D.M. 16 : Optimisation

Pour le vendredi 31 mars, Partie I et II TB pour travailler le D.S. (et les concours de cette année)

Notations : n et m sont des entiers naturels vérifiant $1 \leq m \leq n$ et E et F désignent les espaces vectoriels $\mathbb{R}^n$ et $\mathbb{R}^m$ munis de leur structure euclidienne canonique. On note I_E l'application identité de E. Le produit scalaire est noté $(.|.)$ aussi bien dans E que dans F et la norme euclidienne est notée $\|.\|$. $S^+(E)$ désigne l'ensemble des endomorphismes autoadjoints (ou symétriques) positifs de E, $S^{++}(E)$ le sous-ensemble constitué des endomorphismes autoadjoints définis positifs. On rappelle que, si $u \in S^{++}(E)$, alors $\phi_u : (x, y) \mapsto (u(x) \mid y)$ est un produit scalaire sur E.

1 Produit de deux endomorphismes autoadjoints positifs

On se propose dans cette partie de montrer, en plus de quelques généralités, que si u et v sont des éléments de $S^+(E)$, alors $u \circ v$ est diagonalisable et son spectre est inclus dans $\mathbb{R}^+$.

Généralités

Q1) a) Montrer qu'un endomorphisme symétrique de E est dans $S^+(E)$ (resp. $S^{++}(E)$) si et seulement si son spectre est inclus dans $\mathbb{R}^+$ (resp.$\mathbb{R}^{+*}$).

b) Montrer que si $u \in S^{++}(E)$, alors $u^{-1} \in S^{++}(E)$.

c) Soit $u \in S^+(E)$

(i) Montrer qu'il existe un élément s de $S^+(E)$ tel que $u = s^2$.

(ii) En déduire que :

$$\forall x \in E, (u(x) \mid x) = 0 \Rightarrow u(x) = 0 \tag{1}$$

Q2) Preuve du résultat : u et v désignent des éléments de $S^+(E)$.

a) On note u_1 et w les endomorphismes de $\mathrm{Im}(u)$ induits par u et $u \circ v$ respectivement.

(i) Montrer que u_1 est un élément de $S^{++}(\mathrm{Im}(u))$.

(ii) Montrer que w est autoadjoint positif relativement à $\phi_{u_1^{-1}}$ où $\phi_{u_1^{-1}}$ est le produit scalaire sur $\mathrm{Im}(u)$ défini dans les notations.

b) Déduire de la question précédente que l'endomorphisme de $\mathrm{Im}(u \circ v)$ induit par $u \circ v$ est diagonalisable et que son spectre est inclus dans $\mathbb{R}^+$.

Q3) Montrer, à l'aide de (1), que :

$$E = \mathrm{Im}(u \circ v) \oplus \mathrm{Ker}(u \circ v)$$

Q4) Conclure.

Cas particulier

Ici a désigne un élément de $S^{++}(E)$ et f un élément de $\mathcal{L}(E, F)$.

Q5) On note f^* l'adjoint de f.

a) Montrer que : $\mathrm{Ker}(f^*) = [\mathrm{Im}(f)]^\perp$.

b) En déduire que si une suite $(z_k)_k$ d'éléments de $\mathrm{Im}(f)$ est telle que la suite $(f^*(z_k))_k$ converge vers 0 , alors la suite $(z_k)_k$ converge vers 0 .

c) Montrer que :

$$f^* \circ f \in S^+(E)$$

Q6) Montrer que $a^{-1} \circ f^* \circ f$ est un endomorphismes diagonalisable de E et que son spectre est inclus dans $\mathbb{R}^+$.

On note ρ sa plus grande valeur propre.

Q7) Montrer que :

$$\forall x \in E, \|f(x)\|^2 \leq \rho(a(x) \mid x)$$

2 Minimisation d'une fonctionnelle quadratique

Désormais a désigne un élément de $S^{++}(E), b$ est un élément fixé de E et f est un élément non nul de $\mathcal{L}(E,F)$.

On note J l'application de E dans $\mathbb{R}$ définie par :

$$\forall x \in E, J(x) = \frac{1}{2}(a(x) \mid x) - (b \mid x)$$

Minimisation théorique

On considère un sous-espace vectoriel V de E et on s'intéresse à la minimisation de la restriction de J à V.

Q8) Montrer que si $\|x\|$ tend vers $+\infty$ et $x \in V$, alors $J(x)$ tend vers $+\infty$

Q9) Déduire de la question précédente l'existence d'un minimum de la restriction de J à V.

Q10) Soit (x,y) un élément de V^2 tel que $x \neq y$.

a) Montrer que

$$J\left(\frac{x+y}{2}\right) < \frac{J(x)+J(y)}{2}$$

b) En déduire que la restriction de J à V atteint son minimum en un seul point.

Q11) Soit $x \in V$ et $(t,h) \in \mathbb{R} \times V$.

a) Calculer $J(x+th) - J(x)$.

b) En déduire que la restriction de J à V est minimale en x si et seulement si

$$a(x) - b \in V^{\perp} \tag{2}$$

Q12) Ici $n = 3$ et ω est l'élément de E en lequel J est minimale. Pour tout réel $k > J(\omega)$, on note $\mathcal{E}_k$ la surface d'équation $J(x) = k$ et on considère un plan vectoriel Π inclus dans E auquel ω n'appartient pas.

a) La surface $\mathcal{E}_k$ est appelée un ellipsoïde. Montrer qu'il existe une unique valeur de k pour laquelle Π est tangent à la surface $\mathcal{E}_k$

b) Déterminer cette valeur de k si $\mathcal{E}_k$ et Π sont d'équations respectives : $x^2 + 2y^2 + 3z^2 - 2x = k$ et $x + y + z = 0$ relativement à la base canonique de E.

Lagrangien augmenté

Soit r un réel positif et L_r est l'application de $E \times F$ dans $\mathbb{R}$ définie par

$$L_r(x,p) = J(x) + \frac{r}{2}\|f(x)\|^2 + (p \mid f(x)).$$

On dit que (x,p) est un **point selle** *de L_r si, et seulement si,*

$$\forall\, (y,q) \in E \times F,\ L_r(x,q) \leq L_r(x,p) \leq L_r(y,p)$$

ou encore ($L_r(x,\cdot)$ est maximale en p et $L_r(\cdot,p)$ est minimale en x.

Q13) Montrer que les propriétés suivantes sont équivalentes :

(i) $L_r(x,\cdot)$ admet un maximum

(ii) $x \in \mathrm{Ker}(f)$

(iii) $L_r(x,\cdot)$ est constante

Q14) Montrer que :
$L_r(.,p)$ est minimale en x si et seulement si $(a + rf^* \circ f)(x) + f^*(p) = b$.

Q15) a) Montrer que (x,p) est un point selle de L_r si et seulement si

$$\big(x \in K\,\mathrm{Ker}(f) \text{ et } a(x) + f^*(p) = b\big).$$

b) En déduire que la restriction de J à $K\mathrm{er}(f)$ est minimale en x si et seulement si il existe un élément p de F tel que (x,p) est un point selle de L_r.

Q16) Soit (x,p) un point selle de L_r.

a) Montrer que (x,p') est encore un point selle de L_r si et seulement si $p' - p$ est un élément de $[\mathrm{Im}(f)]^{\perp}$.

b) Montrer que, parmi les points selle de L_r du type (x,p'), il en existe un et un seul pour lequel $\|p'\|$ est minimale et le caractériser.

3 Deux algorithmes d'optimisation

On reprend les notations de la partie précédente et on note x l'élément de $\ker(f)$ en lequel la restriction de J à $\ker(f)$ est minimale.

On note également p un élément de F tel que (x, p) est un point selle de L_r.

Et ρ désigne la plus grande valeur propre de $a^{-1} \circ f^* \circ f, p_0$ est fixé dans F et $(\gamma_k)_k$ désigne une suite de réels à valeurs dans $[\alpha, \beta]$, où $0 < \alpha < \beta < 2\left(r + \frac{1}{\rho}\right)$.

On considère la suite $(x_k)_k$ d'éléments de E et la suite $(p_k)_k$ d'éléments de F définies de la façon suivante :

$$\forall k \in \mathbb{N}, L_r\left(., p_k\right) \text{ est minimale en } x_k, \text{ et } p_{k+1} = p_k + \gamma_k f\left(x_k\right).$$

Premier algorithme

Q17) On pose, pour tout k de $\mathbb{N}, y_k = x_k - x$ et $r_k = p_k - p$.

a) Montrer que :

$$r_{k+1} = r_k + \gamma_k f\left(y_k\right) \text{ et } \left(a + r f^* \circ f\right)\left(y_k\right) + f^*\left(r_k\right) = 0.$$

b) Montrer que :
$\|r_k\|^2 - \|r_{k+1}\|^2 = \gamma_k\left[2\left(a\left(y_k\right) \mid y_k\right) + \left(2r - \gamma_k\right)\|f\left(y_k\right)\|^2\right] \geq \alpha\left[2\left(r + \frac{1}{\rho}\right) - \beta\right]\|f\left(y_k\right)\|^2.$

c) En déduire la convergence de la suite $\left(\|r_k\|\right)_k$ puis celle de la suite $(x_k)_k$ vers x.

Second algorithme

Q18) On pose, pour tout entier $k, p_k = \overline{p_k} + \overline{q_k}$ où $(\overline{p_k}, \overline{q_k}) \in \mathrm{Im}(f) \times [\mathrm{Im}(f)]^{\perp}$ et, de même, $p = \bar{p} + \bar{q}$ où $p = (\bar{p}, \bar{q}) \in \mathrm{Im}(f) \times [\mathrm{Im}(f)]^{\perp}$.

a) Montrer que la suite $(\overline{q_k})_k$ est constante.

b) Montrer que :

$$f^*\left(\overline{p_k} - \bar{p}\right) \underset{k \to \infty}{\longrightarrow} 0.$$

c) En déduire que la suite $(p_k)_k$ converge vers $\bar{p} + \overline{q_0}$.

Désormais, on choisit $p_0 = 0$ et la suite $(\gamma_k)_k$ constante égale à γ. Dans ces conditions, la suite $((x_k, p_k))_k$ converge vers $(\bar{x}, \bar{p})$ point selle de L_r avec $\|\bar{p}\|$ minimale.

Q19) Montrer que :

$$\forall k \in \mathbb{N}, x_k = \left(\left[I_E - \gamma\left(a + r f^* \circ f\right)^{-1} \circ f^* \circ f\right]^k \circ \left(a + r f^* \circ f\right)^{-1}\right)(b).$$

Q20) On suppose que, relativement à la base canonique de E, la matrice de a est diagonale, soit $\mathrm{diag}\left(\lambda_1, \ldots, \lambda_n\right)$ avec $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$ et que celle de f, relativement aux bases canoniques de E et F, admet pour coefficient générique

$$f_{i,j} = \begin{cases} 1 & \text{si } i = j \text{ et } i \leq m \\ 0 & \text{sinon} \end{cases}$$

a) Montrer que $I_E - \gamma\left(a + r f^* \circ f\right)^{-1} \circ f^* \circ f$ est un endomorphisme autoadjoint de E qui laisse stables $\mathrm{Ker}(f)$ et $[\mathrm{Ker}(f)]^{\perp}$. On note ψ l'endomorphisme induit sur $[K\,\mathrm{er}(f)]^{\perp}$

b) Déterminer la norme de ψ subordonnée à $\|.\|$ on la note ϵ.

c) r est supposé fixé. Comment choisir γ pour que ϵ soit minimal ? Quelle est alors sa valeur ?

d) Quelle est alors l'influence de r sur la rapidité de convergence de la suite $(x_k)_k$?

Q21) On se place toujours dans les bases canoniques de E et F et on se donne les matrices A, B et F de a, b et f par leur coefficient générique :

$$a_{i,j} = \begin{cases} i & \text{si } i = j \\ 1 & \text{sinon} \end{cases}, \quad b_i = 1, \quad f_{i,j} = \begin{cases} 1 & \text{si } i + j = m + 1 \\ 0 & \text{sinon} \end{cases}$$

a) Montrer que a est effectivement un endomorphisme de E défini positif.

b) Écrire une procédure effectuant lorsqu'on choisit $\gamma = 2r$, le calcul de X_k, matrice de x_k relativement à la base canonique de E (on supposera n, m et r définis numériquement mais on définira les matrices A, B et F).