

M1 IM/MPA : TD Optimisation et Éléments Finis

Feuille 1 : Formalisation de problèmes d'optimisation et premiers exemples.

Exercice 1

Design d'un building.

Pour économiser les coûts de l'énergie lors du chauffage ou la climatisation, un architecte envisage de construire un immeuble rectangulaire partiellement enterré. La superficie totale de tous les sols est souhaitée supérieure à 20000 m^2 . On ne souhaite qu'un appartement par étage (un appartement peut être dans la partie enterrée). Les largeurs et longueurs limites des étages de l'immeuble sont de 50 m . Après concertation des experts, il a été décidé que le ratio entre longueur et largeur du plan doit être égal au nombre d'or φ et que chaque appartement doit être haut de 3.5 m . Les coûts de chauffage et de climatisation sont estimés par an à 100 euros par m^2 de la surface exposée de l'immeuble. On supposera que le coût d'excavation est proportionnel au volume à excaver. Le propriétaire a précisé que les coûts annuels ne doivent pas excéder 225000 euros. Formuler le problème de la détermination de la dimension de l'immeuble pour minimiser les coûts d'excavation.

On note

- d la profondeur de la partie de l'immeuble sous terre,
- h la hauteur de la partie de l'immeuble à l'air libre,
- l la longueur horizontale de l'immeuble,
- w la largeur de l'immeuble,
- n le nombre d'appartements.

1. Faire un dessin simple où l'on fera apparaître les quantités d, h, l, w .
2. Exprimer les variables d'optimisation et la fonction coût à minimiser.
3. Exprimer toutes les contraintes extraites du texte.
4. Formuler le problème d'optimisation complet.
5. Est-ce un problème d'optimisation en dimension finie ? Avec ou sans contraintes ? quadratique ?

Exercice 2

Un problème de portfolio.

Un manager de portfolio pour une compagnie d'investissement cherche à prendre des décisions d'investissement de telle façon à ce que les investisseurs aient au moins 10% de rentabilité en minimisant les risques majeurs de pertes. Pour les 6 dernières années, les rentabilités annuelles de 4 types d'investissements majeurs sont données par le tableau suivant :

Année	1	2	3	4	5	6	Moyenne
Blue chip (grosses sociétés cotées)	18.24	12.12	15.23	5.26	2.62	10.42	10.6443
Technologie	12.24	19.16	35.07	23.46	-10.62	-7.43	11.98
Biens immobiliers	8.23	8.96	8.35	9.16	8.05	7.29	8.34
Obligations	8.12	8.26	8.34	9.01	9.11	8.95	8.6317

Le manager du portfolio doit décider quel pourcentage du capital total investir dans chaque type d'investissement (tout le capital sera investit). L'objectif est de minimiser les risques de perte. Une mesure du risque est le montant de fluctuation de la rente par rapport à sa valeur moyenne. On définit la variance de l'investissement j comme suit :

$$\nu_{jj} = \frac{1}{n} \sum_{k=1}^n (r_{jk} - \mu_j)^2, \quad (1)$$

où n est le nombre total d'observations, r_{jk} la rentabilité de l'investissement j pour la k -ième observation (l'année dans l'exemple) et μ_j la valeur moyenne de l'investissement j , où $j \in \{1, \dots, p\}$ avec $p \in \mathbb{N}^*$ le nombre de type d'investissement. La variance mesure donc le risque au sein d'un même investissement.

Pour mesurer le risque entre les différents types d'investissements, on définit la co-variance entre 2 investissements i et j ($(i, j) \in \{1, \dots, p\}^2$), comme suit :

$$\nu_{ij} = \frac{1}{n} \sum_{k=1}^n (r_{ik} - \mu_i)(r_{jk} - \mu_j). \quad (2)$$

On définit alors la matrice de covariance, comme $V = (\nu_{ij})_{(i,j) \in \{1, \dots, p\}^2}$, et le risque R comme $R = {}^t x V x$, avec $x \in \mathbb{R}^p$.

1. Donner les 4 variables d'optimisation.
2. Ecrire le problème d'optimisation complet.
3. Est-ce un problème d'optimisation quadratique ? en dimension finie ? avec ou sans contraintes ?

Exercice 3

Produit scalaire euclidien et matrices.

Soit A une matrice de $M_n(\mathbb{R})$. On rappelle que l'on peut écrire sous forme matricielle le produit scalaire euclidien sur \mathbb{R}^n : pour tout $(u, v) \in \mathbb{R}^n \times \mathbb{R}^n$, $\langle u, v \rangle = {}^t uv = {}^t vu = \langle v, u \rangle$. On notera $\|\cdot\|$ la norme canoniquement associée à ce produit scalaire.

1. Montrer que pour tout $(u, v) \in \mathbb{R}^n \times \mathbb{R}^n$, $\langle Au, v \rangle = \langle u, {}^t A v \rangle$.
2. Qu'en déduire si A est symétrique ?
3. Montrer que, si A est symétrique, alors pour tout $(u, v) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$\langle A(u + v), u + v \rangle = \langle Au, u \rangle + \langle Av, v \rangle + 2\langle Au, v \rangle.$$

4. Comment montre-t-on qu'une matrice symétrique $A \in M_n(\mathbb{R})$ est définie positive ?
5. Soit A une matrice symétrique définie positive et b un vecteur de \mathbb{R}^n . On note $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ les n valeurs propres de A rangées par ordre croissant. Montrer que pour tout $w \in \mathbb{R}^n$, $\langle Aw, w \rangle \geq \lambda_1 \|w\|^2$ et que pour tout $w \in \mathbb{R}^n$, $\|Aw\| \leq \lambda_n \|w\|$.
- On pourra utiliser la décomposition de w sur une base de vecteurs propres de A .*

On se donne maintenant une matrice rectangulaire $A \in M_{np}(\mathbb{R})$ (avec $(n, p) \in (\mathbb{N}^*)^2$, avec $n \geq p$). On suppose que A est de rang maximal p et donc que l'application linéaire canoniquement associée à A est injective.

6. Montrer que si pour $x \in \mathbb{R}^p$, ${}^t A Ax = 0$, alors $x = 0$. On pourra étudier $\langle {}^t A Ax, x \rangle$.
7. En déduire que ${}^t A A \in M_{pp}(\mathbb{R})$ est inversible.

Exercice 4

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto -(x - 4)^2 - (y - 4)^2$ et $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $(x, y) \mapsto (g_1(x, y) = x + y - 4, g_2(x, y) = x + 3y - 9)$. On note

$$K := \{(x, y) \in \mathbb{R}^2, g_1(x, y) \leq 0, g_2(x, y) \leq 0\}.$$

On s'intéresse au problème de maximisation suivant :

$$\max_{(x,y) \in K} f(x, y).$$

Montrer que ce problème admet une solution.

Exercice 5

Soit $d \in \mathbb{N}^*$. On note $\langle \cdot, \cdot \rangle$ le produit scalaire canonique euclidien sur \mathbb{R}^d . On considère la fonctionnelle

$$\mathcal{J} : \mathbb{R}^d \rightarrow \mathbb{R}, v \mapsto \frac{1}{2} \langle Av, v \rangle + \langle b, v \rangle,$$

avec A une matrice symétrique définie positive et b un vecteur de \mathbb{R}^d . On note $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_d$ les d valeurs propres de A rangées par ordre croissant.

1. Écrire mathématiquement le problème d'optimisation consistant à minimiser la fonctionnelle \mathcal{J} sur \mathbb{R}^d . Est-ce un problème d'optimisation quadratique ? en dimension finie ? avec ou sans contraintes ?
2. Calculer le gradient et la Hessienne de \mathcal{J} en tout point de \mathbb{R}^d .
3. Montrer que \mathcal{J} est α -convexe, avec $\alpha > 0$ à déterminer.
4. Qu'en déduire sur le problème d'optimisation considéré en 1.

5. On se donne les fonctions suivantes définies sur \mathbb{R}^2 :

$$f_1(x, y) = x^2 + xy + 2y^2 + 3y,$$

$$f_2(x, y) = x^2 + xy + 2y^2,$$

$$f_3(x, y) = x^2 + xy - 2y^2$$

et définies sur \mathbb{R}^3

$$f_4(x_1, x_2, x_3) = x_1^2 + x_1x_2 - 2x_1x_3 - 17x_2^2 - x_3x_2 + 7x_3^2,$$

$$f_5(x_1, x_2, x_3) = x_1^2 + x_1x_2x_3 - 2x_1x_3 - 17x_2^2 - x_3x_2 + 7x_3^2,$$

$$f_6(x_1, x_2, x_3) = x_1^2x_2 + x_1x_2 - 2x_1x_3 - 17x_2^2 - x_3x_2 + 7x_3^2,$$

$$f_7(x_1, x_2, x_3) = x_1^2 + x_1x_2 - 2x_1x_3 - 17x_2^2 - x_3x_2 + 7x_3^2 + \frac{1}{2}x_1 + 4x_2,$$

a) Parmi les fonctions définies ci-dessus, lesquelles sont des formes quadratiques ? des fonctionnelles quadratiques ?

b) Pour chaque forme quadratique ou fonctionnelle quadratique, identifier les éléments pertinents parmi : $n \in \mathbb{N}^*$, $A \in M_n(\mathbb{R})$ symétrique, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ tel que $\forall x \in \mathbb{R}^n$, $f(x) = \frac{1}{2}x^t Ax + b^t x + c$.

c) Parmi les fonctionnelles quadratiques, identifier celles qui sont convexes ou α -convexes avec $\alpha > 0$.

Exercice 6

Approcher un jeu de valeurs.

On cherche à trouver une surface qui approche au mieux le jeu de valeurs observé suivant :

Numéro du Point	x	y	z_{obs}
1	0	1	1.26
2	0.25	1	2.19
3	0.5	1	0.76
4	0.75	1	1.26
5	1	2	1.86
6	1.25	2	1.43
7	1.5	2	1.29
8	1.75	2	0.65
9	2	2	1.6

Pour trouver une surface qui convienne, on commence par partir d'un a priori sur la forme de cette dernière. Ici, on considère la forme générale :

$$z = c_1x^2 + c_2y^2 + c_3xy, \quad (3)$$

avec $(c_1, c_2, c_3) \in \mathbb{R}^3$.

Le but est maintenant de déterminer les valeurs optimales des paramètres c_1, c_2, c_3 pour minimiser la somme des carrés des erreurs entre les valeurs de z observées et les valeurs de z calculées en utilisant

la forme générale.

1. Déterminer les variables d'optimisation.
2. Déterminer la fonction objectif (ou fonction coût).
3. Écrire le problème d'optimisation. On pourra noter $z_{obs,i}$ la i -ème observation et (x_i, y_i) les points qui y sont associés pour $i \in \{1, \dots, 9\}$.
4. Est-ce un problème d'optimisation quadratique ? en dimension finie ? avec ou sans contraintes ?
5. Montrer que pour tout $c \in \mathbb{R}^3$, $\mathcal{J}(c) = \|\mathbf{z}_{\text{obs}} - A\mathbf{c}\|^2$, où $\mathbf{z}_{\text{obs}} = \begin{pmatrix} z_{obs,1} \\ z_{obs,2} \\ \vdots \\ z_{obs,9} \end{pmatrix}$ et $\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$ et déterminer A . **Un tel problème est dit problèmes aux moindres carrés.**
6. Retrouver que la fonction coût associée, que l'on notera \mathcal{J} , est quadratique.

SOLUTIONS DES EXERCICES

Solution exercice 1

1. Voir figure ci-dessous.
2. Les variables d'optimisation sont : $(n, l, w, h, d) \in \mathbb{R}^5$.

La fonction coût est donnée par : $\mathcal{J} : \mathbb{R}^5 \rightarrow \mathbb{R}$, $u = (n, l, w, h, d) \mapsto \mathcal{J}(n, l, w, h, d) = dlw$ définie sur \mathbb{R}^5 .

3. Il y a plusieurs contraintes qui s'écrivent :

$$100(2(hl + wl) + lw) \leq 225000 \quad (4)$$

$$nlw \geq 20000 \quad (5)$$

$$(d + h) = 3.5n \quad (6)$$

$$l = \varphi w \quad (7)$$

$$l \leq 50 \quad (8)$$

$$w \leq 50 \quad (9)$$

$$l, w, h, d \geq 0 \quad (10)$$

$$n \geq 1, \text{ entier} \quad (11)$$

On peut aussi dire que l'ensemble des contraintes est donné l'ensemble K , défini comme l'ensemble

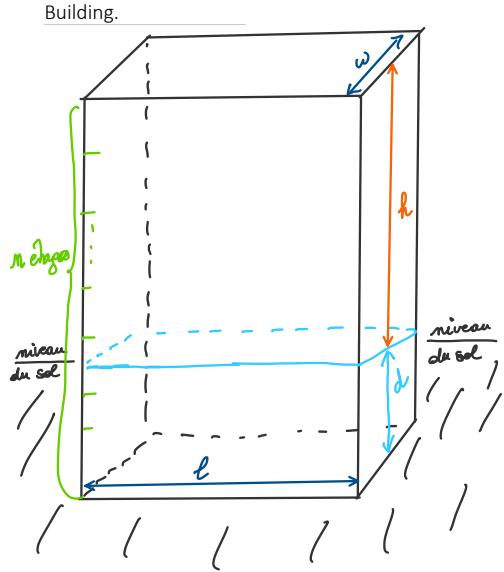


FIGURE 1. Illustration des variables utilisées pour le problème du building.

des points $(n, l, w, h, d) \in \mathbb{R}^5$ tels que

$$100(2(hl + wl) + lw) \leq 225000 \quad (12)$$

$$nlw \geq 20000 \quad (13)$$

$$(d + h) = 3.5n \quad (14)$$

$$l = \varphi w \quad (15)$$

$$l \leq 50 \quad (16)$$

$$w \leq 50 \quad (17)$$

$$l, w, h, d \geq 0 \quad (18)$$

$$n \geq 1, \text{ entier} \quad (19)$$

4. Le problème d'optimisation s'écrit :

Trouver $u^* = (n^*, l^*, w^*, h^*, d^*) \in K$ tel que $\mathcal{J}(u^*) = \inf_{u \in K} \mathcal{J}(u) = \min_{u \in K} \mathcal{J}(u)$.

5. C'est un problème non quadratique, en dimension finie, avec contraintes.

Solution exercice 2

1. Les variables d'optimisation sont les pourcentages du capital total pour chaque type d'investissement : (x_1, x_2, x_3, x_4) , avec $x_i \in \mathbb{R}$, pour $i \in \{1, \dots, 4\}$.

2. On a $n = 6$, les ν_{ij} sont calculables explicitement avec les données et donc la matrice de covariance aussi. Puis la fonction coût est donnée par $J : x = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mapsto^t xVx \in \mathbb{R}$. Tous calculs faits :

$$J(x_1, x_2, x_3, x_4) = 29.0552x_1^2 + 80.7818x_2x_1 - 0.575767x_3x_1 - 3.90639x_4x_1 + 267.344x_2^2 + 0.375933x_3^2 + 0.159714x_4^2 + 13.6673x_2x_3 - 7.39403x_2x_4 - 0.113267x_3x_4. \quad (20)$$

En ce qui concerne les contraintes :

- $\forall i \in \{1, \dots, p\}$, $x_i \in [0, 1]$.
- Ce sont des portions de l'investissement total donc

$$x_1 + x_2 + x_3 + x_4 = 1.$$

- La rentabilité moyenne est souhaitée supérieure à 10%, donc

$$10.6483x_1 + 11.98x_2 + 8.34x_3 + 8.6317x_4 \geq 10.$$

On note

$$K := \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4, (x_1, x_2, x_3, x_4) \in [0, 1]^4, x_1 + x_2 + x_3 + x_4 = 1, 10.6483x_1 + 11.98x_2 + 8.34x_3 + 8.6317x_4 \geq 10\}. \quad (21)$$

Le problème d'optimisation complet s'écrit :

Trouver $u^* = (x_1^*, x_2^*, x_3^*, x_4^*) \in K$ tel que $\mathcal{J}(u^*) = \inf_{u \in K} \mathcal{J}(u) = \min_{u \in K} \mathcal{J}(u)$.

3. C'est un problème d'optimisation quadratique en dimension finie avec contraintes (d'égalité et d'inégalité linéaires).

Solution exercice 3

1. Soient $(u, v) \in \mathbb{R}^n \times \mathbb{R}^n$, on a

$$\langle Au, v \rangle = {}^t(Au)v, \quad (22)$$

$$= {}^t u {}^t A v, \quad (23)$$

$$= {}^t u ({}^t A v), \quad (24)$$

$$= \langle u, {}^t A v \rangle. \quad (25)$$

On remarque cette égalité est généralisable au cas d'une matrice rectangulaire. Dans ce cas, il faut bien préciser les dimensions. Soient $(n, p) \in \mathbb{N}^* \times \mathbb{N}^*$ et $M \in M_{np}(\mathbb{R})$, alors le résultat s'écrit $\forall (u, v) \in \mathbb{R}^p \times \mathbb{R}^n$,

$$\langle Au, v \rangle_{\mathbb{R}^n} = \langle u^t A v \rangle_{\mathbb{R}^p}. \quad (26)$$

2. Si A est symétrique, en utilisant la question précédente, on a pour tout $(u, v) \in \mathbb{R}^d \times \mathbb{R}^d$,

$$\langle Au, v \rangle = \langle u, Av \rangle. \quad (27)$$

En fait, plus généralement, si $\langle Au, v \rangle = \langle u, Av \rangle$ pour tout $(u, v) \in \mathbb{R}^d \times \mathbb{R}^d$, alors A est symétrique (l'application linéaire associée est symétrique, elle est égale à son adjoint, pour ceux qui on déjà vu l'adjoint d'une application linéaire, ici en dimension finie).

3. Il suffit de développer et d'utiliser la question 2.

4. Une matrice A symétrique est définie positive si par définition :

- pour tout $u \in \mathbb{R}^n$, $\langle Au, u \rangle \geq 0$ et
- Si pour $u \in \mathbb{R}^n$, $\langle Au, u \rangle = 0$, alors $u = 0$.

5. Comme A est une matrice symétrique, elle est diagonalisable dans une base orthonormée de vecteurs propres. Il existe donc $(v_i)_{i \in \{1, \dots, n\}}$ base orthonormée de \mathbb{R}^n constituée de vecteurs propres de A : v_i est vecteur propre de A pour la valeur propre λ_i , $i \in \{1, \dots, n\}$. On en déduit que pour tout $u \in \mathbb{R}^n$, il existe α_i tel que

$$u = \sum_{i=1}^n \alpha_i v_i.$$

Et donc,

$$\langle Au, u \rangle = \sum_{i=1}^n \lambda_i \alpha_i^2 \quad (28)$$

Ce qui donne comme $\sum_{i=1}^n \alpha_i^2 = \|u\|^2$,

$$\lambda_1 \|u\|^2 \leq \langle Au, u \rangle \leq \lambda_d \|u\|^2 \quad (29)$$

De la même façon, on montre la seconde inégalité demandée en étudiant directement $\|Au\|^2$ et en exprimant u sur la base de vecteurs propres.

6. Soit $x \in \mathbb{R}^p$, tel que ${}^t A A x = 0$. On a

$$\langle {}^t A A x, x \rangle = \langle A x, A x \rangle, \quad (30)$$

$$= \|A x\|^2. \quad (31)$$

On en déduit donc que $\|A x\| = 0$, et donc $A x = 0$. Et comme A est injective (au sens que son application linéaire canoniquement associée est injective), on en déduit que $x = 0$.

7. Comme ${}^t A A$ est une matrice carrée, on déduit de 3. qu'elle est inversible.

Solution exercice 4

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto -(x-4)^2 - (y-4)^2$ et $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $(x, y) \mapsto (g_1(x, y) = x+y-4, g_2(x, y) = x+3y-9)$. On note

$$K := \{(x, y) \in \mathbb{R}^2, g_1(x, y) \leq 0, g_2(x, y) \leq 0\}.$$

On cherche à résoudre le problème de maximisation suivant :

$$\max_{(x,y) \in K} f(x, y).$$

C'est un problème de maximisation en dimension finie, avec contraintes. Montrons que ce problème admet une solution.

Comme c'est un problème de maximisation, pour appliquer le cours, on se ramène à un problème de minimisation, en passant par $-f$.

On a :

- K est fermé, comme image réciproque d'un fermé par une application continue. On remarque que $K = g^{-1}([-\infty, 0] \times [-\infty, 0])$ et g est continue sur \mathbb{R}^2 et $[-\infty, 0] \times [-\infty, 0]$ est un fermé de \mathbb{R}^2 . De plus K est non vide puisque $(0, 0) \in K$.
- $-f$ est continue sur \mathbb{R}^2 (elle est polynomiale en ses variables). Montrons que $-f$ est infinie à l'infini sur \mathbb{R}^2 . En effet, ici $-f$ est définie sur tout \mathbb{R}^2 et est \mathcal{C}^2 sur \mathbb{R}^2 . On peut par exemple calculer la Hessienne de $-f$ en tout point de \mathbb{R}^2 , on voit que c'est $2Id$ et on a donc $\forall (u, v) \in \mathbb{R}^2 \times \mathbb{R}^2$, $\langle \text{Hess}_f(u), v \rangle = \langle 2Id v, v \rangle = 2\|v\|^2$. Par la caractérisation de l' α -convexité par la Hessienne vue en cours, on en déduit que $-f$ est 2-convexe sur \mathbb{R}^2 . De plus on peut montrer, en revenant à la définition d'un ensemble convexe que K est lui-même un ensemble convexe (à vous de le détailler, cf. par exemple exercices de soutien). On en conclut que $-f$ est donc 2-convexe sur l'ensemble convexe K également. On sait donc que $-f$ est infinie à l'infini sur K , par théorème du cours.

Le théorème I.4.2. du cours s'applique et montre que $-f$ admet au moins un point de minimum sur K et donc f admet au moins un point de maximum sur K .

De plus le fait que $-f$ soit strictement convexe (car 2-convexe) et que K soit convexe permet d'affirmer qu'il existe au plus un point de minimum sur K . Et donc en combinant cela avec le fait qu'il y a au moins un point de minimum, on a la conclusion. Il existe un unique point de minimum de $-f$ sur K .

On pouvait aussi démontrer "à la main" que $-f$ est infinie à l'infini. En effet si on note $v = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$ et $u = \begin{pmatrix} x \\ y \end{pmatrix}$, on trouve

$$f(u) = -\|u - v\|^2.$$

Or

$$|\|u\| - \|v\|| \leq \|u - v\|. \quad (32)$$

Donc $-f(u) \geq \|\|u\| - \|v\|\|^2$. Ce qui donne $f(u) \rightarrow +\infty$, lorsque $\|u\| \rightarrow +\infty$.

Solution exercice 5

1. Le problème d'optimisation s'écrit : Trouver $u^* \in \mathbb{R}^d$, tel que

$$\mathcal{J}(u^*) = \min_{u \in \mathbb{R}^d} \mathcal{J}(u).$$

C'est un problème d'optimisation en dimension finie sans contraintes.

2. On peut répondre à cette question en calculant directement les dérivées partielles. On peut aussi remarquer que si A est la matrice symétrique associée à la forme quadratique f sur \mathbb{R}^d ($f(x) = \frac{1}{2}\langle Ax, x \rangle - \langle b, x \rangle + c$), alors le gradient est donné par $\nabla f(x) = Ax - b$ et la matrice Hessienne est donnée par A . En effet, on peut écrire pour tout $x \in \mathbb{R}^d$, pour tout $h \in \mathbb{R}^d$,

$$f(x + h) = \frac{1}{2}\langle A(x + h), (x + h) \rangle - \langle b, x + h \rangle + c, \quad (33)$$

$$= \frac{1}{2}(\langle Ax, x \rangle + \langle Ah, x \rangle + \langle Ax, h \rangle + \langle Ah, h \rangle) - \langle b, x \rangle - \langle b, h \rangle + c. \quad (34)$$

(35)

On rappelle que par définition du produit scalaire pour tout $(v, w) \in \mathbb{R}^d \times \mathbb{R}^d$, $\langle v, w \rangle = {}^t v w = {}^t w v = \langle w, v \rangle$, et donc $\langle Ah, x \rangle = {}^t (Ah)x = {}^t h {}^t Ax = \langle h, {}^t Ax \rangle = \langle {}^t Ax, h \rangle$. Et donc puisque A est symétrique :

$$f(x + h) = f(x) + \langle Ax - b, h \rangle + \frac{1}{2}\langle Ah, h \rangle. \quad (36)$$

L'application $h \mapsto (Ax, h)$ est une application linéaire sur \mathbb{R}^d (et continue). L'application $h \mapsto (Ah, h)$ est une forme quadratique sur \mathbb{R}^d . On vient donc de montrer que f est différentiable sur \mathbb{R}^d , et on a l'expression de sa différentielle (et de son gradient). On a pour tout $x \in \mathbb{R}^d$, $Df(x) : h \mapsto \langle Ax - b, h \rangle$ et $\nabla f(x) = Ax - b$. De plus, on a en même temps qu'elle est différentiable deux fois et on a l'expression de sa différentielle seconde, que l'on peut caractériser par la matrice Hessienne, ici A . On pourra maintenant utiliser directement ce résultat.

3. Comme \mathcal{J} est $\mathcal{C}^2(\mathbb{R}^d)$ et $Hess(\mathcal{J})(u) = A$ pour tout $u \in \mathbb{R}^d$, on en déduit que pour tout $(u, v) \in \mathbb{R}^d \times \mathbb{R}^d$, $\langle Hess(\mathcal{J})(u)v, v \rangle \geq \lambda_1 \|v\|^2$, par le résultat d'algèbre linéaire revu à l'exercice précédent. En rappelant que \mathbb{R}^d est un ensemble convexe, on déduit que \mathcal{J} est λ_1 -convexe sur \mathbb{R}^d , par caractérisation de l' α -convexité à l'aide de la Hessienne vu en cours. Ici A est définie positive, donc $\lambda_1 > 0$.

4. Comme \mathcal{J} est α -convexe avec $\alpha > 0$ sur \mathbb{R}^d , un résultat du cours nous assure que \mathcal{J} est en particulier "infinie à l'infini". De plus \mathbb{R}^d est un ensemble fermé non vide et \mathcal{J} est continue sur \mathbb{R}^d ; un autre théorème du cours nous assure donc qu'il existe au moins une solution au problème d'optimisation donné en 1. De plus, comme \mathcal{J} est α -convexe sur \mathbb{R}^d avec $\alpha > 0$, on en déduit qu'il existe au plus un point de minimum au problème d'optimisation donné en 1. En combinant ces deux résultats, on en déduit qu'il existe un unique point de minimum au problème d'optimisation donné en 1.

5. a) On a dans le cas général d'une fonctionnelle quadratique :

$$f(x) = \underbrace{\frac{1}{2}\langle Ax, x \rangle}_{\text{partie quadratique}} + \underbrace{\langle b, x \rangle}_{\text{partie linéaire}} + \underbrace{c}_{\text{partie constante}} \quad (37)$$

Si $b = 0$ et $c = 0$, c'est une forme quadratique.

Les applications f_2, f_3, f_4 sont des formes quadratiques, donc aussi des fonctionnelles quadratiques. Les fonctions f_1, f_7 sont des fonctionnelles quadratiques mais pas des formes quadratiques. La fonction f_5 n'est pas une fonctionnelle quadratique à cause du terme cubique $x_1x_2x_3$, la fonction f_6 n'est pas une fonctionnelle quadratique à cause du terme cubique $x_1^2x_2$.

5. b) On cherche à identifier les matrices symétriques A dans chaque cas et les vecteurs b et constante c (si on ne les précise pas ici, c'est qu'ils sont nuls).

Pour $f_1 : A_1 = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$ et $b_1 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$.

Pour $f_2 : A_2 = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$.

Pour $f_3 : A_3 = \begin{pmatrix} 2 & 1 \\ 1 & -4 \end{pmatrix}$.

Pour $f_4 : A_4 = \begin{pmatrix} 2 & 1 & -2 \\ 1 & -34 & -1 \\ -2 & -1 & 14 \end{pmatrix}$.

Pour $f_7 : A_7 = \begin{pmatrix} 2 & 1 & -2 \\ 1 & -34 & -1 \\ -2 & -1 & 14 \end{pmatrix}$ et $b_7 = \begin{pmatrix} \frac{1}{2} \\ 4 \\ 0 \end{pmatrix}$.

5. c) Pour montrer qu'une fonctionnelle définie sur \mathbb{R}^d et \mathcal{C}^∞ sur \mathbb{R}^d est convexe sur \mathbb{R}^d , on peut se concentrer sur la Hessienne de cette fonctionnelle. On dispose de la caractérisation donnée en cours. On dispose donc d'une stratégie qui repose sur l'étude des valeurs propres de la matrice A (qui est symétrique). Si toutes ses valeurs propres sont positives, la fonctionnelle associée est 0-convexe, donc convexe. Si toutes ses valeurs propres sont strictement positives, la fonctionnelle est λ_1 -convexe, avec λ_1 la plus petite valeur propre de la matrice A . Dès qu'une valeur propre est négative (strictement), on ne peut pas montrer de l' α -convexité ($\alpha \geq 0$).

On trouve

- pour A_1, A_2 : symétrique définie positive, il y a α -convexité avec $\alpha > 0$.
- pour A_3 : ni symétrique positive ni symétrique négative (une valeur propre strictement positive, une valeur propre strictement négative),
- Pour A_4, A_7 : il y a une valeur propre négative et deux valeurs propres positives : donc ni symétrique définie positive, ni symétrique définie négative.

Solution exercice 6

1. Les variables d'optimisation sont : $(c_1, c_2, c_3) \in \mathbb{R}^3$.

2. On a un jeu de 9 observations $z_{obs,i}, i \in \{1, \dots, 9\}$ associées chacunes à un couple de valeurs (x_i, y_i) pour $i \in \{1, \dots, 9\}$. Pour chacune de ces 9 observations, on peut calculer la valeur que prédit la forme

générale, i.e. pour chaque $i \in \{1, \dots, 9\}$, on calcule $z_i = c_1x_i^2 + c_2y_i^2 + c_3x_iy_i$. L'erreur entre les valeurs de z observées ($z_{obs,i}$) et les valeurs de z calculées (z_i) sont données par $z_{obs,i} - z_i$. On nous dit que l'on cherche à minimiser la somme des carrés des erreurs. La fonction objectif est donc donnée par :

$$\mathcal{J} : \mathbb{R}^3 \rightarrow \mathbb{R}, (c_1, c_2, c_3) \mapsto \mathcal{J}(c_1, c_2, c_3) = \sum_{i=1}^9 (z_{obs,i} - (c_1x_i^2 + c_2y_i^2 + c_3x_iy_i))^2.$$

Tous calculs faits :

$$J(c_1, c_2, c_3) = 18.7 - 32.8462c_1 + 34.2656c_1^2 - 65.58c_2 + 96.75c_1c_2 + 84c_2^2 - 43.425c_3 + 79.875c_1c_3 + 123c_2c_3 + 48.374c_3^2.$$

3. Le problème de minimisation s'écrit donc :

$$\min_{(c_1, c_2, c_3) \in \mathbb{R}^3} \mathcal{J}(c_1, c_2, c_3).$$

4. C'est un problème d'optimisation non linéaire, quadratique en dimension finie et sans contraintes.

5. En posant $A = \begin{pmatrix} x_1^2 & y_1^2 & x_1y_1 \\ \vdots & \vdots & \vdots \\ x_i^2 & y_i^2 & x_iy_i \\ \vdots & \vdots & \vdots \\ x_9^2 & y_9^2 & x_9y_9 \end{pmatrix} \in M_{9,3}(\mathbb{R})$, on trouve le résultat voulu.

6. On écrit $\mathcal{J}(\mathbf{c}) = \|\mathbf{z}_{obs} - A\mathbf{c}\|_{\mathbb{R}^9}^2$, pour tout $\mathbf{c} \in \mathbb{R}^3$. Cela donne

$$\mathcal{J}(\mathbf{c}) = \|\mathbf{z}_{obs}\|_{\mathbb{R}^9}^2 + \|A\mathbf{c}\|_{\mathbb{R}^9}^2 - 2\langle \mathbf{z}_{obs}, A\mathbf{c} \rangle_{\mathbb{R}^9} \quad (38)$$

$$= \|\mathbf{z}_{obs}\|_{\mathbb{R}^9}^2 + \langle A\mathbf{c}, A\mathbf{c} \rangle_{\mathbb{R}^9} - 2\langle \mathbf{z}_{obs}, A\mathbf{c} \rangle_{\mathbb{R}^9} \quad (39)$$

$$= \|\mathbf{z}_{obs}\|_{\mathbb{R}^9}^2 + \langle {}^t A A \mathbf{c}, \mathbf{c} \rangle_{\mathbb{R}^3} - 2\langle {}^t A \mathbf{z}_{obs}, \mathbf{c} \rangle_{\mathbb{R}^3} \quad (40)$$

$$(41)$$

On reconnaît donc une fonctionnelle quadratique la matrice de la partie quadratique

$$\mathcal{J}(\mathbf{c}) = \underbrace{\frac{1}{2} \langle B\mathbf{c}, \mathbf{c} \rangle}_{\text{partie quadratique}} \underbrace{- \langle g, \mathbf{c} \rangle}_{\text{partie linéaire}} + \underbrace{l}_{\text{partie constante}} \quad (42)$$

étant $B = {}^t A A$, $g = {}^t A \mathbf{z}_{obs}$ et $l = \|\mathbf{z}_{obs}\|^2$.