Convexité

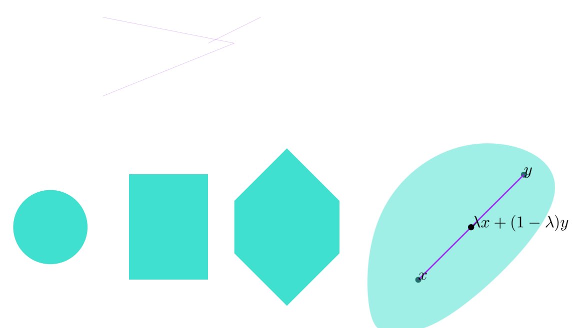
Hamza Ennaji

5 mars 2024

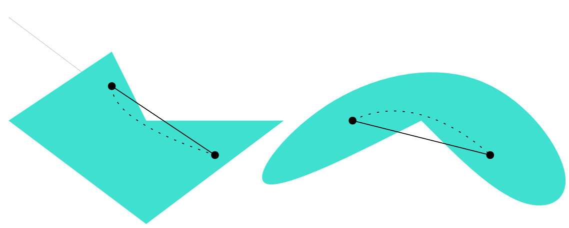
1 Ensembles convexes

Définition 1. *Soit $C \subset \mathbb{R}^n$. On dit que C est convexe si pour tout $x, y \in C$ et tout $\lambda \in [0, 1]$, $\lambda x + (1 - \lambda)y \in C$.*

Remarque. *La définition est équivalent à dire que le ségment $[x, y] \subset C$. La Figures-1 illustre quelques exemples d'ensembles convexes et non-convexes.*



(a)



(b)

Fig. 1 : La Figure-1a des ensembles convexes de $\mathbb{R}^2$. Tandis que Figure-1b montre des examples de d'ensembles non convexes.

Exemple (Voir Td)**.** *Voici quelques exemples d'ensembles convexes :*

- *Les boules (ouvertes, fermées) :*

$$B(a,r) = \{x \in \mathbb{R}^n : \|x-a\| < r\} \ \text{et} \ B(a,r) = \{x \in \mathbb{R}^n : \|x-a\| \leq r\}$$

avec $a \in \mathbb{R}^n, r > 0$.

- *Les hyperplans :*

$$H = \{x \in \mathbb{R}^n : a^T x = b\}, \ a \in \mathbb{R}^n, b \in \mathbb{R}.$$

- *Les demis-plans, les ellipsoïdes etc.*

Proposition 1 (Intersection de convexes)**.** *Soient* $C_i \in \mathbb{R}^n$ *avec* $i \in I$, *des ensembles convexes, alors* $C \overset{def}{=} \cap_{i\in I} C_i$ *est convexe.*

Preuve 1. *Laissée en exercice.*

Remarque. *Dans la Proposition-1, on remarque que l'ensemble des indices I est quelconque, i.e., possiblement infini.*

Le résultat suivant récapitule quelques opérations préservant la convexité.

Théorème 1. 1. *Soient* $C_1, \ldots, C_m$ *des convexes de* $\mathbb{R}^n$ *et* $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$, *alors*

$$\sum_{i=1}^m \lambda_i C_i = \left\{ \sum_{i=1}^m \lambda_i x_i : \ x_i \in C_i, \ \forall i = 1, \ldots, m \right\}$$

est convexe.

2. *Soient* $C_1, \ldots, C_m$ *des convexes de* $\mathbb{R}^n$, *alors*

$$\prod_{i=1}^m C_i = \left\{ (x_1, \ldots, x_m) : \ x_i \in C_i, \ \forall i = 1, \ldots, m \right\}$$

est convexe.

3. *Si* $C \subset \mathbb{R}^n$ *est convexe et* $A \in \mathbb{R}^{m\times n}$, *alors*

$$A(C) = \left\{ Ax : \ x \in C \right\}$$

est convexe.

Définition 2. *Soient* $x_1, \ldots, x_m \in \mathbb{R}^n$. *Une combinaison convexe des* x_i *est un vecteur de la forme* $\sum_{i=1}^m \lambda_i x_i$ *avec* $\lambda_i \geq 0$ *pour tout* $i = 1, \ldots, m$ *et* $\sum_{i=1}^m \lambda_i = 1$. *On écrit* $\lambda = (\lambda_1, \ldots, \lambda_m) \in \Delta_m$ *où*

$$\Delta_m = \left\{ \lambda = (\lambda_1, \ldots, \lambda_m), \ \lambda_i \geq 0, \ \forall i = 1, \ldots, m \ \text{et} \ \sum_{i=1}^m \lambda_i = 1 \right\}$$

.

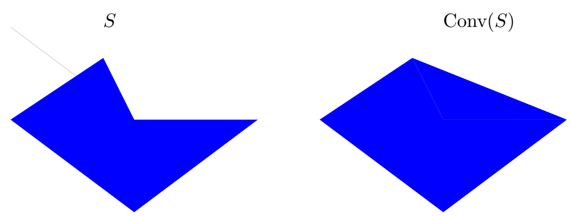
Théorème 2. *Soit $C \subset \mathbb{R}^n$ un convexe et $x_1, \dots, x_m \in C$. Alors pour tout $\lambda \in \Delta_m$ on a $\sum_{i=1}^m \lambda_i x_i \in C$.*

Preuve 2. *Laissée en exercice.*

Définition 3. *Soit $S \subset \mathbb{R}^n$. On appelle enveloppe convexe de S, et on note* $\mathrm{Conv}(S)$ *l'ensemble des combinaisons convexes de S :*

$$\mathrm{Conv}(S) = \left\{ \sum_{i=1}^m \lambda_i x_i : \quad x_i \in S \ \forall i = 1, \dots, m \text{ et } \lambda = (\lambda_1, \dots, \lambda_m) \in \Delta_m \right\}.$$

Remarque. *L'enveloppe convexe d'un ensemble S est le plus petit convexe qui le contient : Si $S \subset U$ avec U convexe alors* $\mathrm{Conv}(S) \subset U$.



(a)

FIG. 2 : Exemple d'un ensemble non-convexe et son enveloppe convexe.

2 Fonctions convexes

Dans cette section, on considère un convexe C de $\mathbb{R}^n$.

Définition 4. *On dit qu'une fonction $f : C \longrightarrow \mathbb{R}$ est convexe si*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y), \text{ pour tout } x, y \in C, \lambda \in [0,1].$$

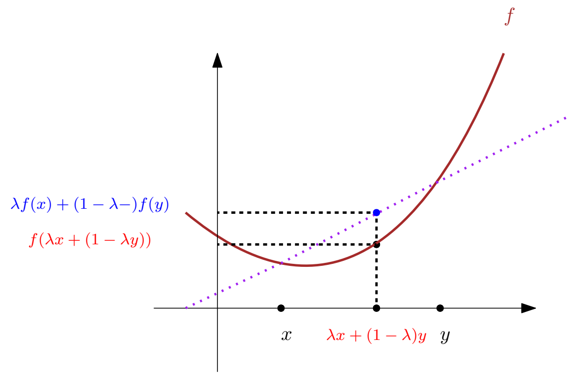
On dira que f est strictement convexe si

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y), \text{ pour tout } x \neq y \in C, \lambda \in (0,1).$$

On dira que f est concave si $-f$ est convexe.

Exemple (Voir Td)**.** *Voici quelques examples de fonctions convexes :*

- *Les normes : $f(x) = \|x\|$ sur $\mathbb{R}^n$.*
- *Les fonctions affines : $f(x) = a^T x + b$, $a \in \mathbb{R}^n$ et $b \in \mathbb{R}$.*



(a)

FIG. 3 : Illustration de l'inégalité de convexité dans la Définition-4

- *La fonction* $f(x) = -\log(x)$.

D'après la Définition-4, une fonction f est convexe si l'image par f, d'une combinaison convexe de deux points x et y, par f est plus petite que la combinaison convexe des valeurs $f(x)$ et $f(y)$. Cette propriété s'étant à la combinaison convexe de n'importe quel nombre de vecteurs.

Théorème 3 (Inégalité de Jensen)**.** *Soit* $f : C \longrightarrow \mathbb{R}$ *une fonction convexe. Alors pour tout* $x_1, \ldots, x_m \in C$ *et* $\lambda \in \Delta_m$ *:*

$$f\Big(\sum_{i=1}^{m} \lambda_i x_i\Big) \leq \sum_{i=1}^{m} \lambda_i f(x_i).$$

Preuve 3. *Par récurrence. Laissée en exercice.*

Le résultat suivant récapitule quelques opérations préservant la convexité de fontions.

Théorème 4.

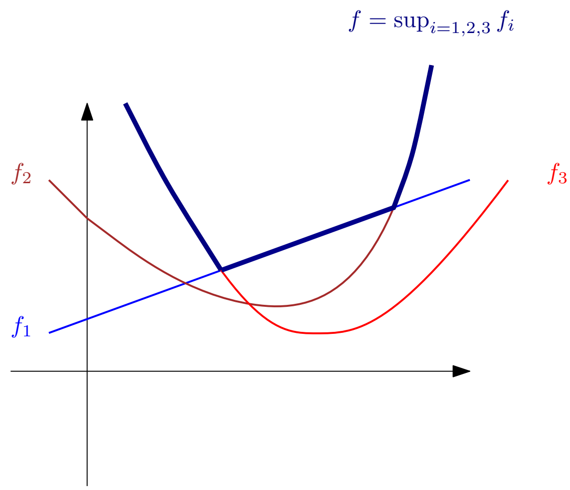
1. *Soit* f *une fonction convexe sur* C *et* $\alpha \in \mathbb{R}^+$. *Alors* αf *est convexe sur* C.

2. *Soient* $f_1, \ldots, f_m$ *des fonctions convexes sur* C. *Alors* $\sum_{i=1}^{m} f_i$ *est convexe sur* C.

3. *Soient* $f_1, \ldots, f_m$ *des fonctions convexes sur* C. *Alors* $f = \sup_{i=1,\ldots,m} f_i$ *est convexe sur* C.

4. *Si* $f : C \longrightarrow \mathbb{R}$ *est convexe et* $A \in \mathbb{R}^{n\times m}, b \in \mathbb{R}^n$. *Alors* $f(Ax+b)$ *est convexe sur*

$$D = \{y \in \mathbb{R}^m : \; Ay + b \in C\}.$$



(a)

Fig. 4 : Convexité du la fonction sup de trois fonctions f_1, f_2, f_3.

3 Caractérisation des fonctions convexes différentiables

Théorème 5. *Soit $f : C \longrightarrow \mathbb{R}$ une fonction différentiable. Les propriétés suivantes sont équivalentes*

1. *f est convexe.*
2. *$f(y) \geq f(x) + \nabla f(x)^T(y-x)$ pour tout $x, y \in C$.*
3. *Monotonie du gradient : $\Big(\nabla f(x) - \nabla f(y)\Big)^T (y-x) \geq 0$ pour tout $x, y \in C$.*

Quand la fonction est deux fois différentiable, alors la convexité de f est équivalente au fait que la matrice Hessienne est semi-définie positive.

Théorème 6. *Soit f une fonction deux fois différentiable sur un ouvert convexe C de $\mathbb{R}^n$ alors f est convexe si et seulement si $\nabla^2 f(x) \geq 0$ pour tout $x \in C$.*

Remarque. *Dans le Théorème-6, $\nabla^2 f(x) \geq 0$ est à comprendre dans le sens : $u^T \nabla^2 f(x) u \geq 0$ pour tout vecteur u de $\mathbb{R}^n$.*

Remarque. *Jusqu'ici on a travaillé avec des fonctions f à valeurs dans $\mathbb{R}$. Les fonctions qu'on va rencontrer dans la pratique peuvent prendre des valeurs infinie. Il se trouve que la définition de convexité introduite précédemment marche aussi pour des fonctions à valeurs dans $\mathbb{R} \cup \{+\infty\}$. À cette définition, il faut rajouter les règles suivantes : $\alpha + \infty = \infty$ pour tout $\alpha \in \mathbb{R}$, $\alpha.\infty = \infty$ pour tout $\alpha > 0$ et $0.\infty = 0$. Cela revient à dire que que le domaine de f définie par*

$$\mathrm{dom}(f) = \{x \in \mathbb{R}^n : f(x) < \infty\},$$

est convexe et que $f_{|\mathrm{dom}(f)}$ est convexe. L'exemple typique de telles fonctions est l'indicatrice d'un ensemble $S \subset \mathbb{R}^n$ définie par

$$\delta_S(x) = \begin{cases} 0 & si\ x \in S, \\ +\infty & sinon, \end{cases}$$

et on peut démontrer que δ_S est convexe si et seulement si S est convexe.

Exemple (Voir Td).

1. *$f(x) = \frac{1}{2}x^T Ax$ avec $A \in \mathcal{S}^n$. Montrer que f est convexe si et seulement si $A \in \mathcal{S}^n_+$.*
2. *On appelle fonction support ou d'appui de $S \subset \mathbb{R}^n$ la fonction*
$$\sigma_S(x) = \sup_{y \in S} x^T y.$$
Montrer que σ_S est convexe.
3. *Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$. On définit la conjuguée de f par*
$$f^*(x) = \sup_y y^T x - f(y).$$
Montrer que f^* est convexe.

4 Sous-ensembles de niveau de fonctions convexes :

Définition 5. *Soit $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction. Le sous-ensemble de f de niveau $\alpha \in \mathbb{R}$ est l'ensembles*

$$\boldsymbol{Lev}(f, \alpha) = \{x \in S : \ f(x) \leq \alpha\}.$$

On a le résultat suivant

Théorème 7. *Soit $f : C \longrightarrow \mathbb{R}$ une fonction convexe et avec C un ensemble convexe. Alors pour tout $\alpha \in \mathbb{R}$, l'ensembles $\boldsymbol{Lev}(f, \alpha)$ est convexe.*

Preuve 4. *Laissée en exercice.*

Remarque. *La réciproque dans le Théorème-7 est fausse comme le montre la fonction $f(x) = \sqrt{|x|}$ (voir Figure-5) qui n'est pas convexe mais dont tout les sous-ensembles de niveau sont convexe. En effet, d'une part, pourt $\alpha < 0, \boldsymbol{Lev}(f, \alpha) = \emptyset$. D'autre part, pour $\alpha \geq 0$, on $\boldsymbol{Lev}(f, \alpha) = [-\alpha^2, \alpha^2]$ qui est convexe. Une telle fonction est dite quasi-convexe.*

Définition 6. *Soit $C \subset \mathbb{R}$ un convexe. On dit qu'une fonction $f : C \to \mathbb{R}$ est quasi-convexe si pour tout $\alpha \in \mathbb{R}$, l'ensemble $\boldsymbol{Lev}(f, \alpha)$ est convexe.*

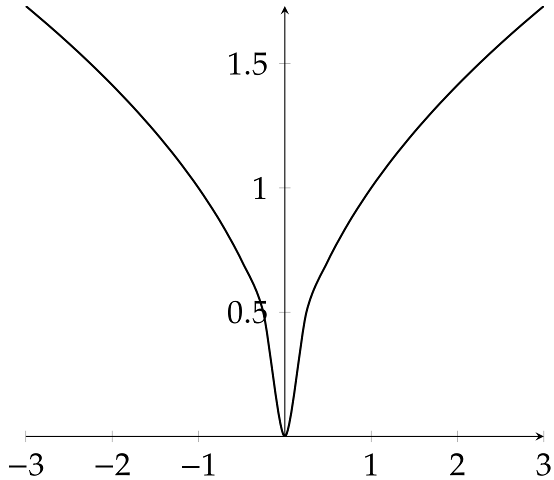


FIG. 5 : $f(x) = \sqrt{|x|}$ comme exemple de fonction quasi-convexe.

5 Optimisation convexe

Pour finir ce chapitre, on va dégager quelques propriétés des problèmes d'optimisation convexe. Étant donné un convexe $C \subset \mathbb{R}^n$ et une fonction convexe on considère le problème de minimisation suivant

$$\min_{x \in C} f(x) \tag{1}$$

1. On dit que $x^* \in C$ est un minimum local de f sur C s'il existe $r > 0$ tel que $f(x^*) \leq f(y)$ pour tout $y \in C \cap B_f(x^*, r)$.

2. On dit que $x^* \in C$ est un minimum global de f sur C si $f(x^*) \leq f(y)$ pour tout $y \in C$.

On a le résultat suivant :

Théorème 8. *Soit $f : C \longrightarrow \mathbb{R}$ une fonction convexe définie sur un convexe C. Soit $x^* \in C$ est un minimum local de f sur C. Alors x^* est un minimum global de f sur C*

Preuve 5. *Soit x^* un minimum local de f sur C, il existe alors $r > 0$ tel que $f(x^*) \leq f(x)$ pour tout $y \in C$ avec $\|x^* - x\| \leq r$. Pour montre que $f(x^*) \leq f(y)$ pour tout $y \in C \setminus \{x^*\}$, on considère un $\bar{x} = (1-t)x^* + ty$ pour un certain $t \in (0, 1]$ de telle façon que $\|\bar{x} - x^*\| \leq r$. On peut prendre par exemple $t = r/\|x^* - y\|$. Comme x^* est minimum local, il sent suit, par convexité de f*

$$f(x^*) \leq f(\bar{x}) = f(ty + (1-t)x^*) \leq tf(y) + (1-t)f(x^*),$$

soit $tf(x^*) \leq tf(y)$ et donc $f(x^*) \leq f(y)$.

De même on démontre que

Théorème 9. *Soit $f : C \longrightarrow \mathbb{R}$ une fonction strictement convexe définie sur un convexe C. Soit $x^* \in C$ est un minimum local de f sur C. Alors x^* est un minimum global stricte de f sur C.*

Notation. On note par $\underset{x \in C}{\operatorname{argmin}} f(x)$ (des fois S) l'ensemble des minimiseurs de f, i.e., les solutions de (1).

Théorème 10. *Soit $f : C \longrightarrow \mathbb{R}$ une fonction convexe, alors S est convexe. Si de f est strictement convexe, alors S contient au plus un élément.*

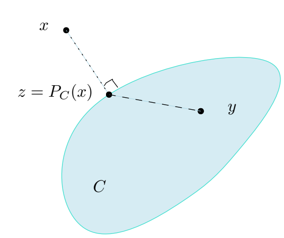
Preuve 6. *Si $S = \emptyset$, rient à démontrer. Sinon, soient $x, y \in S$ et $\lambda \in [0, 1]$. Alors*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) = \lambda f^* + (1-\lambda) f^* = f^*.$$

Ce qui prouve la convexité. Sinon, on peut remarquer que $S = \{x : f(x) \leq f^*\} \cap C$. Maintenant, si f est strictement convexe et supposons qu'il existe $x, y \in S$ avec $x \neq y$. Par convexité de C on a $\frac{1}{2}x + \frac{1}{2}y \in C$ et par convexité de f :

$$f\left(\frac{1}{2}x + \frac{1}{2}y\right) < \frac{1}{2}(x) + \frac{1}{2}f(y) = f^*,$$

cela contredit le fait que f^* est la valeur optimale.



(a)

Fig. 6 : Projection orthogonale.

5.1 Projection orthogonale

Soit $C \subset \mathbb{R}^n$ un convexe fermé.

Définition 7. *L'opérateur de projection orthogonale sur C est l'application $P_C : \mathbb{R}^n \longrightarrow C$ qui à $x \in \mathbb{R}^n$ associe*

$$P_C(x) = \underset{y \in C}{\operatorname{argmin}} \|y - x\|^2 \tag{2}$$

On a le résultat (admis pour l'instant)

Théorème 11 (Premier théorème de la projection)**.** *Soit C un convexe fermé non vide. Alors* (2) *admet une unique solution.*

Le résultat suivant donne une caractérisation géométrique de la projection.

Théorème 12 (Deuxième théorème de la projection)**.** *Soit C un convexe fermé non vide et $x \in \mathbb{R}^n$. Alors $z = P_C(x)$ si et seulement si $(x-z)^T(y-z) \leq 0$ pour tout $y \in C$.*

Le Théorème-12 affirme que pour $x \in \mathbb{R}^n$ quelconque, l'angle entre $x - P_C(x)$ et $y - P_C(x)$, avec $y \in C$, est supérieur à $\pi/2$. Cela est illustré dans la Figure-6.