Pôle Universitaire Léonard de Vinci — Année 2001-2002
CS 301 – Optimisation — Semestre 5

Devoir 2

correction

Exercice I

1. On a

$$\begin{aligned}\nabla f(a,b) &= \begin{pmatrix} \sum_{i=1}^n 2(-x_i)(y_i - ax_i - b) \\ \sum_{i=1}^n -2(y_i - ax_i - b) \end{pmatrix} \\ &= \begin{pmatrix} 2\left(\sum_{i=1}^n x_i y_i\right) - 2a\left(\sum_{i=1}^n x_i^2\right) - 2b\left(\sum_{i=1}^n x_i\right) \\ -2\left(\sum_{i=1}^n y_i\right) + 2a\left(\sum_{i=1}^n x_i\right) + 2bn \end{pmatrix}\end{aligned}$$

et

$$Hf(a,b) = \begin{pmatrix} 2\left(\sum_{i=1}^n x_i^2\right) & 2\left(\sum_{i=1}^n x_i\right) \\ 2\left(\sum_{i=1}^n x_i\right) & 2n \end{pmatrix}$$

la trace de cette matrice est positive, montrons que le determinant est positif. Soit (P_n) la proposition $n\left(\sum_{i=1}^n x_i^2\right) \geq \left(\sum_{i=1}^n x_i\right)^2$.

- (P_1) est vraie
- Supposons (P_n) vraie, alors $n\left(\sum_{i=1}^n x_i^2\right) \geq \left(\sum_{i=1}^n x_i\right)^2$ donc

$$n\left(\sum_{i=1}^n x_i^2\right) + \left(\sum_{i=1}^n x_i^2\right) + (n+1)x_{i+1}^2 \geq \left(\sum_{i=1}^n x_i\right)^2 + \left(\sum_{i=1}^n x_i^2\right) + (n+1)x_{i+1}^2$$

par suite

$$\left(\sum_{i=1}^{n+1} x_i^2\right) \geq \left(\sum_{i=1}^n x_i\right)^2 + \left(\sum_{i=1}^n x_i^2 + x_{i+1}^2\right) + x_{i+1}^2$$

or

$$\sum_{i=1}^n x_i^2 + x_{i+1}^2 \geq 2\sum_{i=1}^n x_i x_{i+1}$$

puisque

$$\sum_{i=1}^n x_i^2 - 2\sum_{i=1}^n x_i x_{i+1} + x_{i+1}^2 = \left(\sum_{i=1}^n x_i^2 - x_{i+1}\right)^2 \geq 0$$

donc

$$\left(\sum_{i=1}^{n+1} x_i^2\right) \geq \left(\sum_{i=1}^n x_i\right)^2 + 2\left(\sum_{i=1}^n x_i^2 x_{i+1}^2\right) + x_{i+1}^2$$

donc

$$\left(\sum_{i=1}^{n+1} x_i^2\right) \geq \left(\sum_{i=1}^n x_i + x_{i+1}\right)^2$$

d'ou finalement

$$\left(\sum_{i=1}^{n+1} x_i^2\right) \geq \left(\sum_{i=1}^{n+1} x_i\right)^2$$

ce qui établit (P_{n+1}).

Par suite le determinant et la trace sont positive, les valeurs propres de la matrice hessienne sont positives en tout (a,b) donc f est 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. Cela se démondre de manière analogue à $(P_n) \Rightarrow (P_{n+1})$ de la question précédente. Nous noterons

$$\begin{aligned}
\bar{x} &= \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right) && (1)\\
\bar{y} &= \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right) && (2)\\
\overline{x^2} &= \frac{1}{n}\left(\sum_{i=1}^{n} x_i^2\right) && (3)\\
\overline{y^2} &= \frac{1}{n}\left(\sum_{i=1}^{n} y_i^2\right) && (4)\\
\overline{xy} &= \frac{1}{n}\left(\sum_{i=1}^{n} x_i y_i\right) && (5)\\
& && (6)
\end{aligned}$$

Le gradient de f s'annulle si et seulement si

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

si et seulement si

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

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 :

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

4. On a

$$\begin{aligned}
\bar{x} &= 1\\
\bar{y} &= \frac{4}{3}\\
\overline{x^2} &= \frac{5}{3}\\
\overline{y^2} &= \frac{14}{3}\\
\overline{xy} &= \frac{8}{3}
\end{aligned}$$

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

Exercice II

La courbe de Bézier associée aux points P_0, P_1 et P_2 dont l'origine est P_0 et l'extrémité est P_2 est donnée par la paramétrisation suivante

$$\gamma(t) = B_0^2(t)P_0 + B_1^2(t)P_1 + B_2^2(t)P_2$$

où B_i^k est le i-ième polynôme de Bernstein de degré k. On a donc

$$\gamma(\frac{1}{2}) = B_0^2(\frac{1}{2})P_0 + B_1^2(\frac{1}{2})P_1 + B_2^2(\frac{1}{2})P_2$$

ce qui donne

$$\begin{array}{rcl} \gamma(\frac{1}{2}) & = & B_0^2(\frac{1}{2})P_0 + B_1^2(\frac{1}{2})P_1 + B_2^2(\frac{1}{2})P_2 \\ & = & \frac{1}{4}\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} a \\ b \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 2 \\ 0 \end{pmatrix} \\ & = & \begin{pmatrix} \frac{a}{2} + \frac{1}{2} \\ \frac{b}{2} \end{pmatrix} \end{array}$$

Ainsi la distance de $\gamma(\frac{1}{2})$ à l'origine est

$$\sqrt{\left(\frac{a}{2} + \frac{1}{2}\right)^2 + \left(\frac{b}{2}\right)^2}$$

comme $t \mapsto t^2$ est croissante sur $\mathbb{R}^+$ il est équivalent de maximiser la fonction

$$f \ : \ (a, b) \mapsto \left(\frac{a}{2} + \frac{1}{2}\right)^2 + \left(\frac{b}{2}\right)^2$$

D'autre part la longueur de $[P_0P_1]$ est 1 ce qui est équivalent à $a^2 + b^2 = 1$. Posons

$$g \ : \ (a, b) \mapsto a^2 + b^2 - 1$$

et $L(a, b, \lambda) = f(a, b) - \lambda g(a, b)$. On a

$$L(a, b, \lambda) = \left(\frac{a}{2} + \frac{1}{2}\right)^2 + \frac{b^2}{4} - \lambda(a^2 + b^2 - 1)$$

$$\nabla L(a, b, \lambda) = \begin{pmatrix} \frac{a}{2} + \frac{1}{2} + 2\lambda a \\ \frac{b}{2} + 2\lambda b \\ a^2 + b^2 - 1 \end{pmatrix}$$

Ainsi (a, b, λ) est un point critique de L si et seulement si

$$\left\{ \begin{array}{l} (1 + 4\lambda)a = -1 \\ (1 + 4\lambda)b = 0 \\ a^2 + b^2 = 1 \end{array} \right.$$

si et seulement si

$$\left\{ \begin{array}{l} b = 0 \\ a^2 = 1 \\ (1 + 4\lambda)a = -1 \end{array} \right.$$

c'est-à-dire $a = 1,\ b = 0,\ \lambda = -\frac{1}{2}$ ou $a = -1,\ b = 0,\ \lambda = 0$. En outre

$$L_{XX}(a, b, \lambda) = \begin{pmatrix} \frac{1}{2} + 2\lambda & 0 \\ 0 & \frac{1}{2} + 2\lambda \end{pmatrix}$$

Il apparait ainsi que $P_1 = (-1, 0)$ minimise la distance de $\gamma(\frac{1}{2})$ à l'origine et que $P_1 = (1, 0)$ la maximise.

Exercice III

1. La première des choses a faire est d'exprimer le prix en fonction de la quantité. Une fois cela fait, la fonction objectif est simplement la fonction de profit, soit

$$\Pi(q_1, q_2) = [120 - (q_1 + q_2)][q_1 + q_2] - [2q_1^2 + 3q_1 + 3q_2^2 + q_2]$$

2. Le vecteur qui optimise (i.e., le point critique) la fonction de profit est $Q^* = (q_1^* = 11.5, q_2^* = 24)$. Le prix optimal est 84.

3. Pour montrer que le vecteur maximise localement (en fait ici globalement), il faut montrer que les valeurs propres de la matrice hessienne au point critique sont toutes les deux négatives. Les valeurs propres sont $\lambda_1 = \frac{-10+\sqrt{20}}{2} < 0$ et $\lambda_2 = \frac{-10-\sqrt{20}}{2} < 0$ il s'agit donc bien d'un maximum local.

4. Les dérivées d'ordre supérieur à deux de la fonction sont nulles donc l'erreur commise est nulle.