dfrac{\varphi(b)-\varphi(a)}{b-a}\geq \varphi'(a)$ et $\dfrac{\varphi(b)-\varphi(a)}{b-a}\leq \dfrac{-\varphi'(b)(a-b)}{b-a}=\varphi'(b)$

Donc $\varphi'(a)\leq \dfrac{\varphi(b)-\varphi(a)}{b-a}\leq \varphi'(b)$. φ' est croissante sur I et φ est convexe sur I.

Donc φ est convexe sur I si et seulement si : $\forall a\in I, \forall x\in I, \varphi(x)\geq \varphi'(a)(x-a)+\varphi(a)$.

__Remarque..__ cette dernière propriété signifie que la courbe représentative de φ est au-dessus de toutes ses tangentes. Elle permet d'obtenir des inégalités classiques.

(Q3) Supposons $\varphi'(x_0)=0$. D'après ce qui précède : $\forall x\in I$, $\varphi(x)\geq \varphi'(x_0)(x-x_0)+\varphi(x_0)$.
Donc $\forall x\in I$, $\varphi(x)\geq\varphi(x_0)$. __$\varphi$ admet un minimum global en x_0.__

__Remarque..__ Si φ est concave sur I et si $\varphi'(x_0)=0$ alors φ admet un maximum global en x_0.

FONCTION CONVEXE DE PLUSIEURS VARIABLES

(Q1) Supposons f convexe sur Ω. Soit $(A,B)\in\Omega^2$.
Notons que $\forall t\in[0,1]$, $B+t(A-B)=tA+(1-t)B\in\Omega$ car Ω est convexe.
Posons alors : $\forall t\in[0,1]$, $\varphi_{A,B}(t)=f(B+t(A-B))$. Montrons que $\varphi_{A,B}$ est convexe sur $[0,1]$. Soit $(a,b)\in[0,1]^2$ et soit $\lambda\in[0,1]$.

$\varphi_{A,B}(\lambda a+(1-\lambda)b)=f(B+(\lambda a+(1-\lambda)b)(A-B))=f(\lambda B+(1-\lambda)B+\lambda a(A-B)+(1-\lambda)b(A-B))$

$\varphi_{A,B}(\lambda a+(1-\lambda)b) = f(\lambda(B+a(A-B))+(1-\lambda)(B+b(A-B)))\leq \lambda f(B+a(A-B))+(1-\lambda)f(B+b(A-B))$

$\begin{cases} f \text{ convexe} \\ B+a(A-B)\in\Omega \text{ car } a\in[0,1] \\ B+b(A-B)\in\Omega \text{ car } b\in[0,1]\end{cases}$

Donc $\varphi_{A,B}(\lambda a+(1-\lambda)b)\leq\lambda\varphi_{A,B}(a)+(1-\lambda)\varphi_{A,B}(b)$.

Ceci achève de montrer que $\varphi_{A,B}$ est convexe sur $[0,1]$.

Réciproquement supposons que pour tout $(A,B)\in\Omega^2$, $\varphi_{A,B}$ est définie et convexe sur $[0,1]$.
Montrons que f est convexe sur Ω. Soit $(A,B)\in\Omega^2$ et $\lambda\in[0,1]$. Posons $\mu=1-\lambda$, $\mu\in[0,1]$.

$f(\lambda A+(1-\lambda)B)=\varphi_{A,B}(\lambda)=\varphi_{A,B}(1-\mu)=\varphi_{A,B}(\mu\times 0+(1-\mu)\times 1)\leq \mu\varphi_{A,B}(0)+(1-\mu)\varphi_{A,B}(1)$

$f(\lambda A+(1-\lambda)B)\leq\mu\varphi_{A,B}(0)+(1-\mu)\varphi_{A,B}(1)=(1-\lambda)f(B)+\lambda f(B+1(A-B))=\lambda f(A)+(1-\lambda)f(B)$

Ceci achève de montrer que f est convexe sur Ω.

Donc __f convexe sur $\Omega \iff \forall (A,B)\in\Omega^2$, $\varphi_{A,B}$ convexe sur $[0,1]$.__

(Q2) Soient $A=(a_1,a_2,\dots,a_n)\in\Omega$ et $B=(b_1,b_2,\dots,b_n)\in\Omega$.
$\forall t\in[0,1]$, $\varphi_{A,B}(t)=f(b_1+t(a_1-b_1), b_2+t(a_2-b_2),\dots,b_n+t(a_n-b_n))$.

Pour $\forall k\in[\![1,n]\!]$, $\forall t\in[0,1]$, $u_k(t)=b_k+t(a_k-b_k)=ta_k+(1-t)b_k$.
→ $u_1,u_2,\dots,u_n$ sont de classe C^1 sur $[0,1]$;
→ $\forall t\in[0,1]$, $(u_1(t),u_2(t),\dots,u_n(t))=B+t(A-B)=tA+(1-t)B\in\Omega$;
→ f est de classe C^1 sur Ω ;
Ainsi $\varphi_{AB}: t\mapsto f(u_1(t),u_2(t),\dots,u_n(t))$ est de classe C^1 sur $[0,1]$.
En particulier __$\varphi_{A,B}$ est dérivable sur $[0,1]$.__

$\forall t\in[0,1]$, $\varphi'_{A,B}(t)=\sum_{k=1}^{n}u'_k(t)\dfrac{\partial f}{\partial x_k}(u_1(t),u_2(t),\dots,u_n(t))=\sum_{k=1}^{n}(a_k-b_k)\dfrac{\partial f}{\partial x_k}(B+t(A-B))$

Ainsi __$\forall t\in[0,1]$, $\varphi'_{A,B}(t)=\langle \operatorname{grad} f(B+t(B-A)), A-B\rangle = df_{B+t(A-B)}(A-B)$.__

(Q3) * Supposons i) ; c'est à dire que f est convexe sur Ω.
Soit $(A,X)\in\Omega^2$. $\varphi_{A,X}$ est convexe sur $[0,1]$. Rappelons que $\forall t\in[0,1]$, $\varphi_{A,X}(t)=f(X+t(A-X))$
En particulier $\varphi_{A,X}(0)\geq\varphi'_{A,X}(1)(0-1)+\varphi_{A,X}(1)$

Donc $f(X)\geq df_{X+1(A-X)}(A-X)(0-1)+f(X+A-X)=df_A(X-A)+f(A)$

Ainsi $f(X)-f(A)\geq df_A(X-A)$; d'où ii).

* Supposons ii) et montrons i). Pour cela posons $(A,B)\in\Omega^2$ et montrons que $\varphi_{A,B}$ est convexe. Prenons alors u et v dans $[0,1]$ et montrons que :
$\varphi_{A,B}(v)\geq\varphi'_{A,B}(u)(v-u)+\varphi_{A,B}(u)$. Soit encore à montrer que :

$f(B+v(A-B))\geq df_{B+u(A-B)}(A-B)\times(v-u)+f(B+u(A-B))$

Ceci est presque clair. En effet $B+v(A-B)$ et $B+u(A-B)$ sont deux éléments de Ω donc en appliquant ii) nous pouvons écrire que :

$f(B+v(A-B)) \geq f(B+u(A-B)) + df_{B+u(A-B)}(B+v(A-B)-(B+u(A-B)))$ ce qui

avec la linéarité de $df_{B+u(A-B)}$ donne encore :

$f(B+v(A-B)) \geq df_{B+u(A-B)}(A-B)\times(v-u) + f(B+u(A-B))$; ainsi :

$\varphi_{A,B}(v) \geq \varphi'_{A,B}(u)(v-u)+\varphi_{A,B}(u)$; $\varphi_{A,B}$ est convexe ... f aussi.

* Supposons f convexe et montrons iii). Soit $(A,X)\in\Omega^2$. $\varphi_{A,X}$ est convexe sur $[0,1]$ donc $\varphi'_{A,X}$ est croissante sur $[0,1]$; en particulier $\varphi'_{A,X}(1) \geq \varphi'_{A,X}(0)$.

Ce qui s'écrit : $df_{X+1(A-X)}(A-X) \geq df_{X+0(A-X)}(A-X)$; c'est à dire que :

$df_A(A-X) \geq df_X(A-X)$; donc $df_A(X-A) = -df_A(A-X) \leq -df_X(A-X) = df_X(X-A)$.

C'est ce qu'il fallait prouver.

* Supposons iii) et montrons que f est convexe sur Ω.

Soit $(A,B)\in\Omega^2$. montrons que $\varphi_{A,B}$ est convexe sur $[0,1]$ en montrant que $\varphi'_{A,B}$ est croissante sur $[0,1]$.

Soit $(u,v)\in[0,1]^2$ tel que : $u<v$.

$\varphi'_{A,B}(u) = df_{B+u(A-B)}(A-B)$ et $\varphi'_{A,B}(v) = df_{B+v(A-B)}(A-B)$.

Appliquons iii).

$df_{B+v(A-B)}(B+v(A-B)-B-u(A-B)) \geq df_{B+u(A-B)}(B+v(A-B)-B-u(A-B))$

Ainsi $df_{B+v(A-B)}((v-u)(A-B)) \geq df_{B+u(A-B)}((v-u)(A-B))$.

donc $(v-u)\,df_{B+v(A-B)}(A-B) \geq (v-u)\,df_{B+u(A-B)}(A-B)$.

comme $v-u$ est strictement positif il vient $\varphi'_{A,B}(v) = df_{B+v(A-B)}(A-B) \geq df_{B+u(A-B)}(A-B) = \varphi'_{A,B}(u)$

Ceci montre alors la convexité de $\varphi_{A,B}$ et donc celle de f.

Q4 a) Soit $(A,B)\in S^2$ et $\lambda\in[0,1]$. Montrons que $\lambda A+(1-\lambda)B\in S$.

$\forall X\in\Omega$, $f(\lambda A+(1-\lambda)B) \underset{f\text{ convexe}}{\leq} \lambda f(A)+(1-\lambda)f(B) \underset{\lambda\in[0,1],A\in S,B\in S}{\leq} \lambda f(X)+(1-\lambda)f(X) = f(X)$; $\lambda A+(1-\lambda)B\in S$.

donc <u>S est convexe.</u>

b) Si $A\in S$, A est un point de l'ouvert Ω où la fonction f possède un minimum global .. donc local ; donc si $A\in S$: $\operatorname{grad} f(A)=0$ et A est un point critique de f.

Réciproquement supposons que A est un point de Ω tel que $\operatorname{grad} f(A)=0$.

Comme f est convexe sur Ω : $\forall X\in\Omega$, $f(X)\geq f(A)+df_A(X-A) = f(A)+\langle \operatorname{grad} f(A), X-A\rangle = f(A)$.

$\forall X\in\Omega$, $f(X)\geq f(A)$; $A\in S$.

Ainsi <u>$S=\{A\in\Omega \mid \operatorname{grad} f(A)=0\}$.</u>

Q5 Ici f est de classe C^2 sur Ω. Soit $(X,Y)\in\Omega^2$.

$\forall t\in[0,1]$, $\varphi_{X,Y}(t) = f(Y+t(X-Y))$ et $\varphi'_{X,Y}(t) = \sum_{k=1}^{n}(x_k-y_k)\frac{\partial f}{\partial x_k}(Y+t(X-Y))$

montrons alors que $\varphi'_{X,Y}$ est dérivable sur $[0,1]$. Il suffit de fixer k dans $[\![1,n]\!]$ et de montrer que $t\mapsto\frac{\partial f}{\partial x_k}(Y+t(X-Y))$ est dérivable sur $[0,1]$.

Comme $\frac{\partial f}{\partial x_k}$ est de classe C^1 sur Ω, en reprenant le raisonnement fait en II Q2 on obtient 1°.. la dérivabilité de $t\mapsto\frac{\partial f}{\partial x_k}(Y+t(X-Y))$ sur $[0,1]$

2°.. et le fait que sa dérivée est $t\mapsto\sum_{i=1}^{n}(x_i-y_i)\frac{\partial^2 f}{\partial x_i\partial x_k}(Y+t(X-Y))$

Ainsi $\varphi'_{X,Y}$ est dérivable sur $[0,1]$ comme combinaison linéaire de fonctions dérivables sur $[0,1]$.

Ainsi $\varphi_{X,Y}$ est deux fois dérivable sur $[0,1]$.

$\forall t\in[0,1]$, $\varphi''_{X,Y}(t) = \sum_{k=1}^{n}(x_k-y_k)\sum_{i=1}^{n}(x_i-y_i)\frac{\partial^2 f}{\partial x_i\partial x_k}(Y+t(X-Y))$

$\forall t\in[0,1]$, $\varphi''_{X,Y}(t) = \sum_{k=1}^{n}\sum_{i=1}^{n}(x_k-y_k)\frac{\partial^2 f}{\partial x_i\partial x_k}(Y+t(X-Y))(x_i-y_i) = {}^t(X-Y)\,H(f, Y+t(X-Y))(X-Y)$

. donc $\forall (X,Y)\in\Omega^2, \forall t\in[0,1],\ \varphi''_{X,Y}(t) = {}^t(X-Y)\,H(f, Y+t(X-Y))\,(X-Y)$.

b) Rappelons que f est convexe sur Ω si et seulement si pour tout $(X,Y)\in\Omega^2$, $\varphi_{X,Y}$ est convexe sur $[0,1]$. Ainsi, en utilisant la partie I on peut dire que f est convexe sur Ω si et seulement si pour tout $(X,Y)\in\Omega^2$, $\varphi''_{X,Y}$ est positive sur $[0,1]$. Il ne reste donc plus qu'à montrer que :

$\forall (X,Y)\in\Omega^2, \forall t\in[0,1],\ {}^t(X-Y)\,H(f, Y+t(X-Y))\,(X-Y)\geq 0 \iff \forall A\in\Omega, \forall H\in\mathbb{R}^n,\ {}^tH\,H(f,A)\,H\geq 0$.

$\Leftarrow$ est claire (prendre $H = X-Y$ et $A = Y+t(X-Y)$). Montrons l'implication inverse.

$\Rightarrow$ Supposons donc que : $\forall (X,Y)\in\Omega^2, \forall t\in[0,1],\ {}^t(X-Y)H(f,Y+t(X-Y))(X-Y)\geq 0$

Soit $A\in\Omega$ et $H\in\mathbb{R}^n$. Si $H=0$ alors ${}^tH\,H(f,A)\,H = 0\geq 0$; supposons $H\neq 0$.

Ω étant ouvert, $\exists r\in\mathbb{R}_+^*$, $B(A,r)\subset\Omega$.

Posons $\sigma = \dfrac{r}{2\|H\|}$, $X = A+\sigma H$ et $Y = A-\sigma H$.

$\|X-A\| = \sigma\|H\| \leq \dfrac{r\,\|H\|}{2\|H\|} = \dfrac{r}{2} < r$; $X\in B(A,r)$ donc $X\in\Omega$.

On obtient de même : $Y\in\Omega$.

Notons que : $Y+\frac{1}{2}(X-Y) = A-\sigma H+\frac{1}{2}(A+\sigma H-A+\sigma H) = A$ et $X-Y = 2\sigma H$.

$X\in\Omega$, $Y\in\Omega$ donc ${}^t(X-Y)\,H(f, Y+t(X-Y))\,(X-Y)\geq 0$. Ce qui s'écrit encore :

${}^t(2\sigma H)\,H(f,A)\,(2\sigma H)\geq 0$ ou $4\sigma^2({}^tH\,H(f,A)\,H)\geq 0$; ou donc encore

${}^tH\,H(f,A)\,H\geq 0$ ce qui achève de montrer l'équivalence.

f est convexe sur Ω ssi $\forall A\in\Omega, \forall H\in\mathbb{R}^n$, ${}^tH\,H(A,f)\,H\geq 0$.

Remarque. Notons $\forall H\in\mathbb{R}^n$, ${}^tH\,H(A,f)\,H\geq 0$ signifie que la matrice symétrique $H(A,f)$ est positive ; c'est-à-dire que ses valeurs propres sont positives.

Q6 $\forall (x,y)\in\mathbb{R}^2$, $f(x,y) = x^4+2x^2+2y^2-4xy+4x-4y+3$.

f est de classe C^2 sur $\mathbb{R}^2$ car f est une fonction polynôme.

$\forall (x,y)\in\mathbb{R}^2$, $\dfrac{\partial f}{\partial x}(x,y) = 4x^3+4x-4y+4$, $\dfrac{\partial f}{\partial y}(x,y) = 4y-4x-4$, $\dfrac{\partial^2 f}{\partial x^2}(x,y) = 12x^2+4$,

$\dfrac{\partial^2 f}{\partial y^2}(x,y) = 4$ et $\dfrac{\partial^2 f}{\partial x\partial y}(x,y) = -4$

Soit $X = (x,y)\in\mathbb{R}^2$. $H(f,X) = \begin{pmatrix} 12x^2+4 & -4\\ -4 & 4\end{pmatrix}$

Soit $H = (\alpha,\beta)\in\mathbb{R}^2$. ${}^tH\,H(f,X)\,H = (12x^2+4)\alpha^2+4\beta^2-8\alpha\beta = 12x^2\alpha^2+4(\alpha-\beta)^2\geq 0$

Ainsi f est convexe sur $\mathbb{R}^2$.

$\dfrac{\partial f}{\partial x}(X) = \dfrac{\partial f}{\partial y}(X) = 0 \iff \begin{cases} 4x^3+4x-4y+4 = 0\\ 4y-4x-4 = 0\end{cases} \iff \begin{cases} 4x^3 = 0\\ y = x+1\end{cases} \iff \begin{cases} x = 0\\ y = 1\end{cases}$.

f est convexe sur $\mathbb{R}^2$ et $A = (0,1)$ est le seul point critique de f.

Ainsi f admet un minimum global en $A = (0,1)$ qui est : 1.

Ici : $\forall (x,y)\in\mathbb{R}^2$, $f(x,y) = x^2+2xy+4y^2+x-2y$.

f est de classe C^2 sur

$\forall (x,y)\in\mathbb{R}^2$, $\dfrac{\partial f}{\partial x}(x,y) = 2x+2y+1$, $\dfrac{\partial f}{\partial y}(x,y) = 2x+8y-2$, $\dfrac{\partial^2 f}{\partial x^2}(x,y) = 2$, $\dfrac{\partial^2 f}{\partial y^2}(x,y) = 8$ et

$\dfrac{\partial^2 f}{\partial x\partial y}(x,y) = 2$. $H(f,X) = \begin{pmatrix} 2 & 2\\ 2 & 8\end{pmatrix}$

Soit $X = (x,y)\in\mathbb{R}^2$. $\forall (\alpha,\beta)\in\mathbb{R}^2$, $(\alpha,\beta)\,H(f,X)\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = 2\alpha^2+8\beta^2+4\alpha\beta = 2(\alpha+\beta)^2+6\beta^2\geq 0$.

f est donc convexe sur $\mathbb{R}^2$.

$\dfrac{\partial f}{\partial x}(X) = \dfrac{\partial f}{\partial y}(X) = 0 \iff \begin{cases} 2x+2y+1 = 0\\ 2x+8y-2 = 0\end{cases} \iff \begin{cases} 6y-3 = 0 & (L_2-L_1)\\ x = 1-4y\end{cases} \iff \begin{cases} y = \frac{1}{2}\\ x = -1\end{cases}$

$A = (-1,\frac{1}{2})$ est le seul point critique de f.

Par conséquent f admet un minimum global en $A = (-1,\frac{1}{2})$ qui est : 1.

III FONCTION FORTEMENT CONVEXE

(Q1) Soit D un convexe de $\mathbb{R}^n$ et f une application de D dans $\mathbb{R}$ fortement convexe.

$\exists \alpha \in \mathbb{R}_+^*$, $\forall \lambda \in [0,1]$, $\forall (A,B) \in D^2$, $f(\lambda A + (1-\lambda)B) \leq \lambda f(A) + (1-\lambda) f(B) - \alpha\lambda(1-\lambda)\|A-B\|^2$

Notons que : $\forall \lambda \in [0,1]$, $\forall (A,B) \in D^2$, $-\alpha\lambda(1-\lambda)\|A-B\|^2 \leq 0$.

Par conséquent : $\forall \lambda \in [0,1]$, $\forall (A,B) \in D^2$, $f(\lambda A + (1-\lambda)B) \leq \lambda f(A) + (1-\lambda) f(B)$; f est convexe sur D.

Toute fonction fortement convexe est convexe.

Remarque.. La réciproque est fausse. Par exemple toute application affine de $\mathbb{R}$ dans $\mathbb{R}$ est convexe sans être fortement convexe.

Exercice.. $f : x \mapsto x^2$ est-elle fortement convexe sur $\mathbb{R}^2$?

Dans ce qui suit Ω est un ouvert convexe de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$.

(Q2) Ici f est fortement convexe et α est un paramètre de convexité forte de f.

a) Soit $(A,B) \in \Omega^2$. Soit C le milieu de $[A,B]$. $C \in \Omega$ et $C = \frac{1}{2}A + \frac{1}{2}B$.

f étant convexe sur Ω on a, d'après II, :

$$\begin{cases} f(C) \geq f(A) + df_A(C-A) = f(A) + df_A(\frac{1}{2}A + \frac{1}{2}B - A) = f(A) + \frac{1}{2} df_A(B-A) \\ f(C) \geq f(B) + df_B(C-B) = f(B) + df_B(\frac{1}{2}A + \frac{1}{2}B - B) = f(B) + \frac{1}{2} df_B(A-B) \end{cases}$$

Par addition on obtient :

$\frac{1}{2} df_A(B-A) + \frac{1}{2} df_B(A-B) + f(A) + f(B) \leq 2f(C) = 2f(\frac{1}{2}A + \frac{1}{2}B)$. Par forte convexité il vient :

$-\frac{1}{2} df_A(A-B) + \frac{1}{2} df_B(A-B) + f(A) + f(B) \leq 2\left(\frac{1}{2}f(A) + \frac{1}{2}f(B) - \alpha \frac{1}{2} \times \frac{1}{2} \times \|A-B\|^2\right)$

Donc $-\frac{1}{2} df_A(A-B) + \frac{1}{2} df_B(A-B) + f(A) + f(B) \leq f(A) + f(B) - \frac{1}{2}\alpha\|A-B\|^2$

Ainsi $\frac{1}{2}\alpha\|A-B\|^2 \leq \frac{1}{2} df_A(A-B) - \frac{1}{2} df_B(A-B)$.

Soit encore : $\alpha\|A-B\|^2 \leq df_A(A-B) - df_B(A-B)$.

$\forall (A,B) \in \Omega^2, \quad \alpha\|A-B\|^2 \leq df_A(A-B) - df_B(B-B)$.

b) Soit $(A,B) \in \Omega^2$. Montrons que : $df_A(B-A) + \alpha\|A-B\|^2 \leq f(B) - f(A)$.

Notons que : $\underbrace{-df_A(B-A)}_{df_A(A-B)} = \varphi'_{A,B}(1) = \lim_{t \to 1^-} \frac{\varphi_{A,B}(t) - \varphi_{A,B}(1)}{t-1} = \lim_{t \to 1^-} \frac{f(B + t(A-B)) - f(A)}{t-1}$.

Soit $t \in [0,1[$. $f(tA + (1-t)B) \leq tf(A) + (1-t)f(B) - \alpha t(1-t)\|A-B\|^2$

$\varphi_{A,B}(t) \leq (1-t)(f(B) - f(A)) + \underbrace{f(A)}_{\varphi_{A,B}(1)} - \alpha t(1-t)\|A-B\|^2$

Par conséquent : $\varphi_{A,B}(t) - \varphi_{A,B}(1) \leq (1-t)\left[f(B) - f(A) - \alpha t\|A-B\|^2\right]$.

En divisant par $1-t$ qui est strictement positif il vient :

$$-\frac{\varphi_{A,B}(t) - \varphi_{A,B}(1)}{t-1} = \frac{\varphi_{A,B}(t) - \varphi_{A,B}(1)}{1-t} \leq f(B) - f(A) - \alpha t\|A-B\|^2.$$

Il ne reste plus qu'à faire tendre t vers 1 pour obtenir : $-\varphi'_{A,B}(1) \leq f(B) - f(A) - \alpha\|A-B\|^2$

Ainsi $df_A(A-B) \leq f(B) - f(A) - \alpha\|A-B\|^2$.

Finalement : $\forall (A,B) \in \Omega^2$, $df_A(A-B) + \alpha\|A-B\|^2 \leq f(B) - f(A)$.

c) Fixons A dans Ω.

$\forall X \in \Omega$, $f(X) \geq f(A) + df_A(A-X) + \alpha\|A-X\|^2$ d'après ce qui précède.

$\forall X \in \Omega$, $|df_A(A-X)| = |\langle \operatorname{grad} f(A), A-X \rangle| \leq \|\operatorname{grad} f(A)\| \, \|A-X\|$

↖ Cauchy-Schwarz.

Donc $\forall X \in \Omega$, $-df_A(A-X) \leq \|\operatorname{grad} f(A)\| \, \|A-X\|$

Ainsi $\forall X \in \Omega$, $f(X) \geq f(A) + df_A(A-X) + \alpha\|A-X\|^2 \geq f(A) - \|\operatorname{grad} f(A)\| \, \|A-X\| + \|A-X\|^2 \alpha$

Posons $\rho = \|\operatorname{grad} f(A)\|$

Il vient alors : $\forall X \in \Omega$, $f(X) \geq f(A) - \rho\|A-X\| + \alpha\|A-X\|^2$

"Intuitivement" le résultat est maintenant clair.

$\lim_{\|X\| \to +\infty} (\|X-A\|) = +\infty$ et $\lim_{\|X\| \to +\infty} \left(f(A) - \rho\|A-X\| + \alpha\|A-X\|^2\right) = +\infty$ car $\alpha > 0$

Il ne fait alors plus de doute que $\lim_{\|X\| \to +\infty} f(X) = +\infty$. Reste à rédiger au propre.

Fixons M dans $\mathbb{R}_+^*$. $\lim_{t \to +\infty} (f(A) - \rho t + \alpha t^2) = +\infty$ car $\alpha > 0$.

Donc $\exists L' \in \mathbb{R}_+^*$, $\forall t \in \mathbb{R}$, $t \geq L' \Rightarrow f(A) - \rho t + \alpha t^2 \geq M$.

Posons $L = L' + \|A\|$

Soit $X \in \Omega$ tel que : $\|X\| \geq L$. Alors $\|X-A\| \geq \big|\|X\| - \|A\|\big| \geq \|X\| - \|A\| \geq L - \|A\| = L'$

Ainsi $\|X-A\| \geq L'$. On peut alors dire que $f(A) - f\,\|X-A\| + \alpha \|X-A\|^2 \geq M$.
Donc $f(X) \geq f(A) - f\,\|A-X\| + \alpha\|A-X\|^2 = f(A) - f\,\|X-A\| + \alpha\|X-A\|^2 \geq M$.
Ainsi $\|X\| \geq L$ donne $f(X) \geq M$.
$\forall M \in \mathbb{R}^*_+, \exists L \in \mathbb{R}^*_+, \forall X \in \Omega, \|X\| \geq L \Rightarrow f(X) \geq M$. $\underline{\underline{\lim_{\|X\| \to +\infty} f(X) = +\infty}}$.

<u>d)</u> Ici $\Omega = \mathbb{R}^n$.

$\lim_{\|X\| \to +\infty} f(X) = +\infty$. Ainsi $\exists L \in \mathbb{R}^*_+, \forall X \in \mathbb{R}^n, \|X\| \geq L \Rightarrow f(X) \geq \overbrace{|f(0)| + 1}^{\in \mathbb{R}^*_+}$!!

Donc $\forall X \in \mathbb{R}^n$, $\|X\| \geq L \Rightarrow f(X) \geq f(0)$!! ce qui donne en général :

$\forall X \in \mathbb{R}^n$, $\|X\| > L \Rightarrow f(X) \geq f(0)$ non ?

La boule fermée $B'(0,L)$ de centre 0 et de rayon L est un fermé borné et la restriction g de f à cette boule est continue (f est C^1 sur $\mathbb{R}^n$).
Ainsi g possède un minimum. Il existe donc $A \in B'(0,L)$ tel que :
$\forall X \in B'(0,L)$, $f(X) = g(X) \geq g(A) = f(A)$.
Donc $\forall X \in B'(0,L)$, $f(X) \geq f(A)$. Montrons en fait que : $\forall X \in \mathbb{R}^n$, $f(X) \geq f(A)$.
Soit donc X un élément de $\mathbb{R}^n$.
Si $X \in B'(0,L)$, $f(A) \geq f(A)$! Supposons que : $X \notin B'(0,L)$. Alors $\|X\| > L$
Ainsi $f(X) \geq f(0) \geq f(A)$ car $0 \in B'(0,L)$.
Par conséquent : $\forall X \in \mathbb{R}^n$, $f(X) \geq f(A)$.
Nous avons donc prouvé que : $\exists A \in \Omega$, $\forall X \in \Omega$, $f(A) \leq f(X)$. Il reste plus qu'à montrer l'unicité de A.
Supposons que A' soit un point de Ω vérifiant : $\forall X \in \Omega$, $f(A') \leq f(X)$.
Alors $f(A') \leq f(A) \leq f(A')$ donc $f(A) = f(A')$.
$\forall \lambda \in [0,1]$, $f(A) \leq f(\lambda A + (1-\lambda)A') \leq \lambda f(A) + (1-\lambda) f(A') - \alpha\lambda(1-\lambda)\|A-A'\|^2$
$\forall \lambda \in [0,1]$, $\alpha\lambda(1-\lambda)\|A-A'\|^2 \leq \lambda f(A) + (1-\lambda) f(A') - f(A) = \lambda f(A) + (1-\lambda) f(A) - f(A) = 0$
$\forall \lambda \in [0,1]$, $\alpha\lambda(1-\lambda)\|A-A'\|^2 \leq 0$.
Ceci donne pour $\lambda = 1/2$: $\frac{\alpha}{4}\|A-A'\|^2 \leq 0$ donc $\|A-A'\|^2 \leq 0$.
Ainsi $\|A-A'\|^2 = 0$ et $A' = A$.
$\exists!\, A \in \Omega$, $\forall X \in \Omega$, $f(A) \leq f(X)$.

(Q3) Notons que s'il existe α dans $\mathbb{R}^*_+$ tel que :
$\forall (A,B) \in \Omega^2$, $df_A(B-A) + \alpha\|A-B\|^2 \leq f(B) - f(A)$ alors f est fortement convexe sur Ω.
Soit $(A,B) \in \Omega^2$ et $\lambda \in [0,1]$. L'hypothèse donne :
$f(X) - f(\lambda X + (1-\lambda)Y) \geq df_{\lambda X + (1-\lambda)Y}(X - (\lambda X + (1-\lambda)Y)) + \alpha\|\lambda X + (1-\lambda)Y - X\|^2$
et
$f(Y) - f(\lambda X + (1-\lambda)Y) \geq df_{\lambda X + (1-\lambda)Y}(Y - (\lambda X + (1-\lambda)Y)) + \alpha\|\lambda X + (1-\lambda)Y - Y\|^2$

Ainsi $\begin{cases} f(X) - f(\lambda X + (1-\lambda)Y) \geq (1-\lambda)\, df_{\lambda X + (1-\lambda)Y}(X-Y) + \alpha(1-\lambda)^2\|Y-X\|^2 \\ f(Y) - f(\lambda X + (1-\lambda)Y) \geq -\lambda\, df_{\lambda X + (1-\lambda)Y}(X-Y) + \alpha\lambda^2\|X-Y\|^2 \end{cases}$

En multipliant la première inégalité par λ, la seconde par $1-\lambda$ et en ajoutant on obtient
$\lambda f(X) - \lambda f(\lambda X + (1-\lambda)Y) + (1-\lambda) f(Y) - (1-\lambda) f(\lambda X + (1-\lambda)Y) \geq \alpha\|X-Y\|^2[\lambda(1-\lambda)^2 + \lambda^2(1-\lambda)]$
$\lambda f(X) + (1-\lambda) f(Y) - f(\lambda X + (1-\lambda)Y) \geq \alpha\|X-Y\|^2 \lambda(1-\lambda)(1-\lambda+\lambda)$
Ainsi $f(\lambda X + (1-\lambda)Y) \leq \lambda f(X) + (1-\lambda) f(Y) - \alpha\lambda(1-\lambda)\|X-Y\|^2$ et ... f est fortement convexe.

(Q4) <u>a)</u> Soit $p \in \mathbb{N}$

$\|X_{p+1} - A\|^2 = \|X_p - \lambda_p \operatorname{grad} f(X_p) - A + \lambda_p \underbrace{\operatorname{grad} f(A)}_{=0}\|^2$

$\|X_{p+1} - A\|^2 = \|X_p - A\|^2 - 2\lambda_p \langle X_p - A, \operatorname{grad} f(X_p) - \operatorname{grad} f(A)\rangle + \|\operatorname{grad} f(X_p) - \operatorname{grad} f(A)\|^2$

$\|X_{p+1} - A\|^2 \leq \|X_p - A\|^2 - 2\lambda_p\, \alpha\|X_p - A\|^2 + K^2\|X_p - A\|^2\lambda_p^2$

$\uparrow$ $\langle X_p - A, \operatorname{grad} f(X_p) - \operatorname{grad} f(A)\rangle = df_{X_p}(X_p - A) - df_A(X_p - A) \geq \alpha\|X_p - A\|^2$ (Q2)

$\|X_{p+1} - A\|^2 \leq (K^2\lambda_p - 2\alpha\lambda_p + 1)\|X_p - A\|^2$.

$\underline{\underline{\forall p \in \mathbb{N},\ \|X_{p+1} - A\|^2 \leq (K^2\lambda_p^2 - 2\alpha\lambda_p + 1)\|X_p - A\|^2.}}$

<u>b)</u> $\forall t \in \mathbb{R}$, $\psi'(t) = 2K^2 t - 2\alpha = 2(K^2 t - \alpha)$

t		α/K^2	
$\psi'(t)$	$-$	0	$+$
$\psi(t)$	$\searrow$	m	$\nearrow$

$\begin{cases} m = f\left(\frac{\alpha}{K^2}\right) = K^2\frac{\alpha^2}{K^4} - 2\alpha\frac{\alpha}{K^2} + 1 = 1 - \frac{\alpha^2}{K^2} \\ \psi(0) = 1 \end{cases}$

1er cas. $1 - \frac{\alpha^2}{K^2} \geq 0$. ψ est strictement décroissante sur $[0, \frac{\alpha}{K^2}]$, $\psi(0) = 1$ et $\psi\left(\frac{\alpha}{K^2}\right) \geq 0$

Si a et b sont deux éléments de $]0, \frac{\alpha}{K^2}]$ tels que $a<b$ alors : $\forall t\in[a,b]$, $0\leq\psi(t)<1$.

De plus $1-\frac{\alpha^2}{K}<0$. ψ est strictement décroissante et continue sur $[0,\frac{\alpha}{K^2}]$, $\psi(0)=1$ et $\psi(\frac{\alpha}{K^2})<0$. $\exists b\in]0,\frac{\alpha}{K^2}]$, $\psi(b)=0$.
En prenant a dans $]0,b[$ on a bien $a<b$ et $\forall t\in[a,b]$, $0\leq\psi(t)<1$.

Ainsi $\exists(a,b)\in\mathbb{R}_+^2$, $a<b$ et $\forall t\in[a,b]$, $0\leq\psi(t)<1$.

ψ est continue sur le segment $[a,b]$.

c) $\forall p\in\mathbb{N}$, $\|X_{p+1}-A\|^2\leq\psi(\lambda_p)\|X_p-A\|^2\leq\max_{u\in[a,b]}\psi(u)\,\|X_p-A\|^2$

Posons $\beta'=\sup_{u\in[a,b]}\psi(u)$. $\beta'\in[0,1[$ car $\forall u\in[a,b]$, $0\leq\psi(u)<1$.

$\forall p\in\mathbb{N}$, $\|X_{p+1}-A\|^2\leq\beta'\|X_p-A\|^2$. Posons alors $\beta=\sqrt{\beta'}$. $\beta\in[0,1[$ car $\beta'\in[0,1[$.

Ainsi $\beta\in[0,1[$ et $\forall p\in\mathbb{N}$, $\|X_{p+1}-A\|\leq\beta\|X_p-A\|$.

Une récurrence simple donne alors : $\forall p\in\mathbb{N}$, $\|X_p-A\|\leq\beta^p\|X_0-A\|$.

Or $\lim_{p\to+\infty}\beta^p=0$ car $\beta\in[0,1[$. Ainsi $\lim_{p\to+\infty}\|X_p-A\|=0$

Donc $(X_p)_{p\geq0}$ converge vers A.

d) Soit $p\in\mathbb{N}$. D'après Q2 :
$$df_{X_p}(A-X_p)+\alpha\|X_p-A\|^2\leq f(A)-f(X_p).\ \text{Par conséquent :}$$

$f(X_p)-f(A)\leq df_{X_p}(X_p-A)-\alpha\|X_p-A\|^2$

$f(X_p)-f(A)\leq\ <\text{grad} f(X_p)-\underbrace{\text{grad} f(A)}_{=0}, X_p-A>-\alpha\|X_p-A\|^2$

$f(X_p)-f(A)\underset{\uparrow}{\leq}\|\text{grad} f(X_p)-\text{grad} f(A)\|\,\|X_p-A\|-\alpha\|X_p-A\|^2\leq K\|X_p-A\|^2-\alpha\|X_p-A\|^2$

Cauchy-Schwarz

Ainsi $\forall p\in\mathbb{N}$, $0\leq f(X_p)-f(A)\leq(K-\alpha)\|X_p-A\|^2$.

```
(*    f(x,y)=x*x+2xy+4y*y+x-2y    *)

program convexite_forte;
uses crt;

var i,n:integer;l,x,y,x1:real;

begin
clrscr;
write('Donnez lambda. labda=');readln(l);
writeln;

x:=0;y:=0;
for i:=1 to 200 do
    begin
    x1:=x-l*(2*x+2*y+1);
    y:=y-l*(2*x+8*y-2);
    x:=x1;
    if i mod 10=0 then writeln('x(',i,')=',x,'  ','y(',i,')=',y);
    end;
end.
Donnez lambda. labda=0.1

x(10)=-7.6508216320E-01  y(10)= 4.2887260160E-01
x(20)=-9.4767575627E-01  y(20)= 4.8415749374E-01
x(30)=-9.8834559982E-01  y(30)= 4.9647133155E-01
x(40)=-9.9740416613E-01  y(40)= 4.9921404475E-01
x(50)=-9.9942181894E-01  y(50)= 4.9982494086E-01
x(60)=-9.9987121928E-01  y(60)= 4.9996100834E-01
x(70)=-9.9997131612E-01  y(70)= 4.9999131522E-01
x(80)=-9.9999361111E-01  y(80)= 4.9999806560E-01
x(90)=-9.9999857697E-01  y(90)= 4.9999956914E-01
x(100)=-9.9999968304E-01  y(100)= 4.9999990403E-01
x(110)=-9.9999992940E-01  y(110)= 4.9999997862E-01
x(120)=-9.9999998427E-01  y(120)= 4.9999999524E-01
x(130)=-9.9999999649E-01  y(130)= 4.9999999894E-01
x(140)=-9.9999999921E-01  y(140)= 4.9999999976E-01
x(150)=-9.9999999982E-01  y(150)= 4.9999999995E-01
x(160)=-9.9999999996E-01  y(160)= 4.9999999999E-01
x(170)=-9.9999999999E-01  y(170)= 5.0000000000E-01
x(180)=-9.9999999999E-01  y(180)= 5.0000000000E-01
x(190)=-9.9999999999E-01  y(190)= 5.0000000000E-01
x(200)=-9.9999999999E-01  y(200)= 5.0000000000E-01

Press any key to return to Turbo Pascal
```