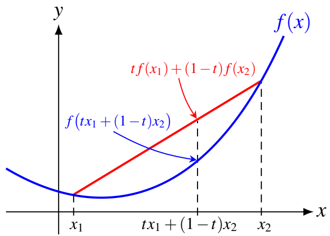
Convexité à l'ordre 0

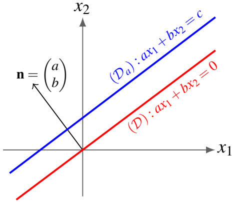
Guillaume Tochon

Laboratoire de Recherche de l'EPITA



Les espaces affines ne permettent qu'une description très restreinte des sous-parties de $\mathbb{R}^n$ que l'on rencontre dans les problèmes d'optimisation en pratique :

→ Dans $\mathbb{R}^2$, les sous-espaces affines ne permettent que de décrire des points et des droites.



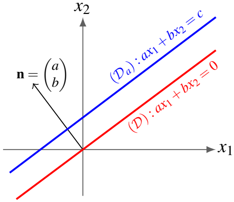
Sous-partie affine

Sous-parties de $\mathbb{R}^n$

Les espaces affines ne permettent qu'une description très restreinte des sous-parties de $\mathbb{R}^n$ que l'on rencontre dans les problèmes d'optimisation en pratique :

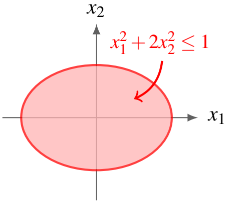
$\rightarrow$ Dans $\mathbb{R}^2$, les sous-espaces affines ne permettent que de décrire des points et des droites.

$\Rightarrow$ Besoin d'une description plus générale pour des sous-parties plus complexes.



Sous-partie affine

vs

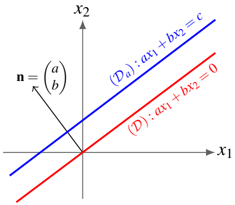


Sous-parties de $\mathbb{R}^n$

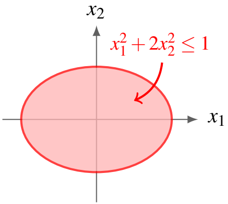
Les espaces affines ne permettent qu'une description très restreinte des sous-parties de $\mathbb{R}^n$ que l'on rencontre dans les problèmes d'optimisation en pratique :

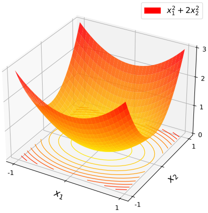
$\rightarrow$ Dans $\mathbb{R}^2$, les sous-espaces affines ne permettent que de décrire des points et des droites.

$\Rightarrow$ Besoin d'une description plus générale pour des sous-parties plus complexes.



vs





Sous-partie affine

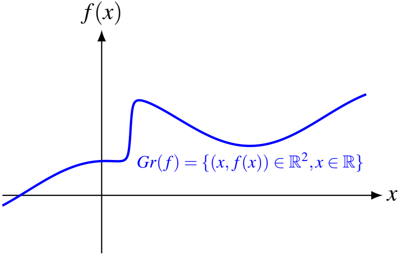
Lieu de sous-niveau 1 de $f(x_1, x_2) = x_1^2 + 2x_2^2$

Solution : construire des sous-parties de $\mathbb{R}^n$ comme lieux particuliers associés à certaines fonctions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ bien choisies.

Étant donné une fonction $f : \mathbb{R}^n \to \mathbb{R}$.

$\rightarrow$ On appelle *graphe* de f la partie de $\mathbb{R}^{n+1}$:

$$Gr(f) = \left\{(\mathbf{x}, f(\mathbf{x})), \mathbf{x} \in \mathbb{R}^n\right\} \subset \mathbb{R}^{n+1}$$



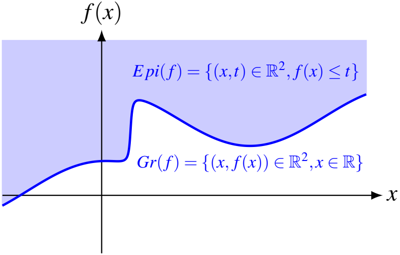
Étant donné une fonction $f : \mathbb{R}^n \to \mathbb{R}$.

→ On appelle *graphe* de f la partie de $\mathbb{R}^{n+1}$:

$$Gr(f) = \left\{(\mathbf{x}, f(\mathbf{x})), \mathbf{x} \in \mathbb{R}^n\right\} \subset \mathbb{R}^{n+1}$$

→ On appelle *épigraphe* de f la partie de $\mathbb{R}^{n+1}$:

$$Epi(f) = \left\{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}, f(\mathbf{x}) \leq t\right\} \subset \mathbb{R}^{n+1}$$



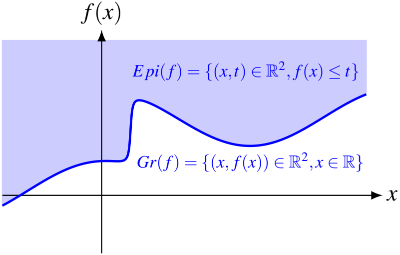
Étant donné une fonction $f : \mathbb{R}^n \to \mathbb{R}$.

$\to$ On appelle *graphe* de f la partie de $\mathbb{R}^{n+1}$:

$$Gr(f) = \left\{(\mathbf{x}, f(\mathbf{x})), \mathbf{x} \in \mathbb{R}^n\right\} \subset \mathbb{R}^{n+1}$$

$\to$ On appelle *épigraphe* de f la partie de $\mathbb{R}^{n+1}$:

$$Epi(f) = \left\{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}, f(\mathbf{x}) \leq t\right\} \subset \mathbb{R}^{n+1}$$

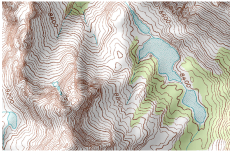


Ces deux descriptions sont dites *paramétriques* : lorsque $\mathbf{x}$ parcourt $\mathbb{R}^n$, son image par f génère les parties de $\boxed{\mathbb{R}^{n+1}}$ $Gr(f)$ et $Epi(f)$.

Étant donné une fonction $f : \mathbb{R}^n \to \mathbb{R}$.

$\rightarrow$ On appelle *courbe de niveau* r de f la partie de $\mathbb{R}^n$:

$$\mathcal{C}_r(f) = \Big\{ \mathbf{x} \in \mathbb{R}^n, f(\mathbf{x}) = r \Big\} \subset \mathbb{R}^n$$



carte topographique
$\equiv$ lignes de niveaux

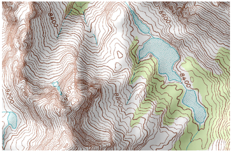
Étant donné une fonction $f : \mathbb{R}^n \to \mathbb{R}$.

$\rightarrow$ On appelle *courbe de niveau* r de f la partie de $\mathbb{R}^n$:

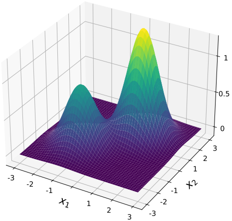
$$\mathcal{C}_r(f) = \left\{\mathbf{x} \in \mathbb{R}^n, f(\mathbf{x}) = r\right\} \subset \mathbb{R}^n$$

$\rightarrow$ On appelle *lieu de sous-niveau* r de f la partie de $\mathbb{R}^n$:

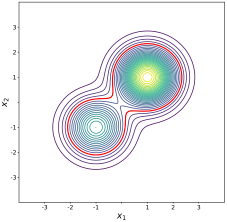
$$\mathcal{C}_{\leq r}(f) = \left\{\mathbf{x} \in \mathbb{R}^n, f(\mathbf{x}) \leq r\right\} \subset \mathbb{R}^n$$



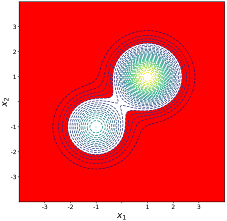
carte topographique
$\equiv$ lignes de niveaux



$Gr(f)$



$\mathcal{C}_r(f)$



$\mathcal{C}_{\leq r}(f)$

Représentation implicite

Courbe de niveau et lieu de sous niveau d'une fonction

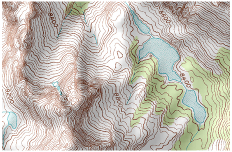
Étant donné une fonction $f : \mathbb{R}^n \to \mathbb{R}$.

$\to$ On appelle *courbe de niveau* r de f la partie de $\mathbb{R}^n$:

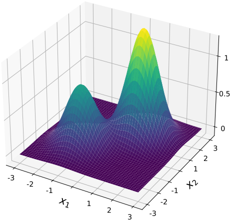
$$\mathcal{C}_r(f) = \Big\{ \mathbf{x} \in \mathbb{R}^n, f(\mathbf{x}) = r \Big\} \subset \mathbb{R}^n$$

$\to$ On appelle *lieu de sous-niveau* r de f la partie de $\mathbb{R}^n$:

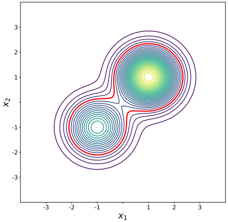
$$\mathcal{C}_{\leq r}(f) = \Big\{ \mathbf{x} \in \mathbb{R}^n, f(\mathbf{x}) \leq r \Big\} \subset \mathbb{R}^n$$



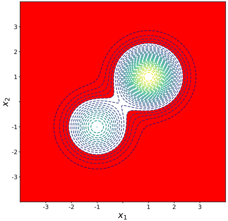
carte topographique
$\equiv$ lignes de niveaux



$Gr(f)$



$\mathcal{C}_r(f)$



$\mathcal{C}_{\leq r}(f)$

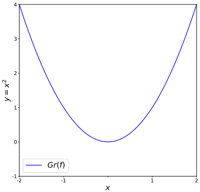
Ces deux descriptions sont dites *implicites* : $\mathcal{C}_r(f)$ est le lieu dans $\boxed{\mathbb{R}^n}$ des zéros de la fonction $g(\mathbf{x}) = f(\mathbf{x}) - r$

Lien entre les représentations paramétriques et implicites

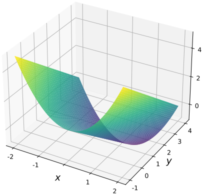
paramétrique → implicite

Toute partie de $\mathbb{R}^{n+1}$ décrite paramétriquement comme le graphe $Gr(f)$ d'une fonction f peut s'écrire implicitement comme la courbe de niveau 0 de la fonction $g : (\mathbf{x}, y) \mapsto f(\mathbf{x}) - y$

$$\mathcal{C}_0(g) = \left\{(\mathbf{x}, y) \in \mathbb{R}^{n+1}, g(\mathbf{x}, y) = 0\right\} = \left\{(\mathbf{x}, y) \in \mathbb{R}^n \times \mathbb{R}, y = f(\mathbf{x})\right\} = Gr(f)$$



$Gr(f(x) = x^2)$

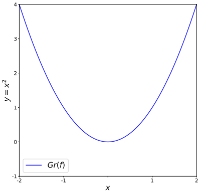


$Gr(g(x, y) = x^2 - y)$

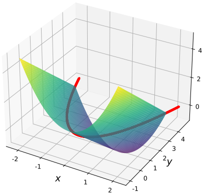
paramétrique $\rightarrow$ implicite

Toute partie de $\mathbb{R}^{n+1}$ décrite paramétriquement comme le graphe $Gr(f)$ d'une fonction f peut s'écrire implicitement comme la courbe de niveau 0 de la fonction $g : (\mathbf{x}, y) \mapsto f(\mathbf{x}) - y$

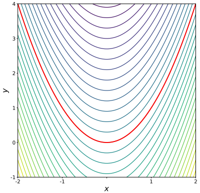
$$\mathcal{C}_0(g) = \left\{(\mathbf{x}, y) \in \mathbb{R}^{n+1}, g(\mathbf{x}, y) = 0\right\} = \left\{(\mathbf{x}, y) \in \mathbb{R}^n \times \mathbb{R}, y = f(\mathbf{x})\right\} = Gr(f)$$



$Gr(f(x) = x^2)$



$\mathcal{C}_0(g)$



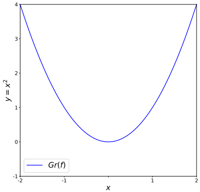
$\mathcal{C}_0(g) = Gr(f)$

Lien entre les représentations paramétriques et implicites

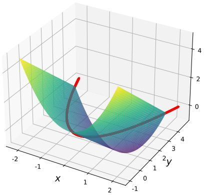
paramétrique $\rightarrow$ implicite

Toute partie de $\mathbb{R}^{n+1}$ décrite paramétriquement comme le graphe $Gr(f)$ d'une fonction f peut s'écrire implicitement comme la courbe de niveau 0 de la fonction $g : (\mathbf{x}, y) \mapsto f(\mathbf{x}) - y$

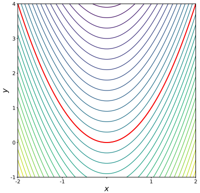
$$\mathcal{C}_0(g) = \left\{(\mathbf{x}, y) \in \mathbb{R}^{n+1}, g(\mathbf{x}, y) = 0\right\} = \left\{(\mathbf{x}, y) \in \mathbb{R}^n \times \mathbb{R}, y = f(\mathbf{x})\right\} = Gr(f)$$



$Gr(f(x) = x^2)$



$\mathcal{C}_0(g)$



$\mathcal{C}_0(g) = Gr(f)$

implicite $\rightarrow$ paramétrique

⛔ L'écriture paramétrique d'une partie donnée implicitement n'est en général pas possible (cf théorème des fonctions implicites).

Une partie $A \subset \mathbb{R}^n$ est dite *convexe* si et seulement si

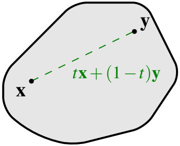
$$\boxed{\forall \mathbf{x}, \mathbf{y} \in A, \forall t \in [0,1], t\mathbf{x} + (1-t)\mathbf{y} \in A}$$

A est convexe si on peut relier n'importe quelle paire de points $\mathbf{x}, \mathbf{y} \in A$ par un segment qui reste intégralement inclus dans A

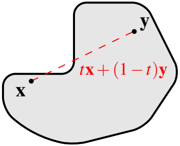
Une partie $A \subset \mathbb{R}^n$ est dite *convexe* si et seulement si

$$\boxed{\forall \mathbf{x}, \mathbf{y} \in A, \forall t \in [0,1], t\mathbf{x} + (1-t)\mathbf{y} \in A}$$

A est convexe si on peut relier n'importe quelle paire de points $\mathbf{x}, \mathbf{y} \in A$ par un segment qui reste intégralement inclus dans A



$\rightarrow$ convexe ✓



$\rightarrow$ pas convexe ✗

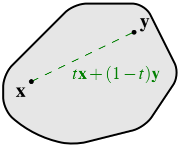
Ensemble convexe

Définition

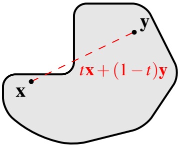
Une partie $A \subset \mathbb{R}^n$ est dite *convexe* si et seulement si

$$\boxed{\forall \mathbf{x}, \mathbf{y} \in A, \forall t \in [0,1], t\mathbf{x} + (1-t)\mathbf{y} \in A}$$

A est convexe si on peut relier n'importe quelle paire de points $\mathbf{x}, \mathbf{y} \in A$ par un segment qui reste intégralement inclus dans A



→ convexe ✓



→ pas convexe ✗

Exemples de parties convexes

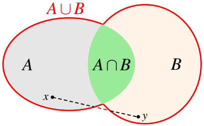
→ les intervalles $[a, b] \in \mathbb{R}$ sont convexes, $(a, b \in \mathbb{R} \cup \{-\infty, +\infty\})$

→ les sous-espaces affines de $\mathbb{R}^n$ sont convexes (droites, plans, hyperplans ...)

→ les demi-espaces de $\mathbb{R}^n$ sont convexes

Ensemble convexe

Propriétés

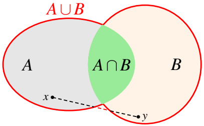


- $\rightarrow$ Si $A_1, \dots A_n$ sont convexes, alors $\bigcap_{i=1}^{n} A_i$ est convexe : la convexité est stable par intersection.
- $\rightarrow$ L'union de parties convexes n'est en général pas convexe.
- $\rightarrow$ A est convexe si et seulement si toute combinaison convexe d'une famille finie de points $\mathbf{x}_1, \dots, \mathbf{x}_p$ de A est dans A : $\sum_{i=1}^{p} \lambda_i \mathbf{x}_i \in A$ avec $\sum_{i=1}^{p} \lambda_i = 1$.

Propriétés

→ Si $A_1, \dots A_n$ sont convexes, alors $\bigcap_{i=1}^{n} A_i$ est convexe : la convexité est stable par intersection.

→ L'union de parties convexes n'est en général pas convexe.

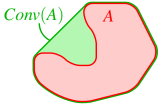


→ A est convexe si et seulement si toute combinaison convexe d'une famille finie de points $\mathbf{x}_1, \dots, \mathbf{x}_p$ de A est dans A : $\sum_{i=1}^{p} \lambda_i \mathbf{x}_i \in A$ avec $\sum_{i=1}^{p} \lambda_i = 1$.

Définitions

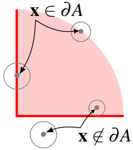
→ On appelle *dimension* d'un ensemble convexe A la dimension du plus petit sous-espace affine contenant A.

→ On appelle *enveloppe convexe* $Conv(A)$ d'une partie $A \subset \mathbb{R}^n$ l'intersection de toutes les parties convexes contenant A. C'est la plus petite sous-partie convexe de R^n qui contient A.

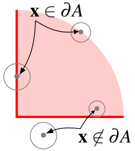


On dit que $\mathbf{x}$ est un *point du bord* de A si $\forall\varepsilon > 0, \mathcal{B}(\mathbf{x},\varepsilon) \cap A \neq \emptyset$ et $\mathcal{B}(\mathbf{x},\varepsilon) \cap \complement_A \neq \emptyset$

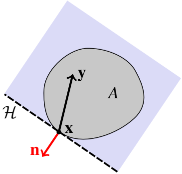
On appelle *bord* (ou *frontière*) de A l'ensemble ∂A des points de son bord



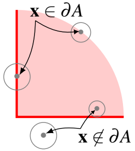
On dit que $\mathbf{x}$ est un *point du bord* de A si $\forall \varepsilon > 0, \mathcal{B}(\mathbf{x}, \varepsilon) \cap A \neq \emptyset$ et $\mathcal{B}(\mathbf{x}, \varepsilon) \cap \complement_A \neq \emptyset$
On appelle *bord* (ou *frontière*) de A l'ensemble ∂A des points de son bord



Soit $A \in \mathbb{R}^n$ et $\mathbf{x} \in A$. On appelle *hyperplan d'appui* à A en $\mathbf{x}$ un hyperplan $\mathcal{H}$ de vecteur normal $\mathbf{n}$ tel que $\mathbf{x} \in \mathcal{H}$, et $\forall \mathbf{y} \in A, \langle \mathbf{n}, \mathbf{y} - \mathbf{x} \rangle = \mathbf{n}^T(\mathbf{y} - \mathbf{x}) \leq 0$

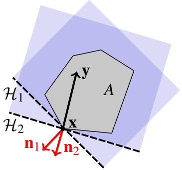


On dit que $\mathbf{x}$ est un *point du bord* de A si $\forall \varepsilon > 0, \mathcal{B}(\mathbf{x}, \varepsilon) \cap A \neq \emptyset$ et $\mathcal{B}(\mathbf{x}, \varepsilon) \cap \complement_A \neq \emptyset$
On appelle *bord* (ou *frontière*) de A l'ensemble ∂A des points de son bord

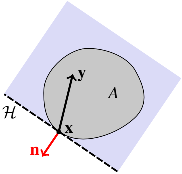


Soit $A \in \mathbb{R}^n$ et $\mathbf{x} \in A$. On appelle *hyperplan d'appui* à A en $\mathbf{x}$ un hyperplan $\mathcal{H}$ de vecteur normal $\mathbf{n}$ tel que $\mathbf{x} \in \mathcal{H}$, et $\forall \mathbf{y} \in A, \langle \mathbf{n}, \mathbf{y} - \mathbf{x} \rangle = \mathbf{n}^T(\mathbf{y} - \mathbf{x}) \leq 0$

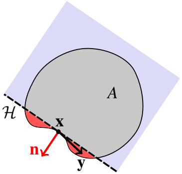
En un point $\mathbf{x}$ donné, un ensemble A peut admettre :



plusieurs hyperplans d'appui

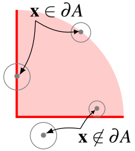


un hyperplan d'appui



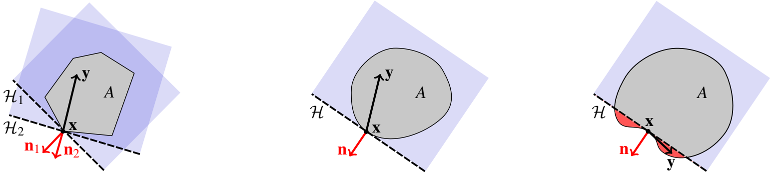
aucun hyperplan d'appui

On dit que $\mathbf{x}$ est un *point du bord* de A si $\forall \varepsilon > 0, \mathcal{B}(\mathbf{x}, \varepsilon) \cap A \neq \emptyset$ et $\mathcal{B}(\mathbf{x}, \varepsilon) \cap \complement_A \neq \emptyset$
On appelle *bord* (ou *frontière*) de A l'ensemble ∂A des points de son bord



Soit $A \in \mathbb{R}^n$ et $\mathbf{x} \in A$. On appelle *hyperplan d'appui* à A en $\mathbf{x}$ un hyperplan $\mathcal{H}$ de vecteur normal $\mathbf{n}$ tel que $\mathbf{x} \in \mathcal{H}$, et $\forall \mathbf{y} \in A, \langle \mathbf{n}, \mathbf{y} - \mathbf{x} \rangle = \mathbf{n}^T(\mathbf{y} - \mathbf{x}) \leq 0$

En un point $\mathbf{x}$ donné, un ensemble A peut admettre :



plusieurs hyperplans d'appui | un hyperplan d'appui | aucun hyperplan d'appui

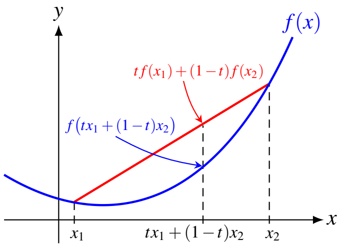
Une partie convexe admet un hyperplan d'appui en tout point de son bord

Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est dite *convexe* si

- son domaine de définition D_f est un ensemble convexe
- $\forall \mathbf{x}_1, \mathbf{x}_2 \in D_f, \forall t \in [0,1],\ \boxed{f\big(t\mathbf{x}_1 + (1-t)\mathbf{x}_2\big) \leq tf(\mathbf{x}_1) + (1-t)f(\mathbf{x}_2)}$

Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est dite *convexe* si

- son domaine de définition D_f est un ensemble convexe
- $\forall \mathbf{x}_1, \mathbf{x}_2 \in D_f, \forall t \in [0,1],\ \boxed{f\big(t\mathbf{x}_1 + (1-t)\mathbf{x}_2\big) \leq tf(\mathbf{x}_1) + (1-t)f(\mathbf{x}_2)}$

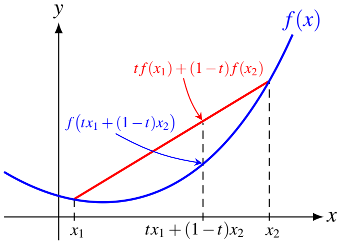


Caractérisation à l'ordre 0 de la convexité

Le graphe d'une fonction f convexe est *en dessous* de n'importe quel segment reliant les points $(\mathbf{x}_1, f(\mathbf{x}_1))$ et $(\mathbf{x}_2, f(\mathbf{x}_2))$, peu importe ces points.

Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est dite *convexe* si

- son domaine de définition D_f est un ensemble convexe
- $\forall \mathbf{x}_1, \mathbf{x}_2 \in D_f, \forall t \in [0,1],$ $\boxed{f\big(t\mathbf{x}_1 + (1-t)\mathbf{x}_2\big) \leq tf(\mathbf{x}_1) + (1-t)f(\mathbf{x}_2)}$



Caractérisation à l'ordre 0 de la convexité

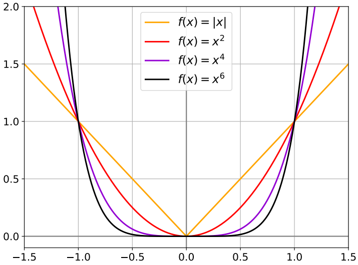
Le graphe d'une fonction f convexe est *en dessous* de n'importe quel segment reliant les points $(\mathbf{x}_1, f(\mathbf{x}_1))$ et $(\mathbf{x}_2, f(\mathbf{x}_2))$, peu importe ces points.

$\to$ On parle de *convexité stricte* si $f(t\mathbf{x}_1 + (1-t)\mathbf{x}_2) < tf(\mathbf{x}_1) + (1-t)f(\mathbf{x}_2)$

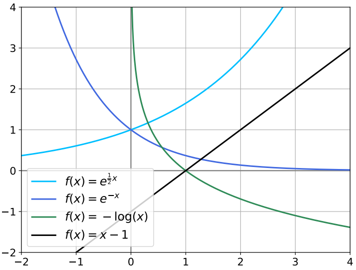
$\to$ On dit que f est *concave* si son opposé $-f$ est convexe.

Exemples de fonctions convexes

- $x \mapsto |x|$
- $x \mapsto x^n$, n pair
- $x \mapsto e^{\alpha x}$, $\forall \alpha \in \mathbb{R}$



- $x \mapsto -\log(x)$ ($\log(x)$ est concave)
- $x \mapsto ax^2 + bx + c$, $a > 0$
- $x \mapsto bx + c$, $\forall b, c \in \mathbb{R}$

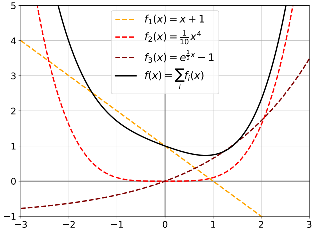


Une fonction affine $x \mapsto bx + c$ est convexe **et** concave.

$\Rightarrow$ Les seules fonctions convexes et concaves sont les fonctions affines.

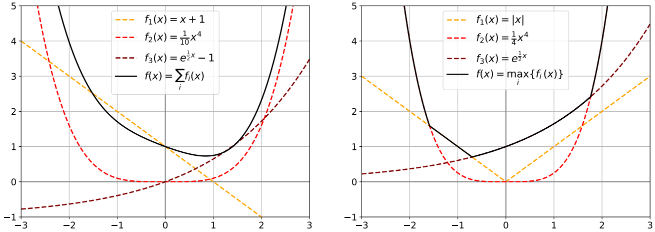
Composition de fonctions convexes

- La somme pondérée $\sum_i \omega_i f_i$ de poids positifs $\omega_i \geq 0$ de fonctions convexes f_i est convexe.



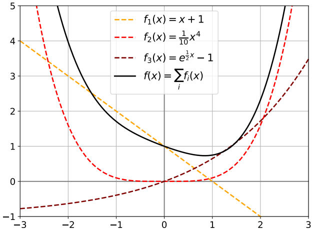
Composition de fonctions convexes

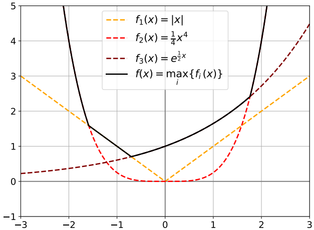
- La somme pondérée $\sum_i \omega_i f_i$ de poids positifs $\omega_i \geq 0$ de fonctions convexes f_i est convexe.
- Le maximum $\max_i f_i$ de fonctions convexes f_i est convexe.



Composition de fonctions convexes

- La somme pondérée $\sum_i \omega_i f_i$ de poids positifs $\omega_i \geq 0$ de fonctions convexes f_i est convexe.
- Le maximum $\max_i f_i$ de fonctions convexes f_i est convexe.





- La composition $g \circ f$ d'une fonction convexe f avec une fonction convexe croissante g est convexe.
- La composition $f(\mathbf{Ax} + \mathbf{b})$ d'une fonction convexe f avec une fonction affine $\mathbf{Ax} + \mathbf{b}$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, est convexe.

Norme et convexité

On appelle fonction *sous-linéaire* tout fonction $f : \mathbb{R}^n \to \mathbb{R}$ telle que :

- $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, f(\mathbf{x}+\mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y})$ (sous-additivité)
- $\forall \lambda > 0, f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$ (homogénéité positive)

Propriété : toute fonction sous-linéaire est convexe

Norme et convexité

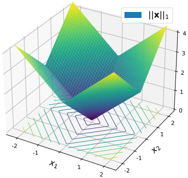
On appelle fonction *sous-linéaire* tout fonction $f : \mathbb{R}^n \to \mathbb{R}$ telle que :

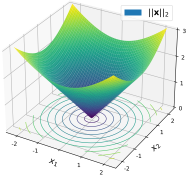
- $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, f(\mathbf{x}+\mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y})$ (sous-additivité)
- $\forall \lambda > 0, f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$ (homogénéité positive)

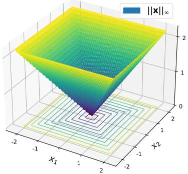
Propriété : toute fonction sous-linéaire est convexe

En particulier, toutes les normes sur $\mathbb{R}^n$ sont des applications sous-linéaires.
$\to \|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ = inégalité triangulaire.

Toutes les normes sont des fonctions convexes







Norme et convexité

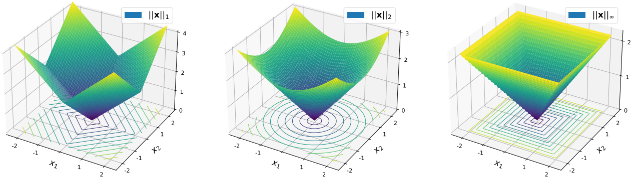
On appelle fonction *sous-linéaire* tout fonction $f : \mathbb{R}^n \to \mathbb{R}$ telle que :

- $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, f(\mathbf{x} + \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y})$ (sous-additivité)
- $\forall \lambda > 0, f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$ (homogénéité positive)

Propriété : toute fonction sous-linéaire est convexe

En particulier, toutes les normes sur $\mathbb{R}^n$ sont des applications sous-linéaires.
$\to \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| =$ inégalité triangulaire.

Toutes les normes sont des fonctions convexes

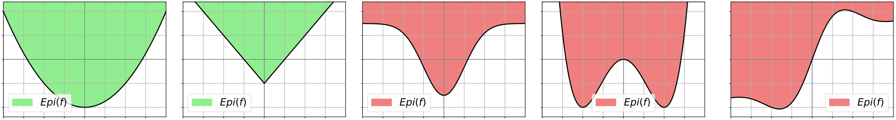


$\Rightarrow$ $\boxed{\min_{x} \|\mathbf{A}\mathbf{x} - \mathbf{b}\| \text{ est un problème convexe}}$ (peu importe la norme $\|\cdot\|$)

Lien entre fonction convexe et partie convexe

Lien entre graphe et épigraphe

Une fonction f est convexe si et seulement si son épigraphe $Epi(f)$ est une partie convexe

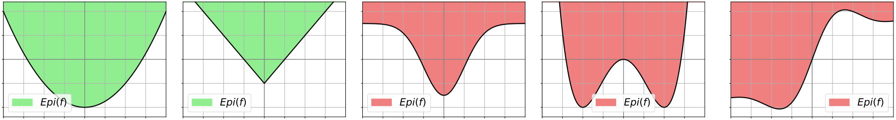


$f(x) = \frac{1}{2}x^2 - 1$ $f(x) = |x| - \frac{1}{2}$ $f(x) = \frac{3}{4} - \frac{3}{2}e^{-2x^2}$ $f(x) = x^4 - 2x^2$ $f(x) = \frac{2x}{1+x^2} + \frac{1}{10}\sin(\pi x)$

Lien entre fonction convexe et partie convexe

Lien entre graphe et épigraphe

Une fonction f est convexe si et seulement si son épigraphe $Epi(f)$ est une partie convexe

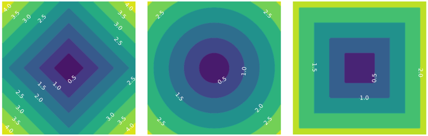


$f(x) = \frac{1}{2}x^2 - 1$ $\quad$ $f(x) = |x| - \frac{1}{2}$ $\quad$ $f(x) = \frac{3}{4} - \frac{3}{2}e^{-2x^2}$ $\quad$ $f(x) = x^4 - 2x^2$ $\quad$ $f(x) = \frac{2x}{1+x^2} + \frac{1}{10}\sin(\pi x)$

Lien entre graphe et lieux de sous niveau

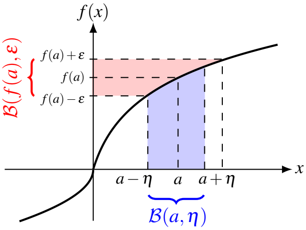
Les lieux de sous niveau d'une fonction convexe sont des parties convexes (⚠ la réciproque n'est pas vraie)

$\Rightarrow$ Puisque les normes $\|\cdot\|$ sont des fonctions convexes, les boules $\mathcal{B}_{\|\cdot\|}(\mathbf{0}, r) = \mathcal{C}_{\leq r}(\|\cdot\|)$ qui en découlent sont des parties convexes



Continuité d'une fonction $f : \mathbb{R} \to \mathbb{R}$

On dit que f est *continue* en $a \in D_f$ si et seulement si $f(x) \xrightarrow[x \to a]{} f(a)$



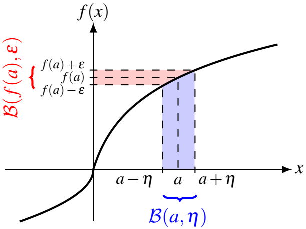
La continuité se réécrit en termes de voisinages :

$$\forall \varepsilon > 0, \exists\, \eta > 0, |x - a| < \eta \;\Rightarrow |f(x) - f(a)| < \varepsilon$$

$$\Leftrightarrow \forall \varepsilon > 0, \exists\, \eta > 0, x \in \mathcal{B}(a, \eta) \;\Rightarrow f(x) \in \mathcal{B}(f(a), \varepsilon)$$

Continuité d'une fonction $f : \mathbb{R} \to \mathbb{R}$

On dit que f est *continue* en $a \in D_f$ si et seulement si $f(x) \underset{x \to a}{\longrightarrow} f(a)$



La continuité se réécrit en termes de voisinages :

$$\forall \varepsilon > 0, \exists\, \eta > 0, |x - a| < \eta \;\Rightarrow |f(x) - f(a)| < \varepsilon$$

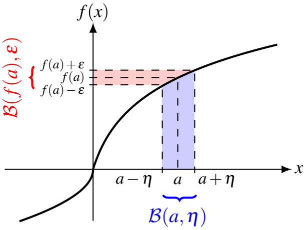
$$\Leftrightarrow \forall \varepsilon > 0, \exists\, \eta > 0, x \in \mathcal{B}(a, \eta) \;\Rightarrow f(x) \in \mathcal{B}(f(a), \varepsilon)$$

Pour tout voisinage $\mathcal{B}(f(a), \varepsilon)$ de $f(a)$, il existe un voisinage $\mathcal{B}(a, \eta)$ de a tel que tout élément $x \in \mathcal{B}(a, \eta)$ a son image $f(x) \in \mathcal{B}(f(a), \varepsilon)$.

Rappels sur la continuité d'une fonction

Continuité d'une fonction $f : \mathbb{R} \to \mathbb{R}$

On dit que f est *continue* en $a \in D_f$ si et seulement si $f(x) \underset{x \to a}{\longrightarrow} f(a)$



La continuité se réécrit en termes de voisinages :

$$\forall \varepsilon > 0, \exists \eta > 0, |x-a| < \eta \Rightarrow |f(x) - f(a)| < \varepsilon$$

$$\Leftrightarrow \forall \varepsilon > 0, \exists \eta > 0, x \in \mathcal{B}(a,\eta) \Rightarrow f(x) \in \mathcal{B}(f(a),\varepsilon)$$

Pour tout voisinage $\mathcal{B}(f(a),\varepsilon)$ de $f(a)$, il existe un voisinage $\mathcal{B}(a,\eta)$ de a tel que tout élément $x \in \mathcal{B}(a,\eta)$ a son image $f(x) \in \mathcal{B}(f(a),\varepsilon)$.

Continuité d'une fonction $f : \mathbb{R}^n \to \mathbb{R}^m$

La définition avec les voisinages s'adapte (et ne dépend pas de la norme $\|\cdot\|$ choisie) :

$$\forall \varepsilon > 0, \exists \eta > 0, \mathbf{x} \in \mathcal{B}_{\|\cdot\|}(\mathbf{a},\eta) \Rightarrow f(\mathbf{x}) \in \mathcal{B}_{\|\cdot\|}(f(\mathbf{a}),\varepsilon)$$

Règles de continuité d'une fonction

Comment déterminer qu'une certaine fonction est continue :

- Une fonction $f : \mathbb{R}^n \to R^m$, $f(\mathbf{x}) = \big(f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})\big)$ est continue si chaque "fonction-coordonnée" $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$ est continue.
- Toute fonction $f : \mathbb{R}^n \to \mathbb{R}$ polynomiale est continue.
- Toute fonction rationnelle $r = \frac{f}{g}$ avec $f, g : \mathbb{R}^n \to \mathbb{R}$ polynomiales est continue partout ou g ne s'annule pas $(g(\mathbf{x}) \neq 0)$.
- Si f et g sont continues, alors $\lambda f + \mu g$ est continue $\forall \lambda, \mu \in \mathbb{R}$ (structure d'espace vectoriel).
- Si $f, g : \mathbb{R}^n \to \mathbb{R}$ sont continues, alors fg est continue et $\frac{f}{g}$ continue partout ou g ne s'annule pas $(g(x) \neq 0)$.
- Si $f : \mathbb{R}^n \to \mathbb{R}^p$ et $g : \mathbb{R}^m \to \mathbb{R}^n$ sont continues, alors $g \circ f$ est continue.

Continuité d'une fonction à plusieurs variables

Il est tentant d'évaluer la continuité d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$, en regardant les fonctions partielles selon les directions de la base canonique uniquement.

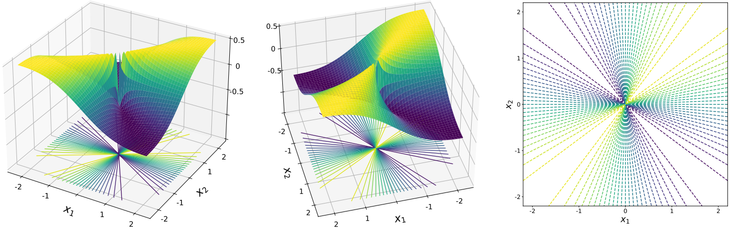
Malheureusement, il ne suffit pas que chaque fonction $t \mapsto f(\mathbf{a} + t\mathbf{e}_i)$ (restriction de f le long de l'axe engendré par $\mathbf{e}_i$) soit continue en $\mathbf{a}$ pour que f le soit.

Il est tentant d'évaluer la continuité d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$, en regardant les fonctions partielles selon les directions de la base canonique uniquement.

Malheureusement, il ne suffit pas que chaque fonction $t \mapsto f(\mathbf{a} + t\mathbf{e}_i)$ (restriction de f le long de l'axe engendré par $\mathbf{e}_i$) soit continue en $\mathbf{a}$ pour que f le soit.

Exemple pathologique classique

Soit $f : \mathbb{R}^2 \to \mathbb{R}$, $(x_1, x_2) \mapsto \dfrac{x_1 x_2}{x_1^2 + x_2^2}$ si $(x_1, x_2) \neq (0,0)$, 0 si $(x_1, x_2) = (0,0)$

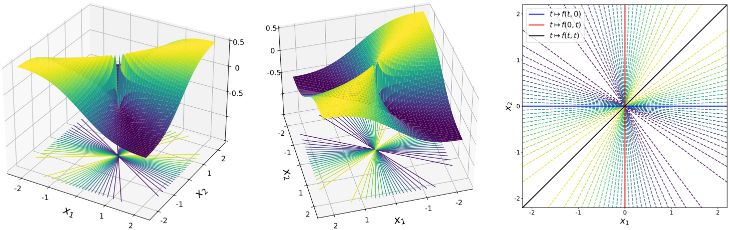


Il est tentant d'évaluer la continuité d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$, en regardant les fonctions partielles selon les directions de la base canonique uniquement.

Malheureusement, il ne suffit pas que chaque fonction $t \mapsto f(\mathbf{a} + t\mathbf{e}_i)$ (restriction de f le long de l'axe engendré par $\mathbf{e}_i$) soit continue en $\mathbf{a}$ pour que f le soit.

Exemple pathologique classique

Soit $f : \mathbb{R}^2 \to \mathbb{R}$, $(x_1, x_2) \mapsto \dfrac{x_1 x_2}{x_1^2 + x_2^2}$ si $(x_1, x_2) \neq (0,0)$, 0 si $(x_1, x_2) = (0,0)$

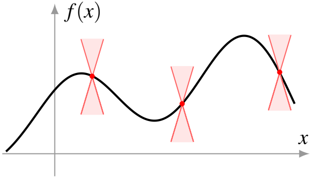


$\forall t \in \mathbb{R}$, $t \mapsto f(t,0) = 0$ et $t \mapsto f(0,t) = 0 \Rightarrow$ les fonctions partielles sont continues

Mais $t \mapsto f(t,t) = \frac{1}{2}$ si $t \neq 0$, 0 si $t = 0 \quad \Rightarrow f$ n'est pas continue en $(0,0)$

On dit que f est *k-lipschitzienne* ($k \geq 0$) si et seulement si $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \|f(\mathbf{x}) - f(\mathbf{y})\| \leq k\|\mathbf{x} - \mathbf{y}\|$

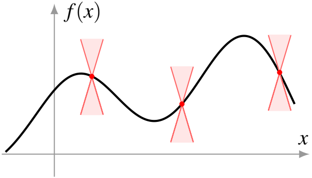
Tout fonction k-lipschitzienne est continue



Fonction k-lipschitzienne

On dit que f est *k-lipschitzienne* ($k \geq 0$) si et seulement si $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \|f(\mathbf{x}) - f(\mathbf{y})\| \leq k\|\mathbf{x} - \mathbf{y}\|$

Tout fonction k-lipschitzienne est continue



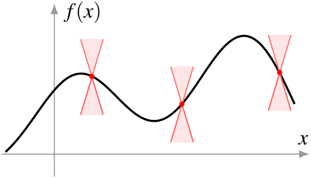
En particulier si $f = \|\cdot\|$: $\|f(\mathbf{x}) - f(\mathbf{y})\| = \big|\|\mathbf{x}\| - \|\mathbf{y}\|\big| \leq \|\mathbf{x} - \mathbf{y}\|$ (inégalité triangulaire inversée)

$\rightarrow$ $\boxed{\text{Toute norme est 1-lipschitzienne, donc continue}}$

Fonction k-lipschitzienne

On dit que f est *k-lipschitzienne* ($k \geq 0$) si et seulement si $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \|f(\mathbf{x}) - f(\mathbf{y})\| \leq k\|\mathbf{x} - \mathbf{y}\|$

Tout fonction k-lipschitzienne est continue



En particulier si $f = \|\cdot\|$: $\|f(\mathbf{x}) - f(\mathbf{y})\| = \big|\|\mathbf{x}\| - \|\mathbf{y}\|\big| \leq \|\mathbf{x} - \mathbf{y}\|$ (inégalité triangulaire inversée)

$\rightarrow$ Toute norme est 1-lipschitzienne, donc continue

Une fonction f k-lipschitzienne est dite:

- *non expansive* si $k \leq 1 \rightarrow$ c'est le cas des normes
- *contractante* si $k < 1$ $\rightarrow$ se rencontrent fréquemment en optimisation dans les algorithmes itératifs de type point fixe

Soit f une fonction contractante avec $0 < k < 1$. Alors f admet un unique point fixe $\tilde{x}$ tel que $\tilde{x} = f(\tilde{x})$. Toute suite d'éléments vérifiant $x_{n+1} = f(x_n)$ vérifie la majoration $\|x_n - \tilde{x}\|_2 \leq \frac{k^n}{1-k}\|x_1 - x_0\|_2$, donc converge vers $\tilde{x}$.