Pôle Universitaire Léonard de Vinci — Année 2002-2003
CS 202 – Introduction à l'optimisation — Semestre 5

Devoir 2

corrigé

Exercice I

1. Soit i un entier compris entre 1 et n. Les fonctions $\varphi_i\ :\ (a,b)\mapsto y_i - ax_i - b$ et $\psi\ :\ t\mapsto t^2$ sont convexes, donc $f_i=\psi\circ\varphi_i\ :\ (a,b)\mapsto (y_i-ax_i-b)^2$ l'est également. Par suite f est une somme de fonctions convexes ; c'est une fonction convexe.

2. En préliminaire à notre étude, remarquons que dès que deux points du nuage n'ont pas la même abcisse, on la moyenne des carrés est différent du carré de la moyenne.

Le gradient de f s'annulle si et seulement si

$$\begin{cases} -2\left(\sum_{i=1}^n x_iy_i\right)+2a\left(\sum_{i=1}^n x_i^2\right)+2b\left(\sum_{i=1}^n x_i\right)=0 \\ -2\left(\sum_{i=1}^n y_i\right)+2a\left(\sum_{i=1}^n x_i\right)+2bn=0 \end{cases}$$

si et seulement si

$$\begin{cases} b=\bar y - a\bar x \\ \left(\sum_{i=1}^n x_i^2\right)a+\left(\sum_{i=1}^n x_i\right)(\bar y-a\bar x)=\left(\sum_{i=1}^n x_iy_i\right) \end{cases}$$

si et seulement si

$$\begin{cases} a=\frac{\overline{xy}-\bar y.\bar x}{\overline{x^2}-\bar x^2} \\ b=\bar y-\bar x\frac{\overline{xy}-\bar y.\bar x}{\overline{x^2}-\bar x^2} \end{cases}$$

3. La droite des moindres carrés a pour équation

$$y=\frac{\overline{xy}-\bar y.\bar x}{\overline{x^2}-\bar x^2}x+\bar y-\bar x\frac{\overline{xy}-\bar y.\bar x}{\overline{x^2}-\bar x^2}$$

on retrouve les formules données en MF 100 :

$$a=\frac{\operatorname{cov}(x,y)}{\operatorname{var}(x)},\quad b=\bar y-a\bar x$$

4. On a

$$\bar x=1,\quad \bar y=\frac{4}{3},\quad \overline{x^2}=\frac{5}{3},\quad \overline{y^2}=\frac{14}{3},\quad \overline{xy}=\frac{8}{3}$$

Ainsi $a=2$ et $b=-\frac{2}{3}$

Exercice II

1. On procède de manière analogue à l'exercice I, f est la somme de composées de fonctions convexe.

2. On a

$$\nabla f_i(c,b,a) = 2\begin{pmatrix} -y_i + ax_i^2 + bx_i + c \\ x_i(-y_i + ax_i^2 + bx_i + c) \\ x_i^2(-y_i + ax_i^2 + bx_i + c) \end{pmatrix}$$

$$= 2\begin{pmatrix} -y_i + ax_i^2 + bx_i + c \\ -x_i y_i + ax_i^3 + bx_i^2 + cx_i \\ -x_i^2 y_i + ax_i^4 + bx_i^3 + cx_i^2 \end{pmatrix}$$

Or $\nabla f = \sum_{i=1}^n \nabla f_i$ donc

$$\nabla f(a,b,c) = 2n\left[\begin{pmatrix} a\overline{x^2} + b\overline{x} + c \\ a\overline{x^3} + b\overline{x^2} + c\overline{x} \\ a\overline{x^4} + b\overline{x^3} + c\overline{x^2} \end{pmatrix} - \begin{pmatrix} \overline{y} \\ \overline{xy} \\ \overline{x^2y} \end{pmatrix}\right]$$

par suite $\nabla f = 0$ si et seulement si

$$M\begin{pmatrix} c \\ b \\ a \end{pmatrix} = \begin{pmatrix} \overline{y} \\ \overline{xy} \\ \overline{x^2y} \end{pmatrix}$$

Comme M est inversible, la fonction f admet un seul point critique (a,b,c). En outre la fonction est convexe, ce point réalise le minimum global.

3. On a $M_1(-3,-4)$, $M_2(-1,2)$, $M_3(1,2)$ et $M_4(3,0)$. donc

$$\overline{x^4} = (81+1+1+81)/4 = 82/2 = 41, \qquad \overline{x^3} = (-27-1+1+27)/4 = 0$$

$$\overline{x^2} = (9+1+1+9)/4 = 5, \qquad \overline{x} = (-3-1+1+3)/4 = 0$$

$$\overline{x^2y} = (-36+2+2)/4 = -32/4 = -8, \qquad \overline{xy} = (12-2+2)/4 = 3$$

$$\overline{y} = (-4+2+2)/4 = 0$$

Or

$$\begin{pmatrix} 41 & 0 & 5 \\ 0 & 5 & 0 \\ 5 & 0 & 1 \end{pmatrix}^{-1} = \frac{1}{80}\begin{pmatrix} 5 & 0 & -25 \\ 0 & 16 & 0 \\ -25 & 0 & 205 \end{pmatrix}$$

donc

$$(a,b,c) = \frac{1}{80}\begin{pmatrix} 5 & 0 & -25 \\ 0 & 16 & 0 \\ -25 & 0 & 205 \end{pmatrix}\begin{pmatrix} -8 \\ -3 \\ 0 \end{pmatrix} = \frac{1}{80}\begin{pmatrix} -40 \\ -48 \\ 8\times 25 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 3/5 \\ 5/2 \end{pmatrix}$$

Ainsi la parabole cherchée est d'équation $y = -\frac{1}{2}x^2 + \frac{3}{5}x - \frac{5}{2}$.

Exercice III

1. Pour tout entier i dans $[1,n]$, soit $f_i(a_0,\dots,a_p) = (y_i - P(x_i))^2$. La fonction f est la somme des fonctions f_i qui sont des composées de fonctions convexes. Donc f est convexe.

2. On a

$$\nabla f_i(a_0,\ldots,a_p) = -2\begin{pmatrix} y_i - \sum_{j=0}^p a_j x_i^j \\ x_i(y_i - \sum_{j=0}^p a_j x_i^j) \\ x_i^2(y_i - \sum_{j=0}^p a_j x_i^j) \\ \vdots \\ x_i^p(y_i - \sum_{j=0}^p a_j x_i^j) \end{pmatrix}$$

donc

$$\nabla f_i(a_0,\ldots,a_p) = 2\begin{pmatrix} y_i - \sum_{j=0}^p a_j x_i^j \\ x_i y_i - \sum_{j=0}^p a_j x_i^{j+1} \\ x_i^2 y_i - \sum_{j=0}^p a_j x_i^{j+2} \\ \vdots \\ x_i^p y_i - \sum_{j=0}^p a_j x_i^{j+p} \end{pmatrix}$$

donc

$$\nabla f(a_0,\ldots,a_p) = 2\begin{pmatrix} \sum_{i=1}^n y_i - \sum_{j=0}^p \left(a_j \sum_{i=1}^n x_i^j\right) \\ \sum_{i=1}^n x_i y_i - \sum_{j=0}^p \left(a_j \sum_{i=1}^n x_i^{j+1}\right) \\ \sum_{i=1}^n x_i^2 y_i - \sum_{j=0}^p \left(a_j \sum_{i=1}^n x_i^{j+2}\right) \\ \vdots \\ \sum_{i=1}^n x_i^p y_i - \sum_{j=0}^p \left(a_j \sum_{i=1}^n x_i^{j+p}\right) \end{pmatrix}$$

donc

$$\nabla f(a_0,\ldots,a_p) = 2n\begin{pmatrix} \overline{y} - \sum_{j=0}^p a_j \overline{x^j} \\ \overline{xy} - \sum_{j=0}^p a_j \overline{x^{j+1}} \\ \overline{x^2 y} - \sum_{j=0}^p a_j \overline{x^{j+2}} \\ \vdots \\ \overline{x^p y} - \sum_{j=0}^p a_j \overline{x^{j+p}} \end{pmatrix}$$

ainsi $\nabla f(a_0,\ldots,a_p) = 0$ équivaut à

$$\begin{pmatrix} \sum_{j=0}^p a_j \overline{x^j} \\ \sum_{j=0}^p a_j \overline{x^{j+1}} \\ \sum_{j=0}^p a_j \overline{x^{j+2}} \\ \vdots \\ \sum_{j=0}^p a_j \overline{x^{j+p}} \end{pmatrix} = \begin{pmatrix} \overline{y} \\ \overline{xy} \\ \overline{x^2 y} \\ \vdots \\ \overline{x^p y} \end{pmatrix}$$

ce qui équivaut à

$$\begin{pmatrix} 1 & \overline{x} & \cdots & \overline{x^p} \\ \overline{x} & \overline{x^2} & \cdots & \overline{x^{p+1}} \\ \vdots & \vdots & & \vdots \\ \overline{x^p} & \overline{x^{p+1}} & \cdots & \overline{x^{2p}} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} \overline{y} \\ \overline{xy} \\ \overline{x^2y} \\ \vdots \\ \overline{x^py} \end{pmatrix}$$

Notons

$$M = \begin{pmatrix} 1 & \overline{x} & \cdots & \overline{x^p} \\ \overline{x} & \overline{x^2} & \cdots & \overline{x^{p+1}} \\ \vdots & \vdots & & \vdots \\ \overline{x^p} & \overline{x^{p+1}} & \cdots & \overline{x^{2p}} \end{pmatrix}$$

et supposons que M est inversible, alors

$$\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} 1 & \overline{x} & \cdots & \overline{x^p} \\ \overline{x} & \overline{x^2} & \cdots & \overline{x^{p+1}} \\ \vdots & \vdots & & \vdots \\ \overline{x^p} & \overline{x^{p+1}} & \cdots & \overline{x^{2p}} \end{pmatrix}^{-1} \begin{pmatrix} \overline{y} \\ \overline{xy} \\ \overline{x^2y} \\ \vdots \\ \overline{x^py} \end{pmatrix} \tag{1}$$

ce qui donne le point critique de f. Comme f est convexe ce point critique est le minimum global de f, ainsi le polynôme cherché est $\sum_{j=0}^{p} a_j x^j$ avec les a_j définis par (1)

3. Lorsque $p = 2$ on trouve

$$M = \begin{pmatrix} 1 & \overline{x} & \overline{x^2} \\ \overline{x} & \overline{x^2} & \overline{x^3} \\ \overline{x^2} & \overline{x^3} & \overline{x^4} \end{pmatrix}$$

est on donc ramené à l'exercice 2. Lorsque $p = 1$ on trouve

$$M = \begin{pmatrix} 1 & \overline{x} \\ \overline{x} & \overline{x^2} \end{pmatrix}$$

on a

$$M^{-1} = \frac{1}{\overline{x^2} - \overline{x}^2} \begin{pmatrix} \overline{x^2} & -\overline{x} \\ -\overline{x} & 1 \end{pmatrix} = \frac{1}{\mathrm{var}(x)} \begin{pmatrix} \overline{x^2} & -\overline{x} \\ -\overline{x} & 1 \end{pmatrix}$$

en multipliant par $\begin{pmatrix} \overline{y} \\ \overline{xy} \end{pmatrix}$ il vient alors

$$a = \frac{\mathrm{cov}(x, y)}{\mathrm{var}(x)}, \quad b = \overline{y} - a\overline{x}$$

ce qui correspond aux résultats trouvés dans l'exercice I.