Exercices d'optimisation et quelques corrigés

1. Optimisation de fonctions, Hessienne, convexité

Exercice 1. Soit l'application $f : \mathbb{R}^2 \to \mathbb{R}$ définie par
$$f(x,y) = x\,y\,e^{-\pi\left(x^2+y^2\right)}$$
Déterminer les points critiques de f ainsi que leur nature : maximum ou minimum local, point-selle, maximum ou minimum global.

Exercice 2. Soit les fonctions f, g et h définies sur $\mathbb{R}^2$ par $f(x,y) = x^4 + y^4$, $g(x,y) = (x-y)^2$ et $h = f - 2\,g$.

(1) Montrer que les fonctions f et g sont convexes sur $\mathbb{R}^2$, mais que $h = f - 2\,g$ n'est ni convexe ni concave sur $\mathbb{R}^2$.
(2) La fonction h admet-elle un minimum ou un maximum sur $\mathbb{R}^2$?
(3) Déterminer les points critiques de h, et préciser leur nature.

Exercice 3. Soit f la fonction définie sur $\mathbb{R}^3$ par $f(x,y,z) = x^4 - 2\,x^2\,y + 2\,y^2 - 2\,y\,z + 2\,z^2 - 4\,z + 5$.

(1) Déterminer les points critiques de f sur $\mathbb{R}^3$
(2) Montrer que l'expression $f(x,y,z)$ peut s'écrire sous forme d'une somme de carrés.
(3) En déduire les extrema de f sur $\mathbb{R}^3$.

Exercice 4. Soit $f : \mathbb{R}^3 \to \mathbb{R}$ donnée par $f(x,y,z) = xy + yz + zx$. Trouver la valeur maximale et la valeur minimale de f sur $B_1 := \{\,(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2 \le 1\,\}$, après avoir démontré qu'elles existent.

Exercice 5. Soit $T \subset \mathbb{R}^2$ l'ensemble
$$T := \left\{\,(x,y) \in \mathbb{R}^2 : x \ge 0, y \ge 0, x+y \le 1\,\right\}.$$
Soit f la fonction donnée par $f(x,y) = -x - 2y - 2xy + \frac{1}{2}x^2 + \frac{1}{2}y^2$.

(1) f est-elle convexe ? concave ?
(2) Démontrer que tout minimum ou maximum local de f sur T se trouve sur la frontière de T.
(3) Démontrer que f admet bien un minimum et un maximum sur T.
(4) Trouver le minimum et le maximum de f sur T.

Solution. 1) La fonction f n'est ni convexe ni concave, puisque elle est C^2 et sa matrice Hessienne est
$$D^2 f(x,y) = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix},$$

de déterminant $\det(D^2 f) = -3 < 0$, ce qui implique que ses valeurs propres sont de signe opposé.

2) Si il y avait un minimum ou maximum local à l'intérieur il devrait annuler ∇f. Or, $\nabla f(x, y) = (-1 - 2y + x, -2 - 2x + y) = 0$ seulement pour $(x, y) = (-\frac{5}{3}, -\frac{4}{3}) \notin C$ donc il n'est pas possible que f ait un minimum ou un maximum local à l'intérieur.

3) La fonction f est continue et C est fermé (par les inégalités larges) et borné (T est inclus dans la boule unité), donc compact. On peut donc utiliser le théorème de Weierstrass.

4) On fait une liste de candidats à l'optimisation : il n'y a pas de points à l'intérieur, ils sont donc sur la frontière. Cette frontière est composée de 6 parties : les trois sommets $(0, 0)$, $(1, 0)$ et $(0, 1)$, et les trois côtés $(0, y)$ avec $y \in]0, 1[$, $(x, 0)$ avec $\in]0, 1[$ et (x, y) avec $x, y \in]0, 1[$ et $x + y = 1$. Sur ce dernier côté nous pouvons appliquer le multiplicateur de Lagrange, en imposant $\nabla f = \lambda (1, 1)$, où $(1, 1)$ est le gradient de la fonction $x + y$. On trouve alors $-1 - 2y + x = \lambda = -2 - 2x + y$, ce qui donne le système

$$\begin{cases} -1 - 2y + x = -2 - 2x + y \\ x + y = 1, \end{cases}$$

qui a pour solution $(x, y) = (\frac{1}{3}, \frac{2}{3})$. Pour les autres côtés il n'est pas nécessaire de faire de multiplicateurs de Lagrange, il suffit d'étudier $f(0, y) = -2y + \frac{1}{2}y^2$, dont la dérivée s'annule pour $y = 2$, c'est-à-dire hors de l'intervalle qui nous intéresse, et $f(x, 0) = -x + \frac{1}{2}x^2$, dont la dérivée s'annule pour $x = 1$, ce qui est aussi hors de l'intervalle d'intérêt (c'est sur son bord). Les points intéressants sont donc $(0, 0)$, $(1, 0)$, $(0, 1)$ et $(\frac{1}{3}, \frac{2}{3})$ et on a

$$f(0, 0) = 0, \quad f(1, 0) = -\tfrac{1}{2}, \quad f(0, 1) = -\tfrac{3}{2}, \quad f(\tfrac{2}{3}, \tfrac{1}{3}) = -\tfrac{11}{6}.$$

Le minimum est donc réalisé en $(\frac{1}{3}, \frac{2}{3})$ et vaut $-\frac{11}{6}$ et le maximum est réalisé en $(0, 0)$ et vaut 0.

Exercice 6. Déterminer les extrema locaux des fonctions suivantes sur $\mathbb{R}^2$

$$\begin{aligned} f_1(x, y) &= x^3 + 3\,x\,y^2 - 15\,x - 12\,y \\ f_2(x, y) &= 3\,x^3 + x\,y^2 - 3\,a\,x\,y \\ f_3(x, y) &= x^4 + y^3/3 - 4\,y - 2 \\ f_4(x, y) &= x^3 + x\,y^2 - x^2\,y - y^3 \end{aligned}$$

Pour chaque fonction, montrer que les extrema locaux ne sont pas globaux.

Exercice 7. Les sous-ensembles de $\mathbb{R}^2$ suivants sont-ils convexes ?

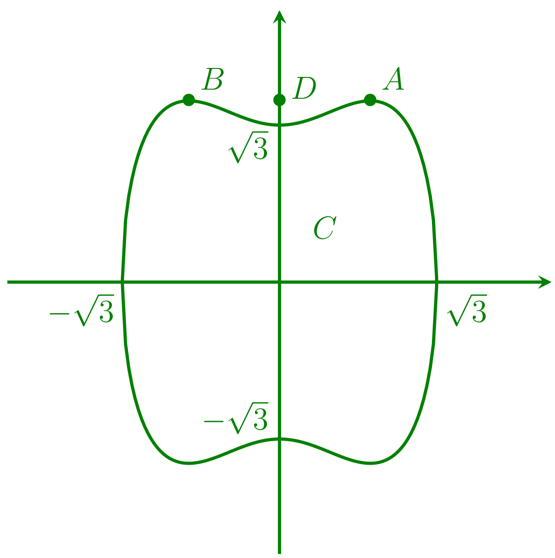
$$A_1 = \left\{ (x,y) \in \mathbb{R}^2 : xy \geq 0 \right\}$$
$$A_2 = \left\{ (x,y) \in \mathbb{R}^2 : x^2 + y^2 - x + y + 1/4 < 0 \right\}$$
$$A_3 = \left\{ (x,y) \in \mathbb{R}^2 : x + 2y + 1 < 0 \text{ ou } y \geq 0 \right\}$$
$$A_4 = \left\{ (x,y) \in \mathbb{R}^2 : x^2 + y^2 - 6x + 4y + 9 < 0 \text{ et } -2 < x + y \leq 2 \right\}$$

Exercice 8. Soit $C \subset \mathbb{R}^2$ l'ensemble donné par

$$C := \left\{ (x,y) \in \mathbb{R}^2 : \left(x^2 - 1\right)^2 + y^2 \leq 4 \right\}.$$

(1) Dessiner l'ensemble C.
(2) S'agit-il d'un ensemble compact ? convexe ?
(3) Considérer la fonction f donnée par $f(x,y) = xy$. Admet-elle un minimum et un maximum sur C ?
(4) Calculer $\inf \{ f(x,y) : (x,y) \in C \}$ et $\sup \{ f(x,y) : (x,y) \in C \}$.

Solution. 1) L'ensemble C est dessiné à la figure suivante.



2) L'ensemble C est bien compact, puisqu'il est fermé (à cause de l'inégalité large dans la définition) et borné (pour tout $(x,y) \in C$ nous avons $|x|, |y| \leq 2$). Il n'est pas convexe, parce que les point $A = (1,2)$ et $B = (-1,2)$ appartiennent à C mais leur point intermédiaire $D = (0,2)$ n'y appartient pas puisque $2^2 + (0^2-1)^2 = 5 > 4$.

3) La fonction f est continue sur un ensemble compact et elle admet donc un minimum et un maximum sur C par le théorème de Weierstrass.

4) Si on calcule ∇f on obtient $\nabla f(x,y) = (y,x)$. La contrainte peut s'écrire sous la forme $g(x,y) \leq 0$ avec $g(x,y) = y^2 + (x^2-1)^2 - 4$ et $\nabla g(x,y) = (4\,x\,(x^2-1)\,, 2y)$. Les deux fonctions f et g sont bien dérivables partout (elles sont C^1). La fonction g

permet d'appliquer le théorème des multiplicateurs de Lagrange parce que $\nabla g = 0$ seulement en $(0,0), (1,0)$ et $(-1,0)$, qui n'appartiennent pas à l'ensemble où $g = 0$ (le bord de C).
Les points candidats à la minimisation et à la maximisation sont donc

- les points où $\nabla f = 0$, c'est-à-dire juste $(0,0)$,
- les points où le système donné par le multiplicateur de Lagrange est satisfait

$$\begin{cases} y = 4\,\lambda\,x\left(x^2-1\right), \\ x = 2\,\lambda\,y, \\ y^2 + \left(x^2-1\right)^2 = 4. \end{cases}$$

Pour résoudre le système on remarque d'abord que λ ne peut pas être nul (sinon on aurait $x = y = 0$ mais la troisième équation n'est pas satisfaite). En multipliant les deux premières équations entre elles après avoir réécrit la deuxième comme $2\,\lambda\,y = x$, on obtient

$$2\,\lambda\,y^2 = 4\,\lambda\,x^2\left(x^2-1\right).$$

Avec la troisième équation, et en divisant par $2\,\lambda \neq 0$, cela donne

$$4 - (x^2-1)^2 = 2x^2(x^2-1).$$

En posant $t = x^2 - 1$, on a $t \geq -1$ et

$$4 - t^2 = 2\,t\,(t+1)\,,$$

qui est une équation d'ordre deux dont les solutions sont

$$t = \frac{-1 \pm \sqrt{13}}{3},$$

mais seulement $t = (-1+\sqrt{13})/3$ satisfait $t \geq -1$. On trouve donc $x = \pm\sqrt{t+1} = \pm\sqrt{(2+\sqrt{13})/3}$. Grâce à la troisième équation du système des multiplicateurs de Lagrange, on obtient $y = \pm\sqrt{4-t^2} = \pm\sqrt{22+2\sqrt{13}}/3$. Les point sur les bords qui sont candidats à l'optimisation sont donc

$$(x,y) = \left(\pm\sqrt{\frac{2+\sqrt{13}}{3}}, \pm\frac{\sqrt{22+2\sqrt{13}}}{3}\right).$$

Les deux choix où x et y ont le même signe donnent une valeur de f égale et positive, les deux avec signes opposés égale et négative, et on compare cela avec $f(0,0) = 0$. On en déduit que

$$\max_C f = \sqrt{\frac{2+\sqrt{13}}{3}}\,\frac{\sqrt{22+2\sqrt{13}}}{3}, \quad \min_C f = -\sqrt{\frac{2+\sqrt{13}}{3}}\,\frac{\sqrt{22+2\sqrt{13}}}{3}.$$

Exercice 9. Soit $C \subset \mathbb{R}^2$ l'ensemble donné par

$$C := \left\{\, (x,y) \in \mathbb{R}^2 : \left(x^2-1\right)^2 + |y| \leq 4 \,\right\}.$$

(1) Dessiner l'ensemble C.

(2) S'agit-il d'un ensemble compact ? S'agit-il d'un ensemble convexe ?
(3) Considérer la fonction f donnée par $f(x,y)=x^2+y$. Admet-elle un minimum et un maximum sur C ?
(4) Calculer $\inf\{\, f(x,y) : (x,y) \in C \,\}$ et $\sup\{\, f(x,y) : (x,y) \in C \,\}$.

Exercice 10. Optimiser les fonctions suivantes sur leurs domaines :

$$f(x,y,z) = x^4 + 2\,y^2 + 3\,z^2 - y\,z - 23\,y + 4\,x - 5$$
$$g(x_1,\dots,x_n) = \sum_{j=1}^{n} x_j \ln\left(\frac{1}{x_j}\right)$$
$$h(x_1,\dots,x_n) = \prod_{j=1}^{n} x_j^{x_j}$$

Exercice 11. Soit $f : \mathbb{R} \to \mathbb{R}$ strictement convexe, $c \in \mathbb{R}$ et $n \in \mathbb{N}$. Sous la contrainte $\sum_{j=1}^{n} x_j = c$, trouver l'infimum et le supremum de

$$\sum_{j=1}^{n} f(x_j).$$

2. Algorithmes d'optimisation et vitesse de convergence

Exercice 12 (Dichotomie)**.** Soit $f \in C^1(]a,b[,\mathbb{R})$ qui ne possède qu'un seul optimum.

(1) Montrer qu'il existe $(a_0,b_0) \in]a,b[^2$ tels que $a_0 < b_0$ et $f'(a_0)$ et $f'(b_0)$ sont de signes différents.
(2) On suppose que $f'(a_0) < 0 < f'(b_0)$ et que l'optimum est le seul point critique et on effectue l'algorithme par récurrence suivant. Pour tout $n \geq 0$, on définit $c_n = \frac{1}{2}\,(a_n + b_n)$, puis on pose

$$(a_{n+1}, b_{n+1}) = \begin{cases} (a_n, c_n) & \text{si } f'(c_n) > 0 \\ (c_n, b_n) & \text{si } f'(c_n) < 0 \\ (c_n, c_n) & \text{si } f'(c_n) = 0 \end{cases}$$

Montrer que l'algorithme converge vers le minimum global de f.
(3) Calculer la vitesse de convergence de l'algorithme.
(4) Proposer un algorithme dans le cas où $f'(b_0) < 0 < f'(a_0)$.

Exercice 13 (Algorithme de la section dorée)**.** Soit $f \in C^0([a,b],\mathbb{R})$ dont le minimum est le seul extremum local dans $]a,b[$. On définit pour $n = 0$, $I_n = [a,b]$. On choisit ensuite $\tau \in]0,1[$ et on effectue l'algorithme suivant.

On définit $c = a + \tau\,(b-a)$ puis $d = c + \tau\,(b-c)$. Si $f(c) < f(d)$, on définit $I_{n+1} = [a,d]$, sinon on définit $I_{n+1} = [c,b]$. On recommence alors l'algorithme sur I_{n+1}.

(1) Pour optimiser l'algorithme, on fait en sorte que les deux possibilités d'intervalle à chaque pas sont de même taille. Trouver alors la valeur de τ.
(2) Montrer que l'algorithme converge vers le minimum de f et calculer sa vitesse de convergence.

Exercice 14. Soit $c > 0$ et $B_1 = \left\{ x \in \mathbb{R}^d, |x|^2 = x_1^2 + x_2^2 + \cdots + x_d^2 \leq 1 \right\}$. Calculer le vitesse de convergence de l'algorithme de gradient à pas fixe pour trouver le minimum de la fonction $f : B_1 \to \mathbb{R}$ définie par $f(x) = |x|^{2+c}$ avec un pas $\alpha < 1/(2+c)$. *On pourra regarder la suite $1/\left|x^{(n)}\right|^c$ où $x^{(n)}$ est la n-ième valeur donnée par l'algorithme.*

Solution. Pour obtenir une expression explicite des pas de l'algorithme, on commence par calculer le gradient de f. On peut écrire $f = g \circ h$ avec $h : \mathbb{R}^d \to \mathbb{R}_+$ définie par $h(x) = |x|^2 = x \cdot x$ et $g : \mathbb{R}_+ \to \mathbb{R}$ définie par $g(r) = r^{1+c/2}$. Comme $\nabla h(x) = \nabla x \cdot x = 2\,x$, on en déduit que

$$\nabla f(x) = g'(h(x))\,\nabla h(x) = \frac{c+2}{2}\left(|x|^2\right)^{\frac{c}{2}}\,2\,x = (c+2)\,|x|^c\,x.$$

Si on écrit $x^{(n)}$ le n-ième point calculé par l'algorithme, on a donc

$$x^{(n+1)} = x^{(n)} - \alpha\,\nabla f(x^{(n)}) = \left(1 - \alpha\,(c+2)\left|x^{(n)}\right|^c\right) x^{(n)}.$$

Si $x^{(0)} = 0$ (le point $0 = (0, 0 \ldots, 0) \in \mathbb{R}^d$) alors c'est un point fixe, l'algorithme est fini. On suppose donc que $x^{(0)} \neq 0$. On s'attend à ce que l'algorithme converge vers le minimum de f en 0. On définit donc la suite $u_n := \left|x^{(n)} - 0\right| = \left|x^{(n)}\right|$, pour laquelle on obtient

$$u_{n+1} = |1 - \alpha\,(c+2)\,u_n^c|\,u_n.$$

Comme $u_0 \in (0,1]$ et $\alpha\,(c+2) < 1$, on déduit par une récurrence immédiate que pour tout $n > 0$, $0 < u_{n+1} \leq u_n < 1$, et vérifie

$$u_{n+1} = (1 - \alpha\,(c+2)\,u_n^c)\,u_n.$$

Donc la suite est bornée et décroissante. En particulier, elle admet une limite $\ell \in [0,1]$ qui vérifie $\ell = |1 - \alpha\,(c+2)\,\ell^c|\,\ell$, ce qui donne $\ell = 0$.

On regarde maintenant sa vitesse de convergence. On regarde

$$\frac{1}{u_{n+1}^c} - \frac{1}{u_n^c} = \frac{1}{u_n^c}\left(\frac{1}{(1 - \alpha\,(c+2)\,u_n^c)^c} - 1\right) = \frac{1}{u_n^c}\left(\frac{1 - (1 - \alpha\,(c+2)\,u_n^c)^c}{(1 - \alpha\,(c+2)\,u_n^c)^c}\right).$$

En utilisant le fait que $1 - (1-x)^c \sim c\,x$ et $(1-x)^c \sim 1$ lorsque $x \to 0$, on en déduit que

$$\frac{1}{u_{n+1}^c} - \frac{1}{u_n^c} \sim \frac{1}{u_n^c}\left(\frac{\alpha\,c\,(c+2)\,u_n^c}{1}\right) = \alpha\,c\,(c+2)\,,$$

lorsque $n \to \infty$. En utilisant le fait que l'équivalence de suites positives associées à des séries divergentes entraîne l'équivalence des sommes partielles, on en déduit que

$$\frac{1}{u_n^c} - \frac{1}{u_0^c} = \sum_{k=0}^{n-1} \frac{1}{u_{k+1}^c} - \frac{1}{u_k^c} \sim \sum_{k=0}^{n-1} \alpha\,c\,(c+2) = \alpha\,c\,(c+2)\,n,$$

ce qui entraîne que

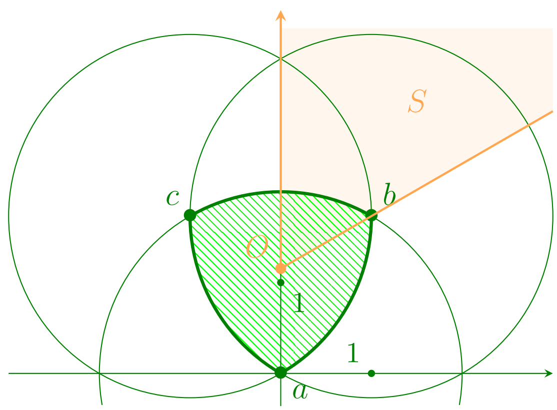
$$\left|x^{(n)} - 0\right| = u_n \underset{n\to\infty}{\sim} \frac{1}{(\alpha\,c\,(c+2)\,n)^{1/c}}.$$

3. Projection

Exercice 15. Considérer l'ensemble $A = \overline{B(a,2)} \cap \overline{B(b,2)} \cap \overline{B(c,2)} \subset \mathbb{R}^2$, où $a = (0,0)$, $b = (1, \sqrt{3})$ et $c = (-1, \sqrt{3})$.

(1) Prouver que A est compact et convexe.
(2) Donner une expression pour la projection sur le convexe A, et dessiner comment cette projection agit sur les différents points du plan.

Solution.



1) L'ensemble A est compact et convexe comme intersection d'un nombre fini d'ensembles fermés, bornés et convexes.

2) On fixe $x = (x_1, x_2) \in \mathbb{R}^2$ et pour obtenir la projection on cherche le minimiseur $P_A(x)$ du problème d'optimisation $\min\{\, |x-y| : y \in A \,\}$, qui a le même minimiseur que le problème

$$\min\{\, |x-y|^2 : y \in A \,\}.$$

Comme A est convexe, on a unicité de la projection et donc unicité du minimiseur. Comme $f_x(y) := |x-y|^2$ est C^2, on commence par regarder les points critiques à l'intérieur du domaine. Or pour $y \in A$, $\nabla f = 2\,(y-x) = 0$ n'est possible que si $x = y$ et donc si $x \in A$. Dans ce cas, on a vérifie bien que $P_A(x) = x$.

On se place donc à présent dans le cas où $x \notin A$, de telle sorte que f_x n'a pas de point critique dans l'intérieur du domaine et donc que son minimum est atteint sur le bord du domaine.

Le triangle (abc) étant équilatéral, l'ensemble A est symétrique par rapport à aux rotations d'angle $2\pi/3$ et de centre $O = (0, 2/\sqrt{3}) = (0, 2\sqrt{3}/3)$ (le centre du triangle) et symétrique par rapport à l'axe vertical. On peut donc se restreindre par exemple à étudier les points x de la région S entre l'axe vertical et la demi-droite d'extrémité O et de direction $[Ob)$, que l'on peut aussi définir par

$$S = \{\, x \in \mathbb{R}^2 : x_1 \geq 0 \text{ et } x_2 > \tfrac{2\sqrt{3}}{3} + \tfrac{\sqrt{3}}{3}\, x_1 \,\}.$$

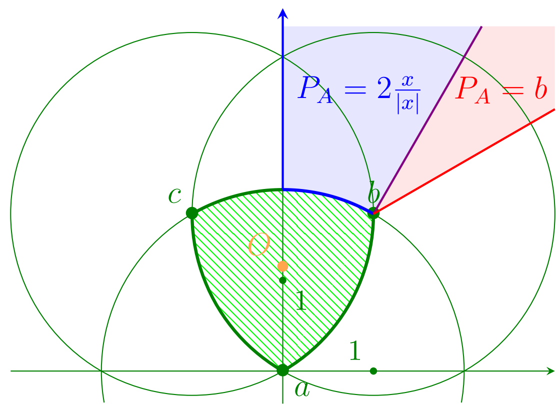
Soit donc $x \in S \setminus A$. Les points candidats à la minimisation sont les points a, b, c ainsi que les points sur les trois arcs de cercle formant le bord du domaine (ab), (ac) et (cb).

- **On montre que le minimiseur n'est pas dans** $(ac) \cup \{ a, c \}$**.** On remarque que $x \in S \setminus A$ implique $x_1 \geq 0$ et $x_2 \geq \sqrt{3}$. Si $y \in (ac)$, alors $y_1 \leq 0$ et $y_2 \leq \sqrt{3}$ et au moins l'une des inégalités est stricte. Donc
$$\begin{aligned} |x-y|^2 &= (x_1 + |y_1|)^2 + \left(x_2 - \sqrt{3} + \sqrt{3} - y_2\right)^2 \\ &> x_1^2 + (x_2 - \sqrt{3})^2 = \left|x - (0, \sqrt{3})\right|^2 \end{aligned}$$
et donc comme $(0, \sqrt{3}) \in A$, le minimiseur ne peut pas être $y \in (ac) \cup \{ a, c \}$.
- **On montre que le minimiseur n'est pas dans** (ab)**.** On utilise le fait que $x \in S \setminus A$ implique que $x_2 \geq \sqrt{3}$. Si $y \in (ab)$, alors $y_1 \in [0, 1[$ et $y_2 < \sqrt{3}$ et donc
$$|x-y|^2 > (x_1 - y_1)^2 + (x_2 - \sqrt{3})^2 = \left|x - (y_1, \sqrt{3})\right|^2$$
et donc comme $(y_1, \sqrt{3}) \in A$, le minimiseur ne peut pas être dans (ab).
- **Sur le côté** (cb), on a $|y - a| = |y| = 2$, et (comme f et g sont C^1 et ∇g ne s'annule pas sur (cb)) on peut donc utiliser les multiplicateurs de Lagrange avec la contrainte $g(y) = |y|^2 - 4$, ce qui donne le système
$$\begin{cases} 2\,(y - x) = 2\,\lambda\, y \\ \qquad |y| = 2. \end{cases}$$
On déduit de la première équation que $(1 - \lambda)\, y = x$. Si $\lambda = 1$, alors $x = (0, 0) \in A$ ce qui n'est pas possible. En prenant la norme de chacun de ces vecteurs et en utilisant la seconde équation du système, on obtient $|x| = |1 - \lambda|\, |y| = |1 - \lambda|\, 2$, c'est-à-dire que $1 - \lambda = \pm \frac{|x|}{2}$, et donc
$$y = \pm 2\, \frac{x}{|x|}.$$
Comme $x \in S$ et $y \in (cb)$, on en déduit que $y_2 > 0$ et $x_2 > 0$, ce qui n'est possible si l'on a un plus dans l'expression ci-dessus, et donc
$$y = 2\, \frac{x}{|x|}.$$
(Remarque: ici, on aurait aussi pu utiliser le fait que l'on projette sur un cercle et utiliser le résultat du cours, que l'on a finalement reprouvé.) Il reste à vérifier dans quels cas un tel point appartient à (cb). Comme
$$(cb) = \{ y \in \mathbb{R}^2 : |y| = 2 \text{ et } y_2 > \sqrt{3}\, |y_1| \},$$
on voit que $2\, \frac{x}{|x|} \in (cb)$ si et seulement si $2\, \frac{x_2}{|x|} > 2\, \sqrt{3}\, \frac{|x_1|}{|x|}$, c'est-à-dire (comme $x_1 \geq 0$) $x_2 > \sqrt{3}\, x_1$.

On déduit de tout cela que pour $x \in S \setminus A$, lorsque $x_2 \leq \sqrt{3}\, x_1$, le minimiseur ne peut être que b, alors que lorsque $x_2 > \sqrt{3}\, x_1$, comme $2\, \frac{x}{|x|}$ est la projection de x sur

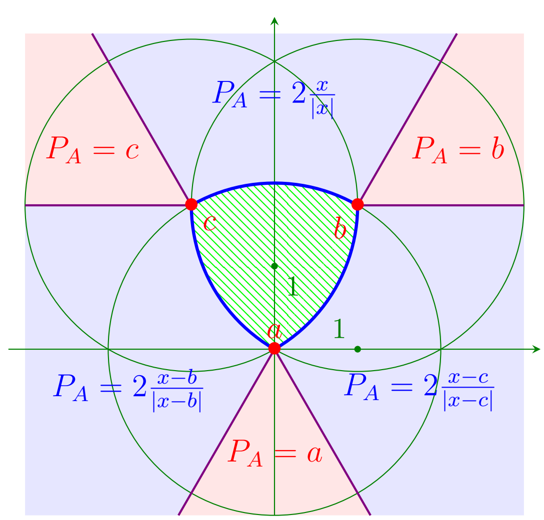
$\overline{B(a,2)}$ et que $b \in \overline{B(a,2)}$, $\left|x - 2\,\frac{x}{|x|}\right| \leq |x-b|$, le minimiseur est $2\,\frac{x}{|x|}$. Cela donne

$$\begin{cases} P_A(x) = 2\,\frac{x}{|x|} & \text{si } x_1 \geq 0 \text{ et } x_2 \geq \sqrt{3}\,x_1 \\ P_A(x) = b & \text{si } x_1 \geq 0 \text{ et } \frac{2\sqrt{3}}{3} + \frac{\sqrt{3}}{3}\,x_1 < x_2 \leq \sqrt{3}x_1 \end{cases}$$



Pour obtenir le cas général, on utilise les symétries comme expliqué au début, pour obtenir $P_A(x) = x$ si $x \in A$ et si $x \notin A$,

$$\begin{cases} P_A(x) = 2\,\frac{x}{|x|} & \text{si } x_2 \geq \sqrt{3}\,|x_1| \\ P_A(x) = 2\,\frac{x-b}{|x-b|} & \text{si } \sqrt{3}\,x_1 \leq x_2 \leq \sqrt{3} \\ P_A(x) = 2\,\frac{x-c}{|x-c|} & \text{si } -\sqrt{3}\,x_1 \leq x_2 \leq \sqrt{3} \\ P_A(x) = a & \text{si } x_2 \leq -\sqrt{3}\,|x_1| \\ P_A(x) = b & \text{si } \sqrt{3} \leq x_2 \leq \sqrt{3}\,x_1 \\ P_A(x) = c & \text{si } \sqrt{3} \leq x_2 \leq -\sqrt{3}\,x_1 \end{cases}$$



Exercice 16. Soit $K \subset \mathbb{R}^n$ un sous-ensemble fermé. Pour $x \in \mathbb{R}^n$ on définit $d(x, K) = \inf \{ |x - y| : y \in K \}$.

(1) La borne inférieure dans la définition de $d(x, K)$ est-elle atteinte ?
(2) Donner un exemple d'ensemble K et de point x telle que cette borne est atteinte mais en plusieurs points $y \in K$.
(3) La fonction $d_K : x \mapsto d(x, K)$ est-elle continue ?
(4) Donner un exemple où la fonction $d_K : x \mapsto d(x, K)$ n'est pas convexe.
(5) Peut-on dire qu'elle est convexe, si on rajoute l'hypothèse que K est convexe ?

Solution. 1) On est en train de minimiser la fonction $f_x(y) = |x - y|$ qui dépend continument de la variable y. Si K est borné, il est compact, et donc le minimum est atteint par le théorème de Weierstrass. Si K n'est pas borné, on remarque que la fonction que l'on minimise est coercive ($f_x(y) \to \infty$ lorsque $y \to \infty$), et donc que la minimisation peut se faire sur un compact : pour $y_0 \in K$, en notant $R = |x - y_0|$ et $B(x, R)$ la boule de rayon R et de centre x, on a en effet

$$\inf_{y \in K} f_x(y) = \inf_{y \in K \cap B} f_x(y)$$

et $K \cap B$ est bien fermé (intersection de deux fermés) et borné (Car $K \cap B \subset B$), donc compact.

2) Si $K = \{ A, B \}$ avec $A \neq B$ le point $(A + B)/2$ est à distance égale de A et de B, il a donc deux projections.

3) Pour tout y la fonction $x \mapsto |x - y|$ est lipschitzienne de constante 1. Par conséquent, la fonction $d_K(x) = \inf \{ |x - y| : y \in K \}$ l'est aussi (comme $\inf$ d'une famille de fonctions lipschitziennes avec la même distance).

On peut aussi retrouver ceci par la définition. Si $x, x' \in \mathbb{R}^n$, alors pour tout $y \in K$, par l'inégalité triangulaire

$$d_K(x) \leq |x - y| \leq |x - x'| + |x' - y| .$$

C'est-à-dire que pour tout $y \in K$, $|x' - y| \geq d_K(x) - |x - x'|$, et donc en prenant l'infimum,

$$d_K(x') \geq d_K(x) - |x - x'| .$$

Par symétrie, la même inégalité est vraie en échangeant x et x', d'où l'on déduit que

$$|d_K(x') - d_K(x)| \leq |x - x'| ,$$

et donc d_K est continue.

4) Si $K = \{ A, B \}$ avec $A \neq B$, le point $(A+B)/2$ on a $d_K(A) = d_K(B) = 0$ alors que $d_K((A + B)/2) = |A - B| /2 > 0$, ce qui démontre que d_K n'est pas convexe.

5) On suppose K convexe. Soit $x_0, x_1 \in \mathbb{R}^n$ et $y_0, y_1 \in K$ leurs projections respectives. Soit $t \in [0, 1]$. Le point $y_t = (1 - t)\, y_0 + t\, y_1$ appartient à K parce que K est convexe. Donc, en posant $x_t := (1 - t)\, x_0 + t\, x_1$ et par l'inégalité triangulaire,

$$\begin{aligned} d_K(x_t) \leq |x_t - y_t| &= |(1 - t)(x_0 - y_0) + t(x_1 - y_1)| \\ &\leq (1 - t)\, |x_0 - y_0| + t\, |x_1 - y_1| = (1 - t)\, d_K(x_0) + t\, d_K(x_1), \end{aligned}$$

ce qui montre la convexité de d_K.

Exercice 17. Considérer les ensembles

$$A = \left\{ (x,y) \in \mathbb{R}^2 : |x+y| \le 1 \right\} \quad \text{et} \quad B = \overline{B_1} = \left\{ (x,y) \in \mathbb{R}^2 : x^2 + y^2 \le 1 \right\}$$

(1) Dessiner A, B et $A \cap B$. A et B sont-ils convexes ?
(2) Démontrer que l'intersection de deux ensembles convexes est toujours convexe.
(3) Écrire une formule pour la projection sur $A \cap B$, du type $P_{A\cap B}(x,y) = \dots$, en distinguant éventuellement des cas (il suffit de la justifier avec un dessin).
(4) Lesquelles des relations suivantes sont-elles vraies ?

$$P_{A\cap B} = P_A \circ P_B, \quad P_A \circ P_B = P_B \circ P_A, \quad P_{A\cap B} = P_{A\cap B} \circ P_B.$$

Exercice 18. Soit $K \subset \mathbb{R}^2$ le polyèdre convexe donné par

$$K = \left\{ x = (x_1, x_2) \in \mathbb{R}^2_+ : x_1 \le 1, x_1 + x_2 \le 2 \right\}.$$

Trouver les sommets de K et écrire une formule (en distinguant éventuellement plusieurs cas) pour la projection P_K sur K.

4. Transformée de Fenchel et sous-différentiel

Exercice 19. Soit $a > 0$, $p \ge 1$ et $f : \mathbb{R}^d \to \mathbb{R}$ la fonction définie par

$$f(x) = a\,|x| + \frac{|x|^p}{p}$$

où $|x|$ désigne la norme euclidienne du vecteur $x \in \mathbb{R}^d$.

(1) La fonction f est-elle convexe ?
(2) Calculer f^* et f^{**}. On pourra commencer par regarder le cas de la dimension $d = 1$.
(3) Calculer ∂f.

Solution. (1) f est convexe car c'est une somme de fonctions convexes.
(2) Comme f est convexe, on a $f^{**} = f$. On calcule à présent f^*. On se place en dimension $d = 1$. Par définition

$$f^*(y) = \sup_x \, xy - f(x) = \sup_x \, xy - a\,|x| - \frac{|x|^p}{p}.$$

Si $p > 1$ ou si $|y| < a$, on remarque que ce $\sup$ est atteint (puisque la fonction a optimiser est continue et $xy - a\,|x| - \frac{|x|^p}{p} \to -\infty$ quand $x \to \pm\infty$) et qu'il est atteint en un point du même signe que y (sinon $-x$ donnerait un résultat meilleur). Considérons d'abord $y \ge 0$. Dans ce cas là on cherche un maximum sur les x positifs, et on cherche donc à maximiser $(y-a)\,x - \frac{|x|^p}{p}$. La dérivée s'annule pour $y - a = x^{p-1}$, ce qui n'est possible que pour $y \ge a$ (sinon on n'a pas $x \ge 0$). Si $0 \le y < 1$ la dérivée ne s'annule pas, le maximum est donc réalisé sur le bord du domaine $\{ x \ge 0 \}$, c'est-à-dire en $x = 0$, et vaut 0. Pour $y \ge a$ le maximum est réalisé en $x = (y-a)^{1/p-1}$, ce qui donne une valeur de $\frac{(y-a)^{p'}}{p'}$ où $p' = \frac{p}{p-1}$. On peut voir que pour $y < 0$ on a une valeur nulle pour $-a < y < 0$ et égale à $\frac{(-y-a)^{p'}}{p'}$ pour $y \leq -a$. On peut résumer le tout en disant que

$$f^*(y) = \frac{(|y| - a)_+^{p'}}{p'}$$

Quant la fonction est définie sur tout $\mathbb{R}^d$, elle est encore convexe comme somme de fonctions convexes. Pour calculer f^* on remarque que pour maximiser $x \cdot y - a\,|x| - \frac{|x|^p}{p}$, parmi tous les vecteurs de même norme, il vaut mieux prendre celui qui est parallèle à y et dans la même direction. Disons $x = r\,\frac{y}{|y|}$ avec $r \geq 0$. Le problème revient ensuite au même que précédemment, puisqu'on doit maximiser $(|y| - a)\,r - \frac{r^p}{p}$, qui est la même fonction d'une variable qu'avant. La réponse est alors encore

$$f^*(y) = \frac{(|y| - a)_+^{p'}}{p'}$$

(3) En $x \neq 0$, f est dérivable donc $\partial f(x) = \{\, \nabla f(x) \,\} = \{\, a\,|x|^{-1}\,x - |x|^{p-2}\,x \,\}$. Le seul cas restant est donc $x = 0$. Dans ce cas, comme $f(0) = 0$, on a par définition

$$\partial f(0) = \{\, z \in \mathbb{R}^d : f(y) \geq z \cdot y \,\}\,.$$

Si $z \in \partial f(0)$, alors en particulier en prenant $y = r\,z$ avec $r > 0$, on voit qu'il faut que

$$a\,r\,|z| - \frac{r^p\,|z|^p}{p} \geq r\;|z|^2$$

ce qui peut s'écrire

$$|z| \leq a - \frac{r^{p-1}\,|z|^{p-1}}{p}.$$

En faisant tendre $r \to 0$, on en déduit que $|z| \leq a$. Réciproquement, si $|z| \leq a$, alors pour tout $y \in \mathbb{R}^d$,

$$z \cdot y \leq |z|\,|y| \leq a\,|y| \leq f(y),$$

et donc on en déduit que

$$\partial f(0) = \overline{B_a} = \{\, z \in \mathbb{R}^d, |z| \leq a \,\}\,.$$

Exercice 20. Calculer le sous-différentiel de la fonction $f(x,y) = \sqrt{x^2 + y^4} + \sqrt{x^4 + y^2}$.

5. Choix d'algorithme

Exercice 21. Soit $B := \{\, x \in \mathbb{R}^d : \|x\|_3 \leq 1 \,\}$ la boule unité pour la norme ℓ^3 définie pour $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ par $\|x\|_3 = \left(|x_1|^3 + \cdots + |x_d|^3\right)^{1/3}$. Décrire de manière détaillée et explicite au moins une méthode numérique pour calculer la projection sur B et justifier sa convergence. Pourquoi ne serait-il pas raisonnable *du tout* de considérer un algorithme de gradient projeté pour répondre à la question précédente ?

Solution. S'agissant d'une optimisation sous contrainte, un algorithme adapté est, par exemple, celui d'Uzawa. L'algorithme du gradient projeté par contre est complètement inadapté parce qu'il demanderait à utiliser la projection sur K, ce qui est exactement ce qu'on cherche à calculer !

La projection d'un point $y \in \mathbb{R}^d \setminus B$ sur l'ensemble B s'obtient en trouvant le minimiseur du problème suivant

$$\min\{|x-y|^2 : \|x\|_3 \leq 1\}.$$

Pour trouver le minimiseur, l'algorithme d'Uzawa utilise la dualité, et cherche à résoudre

$$\max_{\lambda \geq 0} \min_{x \in \mathbb{R}^d} |x-y|^2 + \lambda\, g(x).$$

où l'on peut prendre $g(x) = \|x\|_3^3 - 1$. Si on appelle $F(\lambda) := \min_{x \in \mathbb{R}^d} |x-y|^2 + \lambda\, g(x)$, l'algorithme d'Uzawa est un algorithme de gradient projeté pour la fonction concave F (qu'on cherche à maximiser), soumise à la contrainte $\lambda \geq 0$ (qui est simple à gérer). On sait que $F'(\lambda) = g(x_\lambda)$ où x_λ est le x optimal correspondant à λ, celui qui minimise $|x-y|^2 + \lambda\, g(x)$. Ce x_λ peut être trouvé en imposant $2\,(x-y) + \lambda \nabla g(x) = 0$. La suite des λ produite par l'algorithme doit satisfaire $\lambda_{n+1} = (\lambda_n + \alpha\, g(x_{\lambda_n}))_+$, où $\alpha > 0$ est le pas de l'algorithme de gradient projeté. Ce α doit être suffisamment petit, et éventuellement variable.

• Calcul de x_λ : pour chaque composante j, on doit imposer

$$x_j - y_j + \lambda\, \partial_{x_j} g(x) = 0$$

Comme $\partial_{x_j} g(x) = 3\, \mathrm{sign}(x_j)\, x_j^2$ (le signe étant celui de x_j, il est nécessaire que x_j ait le même signe que y_j), donc pour $y_j > 0$ on cherche $x_j > 0$ tel que $x_j - y_j + \lambda x_j^2 = 0$ et pour $y_j < 0$ on cherche $x_j < 0$ tel que $x_j - y_j - \lambda x_j^2 = 0$, ce qui donne

$$x_j = \frac{-1 + \sqrt{1 + 12\,\lambda\, y_j}}{6\,\lambda} \quad \text{si } y_j \geq 0, \qquad x_j = \frac{1 - \sqrt{1 - 12\,\lambda\, y_j}}{6\,\lambda} \quad \text{si } y_j < 0.$$

L'algorithme itératif est donc obtenu en partant d'un couple $(x^{(0)}, \lambda_0)$ quelconque et en prenant ensuite

$$\begin{aligned} x_j^{(n+1)} &= \frac{-1 + \mathrm{sign}(y_j)\sqrt{1 + 12\,\lambda_n\, |y_j|}}{6\,\lambda_n} \qquad \text{pour tout } j \in [\![1, d]\!] \\ \lambda_{n+1} &= \Big(\lambda_n + \alpha\, g(x^{(n+1)})\Big)_+ \end{aligned}$$

• La suite $x^{(n)}$ obtenue converge alors vers la projection de y sur K. La preuve de la convergence est similaire mais plus technique que celle donnée en cours, et sort du cadre du cours.