Pôle Universitaire Léonard de Vinci — S5 – Année 2006-2007
ESI CS 2505 – Introduction à l'optimisation — Promotion ESILV 2009

Interrogation du 3/10/2006

Corrigé

Exercice I

1. Soit $(x, y) \in \mathbb{R}^2$, on a

$$\nabla f(x,y) = \begin{pmatrix} 3x^2 - 6y \\ 12y - 6x \end{pmatrix}$$

ainsi (x, y) est un point critique de f si et seulement si

$$\begin{cases} 3x^2 - 6y = 0 \\ 12y - 6x = 0 \end{cases}$$

ce qui équivaut à

$$\begin{cases} y = \frac{1}{2}x^2 \\ x^2 - x = 0 \end{cases}$$

ce qui équivaut à

$$\begin{cases} x = 0 \text{ ou } x = 1 \\ y = \frac{1}{2}x^2 \end{cases}$$

ce qui équivaut à $(x, y) \in \{(0, 0), (1, \frac{1}{2})\}$.

2. La matrice hessienne de f en (x, y) est

$$\mathrm{H}f(x,y) = \begin{pmatrix} 6x & -6 \\ -6 & 12 \end{pmatrix}$$

ainsi

$$\mathrm{H}f(0,0) = \begin{pmatrix} 0 & -6 \\ -6 & 12 \end{pmatrix} \quad \text{et} \quad \mathrm{H}f\left(1, \frac{1}{2}\right) = \begin{pmatrix} 6 & -6 \\ -6 & 12 \end{pmatrix}$$

3. Une condition nécessaire pour que $(x, y) \in \mathbb{R}^2$ soit un point-selle de f est que (x, y) soit un point critique de f. Ainsi les deux points candidats à être un point selle de f sont $(0, 0)$ et $(1, \frac{1}{2})$. On a

$$\det \mathrm{H}f(0,0) = -36 < 0$$

donc f admet effectivement un point sele en $(0, 0)$. On a

$$\det \mathrm{H}f\left(1, \frac{1}{2}\right) = 36 > 0$$

donc f admet un extremum local en $(1, \frac{1}{2})$. Ainsi f admet un seul point selle : $(0, 0)$.

4. Le centre du repère est le point-selle que l'on souhaite étudier, il n'est donc pas utile de faire un changement de repère. Les lignes séparatrices de col sont données par

$$^t\begin{pmatrix} x \\ y \end{pmatrix} \mathrm{H}f(0,0) \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

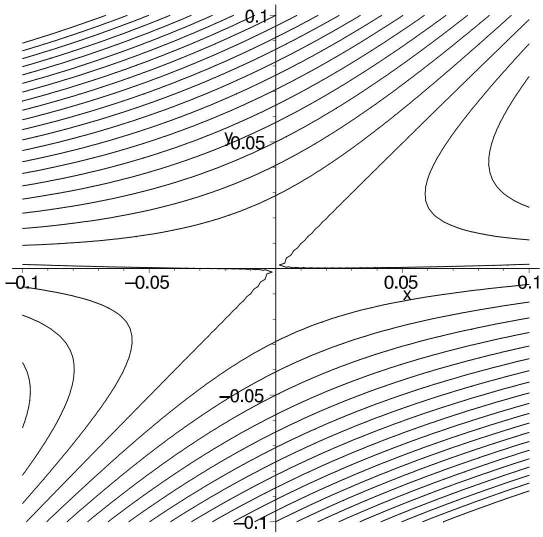
c'est-à-dire

$$-12xy + 12y^2 = 0$$

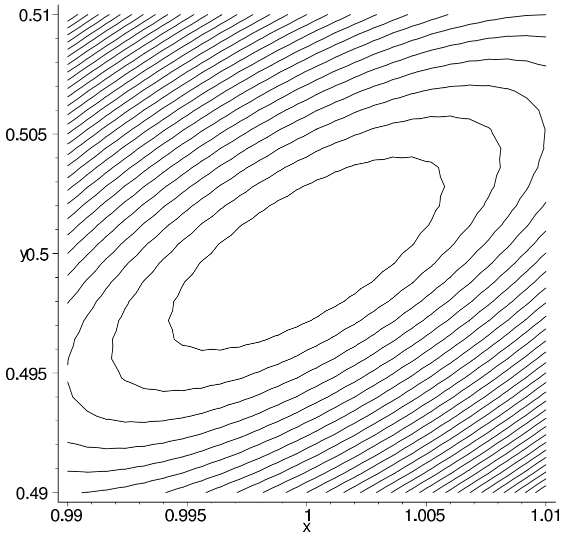
ce qui équivaut à

$$y(y-x) = 0$$

ainsi les lignes séparatrices de col ont pour équations $y = x$ et $y = 0$, cela conduit à esquisser les courbes de f au voisinage de $(0,0)$ de la manière suivante.



Au voisinage de $(1, \frac{1}{2})$ les courbes de niveau sont des courbes concentriques, que l'ont peut représenter en premère approximation par des cercles concentriques. Une étude plus poussée (non demandée dans ce devoir) consisterait à calculer les valeurs propres et vecteurs propres de $\mathrm{H}f(1, \frac{1}{2})$ pour determiner l'orientation de l'ellipse. On obtient les courbes de f au voisinage de $(1, \frac{1}{2})$ de la manière suivante.



Exercice II

1. On a

$$\nabla g = \begin{pmatrix} \frac{\partial g}{\partial x} \\ \frac{\partial g}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial (f^2)}{\partial x} \\ \frac{\partial (f^2)}{\partial y} \end{pmatrix} = \begin{pmatrix} f\frac{\partial f}{\partial x} + \frac{\partial f}{\partial x} f \\ f\frac{\partial f}{\partial y} + \frac{\partial f}{\partial y} f \end{pmatrix} = 2f\nabla f$$

2. Un point critique de f est forcément un point critique de g puisque, en vertu de la question 1, $\nabla f(x,y) = 0$ entraine $\nabla g(x,y) = 0$.

3. La réciproque n'est pas vraie, comme le montre le contre exemple suivant, qui est inspiré par la question 1.

$$f(x,y) = x + y$$
$$g(x,y) = [f(x,y)]^2 = (x+y)^2$$

on a alors

$$\nabla f(x,y) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
$$\nabla g(x,y) = \begin{pmatrix} 2x+2y \\ 2x+2y \end{pmatrix}$$

Le point $(0,0)$ est un point critique de g mais n'est pas un point critique de f.

Exercice III

1. Considérons φ définie de $\mathbb{R}$ dans $\mathbb{R}^2$ par

$$\phi(t) = U_n + t(U_{n+1} - U_n)$$

Cette fonction affine est dérivable et sa dérivée est

$$\frac{d\phi(t)}{dt} = U_{n+1} - U_n$$

Comme f est de classe C^2, la fonction $f \circ t$ est dérivable et sa dérivée est

$$\frac{d(f\circ\phi(t))}{dt} = \left\langle \frac{d\phi(t)}{dt},\, {}^tDf\circ\phi(t) \right\rangle = \langle U_{n+1} - U_n, \nabla f(U_n + t(U_{n+1} - U_n))\rangle$$

D'autre part

$$f(U_{n+1}) - f(U_n) = [f(U_n + t(U_{n+1} - U_n))]_{t=0}^{t=1}$$

donc

$$f(U_{n+1}) = f(U_n) + \int_0^1 \frac{d}{dt} f(U_n + t(U_{n+1} - U_n))dt$$

en vertu de la question précédente, il vient

$$f(U_{n+1}) = f(U_n) + \int_0^1 \langle \nabla f(U_n + t(U_{n+1} - U_n)), U_{n+1} - U_n\rangle\, dt$$

donc

$$f(U_{n+1}) = f(U_n) + \int_0^1 \langle \nabla f(U_n), U_{n+1} - U_n\rangle\, dt$$
$$+ \int_0^1 \langle \nabla f(U_n - t(U_{n+1} - U_n)) - \nabla f(U_n), U_{n+1} - U_n\rangle\, dt$$

d'où finalement

$$f(U_{n+1}) = f(U_n) + \langle \nabla f(U_n), U_{n+1} - U_n\rangle + \int_0^1 \langle \nabla f(U_n + t(U_{n+1} - U_n)) - \nabla f(U_n), U_{n+1} - U_n\rangle\, dt$$

ce qu'il fallait démonter.

2. Par construction de U_{n+1}, il vient

$$\nabla f(U_n) = -\frac{1}{\alpha}(U_{n+1} - U_n)$$

ainsi

$$\langle \nabla f(U_n), U_{n+1} - U_n \rangle = \left\langle -\frac{1}{\alpha}(U_{n+1} - U_n), U_{n+1} - U_n \right\rangle = -M\|U_{n+1} - U_n\|^2$$

d'autre part

$$\langle \nabla f(U_n + t(U_{n+1} - U_n)) - \nabla f(U_n), U_{n+1} - U_n \rangle \leq \|\nabla f(U_n + t(U_{n+1} - U_n)) - \nabla f(U_n)\| \|U_{n+1} - U_n\|$$

ainsi l'égalité établie dans la question précédente donne

$$f(U_{n+1}) \leq f(U_n) - M\|U_{n+1} - U_n\|^2 + \int_0^1 \|\nabla f(U_n + t(U_{n+1} - U_n)) - \nabla f(U_n)\| \|U_{n+1} - U_n\| dt$$

en replaçant $f(U_{n+1})$ par v_{n+1} et $f(U_n)$ par v_n, on obtient l'inégalité voulue.

3. Par hypothèses, ∇f est M-Lipchitzienne, ainsi

$$\|\nabla f(U_n) - \nabla f(U_n + t(U_{n+1} - U_n))\| \leq M\|U_{n+1} - [U_n + t(U_{n+1} - U_n)]\|$$

donc

$$\|\nabla f(U_n) - \nabla f(U_n + t(U_{n+1} - U_n))\| \leq tM\|U_{n+1} - U_n\|$$

ainsi

$$\int_0^1 \|\nabla f(U_n + t(U_{n+1} - U_n)) - \nabla f(U_n)\| \|U_{n+1} - U_n\| dt \leq \int_0^1 M(1-t)\|U_{n+1} - U_n\|^2 dt$$

or $\int_0^1 t dt = \frac{1}{2}$ donc

$$\int_0^1 \|\nabla f(U_n + t(U_{n+1} - U_n)) - \nabla f(U_n)\| \|U_{n+1} - U_n\| dt \leq \frac{M}{2}\|U_{n+1} - U_n\|^2$$

Cette inégalité et celle de la question précédente conduisent à

$$v_{n+1} - v_n \leq -\frac{M}{2}\|U_{n+1} - U_n\|^2$$

4. Ainsi $v_{n+1} - v_n \leq 0$ donc la suite est décroissante. Par ailleurs elle est minorée puisque

$$v_n = f(U_n) \geq \min f$$

La suite (v_n) est donc convergente.

5. Comme $f(U_n)$ converge, toutes ses sous-suites convergent également ; ainsi, la coercivité de f impose que (U_n) ne puisse avoir de sous-suite dont la norme tend vers $+\infty$. Il en résulte que (U_n) est bornée. En vertu du théorème de Bolzano-Weirstrass, la suite (U_n) étant bornée, elle admet au moins une valeur d'adhérence, elle admet donc au moins une sous-suite $(U_{\varphi(n)})$ convergente. Soit Z sa limite. La fonction f étant de classe C^2, la fonction ∇f est continue, donc

$$\nabla f(Z) = \lim_{n\to+\infty} \nabla f(U_{\varphi(n)}) = \lim_{n\to+\infty} \frac{U_{\varphi(n)} - U_{\varphi(n+1)}}{\alpha} = 0$$

Ainsi une sous-suite quelconque de (U_n) converge vers un élément qui annule le gradient de f. Or la fonction f a un minimum global et un seul point critique, donc *un* point critique Z de f est *le* point critique de f est le minimum global de f. Par suite, toutes les sous-suites de (U_n) convergent vers Z, en conséquences (U_n) converge vers le point où f réalise son minimum global.