.

Optimisation avec ou sans contraintes

Eléments de théorie générale

Exercice $\boxed{1}$ $\boxed{\textbf{RàR}}$

1. Soit S une matrice symétrique d'ordre $n \in \mathbb{N}^*$ à coefficients réels vérifiant :

$$\exists r > 0 \quad h \in \mathcal{B}(O_n; r) \Rightarrow \ {}^t hSh \geq 0$$

Soit alors $r > 0$ dont l'existence est assurée par hypothèse. Donnons-nous $x \in \mathbb{R}^n$ et posons successivement : $||x|| = \alpha > 0$ puis $q = \frac{2\alpha}{r} > 0$ et enfin $h = \frac{1}{q}x \in \mathbb{R}^n$.

On observe alors que $||h|| = \left|\frac{1}{q}\right| ||x|| = \frac{r}{2\alpha}\alpha = \frac{r}{2} < r$. Ceci permet d'écrire que $h \in \mathcal{B}(O; r)$ et aussi ${}^t hSh \geq 0$ par l'hypothèse faite sur S. Il vient alors :

$$\begin{aligned} & (q)^2 \ {}^t hSh \geq 0 \quad \text{car } q^2 > 0 \\ \Longrightarrow \quad & {}^t (qh)\, S\, (qh) \geq 0 \\ \Longrightarrow \quad & {}^t xSx \geq 0 \end{aligned}$$

2. Soit f une fonction définie, de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{U}$ de $\mathbb{R}^n$, à valeurs réelles.

(a) En posant $x = x_0 + h$ pris au voisinage de x_0 la formule de Taylor s'écrit :

$$f(x) = f(x_0) + \ {}^t h\nabla f(x_0) + \frac{1}{2}\ {}^t h\nabla^2 f(x_0)h + ||h||^2\varepsilon(||h||)$$

Dans la suite, nous écrirons plus facilement $H_f(x) = \nabla^2 f(x)$.

(b) Par le théorème de Schwartz, $H_f(x)$ est symétrique en tout $x \in \mathcal{U}$ puisque $\mathcal{C}^2(\mathcal{U})$ et, en particulier, est donc aussi diagonalisable. En choisissant $x = x_0$ le résultat demeure. On note alors $H_0 = H_f(x_0)$ dans la suite.

Comme f admet un minimum local en x_0 de $\mathcal{U}$ par hypothèse, on a $\nabla f(x_0) = O_n$ car x_0 est alors un point critique (condition nécessaire des extrema locaux) et, par définition, on peut écrire qu'il existe $\eta > 0$ vérifiant :

$$\forall h \in \mathcal{B}(O; \eta) \ \ f(x_0) \leq f(x_0 + h) = f(x) = f(x_0) + \ {}^t h\nabla f(x_0) + \frac{1}{2}\ {}^t h\nabla^2 f(x_0)h + ||h||^2\varepsilon(||h||)$$

en enchaînant avec la formule de Taylor. Il vient alors, par les remarques qui précèdent :

$$\begin{aligned} \forall h \in \mathcal{B}(O_n; \eta)^* \qquad & 0 \leq \frac{1}{2}\ {}^t hH_0h + ||h||^2\varepsilon(||h||) \\ \Longrightarrow \quad & 0 \leq \frac{1}{2}||h||^2 \left({}^t \left(\frac{1}{||h||}h\right) H_0 \left(\frac{1}{||h||}h\right) + 2\varepsilon(||h||)\right) \\ \Longrightarrow \quad & 0 \leq {}^t \left(\frac{1}{||h||}h\right) H_0 \left(\frac{1}{||h||}h\right) + 2\varepsilon(||h||) \quad \text{car } \frac{1}{2}||h||^2 \geq 0 \end{aligned}$$

et en remarquant que le cas $h = O_n$ ne pose aucun problème (correspond à $f(x_0) \leq f(x_0)$)
Nous observons que $\frac{1}{||h||}h \in \partial\mathcal{B}(O_n; 1)$ et que donc, $h \mapsto \ {}^t H_0 h$ est polynômiale donc de classe $\mathcal{C}^0$ sur $\partial\mathcal{B}(O_n; 1)$ fermé, borné de $\mathbb{R}^n$ donc admet un minimum m atteint en u de norme 1 (et un maximum dont on se moque).
Nous obtenons ainsi :

$$\begin{aligned} \forall h \in \mathcal{B}(O_n; \eta)^* \quad & 0 \leq m + 2\varepsilon(||h||) \\ \Longrightarrow \quad & 0 \leq m \quad \text{par passage à la limite : } \lim_{||h|| \to 0} \end{aligned}$$

Ce qui se réécrit aussi :

$$\begin{array}{rl} \forall h \in \mathcal{B}(O_n;\eta)^* & {}^t\left(\frac{1}{||h||}h\right) H_0 \left(\frac{1}{||h||}h\right) \geq 0 \\ \Longrightarrow & ||h||^2 \left({}^t\left(\frac{1}{||h||}h\right) H_0 \left(\frac{1}{||h||}h\right)\right) \geq 0 \quad \text{car } ||h||^2 \geq 0 \\ \Longrightarrow & {}^t h H_0 h = {}^t h \nabla^2 f(x_0) h \geq 0 \end{array}$$

Nous venons de démontrer que $H_0 = \nabla^2 f(x_0)$ est un matrice symétrique qui vérifie :

$$\exists r > 0 \quad h \in \mathcal{B}(O_n;r) \ \Rightarrow \ {}^t h H_0 h \geq 0$$

Donc, d'après 1° on a $\forall h \in \mathbb{R}^n \quad {}^t h \nabla^2 f(x_0) h \geq 0$.

Notons $Sp(H_0) = \{\lambda_1 \dots \lambda_n\}$ des n valeurs propres de H_0 (qui est bien diagonalisable comme vu supra). Soit $\lambda \in Sp(H_0)$ et fixons $u \in E_\lambda$ un vecteur propre non nul.

Il vient :

$$H_0 h = \lambda h \implies {}^t h H_0 h = {}^t h (H_0 h) = \lambda\, {}^t h h = \lambda \sum_{i=1}^{n} h_i^2$$

Comme $\forall i \leq n \quad h_i^2 \geq 0$ on en déduit que ${}^t h H_0 h$ et λ sont de même signe, soit positif. Donc $Sp(H_0) \subset \mathbb{R}_+$

(c) Reprendre la démarche soi-même en adaptant :
- prouver le résultat de 1° dans le cas négatif
- utiliser la formule de Taylor
- considérer qu'il existe $\eta > 0$ vérifiant $\forall h \in \mathcal{B}(O;\eta) \quad 0 \geq {}^t h H_0 h$
- utiliser le fermé borné $\partial\mathcal{B}(O;\eta)$ pour se donner un maximum M des valeurs ${}^t\left(\frac{1}{||h||}h\right) H_0 \left(\frac{1}{||h||}h\right)$
- se moquer du minimum m de cette fonction
- reconstruire l'expression signée $0 \geq {}^t h H_0 h$
- utiliser un vecteur propre non nul d'un valeur propre λ pour obtenir le signe de $\lambda \leq 0$

Exercice $\boxed{2}$ $\boxed{\textbf{RàR}}$ On se donne une fonction f de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{U}$ de $\mathbb{R}^n$ avec $n \in \mathbb{N}^*$ ainsi qu'un point $x_0 \in \mathcal{U}$ supposé critique pour f.

Par souci de simplicité et de clarté, nous noterons abusivement ∇^2 la matrice hessienne de f en x_0.

1. L'hypothèse effectuée sur ∇^2 se traduit par : $\forall x \in \mathbb{R}^n \quad {}^t x \nabla^2 x \geq 0$ avec égalité seulement si $x = O_n$.

(a) Calculons, par la formule de Taylor, en prenant $x \in \mathcal{U}$ et en posant $h = x - x_0$ avec $x \neq x_0$ et donc $h \neq O_n$:

$$\begin{array}{rcl} f(x_0+h) - f(x_0) & = & f(x_0) + {}^t h \nabla f(x_0) + \frac{1}{2}\, {}^t h \nabla^2 f(x_0) h + ||h||^2 \varepsilon(||h||) - f(x_0) \\ & = & {}^t h \nabla f(x_0) + \frac{1}{2}\, {}^t h \nabla^2 f(x_0) h + ||h||^2 \varepsilon(||h||) \\ & = & \frac{1}{2}\, {}^t h \nabla^2 f(x_0) h + ||h||^2 \varepsilon(||h||) \quad \text{car } x_0 \text{ est critique pour } f \\ & = & \frac{1}{2}||h||^2 \left({}^t\left(\frac{1}{||h||}h\right) \nabla^2 \left(\frac{1}{||h||}h\right) + 2\varepsilon(||h||)\right) \end{array}$$

et donc $f(x_0+h) - f(x_0)$ est du même signe que ${}^t\left(\frac{1}{||h||}h\right) \nabla^2 \left(\frac{1}{||h||}h\right) + 2\varepsilon(||h||)$ par positivité de $\frac{1}{2}||h||^2$.

Nous observons que $\frac{1}{||h||}h \in \partial\mathcal{B}(O_n;1)$ et que donc, $h \mapsto {}^t\nabla^2 h$ est polynômiale donc de classe $\mathcal{C}^0$ sur $\partial\mathcal{B}(O_n;1)$ fermé, borné de $\mathbb{R}^n$ donc admet un minimum m atteint en u de norme 1 (et un maximum dont on se moque). Ce minimum est strictement positif d'après l'hypothèse effectuée sur ∇^2 (et on ne s'en moque pas du tout !)

Par ailleurs, $\lim\limits_{||h|| \to 0} \varepsilon(||h||) = 0$ ce qui signifie qu'il existe $r > 0$ tel que $||h|| < r \implies |\varepsilon(||h||)| < \frac{m}{3}$. Ce qui se réécrit :

$$\forall h \in \mathcal{B}(O_n;r) \ \ h \neq O_n \implies -\frac{m}{3} < \varepsilon(||h||) < \frac{m}{3}$$

En le réinjectant dans l'expression étudiée, on a :

$$\forall h \in \mathcal{B}(O_n; r)\ \ h \neq O_n \quad \Longrightarrow \quad \left({}^t\left(\frac{1}{||h||}h\right) \nabla^2 \left(\frac{1}{||h||}h\right) + 2\varepsilon(||h||) \right) \geq m - 2\frac{m}{3} = \frac{m}{3} > 0$$

et clairement, comme $\frac{1}{2}||h||^2 = 0 \iff h = O_n$ on a $f(x_0 + h) - f(x_0) \geq \frac{m}{6}||h||^2 \geq 0$ pour $h \in \mathcal{B}(O_n; r)$ avec égalité seulement lorsque $h = O_n$.

(b) Sous couvert des hypthèses données, il existe $r > 0$ tel que, pour tout $h \in \mathcal{B}(O_n; r)$ on a $f(x_0 + h) - f(x_0) \geq 0$ d'après ce qui vient d'être prouvé. Comme $||x - x_0|| = ||h||$ si l'on a noté $x = x_0 + h$, ce résultat peut se réécrire :

$$\forall x \in \mathcal{B}(x_0; r)\ \ f(x) \geq f(x_0)$$

avec égalité seulement lorsque $x = x_0$. Nous obtenons la définition d'un minimum local, l'aspect *local* étant donné par la boule $\mathcal{B}(x_0; r)$, et la condition *avec égalité seulement lorsque* $x = x_0$ n'étant pas utilie pour cet énoncé.

2. A vos crayons ou stylos... reprenez la démarche et adaptez les sens des inégalités.
3. *Le lecteur est invité à détailler tous les calculs proposés :*

   On continue de noter $\nabla^2 = \nabla^2 f(x_0)$ par simplicité. Les signes des valeurs propres de ∇^2 fourniront alors une condition suffisante pour obtenir un minimum ou un maximum local en x_0 critique. Ecrivons alors :

$$det(\nabla^2 - X \cdot I_2) = \left| \begin{pmatrix} r - X & s \\ s & t - X \end{pmatrix} \right| = X^2 - (r + t)X + rt - s^2$$

   Le discriminant de calcule et se simplifie en $\Delta = (r - t)^2 + 4s^2 \geq 0$ car ∇^2 est symétrique donc diagonalisable et ainsi l'existence de deux valeurs propres (éventuellemet confondues) est avérée. Les racines (donc valeurs propres de ∇^2) s'écrivent alors :

$$\lambda_1 = \frac{1}{2}\left(r + t + \sqrt{(r-t)^2 + 4s^2}\right) \geq \frac{1}{2}\left(r + t - \sqrt{(r-t)^2 + 4s^2}\right) = \lambda_2$$

   - Cherchons une condition pour obtenir $\lambda_1 \geq \lambda_2 > 0$ (valeurs propres positives)

     Il suffit de vérifier que $\frac{1}{2}\left(r + t - \sqrt{(r-t)^2 + 4s^2}\right) = \lambda_2 > 0$ soit que $r + t > \sqrt{(r-t)^2 + 4s^2}$. Ceci demande donc $r + t > 0$ et par voie de conséquence, soit $r > 0$ soit $t > 0$.

     Supposons $r > 0$. On aura, par élévation au carré : $(r + t)^2 > (r - t)^2 + 4s^2$ qui se ramène à $rt - s^2 > 0$. Nous observons que $rt > 0$ est nécessaire et que donc, $t > 0$ aussi et que $rt - s^2 = det(\nabla^2)$.. Nous pouvons donc énoncer le critère suivant d'obtention d'un minimum local :

     Si x_0 est critique pour f de classe $\mathcal{C}^2$ alors f réalise un minimum local en x_0 si $det(\nabla^2 f(x_0)) > 0$ et que $r > 0$

     en observant qu'on remplacerait r (qui est le premier mineur principal de ∇^2) par t (qui est le dernier mineur principal taille 1 de ∇^2) sans problème dans cet énoncé. En particulier, ce sont des mineurs de même signes (positifs).
   - Cherchons une condition pour obtenir $0 > \lambda_1 \geq \lambda_2$ (valeurs propres négatives)

     Il suffit de vérifier que $\frac{1}{2}\left(r + t + \sqrt{(r-t)^2 + 4s^2}\right) = \lambda_1 < 0$ soit que $r+t < -\sqrt{(r-t)^2 + 4s^2}$. Ceci demande donc $r + t < 0$ et par voie de conséquence, soit $r < 0$ soit $t < 0$.

     Supposons $r < 0$. On aura, par élévation au carré : $(-(r + t))^2 = (r + t)^2 > (r - t)^2 + 4s^2$ qui se ramène à $4s^2 < (r + t)^2 - (r - t)^2$ et donc à $0 < rt - s^2$. Nous observons que $rt > 0$ est nécessaire et que donc, $t < 0$ aussi : les deux mineurs principaux de taille 1 sont négatifs tous deux.

     Nous pouvons donc énoncer le critère suivant d'obtention d'un maximum local :

     Si x_0 est critique pour f de classe $\mathcal{C}^2$ alors f réalise un maximum local en x_0 si $r < 0$ et $det(\nabla^2 f(x_0)) > 0$

     en observant qu'on remplacerait r par t sans problème dans cet énoncé de nouveau. En particulier, ce sont des mineurs qui altèrnent en signes.
4. *Application :* On donne $f(x; y) = x^4 + y^4 - 4x^2 + 8xy - 4y^2$, polynômiale donc de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$.

   On peut déterminer les dérivées partielles de f d'ordre 1 et ainsi constituer le gradient :

$$\forall (x; y) \in \mathbb{R}^2 \quad \nabla f(x; y) = (4x^3 - 8x + 8y\ ;\ 4y^3 + 8x - 8y)$$

L'étude des points critiques conduit à résoudre :

$$\begin{cases} x^3 - 2x + 2y &= 0 \\ y^3 - 2y + 2x &= 0 \end{cases} \iff \begin{cases} x^3 - 2x + 2y &= 0 \\ y^3 + x^3 &= 0 \end{cases} \iff \begin{cases} x^3 - 4x &= 0 \\ y &= -x \end{cases} \iff \begin{cases} x(x^2 - 4) &= 0 \\ y &= -x \end{cases}$$

Les points critiques sont donc $C_0 = (0;0)$; $C_1 = (2;-2)$ et $C_2 = -C_1$. On détermine enfin la matrice hessienne générale en $(x;y) \in \mathbb{R}^2$ par calculs des dérivées partielles d'ordre 2 :

$$\forall (x;y) \in \mathbb{R}^2 \quad \nabla^2 f(x;y) = \begin{pmatrix} 12x^2 - 8 & 8 \\ 8 & 12y^2 - 8 \end{pmatrix}$$

- Etude en C_1 et C_2.

  Les calculs fournissent dans les deux cas $\nabla^2 f(C_i) = \begin{pmatrix} 40 & 8 \\ 8 & 40 \end{pmatrix}$.

  On a $r = 40 > 0$ et $det(\nabla^2 f(C_i)) = 1600 - 64 > 0$ donc, d'après les critères établis dans cet exercice, f réalise en chaque C_i un minimum local.
- Etude en $C_0 = (0;0)$.

  La matricienne hessienne s'écrit alors $\nabla^2 f(0;0) = \begin{pmatrix} -8 & 8 \\ 8 & -8 \end{pmatrix}$ de déterminant nul donc on ne peut pas appliquer ce critère.

  On va juste observer que $\forall t \in \mathbb{R}^* \quad f(t;t) = 2t^4 > 0$ et $\forall t \in \mathbb{R} \quad f(t;0) = t^4 - 4t^2 = -4t^2 + o(t^2)$ et comme $f(t;0) = -4t^2 + o(t^2) < 0$ au voisinage de 0 pour $t \neq 0$, il vient que f réalise en $(0;0) = C_0$ un point-selle (ou point-col)