UNIVERSITÉ DE BOURGOGNE
UFR Droit, Sciences Économique et Politique

LICENCE ÉCONOMIE–GESTION
L1/S1, Année 2022/2023

MATHÉMATIQUES

Cours par M. TRIESTINO
contact : michele.triestino@u-bourgogne.fr
page de cours : https://mtriestino.perso.math.cnrs.fr/eco.html

version du 29 novembre 2022

MODALITÉS D'EXAMEN

- **Le contrôle continu (CC)**
  Il est constitué de deux épreuves écrites organisées en TD.

- **Le contrôle terminal (CT)**
  Il s'agit d'une épreuve écrite de 2h, organisée en amphi.
  La note finale est calculée suivant la formule NF=(CC + 2CT)/3.

- **La session de rattrapage**
  Il s'agit d'une épreuve écrite de 2h qui se déroule dans les mêmes conditions que celles de l'examen de première session.

PLANNING PREVISIONNEL

Semaine	**Cours**	**TD**
5 sept – 9 sept	C1, C2	–
12 sept – 16 sept	C3, C4	–
19 sept – 23 sept	C5	TD1
26 sept – 30 sept	C6	TD2
3 oct – 7 oct	C7	TD3
10 oct – 14 oct	–	TD4
17 oct – 21 oct	C8	TD5 (CC n°1)
24 oct – 28 oct	C9	TD6
31 oct – 4 nov	Interruption pédagogique	
7 nov – 11 nov	C10	TD7
14 nov – 19 nov	C11	TD8
21 nov – 26 nov	C12	TD9 (CC n°2)
28 nov – 2 déc	–	–
5 déc – 9 déc	Première semaine des examens	
12 déc – 16 déc	Deuxième semaine des examens	

RÉSUMÉ DE COURS

Objectifs principaux

Dans ce premier cours de mathématiques, nous allons découvrir des notions fondamentales sur les fonctions, principalement de une ou deux variables, et la géométrie de l'espace euclidien. Voici une liste des objectifs, qui correspondent *grosso modo* aux chapitres de ce cours.

1. Décrire l'allure du graphe d'une fonction de une et deux variables, et notamment étudier la nature des points critiques.
2. Représenter des parties du plan.
3. Éléments de géométrie de l'espace euclidien $\mathbb{R}^n$.
4. Reconnaître des propriétés géométriques des fonctions (fonctions affines, homogènes, convexes/concaves).

Ces notions sont essentielles pour la suite de la formation en mathématiques pour l'économie, notamment pour savoir traiter des *problèmes d'optimisation* (par exemple, savoir réduire les coûts, ou maximiser un gain).

Table des matières

1 Fonctions élémentaires

1.1 Manipulation algébriques élémentaires

Voici quelques règles élémentaires de calcul algébrique, toujours utiles !

1. $(a^n)^m = a^{nm}$ (attention, ceci est différent de a^{n^m}).
2. $a^{-n} = \dfrac{1}{a^n}$.
3. $(a+b)^2 = a^2 + 2ab + b^2$.
4. $a^2 - b^2 = (a-b)(a+b)$.
5. $a^3 - b^3 = (a-b)(a^2 + ab + b^2)$.

1.2 L'opérateur Σ

Une expression du type

$$S := x_1 + x_2 + x_3 + \cdots + x_n$$

est fastidieuse à écrire et occupe une place importante. C'est pourquoi il est d'usage de simplifier l'écriture en employant un opérateur de sommation représenté par la lettre grecque Σ. En utilisant cet opérateur l'expression précédente s'écrit

$$S = \sum_{i=1}^{n} x_i.$$

Par exemple, la quantité $S := 2 + 2^2 + 2^3 + 2^4 + 2^5 + \cdots + 2^{15}$ s'écrit de façon plus compacte

$$S = \sum_{i=1}^{15} 2^i.$$

La lettre choisie pour l'indice de sommation n'a aucune importance. On dit qu'elle est *muette*. Ainsi il revient au même d'écrire

$$S = \sum_{i=1}^{n} x_i \quad \text{ou} \quad S = \sum_{j=1}^{n} x_j.$$

Il arrive aussi fréquemment que l'on soit amené, pour des raisons techniques, notamment des simplifications d'expressions, à faire un changement de variable sur l'indice. On peut ainsi écrire

$$\sum_{i=1}^{n} x_i = \sum_{i=0}^{n-1} x_{i+1}.$$

1.3 Polynômes : définitions et propriétés fondamentales

Les exemples les plus élémentaires de fonctions sont constitués par les *monômes*, à savoir les fonctions de la forme

$$P(x) = ax^n,$$

où a est une constante réelle, appelée le *coefficient* du monôme, et n est un entier positif, appelé le *degré*[1]. En faisant la somme de monômes de différents degrés, on obtient les *polynômes*.

Définition (Polynôme). Un *polynôme* (à une variable) est une fonction de la forme

$$P(x) = a_0 + a_1 x + \cdots + a_n x^n,$$

où les a_i sont des réels, et $a_n \neq 0$ lorsque $n \geq 1$. Le plus haut exposant n s'appele le *degré* du polynôme[2], et les a_i sont ses *coefficients*. Le coefficient a_0 est aussi appelé le coefficient *constant*, et a_n est le coefficient *dominant*.

1. Sauf si $P(x) = 0$ est le monôme nul.
2. Sauf si $P(x) = 0$ est le polynôme nul.

Les polynômes de degré 0 s'appellent *constants*, ceux de degré 1 s'appellent *linéaires*, ceux de degré 2 s'appellent *quadratiques* (ou *trinômes du second degré*), etc.

Proposition. *Deux fonctions polynômes sont égales si et seulement si leurs coefficients respectifs sont égaux.*

On *additionne* deux polynômes en sommant les coefficients des monômes de même degré : si $n \leq m$, et

$$P(x) = a_0 + \cdots + a_n x^n, \quad Q(x) = b_0 + \cdots + b_m x^m,$$

alors

$$(P+Q)(x) = (a_0 + b_0) + \cdots + (a_n + b_n)x^n + b_{n+1}x^{n+1} + \cdots b_m x^m.$$

On peut également *multiplier* les deux polynômes, en utilisant la règle de multiplication $x^i x^j = x^{i+j}$:

$$(PQ)(x) = a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + \cdots + a_nb_mx^{n+m}.$$

Inversement, on *factorise* un polynôme lorsqu'on l'écrit sous la forme de produit de polynômes de degré strictement inférieur.

Définition (Racine d'un polynôme)**.** Soit P un polynôme non nul. Un nombre réel α est une *racine* de P si $P(\alpha) = 0$.

En d'autres termes, une racine est une solution de l'équation $P(x) = 0$. Factoriser facilite la résolution des équations et la détermination du signe d'une expression.

Exemple. Factoriser le polynôme $P(x) = x^2 + x - 2$ sous la forme

$$P(x) = (x-1)(x+2),$$

permet de résoudre l'équation $x^2 + x - 2 = 0$, et de déterminer le signe de la quantité $x^2 + x - 2$ en fonction de x. Ainsi, grâce à la factorisation, on voit que l'équation $x^2 + x - 2 = 0$ a deux solutions $x = 1$ et $x = -2$. On en déduit aussi que la quantité $x^2 + x - 2$ est positive (ou nulle) pour $x \leq -2$ ou $x \geq 1$ et qu'elle est négative (ou nulle) pour $-2 \leq x \leq 1$.

Lorsque l'on cherche à factoriser pour résoudre une équation ou une inéquation, l'objectif est de remplacer l'équation (l'inéquation) donnée par plusieurs équations (inéquations) plus simples. Dans l'exemple précédent, avoir factorisé a permis de remplacer une équation du second degré par deux équations du premier degré en vertu du principe

pour qu'un produit soit nul, il faut et il suffit que l'un de ses facteurs soit nul.

Proposition (Lien entre racine et factorisation)**.** *Si $\alpha \in \mathbb{R}$ est racine du polynôme P, on peut le factoriser sous la forme*

$$P(x) = (x - \alpha)Q(x),$$

où $Q(x)$ est un polynôme de degré $n-1$.

Par conséquent, un polynôme de degré n possède au plus n racines distinctes.

On détermine alors facilement l'expression de $Q(x)$ par identification.

Exemple. Prenons le polynôme $P(x) = 5x^3 - 8x^2 + 5x - 2$. Manifestement $x = 1$ est racine puisque $P(1) = 0$. On a donc

$$\begin{aligned} P(x) &= (x-1)(ax^2 + bx + c) \\ &= ax^3 + (b-a)x^2 + (c-b)x - c. \end{aligned}$$

En identifiant les termes de plus haut et de plus bas degré, on obtient immédiatement $a = 5$ et $c = 2$. Il ne reste alors plus qu'à déterminer b. En identifiant les termes de degré 1, (ou ceux de degré 2) on obtient $b = -3$.

Remarque. Un polynôme peut être factorisable sans admettre de racine. Par exemple, le polynôme $P(x) = 1 + x^4$ se factorise sous la forme

$$P(x) = (x^2 + \sqrt{2}\,x + 1)(x^2 - \sqrt{2}\,x + 1).$$

Il n'admet pas de racine, puisque la quantité $1 + x^4$ est supérieure ou égale à 1, donc strictement positive.

On voit donc avec l'exemple précédent que tout polynôme n'est pas factorisable en produit de polynômes de degré 1. Toutefois, tout polynôme peut être factorisé en produit de polynômes de degré au plus 2.

1.4 Trinôme du second degré

On discute ici le cas particulier des polynômes quadratiques, donc de la forme

$$P(x) = ax^2 + bx + c$$

où $a \neq 0$, b et c sont des coefficients réels.

L'existence d'une racine réelle est liée au signe d'une quantité que l'on appelle le discriminant.

Définition (Discriminant d'un trinôme du second degré)**.** On appelle discriminant du trinôme $ax^2 + bx + c$ la quantité

$$\Delta := b^2 - 4ac.$$

Pour comprendre pourquoi le discriminant joue un rôle clé, il suffit de mettre le trinôme sous *forme canonique*, ce qui consiste à écrire $P(x)$ sous la forme

$$P(x) = a\left(x^2 + 2\,\frac{b}{2a}\,x + \frac{c}{a}\right)$$

et à voir dans cette expression le début d'un carré. En mettant celui-ci en évidence, on obtient

$$\begin{aligned} P(x) &= a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} + \frac{c}{a}\right] \\ &= a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2}\right]. \end{aligned}$$

— Si $\Delta < 0$, l'expression entre crochets reste toujours strictement positive et le trinôme est donc du signe de a, si bien qu'il n'y a pas de racine.

— Si $\Delta = 0$, l'expression entre crochets est un carré et le trinôme s'annule pour

$$x_0 = -\frac{b}{2a}$$

qui est donc l'unique racine. Dans ce cas, le trinôme admet comme factorisation

$$P(x) = a\,(x - x_0)\,2.$$

— Si $\Delta > 0$, on voit apparaître entre crochets une différence de carrés et en utilisant l'une des identités remarquables on peut écrire

$$P(x) = a\left[\left(x + \frac{b}{2a}\right) + \frac{\sqrt{\Delta}}{2a}\right]\left[\left(x + \frac{b}{2a}\right) - \frac{\sqrt{\Delta}}{2a}\right],$$

si bien qu'il y a deux racines

$$x_1 := \frac{-b - \sqrt{\Delta}}{2a} \quad \text{et} \quad x_2 := \frac{-b + \sqrt{\Delta}}{2a}.$$

Comme écrit plus haut, dans ce cas le trinôme admet comme factorisation

$$P(x) = a(x - x_1)(x - x_2).$$

Le calcul des racines nécessite donc, a priori, le calcul du discriminant. Mais en pratique il est conseillé, avant de calculer le discriminant, de regarder si le trinôme n'a pas de racine évidente. On rappelle que si $x = \alpha$ est racine, alors $(x - \alpha)$ se met en facteur si bien que l'on peut écrire le trinôme sous la forme

$$P(x) = (x - \alpha)(dx - e)$$

ce qui permet, par identification, de déterminer immédiatement $d = a$ et e (car $e\alpha = c$). En reformulant cela, on a le fait suivant, qui est très utile dans la pratique.

Proposition (Somme et produit des racines d'un trinôme)**.** *Étant donné le trinôme $ax^2 + bx + c$, si le discriminant Δ est positif ou nul, la somme σ et le produit p des racines sont respectivement donnés par $\sigma = -b/a$ et $p = c/a$.*

1.5 Valeur absolue

Définition (Valeur absolue). Étant donné $x \in \mathbb{R}$, on appelle valeur absolue de x la quantité définie par

$$|x| := \begin{cases} x & \text{si } x \geq 0 \\ -x & \text{si } x < 0. \end{cases}$$

De façon équivalente on peut aussi définir la valeur absolue par l'expression

$$|x| := \max\{x, -x\}.$$

Proposition (Sur une propriété de la valeur absolue). *Si $a \in \mathbb{R}$ et $\epsilon > 0$ sont donnés, on a pour tout $x \in \mathbb{R}$*

$$|x - a| \leq \epsilon \Leftrightarrow a - \epsilon \leq x \leq a + \epsilon.$$

La valeur absolue possède aussi la propriété suivante connue sous le nom d'inégalité triangulaire.

Proposition (Inégalité triangulaire pour la valeur absolue). *La valeur absolue vérifie l'inégalité triangulaire*

$$\forall x, y \in \mathbb{R}, \quad |x + y| \leq |x| + |y|.$$

1.6 D'autres fonctions élémentaires

Toutes les fonctions que nous allons rencontrer, sans que l'on spécifie, sont à valeur dans $\mathbb{R}$, c'est-à-dire à toute variable x, la valeur $f(x)$ est un nombre réel. Remarquons que pour toute fonction, il est nécessaire de spécifier son *domaine de définition*, à savoir l'ensemble des valeurs pour lesquelles la fonction est définie. Dans le cas où le domaine de définition n'est pas spécifié, par convention on considérera le plus grand ensemble de définition possible pour la fonction. Par exemple, tout polynôme est défini sur l'ensemble $\mathbb{R}$ des nombres réels. Si f est une fonction, on notera D_f son domaine de définition.

Voici une liste de fonctions élémentaires qu'on rencontrera dans ce cours.

1. Les *fractions rationnelles* sont les fonctions de la forme
$$f(x) = \frac{P(x)}{Q(x)},$$
où P et Q sont deux polynômes. Le domaine de définition est donné par
$$D_f = \{x \in \mathbb{R} : Q(x) \neq 0\} = \mathbb{R} \setminus \{Q(x) = 0\}.$$
Rappelons que la règle de somme de fractions rationnelles suit la règle de somme de fractions usuelles :
$$\frac{P_1(x)}{Q_1(x)} + \frac{P_2(x)}{Q_2(x)} = \frac{P_1(x)Q_2(x) + P_2(x)Q_1(x)}{Q_1(x)Q_2(x)}.$$
2. La fonction *exponentielle* $f(x) = e^x$ est définie pour tout réel x. Pour l'instant, on retiendra les propriétés suivantes :
   — $e^x > 0$ pour tout $x \in \mathbb{R}$,
   — $e^0 = 1$,
   — $e^1 = e$, où $e \simeq 2,71828\ldots$ est la constante de Néper,
   — $e^{x+y} = e^x e^y$,
   — elle est croissante (si $x < y$ alors $e^x < e^y$).
3. La fonction *logarithme naturel* $f(x) = \ln x$ est la réciproque de la fonction exponentielle, c'est-à-dire elle satisfait
$$\ln e^x = x \quad \text{et} \quad e^{\ln x} = x.$$
Puisque $e^x > 0$, le logarithme naturel n'est défini que sur l'intervalle $]0, +\infty[$. Pour l'instant, on retiendra les propriétés suivantes :

— $\ln 1 = 0$,
— $\ln e = 1$,
— $\ln(xy) = \ln x + \ln y$,
— $\ln(1/x) = -\ln x$,
— elle est croissante.

4. Pour tout réel $a > 0$, on peut définir la fonction exponentielle $f(x) = a^x$ (de base a), qui correspond à la fonction $e^{x \ln a}$. On retiendra la propriété $a^1 = a$.
   Également on peut définir la fonction logarithme $\log_a x$ (en base a). On retiendra la propriété $\log_a a = 1$.
5. En généralisant les monômes, on définit aussi les fonctions puissance $f(x) = x^\alpha$, où α est un réel. La nature de cette fonction dépend de la valeur α.
   — Si $\alpha = n \geq 0$ est un entier positif, on retrouve la fonction monôme. Cette fonction prend des valeurs toujours positives pour n pair.
   — Si $\alpha = -n < 0$ est un entier strictement négatif, on peut écrire $f(x) = 1/x^n$, et on retrouve une fraction rationnelle, définie sur $\mathbb{R} \setminus \{0\}$.
   — Si $\alpha = 1/q$ est un nombre rationnel, la fonction $f(x) = x^{1/q} = \sqrt[q]{x}$ (*racine q-ème*) est la fonction réciproque de x^q. Par multiplication, cela permet de définir aussi la fonction $x^{p/q} = \sqrt[q]{x^p} = (\sqrt[q]{x})^p$, pour p/q rationnel quelconque.
   — En général, si α est un réel quelconque (irrationnel), on définit la fonction puissance $x^\alpha = e^{\alpha \ln x}$. Dans ce cas, puisque la fonction $\ln x$ n'est définie que sur $]0, +\infty[$, on a que l'expression x^α est bien définie uniquement pour $x > 0$. Lorsque $\alpha > 0$, on peut « prolonger par continuité » la fonction x^α en 0, en posant $0^\alpha = 0$.

   Voici un récapitulatif pour le domaine de définition.

α	domaine
entier ≥ 0	$\mathbb{R}$
entier < 0	$\mathbb{R} \setminus \{0\}$
rationnel $p/q > 0$, q pair	$[0, +\infty[$
rationnel $p/q > 0$, q impair	$\mathbb{R}$
rationnel $p/q < 0$, q pair	$]0, +\infty[$
rationnel $p/q < 0$, q impair	$\mathbb{R} \setminus \{0\}$
irrationnel ≥ 0	$[0, +\infty[$
irrationnel < 0	$]0, +\infty[$

1.7 Opérations entre fonctions

À partir des fonctions élémentaires, on obtient d'autres fonctions par moyen d'opérations élémentaires.

1. Les opérations classiques s'étendent aux fonctions : si f, g sont deux fonctions, on peut définir les nouvelles fonctions $f + g$, $f - g$, fg, f/g. Le domaine de définition de ces nouvelles fonctions est donné par l'intersection $D_f \cap D_g$, et pour le dernier cas on a $D_{f/g} = (D_f \cap D_g) \setminus \{g(x) = 0\}$.
2. On peut aussi définir les fonctions $\min\{f, g\}$ et $\max\{f, g\}$, et donc $|f| = \max\{f, -f\}$.
3. Enfin, on peut aussi *composer* deux fonctions, en écrivant $f \circ g$, ce qui correspond à la fonction

$$(f \circ g)(x) = f(g(x)).$$

Le domaine de définition dans ce cas est donné par

$$D_{f\circ g} = \{x \in \mathbb{R} : x \in D_g \text{ et } g(x) \in D_f\}.$$

Définition. Deux fonctions f et g sont *réciproques* si $(f \circ g)(x) = x$ et $(g \circ f)(x) = x$ pour tout x où ces compositions sont définies.

Par exemple, les fonctions x^α et $x^{1/\alpha}$ sont réciproques.

La composition des fonctions est très utile en pratique car elle permet, pour étudier les propriétés d'une fonction complexe, de la décomposer en fonctions plus simples et de se ramener à l'étude des propriétés de celles-ci.

Exemple. La fonction

$$\varphi(x) := e^{2x^2+1}$$

se décompose sous la forme

$$\varphi = g \circ f, \quad \text{avec } f(x) := 2x^2 + 1 \text{ et } g(t) := e^t.$$

Exemple. Considérons la fonction $f(x) := \sqrt{1-x^2} = (1-x^2)^{1/2}$. Cette fonction est obtenue comme la composition $f = g \circ h$, où $g(x) = x^{1/2}$ et $h(x) = 1 - x^2$. La fonction h étant un polynôme, elle est définie sur $\mathbb{R}$. En revanche, le domaine de définition de g est $[0, +\infty[$, et donc

$$D_f = \{x \in \mathbb{R} : h(x) \in [0, +\infty[\} = \{x \in \mathbb{R} : 1 - x^2 \geq 0\} = [-1, 1].$$

1.8 Fonctions de plusieurs variables

Typiquement, les fonctions peuvent dépendre de plusieurs variables. Dans ce cours nous considérons principalement des fonctions de une ou deux variables réelles, c'est-à-dire définies sur une partie de la droite réelle $\mathbb{R}$ ou du plan $\mathbb{R}^2$. Par exemple, on peut vouloir étudier des fonctions du type $f(x, y) = e^{x+y}$, $g(x, y) = x^3y - 2x + y^2 - 3$, ou $h(x, y) = \min\{x, y\}$. À valeur de y fixée, ces fonctions sont des fonctions d'une variable (la variable x). De même, lorsque l'on fixe la valeur de x, on retrouve une fonction d'une variable (la variable y).

Un *monôme* de deux variables est une fonction de la forme $f(x, y) = ax^iy^j$, avec $a \in \mathbb{R}$ et $i, j \in \mathbb{N}$. La somme $i + j$ est le degré du monôme f. Un *polynôme* de deux variables est une somme de monômes, et sont degré est le plus haut degré des monômes qui le composent.

Remarque. Les règles de somme et multiplication entre polynômes de plusieurs variables généralises naturellement les opérations pour une variable. On peut également factoriser un polynôme de plusieurs variables, mais les méthodes sont moins systématiques.

Dans le cas des fonctions de deux variables, le domaine de définition est une partie du plan $\mathbb{R}^2$, et on l'écrira sous la forme $\{(x, y) \in \mathbb{R}^2 : \ldots\}$. Par exemple, la fonction $f(x, y) = \ln(x - y)$, qui est obtenue comme la composition $\ln \circ h$, avec $h(x, y) = x - y$, a pour domaine de définition

$$D_f = \{(x, y) \in \mathbb{R}^2 : x - y \in D_{\ln}\} = \{(x, y) \in \mathbb{R}^2 : x - y > 0\}.$$

Remarque. Pour composer des fonctions de plusieurs variables, il faut que le nombre de variables dans la composition soit cohérent ! Dans l'exemple précédent, la composition $h \circ \ln$ n'est pas possible, car h nécessite de deux variables réelles, alors que l'image de la fonction ln est un nombre réel.

2 Représentations graphiques

2.1 Représenter les fonctions

Définition (Graphe d'une fonction). Si f est une fonction de n variables réelles, on appelle *graphe* de f l'ensemble

$$\mathrm{Gr}(f) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : x \in D_f,\ f(x) = y\}.$$

Les ensembles

$$\{(x, y) \in \mathbb{R}^n \times \mathbb{R} : x \in D_f,\ f(x) \geq y\} \quad \text{et} \quad \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : x \in D_f,\ f(x) \leq y\}$$

sont appelés respectivement *épi-graphe* et *sous-graphe* de la fonction f.

Le graphe d'une fonction d'une variable réelle est une partie du plan $\mathbb{R}^2$, c'est une *courbe*. Le graphe d'une de deux variables est une partie de l'espace $\mathbb{R}^3$, c'est une *surface*. Puisqu'il est parfois difficile de représenter graphiquement une surface dans l'espace, une manière alternative de visualiser le graphe d'une fonction de deux variables est d'utiliser les *lignes de niveau*, comme l'on fait par exemple en cartographie pour représenter le relief d'une région, ou une *échelle de couleurs*, comme par exemple en météorologie pour représenter la température sur une carte.

Définition (Ensemble de niveau d'une fonction). Si f est une fonction de n variables, on appelle *ensemble de niveau* de f un ensemble du type

$$L_\lambda := \{x \in \mathbb{R}^n : x \in D_f,\ f(x) = \lambda\}, \quad \text{où } \lambda \in \mathbb{R}.$$

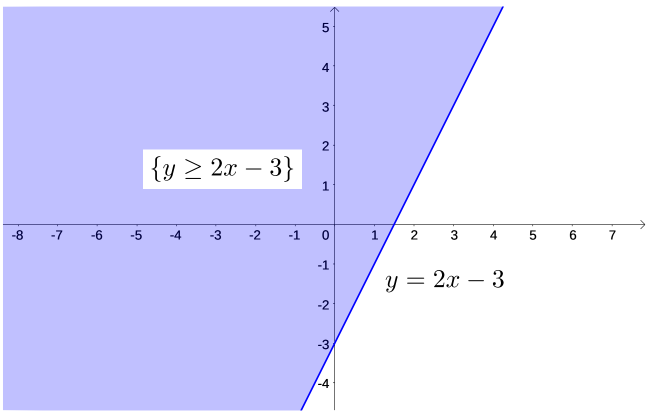
Pour les fonctions de deux variables, on parle également de courbe (ou ligne) de niveau. Il est important de souligner que le mot courbe doit être pris avec réserve car pour certaines fonctions l'ensemble L_λ ne ressemble pas nécessairement à une courbe de par la présence de parties plus ou moins « épaisses ».

2.2 Graphes des fonctions élémentaires

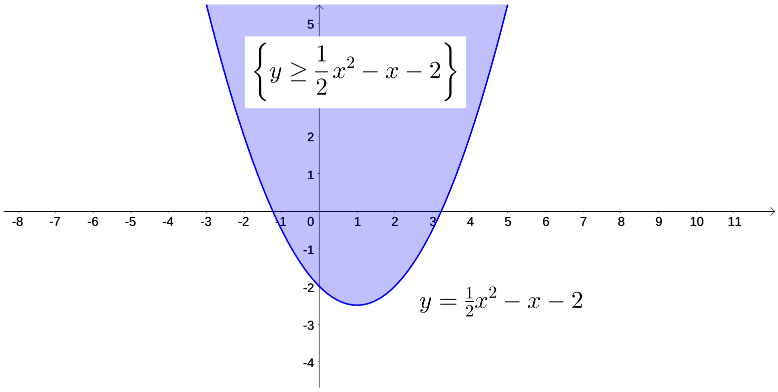
Commençons par les polynômes. Le graphe d'un polynôme constant $f(x) = c$ est représente dans le plan cartésien par la droite horizontale d'équation $y = c$. Sont épi-graphe est le demi-plan supérieur $\{y \geq c\}$, et sont sous-graphe est le demi-plan inférieur $\{y \leq c\}$.

Remarque. La droite verticale $x = c$ n'est pas le graphe d'une fonction en la variable x, mais on peut la voir comme le graphe de la fonction constante en la variable y, $f(y) = c$.

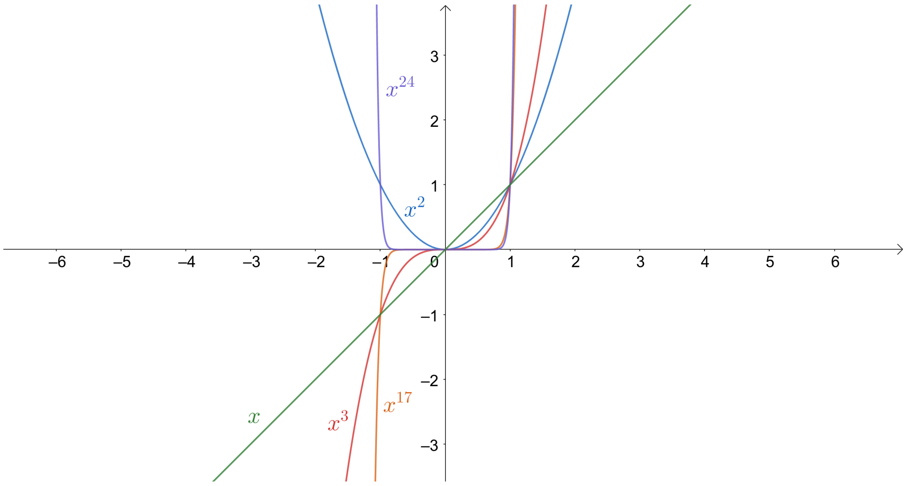
En général, une droite d'équation cartésienne $y = ax + b$ est le graphe du polynôme $f(x) = ax + b$. Son épi-graphe est le demi-plan $\{y \geq ax + b\}$. Rappelons que pour tracer une droite il suffit de déterminer deux points appartenant à la droite, et dessiner la seule droite qui passe par ces deux points.

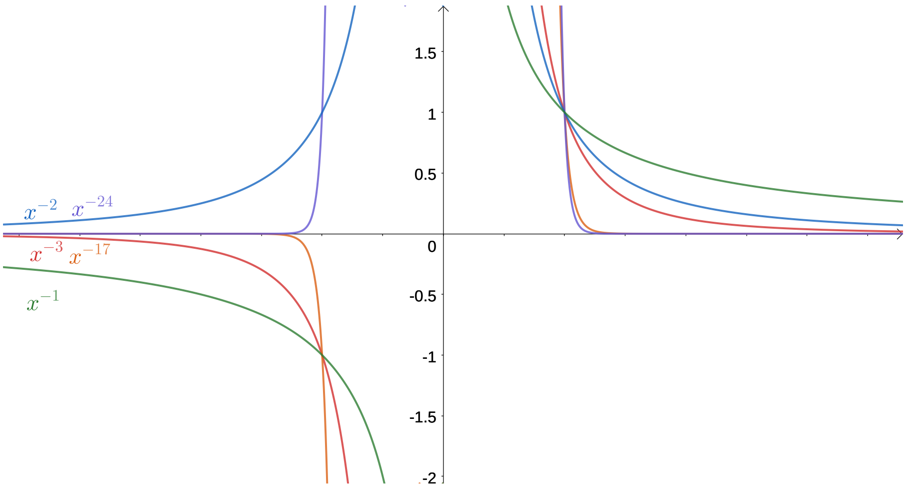


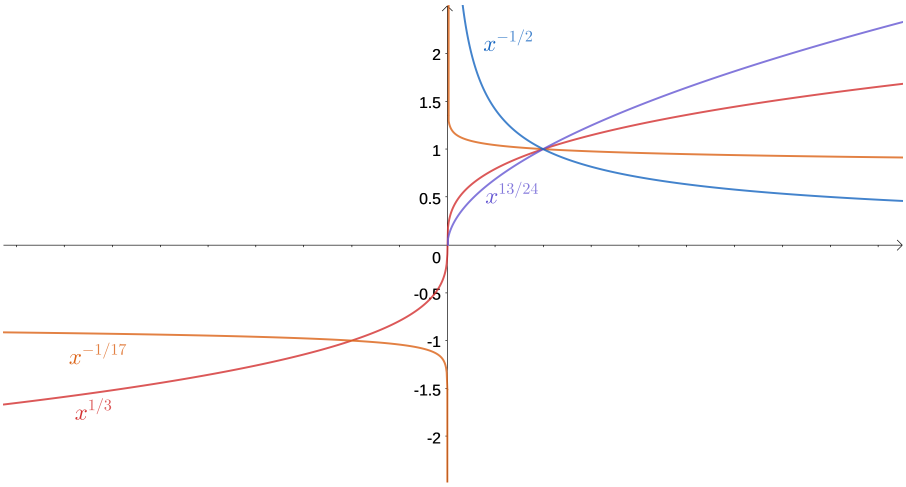
Passons maintenant aux graphes de trinômes du second degré $f(x) = ax^2 + bx + c$. Le graphe s'appele une *parabole*. Elle est dirigée vers le haut si $a > 0$, et vers le bas si $a < 0$. Le point extrémal de la parabole a pour abscisse $\overline{x} = -\frac{b}{2a}$. La parabole rencontre l'axe des abscisses aux deux racines du trinôme (lorsque son discriminant Δ est positif).

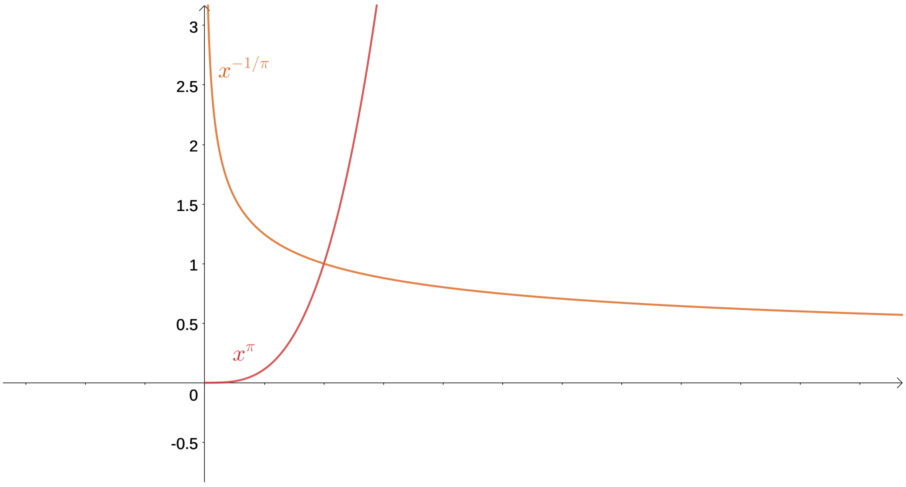


Il est important aussi de retenir l'allure des graphes des fonctions puissance $f(x) = x^\alpha$, pour des différentes valeurs de α.

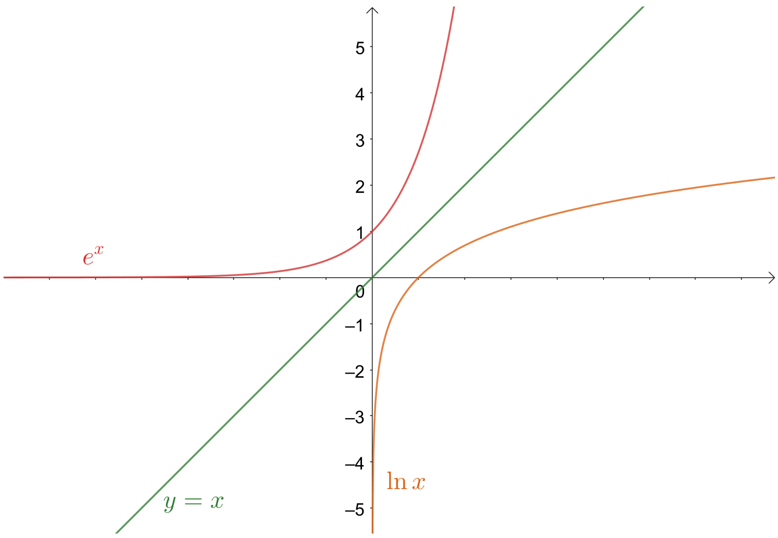




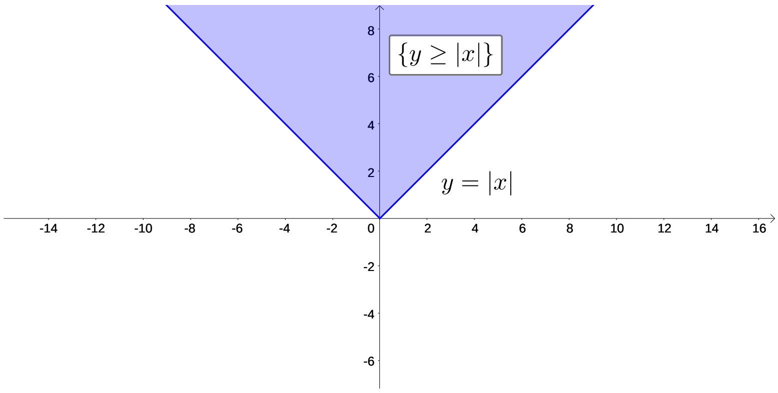




On peut constater dans les images que les graphes des fonctions x^{α} et $x^{1/\alpha}$ sont symétriques par rapport à la droite diagonale $y = x$. Ceci est parce que les deux fonctions sont réciproques. On observe aussi cette symétrie dans les graphes des fonctions e^x et $\ln(x)$, qui sont réciproques par définition.

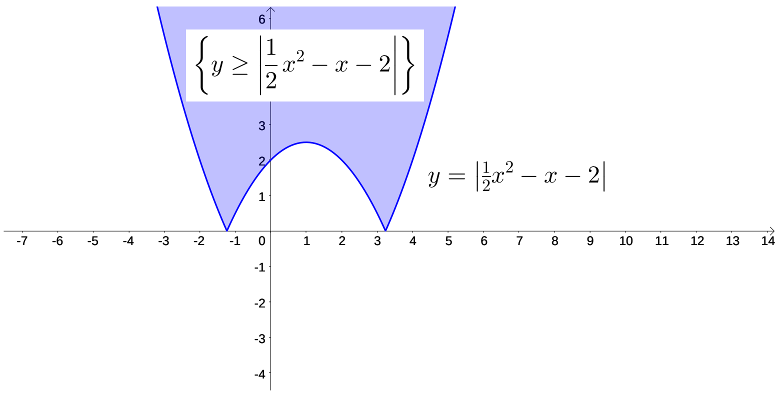


Le graphe de la fonction « valeur absolue » est une droite brisée, symétrique par rapport à l'origine.

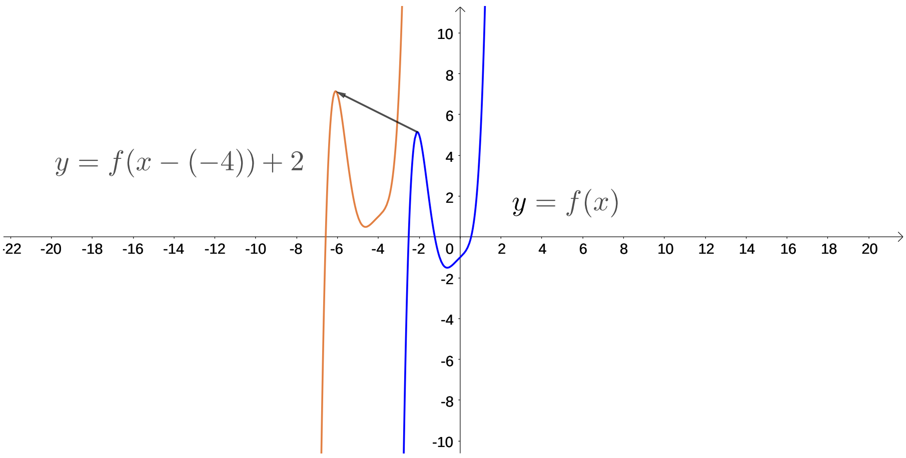


2.3 Opérations sur les graphes

En général, si f est une fonction, le graphe de la fonction $g(x) = |f(x)|$ s'obtient par réflexion de la partie du graphe dans le demi-plan inférieur $\{y \leq 0\}$ dans le demi-plan supérieur $\{y \geq 0\}$.



Étant donné une fonction f, et deux constantes $a, b \in \mathbb{R}$, le graphe de la fonction $g(x) = f(x - a) + b$ et obtenu en déplaçant le graphe de g par le *vecteur* (a, b) (attention au signe de a !). À savoir, on déplace le graphe de f de a vers la droite (donc à droite seulement si a est positif), et de b vers le haut.



2.4 Opérations sur les sous-ensembles

Rappelons ici les opérations que l'on peut effectuer sur des sous-ensemble d'un ensemble donné. On se fixe ici un ensemble Ω (dans la pratique, ce sera l'espace $\mathbb{R}^n$), et on considère deux sous-ensembles $A, B \subset \Omega$. L'*union* $A \cup B = \{x \in \Omega : x \in A \text{ ou } x \in B\}$ consiste à prendre le sous-ensemble des points qui sont dans A *ou* B. L'*intersection* $A \cap B = \{x \in \Omega : x \in A, x \in B\}$ consiste à prendre le sous-ensemble des points qui sont dans A *et* B. On note $A^c = \{x \in \Omega : x \notin A\}$ le *complémentaire* de A, c'est-à-dire l'ensemble des points qui ne sont pas dans A. On peut aussi prendre la *différence* $A \setminus B = \{x \in \Omega : x \in A, x \notin B\} = A \cap B^c$ entre sous-ensembles.

Parfois ces opérations sont cachées dans les conditions numériques que les points d'un ensemble doivent satisfaire. Par exemple, écrire $\{x :\in \mathbb{R} : 0 \leq x \leq 1\}$ signifie considérer l'ensemble $\{x :\in \mathbb{R} : x \geq 0, x \leq 1\} = \{x \geq 0\} \cap \{x \leq 1\}$. Un deuxième exemple est donné par $\{(x, y) \in \mathbb{R}^2 : xy = 0\}$, où la condition $xy = 0$ se traduit en la condition « $x = 0$ ou $y = 0$ », et donc

$$\{(x, y) \in \mathbb{R}^2 : xy = 0\} = \{(x, y) \in \mathbb{R}^2 : x = 0\} \cup \{(x, y) \in \mathbb{R}^2 : y = 0\}.$$

3 Dérivées et dérivées partielles

3.1 Dérivabilité des fonctions d'une variable

La notion de dérivée est capitale car elle permet d'étudier les variations d'une fonction. Elle est aussi centrale dans l'étude des problèmes d'optimisation et des problèmes d'équilibre. Pour les fonctions d'une variable, la dérivée en un point représente la pente de la droite tangente au graphe de la fonction en ce point, c'est-à-dire, de la droite qui approxime le mieux la fonction au voisinage de ce point. Pour être plus précis, si f est une fonction définie dans un intervalle contenant un point $x \in \mathbb{R}$, on dit que la *dérivée* $f'(x)$ au point x existe, si l'on peut écrire

$$f(x+h) = f(x) + f'(x) \times h + \text{termes d'ordre supérieur en } h,$$

pour tout $h > 0$ suffisamment petit. Cela se formalise encore mieux en utilisant la notion de limite (qui dépasse les buts de ce cours).

Définition (Dérivée d'une fonction d'une variable en un point)**.** On dit qu'une fonction f d'une variable est dérivable en un point x si elle est définie sur un intervalle ouvert contenant x (on dit *au voisinage* de x) et si le quotient différentiel, défini uniquement pour $h \neq 0$, par

$$\tau(h) := \frac{f(x+h) - f(x)}{h}$$

admet une limite lorsque h tend vers zéro. Si la limite existe, on la note $f'(x)$ et on dit que c'est la *dérivée* de f en x.

Définition (Fonction dérivée)**.** Si une fonction f définie sur un intervalle ouvert I est dérivable en tout $x \in I$, on dit que f est *dérivable* sur I. La *fonction dérivée* f' est la fonction qui à tout $x \in I$ associe $f'(x)$.

Si la fonction f' est dérivable en x, on dit que f est *deux fois dérivable* en x. La dérivée de f' en $\overline{x}$ s'appelle la *dérivée seconde* de f en $\overline{x}$ et se note $f''(\overline{x})$ ou parfois $f^{(2)}(\overline{x})$. Si f' est dérivable en tout $x \in I$, on peut alors définir la fonction f''. Il s'agit de la fonction qui à tout $x \in I$ associe $f''(x)$. On définit ainsi les dérivées successives de f.

Remarque. Les dérivées n'existes pas toujours. Par exemple, la fonction $f(x) = |x|$ est dérivable pour tout $x \neq 0$, mais elle n'est pas dérivable en $x = 0$. Il existe aussi des fonctions qui ne sont dérivable en presque aucun point, comme par exemple la plupart des fonctions en finance, qui représentent les évolutions des prix sur le temps dans un marché financier (par exemple, le taux de change EUR/USD).

3.2 Dérivées des fonctions usuelles

Les dérivées de fonctions classiques peuvent être obtenue en calculant les limites des quotients différentiels.

Proposition (Dérivées des fonctions usuelles)**.**

f	f'
$x^\alpha \quad (\alpha \in \mathbb{R})$	$\alpha x^{\alpha-1}$
$\ln x$	$\frac{1}{x}$
e^x	e^x

On écrit pour permettre une meilleure mémorisation, les dérivées des fonctions x^α pour des valeurs de α typiques :

$$(1)' = 0, \quad (x)' = 1, \quad \left(x^2\right)' = 2x, \quad \left(\frac{1}{x}\right)' = -\frac{1}{x^2}, \quad \left(\sqrt{x}\right)' = \frac{1}{2\sqrt{x}}.$$

On remarquera que même si la fonction $f(x) = \sqrt{x}$ est définie en $x = 0$, sa dérivée $f'(x) = \frac{1}{2\sqrt{x}}$ ne l'est pas.

3.3 Propriétés des fonctions dérivables

D'après la définition de dérivée, on peut aussi déterminer la dérivée des fonctions obtenues en faisant des opérations sur des fonctions plus élémentaires.

Proposition (Propriétés des fonctions dérivables). *Si f, g sont des fonctions dérivables en un point x, alors :*

1. *la somme $f+g$ est dérivable en x et on a $(f+g)'(x) = f'(x) + g'(x)$,*
2. *le produit fg est dérivable en x et on a $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$,*
3. *si $g(x) \neq 0$, f/g est dérivable en x et $(f/g)'(x) = \dfrac{f'(x)g(x) - g'(x)f(x)}{g^2(x)}$.*

Un cas particulier, mais bien important, de la deuxième propriété est lorsque $g(x) = \lambda$ est une fonction constante. On a alors la propriété

$$(\lambda f)'(x) = \lambda f'(x).$$

Proposition (Règle de dérivation en chaîne). *Soient u une fonction dérivable en un point x et g une fonction dérivable au point $u(x)$. Alors $g \circ u$ est dérivable au point x et on a*

$$(g \circ u)'(x) = g'(u(x))u'(x).$$

Dans les exemples, on doit employer plusieurs fois les règles précédentes pour déterminer une dérivée.

Exemple. Considérons la fonction f définie par

$$f(x) := e^{\sqrt{x^3+1}}.$$

La règle de dérivation en chaîne donne

$$f'(x) = e^{\sqrt{x^3+1}} \times \frac{1}{2\sqrt{x^3+1}} \times (3x^2).$$

3.4 Dérivées partielles

Le graphe d'une fonction de deux variables est une surfaces, et au lieu de parler de droites tangentes pour définir une dérivée, il est plus approprié d'étudier les *plans* tangents. Ceci est plus difficile à exprimer par des fonctions à valeurs réelles.

Définition (Dérivée partielle). Soit f une fonction de deux variables définie au voisinage d'un point $p = (x, y) \in \mathbb{R}^2$. On appelle *dérivée partielle* de f par rapport à la première variable au point p la limite (si elle existe)

$$\frac{\partial f}{\partial x}(x, y) := \lim_{h \to 0} \frac{f(x+h, y) - f(x, y)}{h},$$

qui est la dérivée de la fonction d'une variable $x \mapsto f(x, y)$, la variable y étant fixée. On note aussi parfois cette dérivée partielle $\partial_x f(x, y)$.

De façon similaire on appelle *dérivée partielle* de f par rapport à la seconde variable au point $p = (x, y)$ la limite (si elle existe)

$$\frac{\partial f}{\partial y}(x, y) := \lim_{h \to 0} \frac{f(x, y+h) - f(x, y)}{h},$$

qui est la dérivée de la fonction d'une variable $y \mapsto f(x, y)$, la variable x étant fixée. On note aussi parfois cette dérivée partielle $\partial_y f(x, y)$.

Il est d'usage de regrouper les dérivées partielles en un vecteur que l'on appelle le gradient de la fonction. Le vecteur gradient indique la direction où la fonction varie le plus fortement, et il est perpendiculaire au plan tangent au graphe de la fonction.

Définition (Gradient d'une fonction de 2 variables). Soit f une fonction de deux variables admettant des dérivées partielles en un point $p = (x, y)$. On appelle *gradient* de f en p le vecteur

$$\nabla f(x, y) := \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \right).$$

3.5 Variations d'une fonction

Définition. Soit f une fonction d'une variable réelle définie sur un intervalle I.

1. On dit que f est *(monotone) croissante* sur I si pour tout $x, y \in I$ tels que $x < y$ on a $f(x) \leq f(y)$. La fonction f est *strictement* croissante sur I, si pour tout $x, y \in I$ tels que $x < y$ on a $f(x) < f(y)$.
2. On dit que f est *(monotone) décroissante* sur I si pour tout $x, y \in I$ tels que $x < y$ on a $f(x) \geq f(y)$. La fonction f est *strictement* décroissante sur I, si pour tout $x, y \in I$ tels que $x < y$ on a $f(x) > f(y)$.

En général, on dit que la fonction f est monotone sur I si elle est croissante ou décroissante.

Remarque. Les fonctions constantes sont les seules fonctions à être à la fois croissantes et décroissantes.

Exemple. La fonction $f(x) = x^2$ est strictement décroissante sur $]-\infty, 0]$ et strictement croissante sur $[0, \infty[$. Pareil pour la fonction $g(x) = |x|$.

Décrire les *variations* d'une fonction f consiste à déterminer les intervalles de monotonie de f : on veut déterminer les intervalles pour lesquels la restriction de f est monotone, croissante ou décroissante, comme dans l'exemple précédent. Ces intervalles sont alors délimités par des point où la fonction change son caractère de monotonie. Par exemple, on peut avoir un point $x \in I$ tel que f est croissante à la gauche de x, et décroissante à la droite de x. On dira alors que x est un point de *maximum* pour I. Plus généralement, on a la définition suivante.

Définition. Soit f une fonction d'une variable réelle définie sur un intervalle I.

1. Un point $x \in I$ est un *point de maximum local* s'il existe $\varepsilon > 0$ tel que pour tout $y \in]x - \varepsilon, x + \varepsilon[\cap I$, on a $f(y) \leq f(x)$. Lorsqu'on a $f(y) < f(x)$ pour tout $y \in]x - \varepsilon, x + \varepsilon[\cap I$, $y \neq x$, on dit que x est un point de maximum local *strict*. La valeur $f(x)$ est un *maximum local* de la fonction f.
2. Un point $x \in I$ est un *point de minimum local* s'il existe $\varepsilon > 0$ tel que pour tout $y \in]x - \varepsilon, x + \varepsilon[\cap I$, on a $f(y) \geq f(x)$. Lorsqu'on a $f(y) > f(x)$ pour tout $y \in]x - \varepsilon, x + \varepsilon[\cap I$, $y \neq x$, on dit que x est un point de minimum local *strict*. La valeur $f(x)$ est un *minimum local* de la fonction f.

Si $x \in I$ est un point de maximum ou minimum local de f, on dit que x est un *extremum* de f.

Pour les fonctions dérivables, on peut déterminer les variations à l'aide du *signe* de la dérivée première. Dans la suite lorsqu'on parlera d'intervalle ouvert I, on désignera soit un intervalle ouvert dont les deux bornes a et b sont des nombres réels, soit un intervalle ouvert dont l'une des extrémités est infinie du type $]a, +\infty[$, $]-\infty, b[$ ou $]-\infty, +\infty[$.

Proposition (Caractérisation des fonctions croissantes/décroissantes par le signe de la dérivée). *Soit $I \subset \mathbb{R}$ un intervalle* ouvert *et $f : I \to \mathbb{R}$ une fonction dérivable sur I. Alors f est croissante si et seulement si $f'(x) \geq 0$ pour tout $x \in I$ et f est décroissante si et seulement si $f'(x) \leq 0$ pour tout $x \in I$.*

Il faut prendre garde que la stricte positivité (négativité) de la dérivée d'une fonction sur un intervalle ne caractérise pas sa stricte croissance (décroissance), comme le montre la fonction $x \mapsto x^3$. La stricte positivité (négativité) de la dérivée d'une fonction assure sa stricte croissance (décroissance), mais la réciproque est fausse.

Définition. Soit $I \subset \mathbb{R}$ un intervalle *ouvert* et $f : I \to \mathbb{R}$ une fonction dérivable sur I. Un point $x \in I$ est un *point critique* pour f si $f'(x) = 0$.

Proposition (Condition nécessaire d'optimalité sans contrainte pour les fonctions d'une variable). *Soit $I \subset \mathbb{R}$ un intervalle* ouvert *et $f : I \to \mathbb{R}$ une fonction dérivable sur I. Si un point $x \in I$ est un extremum de f sur I, alors $f'(x) = 0$.*

La condition $f'(x) = 0$ n'est pas en général une condition *suffisante* d'optimalité. Pour s'en convaincre il suffit de prendre la fonction $x \mapsto x^3$, et le point $x = 0$. Il faut aussi prendre garde que si l'intervalle n'est pas *ouvert* la minimalité (maximalité) de la fonction n'implique plus nécessairement la nullité de la dérivée (penser à la fonction $f(x) = x$ sur $[0, 1]$).

Chercher les extrema d'une fonction revient donc à chercher les points critiques via la résolution de l'équation $f'(x) = 0$ et ensuite à trier parmi les points critiques ceux qui sont minimum, ceux qui sont maximum et ceux qui ne sont pas des extrema, et que l'on appelle *points d'inflexion*.

Cette tâche est relativement simple pour les fonctions deux fois dérivables, car le signe de la dérivée seconde en un point critique peut déterminer sa nature.

Proposition. *Soit $I \subset \mathbb{R}$ un intervalle ouvert et $f : I \to \mathbb{R}$ une fonction deux fois dérivable. On suppose que $x \in I$ est un point critique de f.*

1. *Si $f''(x) > 0$, alors x est un point de minimum local pour f.*
2. *Si $f''(x) < 0$, alors x est un point de maximum local pour f.*

Comme précédemment, la condition $f''(x) = 0$ ne permet pas de conclure. Par exemple, $x = 0$ est un point critique pour les fonctions $f(x) = x^4$, et un minimum local pour cette fonction. Mais on a aussi que $x = 0$ est un point de maximum local pour $f(x) = -x^4$, et un point d'inflexion pour $f(x) = x^3$. Dans les trois cas, en $x = 0$ on a $f''(0) = 0$.

3.6 Étude d'une fonction d'une variable

On récapitule ici les étapes pour étudier une fonction d'une variable, et plus précisément, pour comprendre l'allure de son graphe. Cette liste s'enrichira au cours du second semestre, une fois que vous étudierez les limites d'une fonction.

On considère ici une fonction $f : I \to \mathbb{R}$ d'une variable réelle.

1. On détermine le domaine de définition de la fonction.
2. Si la fonction est dérivable, on calcule $f'(x)$.
3. Si la fonction est deux fois dérivable, on calcule $f''(x)$.
4. On détermine les points critiques de f. On pourra regarder le signe de f'' aux points critiques pour déterminer leur nature.
5. On étudie le signe de $f'(x)$. On ordonne les points critiques de f sur la droite réelle, et on indique le signe de f' dans les intervalles entre deux points critiques successifs. Cela permet de déterminer le caractère monotone de f sur ces intervalles, et donc d'étudier ses variations.
   En faisant le tableau du signe de f', on arrive normalement à déterminer la nature des points critiques de f (sans regarder le signe de f'').
6. On trace un graphe approximatif pour la fonction f.

3.7 Points critiques pour les fonctions de plusieurs variables

Pour les fonctions de deux ou plusieurs variables, on peut également définir les extrema locaux. Il suffit alors de remplacer les intervalles par des « boules » définies par la distance. Pour les fonctions dérivables, les extrema sont aussi détectés par l'annulation des dérivées.

Définition (Point critique d'une fonction de 2 variables)**.** Soit f une fonction de 2 variables définie au voisinage d'un point $p = (x, y)$ et admettant en p des dérivées partielles. Si $\nabla f(x, y) = 0$, on dit du point p que c'est un point critique.

Proposition (Condition nécessaire d'optimalité d'une fonction de 2 variables)**.** *Soit f une fonction admettant des dérivées partielles en un point $p = (x, y)$. Si p est minimum ou maximum local, alors son gradient est nul, i.e. $\nabla f(x, y) = 0$.*

La condition $\nabla f(x, y) = 0$ n'est pas une condition *suffisante* d'optimalité. Pour s'en convaincre il suffit de prendre la fonction $f(x, y) = xy$. On a clairement

$$\nabla f(x, y) = (y, x),$$

si bien que l'origine $p = (0, 0)$ est le seul point critique. Il est clair que ce point n'est ni un minimum ni un maximum local puisque la fonction est nulle en ce point et qu'elle prend des valeurs strictement positives et des valeurs strictement négatives autour de p.

4 Géométrie de l'espace $\mathbb{R}^n$

4.1 Points et vecteurs

La plupart des fonctions utilisées en économie sont des fonctions de plusieurs variables. Travailler avec deux variables réelles x_1 et x_2, avec trois variables réelles x_1, x_2 et x_3, et de façon plus générale avec n variables réelles $x_1, \ldots, x_n$ revient à travailler avec des couples (x_1, x_2), des triplets (x_1, x_2, x_3) ou plus généralement avec des n-uplets $(x_1, \ldots, x_n)$. On désigne par $\mathbb{R}^2$ l'ensemble des couples de réels

$$\mathbb{R}^2 := \{(x_1, x_2) : x_1 \in \mathbb{R},\ x_2 \in \mathbb{R}\},$$

et par $\mathbb{R}^3$ l'ensemble des triplets de réels

$$\mathbb{R}^3 := \{(x_1, x_2, x_3) : x_1 \in \mathbb{R},\ x_2 \in \mathbb{R},\ x_3 \in \mathbb{R}\}.$$

De façon plus générale, on note $\mathbb{R}^n$ l'ensemble

$$\mathbb{R}^n := \{(x_1, \ldots, x_n) : x_i \in \mathbb{R},\ i = 1, \ldots, n\}.$$

Le nombre n est la *dimension* de l'espace. Les éléments de $\mathbb{R}^n$ sont des n-uplets de réels, distincts ou non, et placés dans un certain ordre. Il faut bien voir que l'ordre des x_i est *important*. Par exemple, les couples $(1, 2)$ et $(2, 1)$ désignent des éléments différents de $\mathbb{R}^2$. Deux n-uplets $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ sont égaux si, et seulement si, $x_i = y_i$ pour tout $i = 1, \ldots, n$.

Remarque. On utilise aussi la notation (x, y) pour les points de $\mathbb{R}^2$ et (x, y, z) pour les points de $\mathbb{R}^3$.

Pour un élément $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, les x_i sont appelées les *composantes* de x. Suivant le contexte, on dit d'un élément $x \in \mathbb{R}^n$, que c'est un *point* ou un *vecteur*. On utilisera les lettres majuscules pour désigner des points $(A, B, C, \ldots)$, et des lettres minuscules avec une flèche pour désigner des vecteurs $(\vec{x}, \vec{y}, \ldots)$. Quand on parle de points, cela permet alors d'utiliser le langage classique de la géométrie et de parler de droite, de plan, de triangle, de rectangle, etc. Les vecteurs $\vec{x} = (x_1, \ldots, x_n)$ sont représentés comme une flèche issue de l'origine et dont l'extrémité est le point $(x_1, \ldots, x_n)$.

On peut additionner les vecteurs de $\mathbb{R}^n$ et les multiplier par un *scalaire* (c'est-à-dire un nombre réel). En effet, si $\vec{x}, \vec{y} \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$, la *somme* de $\vec{x}$ et $\vec{y}$ est définie par

$$\vec{x} + \vec{y} := (x_1 + y_1, \ldots, x_n + y_n)$$

et le *produit* de $\vec{x}$ par le scalaire λ est défini par

$$\lambda\vec{x} := (\lambda x_1, \ldots, \lambda x_n).$$

L'élément $\vec{0} = (0, \ldots, 0) \in \mathbb{R}^n$, dont toutes les composantes sont nulles, est appelé le *vecteur nul* ou encore l'*origine* de $\mathbb{R}^n$.

Définition (Colinéarité entre deux vecteurs de $\mathbb{R}^n$)**.** On dit que deux vecteurs non-nuls $\vec{x}$ et $\vec{y}$ de $\mathbb{R}^n$ sont *colinéaires* s'il existe un scalaire $t \neq 0$ tel que $\vec{x} = t\vec{y}$.

Au contraire des vecteurs, quand on considère les éléments de $\mathbb{R}^n$ comme des points, on n'a pas le droit de les additionner ou multiplier par un scalaire. En revanche, si $A = (a_1, \ldots, a_n) \in \mathbb{R}^n$ est un point et $\vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ est un vecteur, on peut *déplacer* le point A par le vecteur $\vec{x}$:

$$A + \vec{x} = (a_1 + x_1, \ldots, a_n + x_N).$$

Si $A = (a_1, \ldots, a_n), B = (b_1, \ldots, b_n) \in \mathbb{R}^n$, on note $\overrightarrow{AB}$ le vecteur qui déplace A vers B :

$$A + \overrightarrow{AB} = B.$$

On a alors $\overrightarrow{AB} = (b_1 - a_1, \ldots, b_n - a_n)$.

4.2 Mesurer dans $\mathbb{R}^n$

On peut munir $\mathbb{R}^n$ d'une structure d'espace euclidien en définissant un *produit scalaire* entre les vecteurs.

Définition (Produit scalaire). On appelle *produit scalaire* de deux vecteurs $\vec{x} \in \mathbb{R}^n$ et $\vec{y} \in \mathbb{R}^n$ la quantité $\langle \vec{x}, \vec{y} \rangle := \sum_{i=1}^n x_i y_i$.

Proposition (Propriétés du produit scalaire de deux vecteurs). *Si $\vec{x}$, $\vec{y}$ et $\vec{z}$ sont des vecteurs de $\mathbb{R}^n$ et λ est un scalaire, on a*

— $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$ *symétrie*
— $\langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$, *et* $\langle \lambda\vec{x}, \vec{y} \rangle = \lambda \langle \vec{x}, \vec{y} \rangle$ *linéarité par rapport au premier membre*
— $\langle \vec{x}, \vec{y} + \vec{z} \rangle = \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}, \vec{z} \rangle$, *et* $\langle \vec{x}, \lambda\vec{y} \rangle = \lambda \langle \vec{x}, \vec{y} \rangle$ *linéarité par rapport au second membre.*

La notion de produit scalaire permet de mesurer l'angle entre deux vecteurs. Sans rentrer dans le spécifique, on s'intéressera uniquement à la notion d'orthogonalité.

Définition (Vecteurs orthogonaux). On dit de deux vecteurs $\vec{x}$ et $\vec{y}$ non nuls qu'ils sont *orthogonaux* si $\langle \vec{x}, \vec{y} \rangle = 0$.

Le produit scalaire permet aussi de définir la notion de *norme* (ou *longueur*, ou *taille*) d'un vecteur.

Définition (Norme euclidienne d'un vecteur). On appelle *norme* (euclidienne) d'un vecteur $\vec{x} \in \mathbb{R}^n$ la quantité $||\vec{x}|| := \sqrt{\sum_{i=1}^n x_i^2}$.

On observera que par définition on a $||\vec{x}|| = \sqrt{\langle \vec{x}, \vec{x} \rangle}$. On peut donc exprimer la norme d'un vecteur à l'aide du produit scalaire. On peut montrer que la norme euclidienne possède les propriétés suivantes.

Proposition (Propriétés de la norme euclidienne d'un vecteur). *La norme euclidienne vérifie les propriétés suivantes.*

— $||\vec{x}|| \geq 0$ (non négativité)
— $||\vec{x}|| = 0 \Leftrightarrow x = 0$ (non dégénérescence)
— $||\lambda\vec{x}|| = |\lambda| \, ||\vec{x}||$ *pour* $\lambda \in \mathbb{R}$ (homogénéité de degré 1)
— $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$ (inégalité triangulaire)

Les identités remarquables sur les nombres réels se généralisent dans $\mathbb{R}^n$ sous la forme suivante.

Proposition (Identités remarquables vectorielles). *Si $\vec{x}$ et $\vec{y}$ sont des vecteurs de $\mathbb{R}^n$ on a les identités remarquables*

— $||\vec{x} + \vec{y}||^2 = ||\vec{x}||^2 + ||\vec{y}||^2 + 2\langle \vec{x}, \vec{y} \rangle$,
— $||\vec{x} - \vec{y}||^2 = ||\vec{x}||^2 + ||\vec{y}||^2 - 2\langle \vec{x}, \vec{y} \rangle$,
— $||\vec{x}||^2 - ||\vec{y}||^2 = \langle \vec{x} + \vec{y}, \vec{x} - \vec{y} \rangle$.

Dans $\mathbb{R}^2$, une application du théorème de Pythagore montre que la distance usuelle entre deux points $P = (x_1, x_2)$ et $Q = (y_1, y_2)$ est mesurée par la quantité

$$d(P, Q) := \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} = ||\overrightarrow{PQ}||.$$

On définit sur $\mathbb{R}^n$ la notion de distance euclidienne comme suit.

Définition (Distance euclidienne entre deux points de $\mathbb{R}^n$). On appelle *distance euclidienne* entre deux points A, $B \in \mathbb{R}^n$ la quantité définie par

$$d(A, B) := ||\overrightarrow{AB}|| = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}.$$

Proposition. *La distance euclidienne vérifie les propriétés suivantes.*

— $d(A, B) \geq 0$ (non négativité)
— $d(A, B) = 0 \Leftrightarrow A = B$ (non dégénérescence)
— $d(A, B) \leq d(A, C) + d(C, B)$ (inégalité triangulaire)

4.3 Objets géométriques dans $\mathbb{R}^n$

Définition (Droite de $\mathbb{R}^n$). Soit $P \in \mathbb{R}^n$ un point et $\vec{d} \in \mathbb{R}^n$ un vecteur non-nul. La *droite (affine)* passant par le point P et dirigée par le vecteur $\vec{d}$ est l'ensemble

$$D := \{P + t\,\vec{d} : t \in \mathbb{R}\}.$$

On appelle $\vec{d}$ le vecteur *directeur* de D.

La droite passant par deux points distincts $P, Q \in \mathbb{R}^n$ est dirigée par le vecteur $\overrightarrow{PQ}$. Par conséquent, elle est donnée par

$$D := \{P + t\,\overrightarrow{PQ} : t \in \mathbb{R}\}.$$

Définition (Demi-droite de $\mathbb{R}^n$). Soit $P \in \mathbb{R}^n$ un point et $\vec{d} \in \mathbb{R}^n$ un vecteur non-nul. La *demi-droite (positive)* issue du point P et dirigée par le vecteur $\vec{d}$ est l'ensemble

$$D_+ := \{P + t\,\vec{d} : t \geq 0\}.$$

Pour A, $b \in \mathbb{R}$, le *segment* $[a, b]$ est l'ensemble constitué de tous les réels situés entre a et b :

$$[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\} \quad \text{(intervalle fermé)}.$$

Cette notion se généralise facilement dans $\mathbb{R}^n$. Si A et B sont des points distincts de $\mathbb{R}^n$, on peut considérer la droite qui passe par ces deux points et définir le segment $[A, B]$ comme l'ensemble des points de cette droite situés entre A et B.

Définition (Segment de $\mathbb{R}^n$). Soit $A, B \in \mathbb{R}^n$. On appelle *segment* d'extrémités A et B l'ensemble

$$[A, B] := \{A + t\,\overrightarrow{AB} : t \in [0, 1]\}.$$

Remarque. Si $A = (a_1, \ldots, a_n)$ et $B = (b_1, \ldots, b_n)$, on peut écrire

$$[A, B] = \{(1 - t)(a_1, \ldots, a_n) + t(b_1, \ldots, b_n) : t \in [0, 1]\}.$$

On peut écrire $(1 - t)(a_1, \ldots, a_n) + t(b_1, \ldots, b_n) = (1 - t)A + tB$, et on appelle une telle expression une *combinaison convexe* des points A et B.

La notion de segment nous permet de définir aussi les ensembles convexes, comme les ensembles vérifiant que pour tout deux points, le segment entre ces deux points reste dans l'ensemble.

Définition (Ensemble convexe). On dit qu'un ensemble $\Omega \subset \mathbb{R}^n$ est *convexe* si pour tout $A, B \in \Omega$ on a $[A, B] \subset \Omega$.

Proposition. *Soient Ω et Ω' deux convexes de $\mathbb{R}^n$. Alors $\Omega \cap \Omega'$ est convexe.*

On sera aussi intéressés à la notion de cône, ce qui correspond essentiellement à un ensemble qui est constitué de demi-droites positives issues de l'origine.

Définition (Cône). On dit qu'un ensemble $\Omega \subset \mathbb{R}^n$ est un *cône* (de sommet l'origine) si

$$\forall \vec{x} \in \Omega, \quad \forall t > 0 \quad \text{on a} \quad t\vec{x} \in \Omega.$$

Remarque. De manière équivalente, $\Omega \subset \mathbb{R}^n$ est un cône si pour tout $\vec{x} \in \Omega$, la demi-droite strictement positive $D_+^* = \{t\vec{x} : t > 0\}$ est dans Ω.

Proposition. *Soient Ω et Ω' deux cônes de $\mathbb{R}^n$.*

1. *L'intersection $\Omega \cap \Omega'$ est aussi en cône.*
2. *L'union $\Omega \cup \Omega'$ est aussi un cône.*
3. *Le complémentaire $\Omega \setminus \Omega'$ est aussi un cône.*

Exemple. Dans $\mathbb{R}$, les ensembles convexes sont exactement les intervalles et le (demi-)droites : $[a, b]$, $[a, b[$, $]a, b]$, $]a, b[$, avec $a \leq b$ dans $\mathbb{R} \cup \{-\infty, +\infty\}$.

Dans $\mathbb{R}$, les cônes sont exactement ces ensembles : $\mathbb{R}$, $\emptyset$, $\{0\}$, $\mathbb{R} \setminus \{0\}$, $[0, +\infty[$, $]0, \infty[$, $]-\infty, 0]$, $]-\infty, 0[$

Exemple. Soit $\Omega = \{(x, y) \in \mathbb{R}^2 : x - 2y \geq 0\}$. On va voir que Ω est à la fois un cône et un ensemble convexe.

Soit $(x, y) \in \Omega$ et $t > 0$. Alors pour le point (tx, ty) on a $tx - 2ty = t(x - 2y) \geq 0$ (car $t > 0$ et $x - 2y \geq 0$), donc $(tx, ty) \in \Omega$, ce qui montre que Ω est un cône.

On prend maintenant $A = (x_A, y_A)$ et $B = (x_B, y_B)$, et $t \in [0, 1]$. On doit vérifier que le point

$$(1 - t)A + tB = ((1 - t)x_A + tx_B, (1 - t)y_A + ty_B)$$

est dans Ω, ce qui est équivalent à la condition

$$(1 - t)x_A + tx_B - 2((1 - t)y_A + ty_B) \geq 0.$$

Pour vérifier cela, on écrit

$$(1 - t)x_A + tx_B - 2((1 - t)y_A + ty_B) = (1 - t)(x_A - 2y_A) + t(x_B - 2y_B).$$

Puisque $0 \leq t \leq 1$, on a $(1 - t) \geq 0$ et $t \geq 0$. Puisque $A, B \in \Omega$, on a $x_A - 2y_A \geq 0$ et $x_B - 2y_B \geq 0$. Ces inégalités nous permettent de conclure

$$(1 - t)x_A + tx_B - 2((1 - t)y_A + ty_B) = (1 - t)(x_A - 2y_A) + t(x_B - 2y_B) \geq 0,$$

ce qui implique que $tA + (1 - t)B \in \Omega$, comme l'on voulait.

Exemple. Soit $\Omega = \{(x, y) \in \mathbb{R}^2 : x \geq 1\}$. On va voir que Ω est un ensemble convexe, mais il n'est pas un cône.

Pour voir qu'il n'est pas un cône, on choisit $(x, y) = (1, 0) \in \Omega$ et $t = 1/2 > 0$. Alors le point $t(x, y) = (1/2, 0)$ a sa première coordonnée < 1, est donc n'est pas dans Ω. Cela montre que Ω n'est pas un cône.

Pour montrer que Ω est convexe, on prend maintenant $A = (x_A, y_A)$ et $B = (x_B, y_B)$, et $t \in [0, 1]$. On doit vérifier que le point

$$(1 - t)A + tB = ((1 - t)x_A + tx_B, (1 - t)y_A + ty_B)$$

est dans Ω, ce qui est équivalent à la condition

$$(1 - t)x_A + tx_B \geq 1.$$

Puisque $A, B \in \Omega$, on a $x_A \geq 1$ et $x_B \geq 1$. Ces inégalités nous permettent de conclure

$$(1 - t)x_A + tx_B \geq (1 - t) \times 1 + t \times 1 = 1 - t + t = 1,$$

ce qui implique que $tA + (1 - t)B \in \Omega$, comme l'on voulait.

5 Fonctions et géométrie

Dans ce chapitre nous considérons des fonctions avec des propriétés géométriques définies sur des parties de $\mathbb{R}^n$.

5.1 Fonctions linéaires et affines

Définition (Linéarité d'une fonction)**.** On dit qu'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est *linéaire* si elle vérifie les deux conditions suivantes :

a) $\forall x_1, x_2 \in \mathbb{R}^n,\ f(x_1 + x_2) = f(x_1) + f(x_2)$,

b) $\forall x \in \mathbb{R}^n, \forall \lambda \in \mathbb{R},\ f(\lambda x) = \lambda f(x)$.

On dit encore d'une fonction linéaire que c'est une *forme* linéaire.

Proposition (Caractérisation des formes linéaires)**.** *Soit $f : \mathbb{R}^n \to \mathbb{R}$. Alors f est une forme linéaire si, et seulement si, il existe un vecteur a* unique *tel que $f(x) = \langle a, x\rangle$.*

On dit ainsi qu'une forme linéaire est *représentée* par un produit scalaire. Les courbes de niveau d'une forme linéaire de deux variables sont des droites toutes parallèles et orthogonales au vecteur a qui représente la forme. Quant aux surfaces de niveau d'une forme linéaire de trois variables, ce sont des plans tous parallèles et orthogonaux au vecteur a.

Définition (Fonction affine)**.** On dit d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ qu'elle est *affine* si elle est de la forme :

$$f(x) = \langle a, x\rangle + c$$

où $a \in \mathbb{R}^n$ et c est une constante. La partie $\langle a, x\rangle$ est appelée la *partie linéaire* de f.

Remarque. Une fonction affine $f : \mathbb{R}^n \to \mathbb{R}$ a pour domaine de définition $D_f = \mathbb{R}^n$. Toute fonction linéaire est aussi affine (avec constante $c = 0$).

Si f est une fonction affine, on peut déterminer la constate c en évaluant f en 0 : $c = f(0)$. Le vecteur a qui représente la partie linéaire de f est donné par le gradient de f.

Proposition (Caractérisation des fonctions linéaires)**.** *Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est affine si et seulement si le gradient ∇f est constant. Dans ce cas, en posant $a := \nabla f(x)$ et $c = f(0)$, on a $f(x) = \langle a, x\rangle + c$.*

5.2 Fonctions homogènes

Définition (Fonction homogène)**.** Soit Ω un cône et $f : \Omega \to \mathbb{R}$. On dit que f est *homogène* de degré α (sur Ω) si

$$\forall x \in \Omega, \quad \forall\, t > 0 \quad \text{on a} \quad f(tx) = t^\alpha f(x).$$

Cette définition montre bien pourquoi le domaine de définition d'une fonction homogène doit être un cône. Dès qu'un x appartient au domaine de définition, il doit exister une relation entre la valeur de la fonction en x et les valeurs sur les points de la forme tx avec $t > 0$. Cela impose donc implicitement à tous ces points d'appartenir aussi au domaine de définition.

Pour les fonctions dérivables, on a une caractérisation remarquable des fonctions homogènes.

Théorème (Identité d'Euler)**.** *Soit $\Omega \subset \mathbb{R}^n$ un cône et $f : \Omega \to \mathbb{R}$ une fonction dérivable. Alors f est homogène de degré α si et seulement si l'égalité suivante (*identité d'Euler*) est vérifiée pour tout $x \in \Omega$:*

$$\langle x, \nabla f(x)\rangle = \alpha f(x).$$

Remarque. Il existe des fonctions homogènes qui ne sont pas dérivables. Par exemple, la fonction $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = ||x||$ est homogène de degré 1, mais elle n'est pas dérivable en 0.

5.3 Fonctions convexes et concaves

Une fonction est convexe si entre deux points x et y arbitrairement choisis dans son domaine de définition, le graphe de f se trouve toujours en dessous de la droite passant par les deux points $(x, f(x))$ et $(y, f(y))$.

Définition (Fonction convexe). Soit $\Omega \subset \mathbb{R}^n$ un convexe et $f : \Omega \to \mathbb{R}$. On dit que f est *convexe* sur Ω si

$$\forall\, x, y \in \Omega, \quad \forall\, t \in [0, 1], \text{ on a } \quad f\,(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

La définition précédente impose au domaine de définition d'une fonction convexe d'être un ensemble convexe. Pour tout x et tout y dans le domaine de définition, il doit exister une relation entre les valeurs de la fonction en x et en y et les valeurs sur les points du segment $[x, y]$. Cela impose donc implicitement à tous les points du segment $[x, y]$ d'appartenir aussi au domaine de définition.

En utilisant la définition d'une fonction convexe, on voit facilement que les fonctions suivantes sont convexes :
- la fonction valeur absolue $x \mapsto |x|$ sur $\mathbb{R}$,
- la fonction norme $x \mapsto ||x||$ sur $\mathbb{R}^n$,
- les fonctions affines d'une ou plusieurs variables.

Le fonctions convexes permettent de déterminer que certains ensembles sont convexes, et réciproquement.

Proposition. *Soit $\Omega \subset \mathbb{R}^n$ un ensemble convexe, et $f : \Omega \to \mathbb{R}$ une fonction. Alors f est une fonction convexe si et seulement si l'épi-graphe $\{(x, t) \in \mathbb{R}^n \times \mathbb{R} : t \geq f(x)\}$ est un sous-ensemble convexe de $\mathbb{R}^{n+1}$.*

On verra plus loin qu'une fonction $f : I \to \mathbb{R}$ deux fois dérivable sur l'intervalle I est convexe si et seulement si $f''(x) \geq 0$ pour tout $x \in I$. Avec ce critère, on déduit plusieurs exemples de fonctions usuelles convexes.

Exemple. Les fonctions suivantes sont convexes :

1. la fonction exponentielle $x \mapsto e^x$ (sur tout $\mathbb{R}$),
2. les trinômes du second degré $x \mapsto ax^2 + bx + c$ pour lesquels $a > 0$ (sur tout $\mathbb{R}$),
3. les fonctions puissances $x \mapsto x^\alpha$ pour lesquelles $\alpha \geq 1$ ou $\alpha \leq 0$ (sur $[0, +\infty[$).

Effectuer des opérations entre fonctions convexes ne préserve pas en général la convexité sauf sous des conditions assez restrictives comme celles données ci-dessous.

Proposition (Opérations entre fonctions qui préservent la convexité). *On obtient une fonction convexe*

1) en additionnant deux fonctions convexes,

2) en effectuant une combinaison à coefficients positifs de fonctions convexes

3) en prenant la composée $g \circ a$ d'une fonction affine a et d'une fonction convexe g

4) en prenant la composée $g \circ f$ d'une fonction convexe f et d'une fonction convexe croissante *g.*

En utilisant le point 3) de la proposition ci-dessus, on montre ainsi que la fonction

$$(x_1, x_2) \mapsto (x_1 - x_2)^4$$

est convexe sur $\mathbb{R}^2$. De même, en utilisant le point 4), on vérifie que la fonction

$$x \mapsto e^{2x^2+1}$$

est convexe sur $\mathbb{R}$.

Définition (Fonction concave). Soit $\Omega \subset \mathbb{R}^n$ un convexe et $f : \Omega \to \mathbb{R}$. On dit que f est *concave* sur Ω si

$$\forall\, x, y \in \Omega, \quad \forall\, t \in [0, 1], \text{ on a } \quad f\,(tx + (1-t)y) \geq tf(x) + (1-t)f(y).$$

On observera qu'une fonction f est concave si et seulement si $(-f)$ est convexe. On vérifie immédiatement que les fonctions affines sont concaves. Ces dernières ont donc la propriété d'être à la fois convexes et concaves.

Proposition. *Soit $\Omega \subset \mathbb{R}^n$ un ensemble convexe, et $f : \Omega \to \mathbb{R}$ une fonction. Alors f est une fonction concave si et seulement si le sous-graphe $\{(x, t) \in \mathbb{R}^n \times \mathbb{R} : t \geq f(x)\}$ est un sous-ensemble convexe de $\mathbb{R}^{n+1}$.*

On verra juste après qu'une fonction $f : I \to \mathbb{R}$ deux fois dérivable sur l'intervalle I est concave si et seulement si $f''(x) \leq 0$ pour tout $x \in I$. Avec ce critère, on déduit plusieurs exemples de fonctions usuelles concaves.

Exemple. Les fonctions suivantes sont concaves :

1. la fonction logarithme $x \mapsto \ln x$ (sur $]0, +\infty[$),
2. les trinômes du second degré $x \mapsto ax^2 + bx + c$ pour lesquels $a < 0$ (sur tout $\mathbb{R}$),
3. les fonctions puissances $x \mapsto x^\alpha$ pour lesquelles $\alpha \in [0, 1]$ (sur $[0, +\infty[$).

Proposition (Opérations entre fonctions qui préservent la concavité)**.** *On obtient une fonction concave*

1) en additionnant deux fonctions concaves,

2) en effectuant une combinaison à coefficients positifs de fonctions concaves,

3) en prenant la composée $g \circ a$ d'une fonction affine a et d'une fonction concave g,

4) en prenant la composée $g \circ f$ d'une fonction concave f et d'une fonction concave croissante *g.*

Pour les fonctions deux fois dérivables, la convexité et la concavité sont détectées par la dérivée seconde. On discutera ici uniquement le cas des fonction d'une variable réelle. Les fonctions de deux variables seront traitées au cours du prochain semestre.

Proposition (Caractérisation des fonctions convexes/concaves par le signe de la dérivée seconde)**.** *Soit $I \subset \mathbb{R}$ un intervalle ouvert et $f : I \to \mathbb{R}$ une fonction deux fois dérivable. Alors f est convexe si et seulement si $f^{''}(x) \geq 0$ en tout $x \in I$ et f est concave si et seulement si $f^{''}(x) \leq 0$ en tout $x \in I$.*

On ne peut caractériser les fonctions strictement convexes/concaves deux fois dérivables à partir du signe de la dérivée seconde. On observera que la fonction $x \mapsto x^4$ est bien strictement convexe, alors que sa dérivée seconde est nulle en $x = 0$.

Proposition (Caractérisation des extrema globaux des fonctions convexes/concaves dérivables)**.** *Soit $I \subset \mathbb{R}$ un intervalle ouvert et $f : I \to \mathbb{R}$ une fonction convexe (resp. concave) dérivable. Alors un point $\overline{x}$ est minimum (resp. maximum) global de f sur I si et seulement si $f'(\overline{x}) = 0$.*

Remarque. La convexité d'une fonction n'assure pas l'existence d'un minimum, comme on peut le voir avec la fonction exponentielle. Celle-ci est bien convexe mais n'atteint jamais sa valeur minimale. De même la concavité d'une fonction n'assure pas l'existence d'un maximum comme le montre la fonction logarithme.