

MATHÉMATIQUES

MODALITES 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
6 sept – 10 sept	C1, C2 (1h)	–
13 sept – 17 sept	C3, C4 (1h)	–
20 sept – 24 sept	C5	TD1
27 sept – 1 oct	C6	TD2
4 oct – 8 oct	C7	TD3
11 oct – 15 oct	C8	TD4
18 oct – 22 oct	C9	TD5 (CC n°1)
25 oct – 29 oct	C10	TD6
1 nov – 5 nov	vacances	
8 nov – 12 nov	C11	TD7
15 nov – 20 nov	C12	TD8
22 nov – 27 nov	C13	TD9 (CC n°2)
29 nov – 3 dec	–	–
6 dec – 10 dec	Première semaine des examens	
13 dec – 17 dec	Deuxième semaine des examens	

RÉSUMÉ DE COURS

1 Rappels

1.1 Trinôme du second degré

Définition(d'un trinôme du second degré)

On appelle trinôme du second degré un polynôme du second degré, c'est-à-dire une fonction du type

$$f(x) := ax^2 + bx + c$$

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

On appelle racine (ou zéro) du trinôme f toute valeur $x \in \mathbb{R}$ telle que $ax^2 + bx + c = 0$.

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

Définition(du 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 $f(x)$ sous la forme

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

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

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

- 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

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

qui est donc l'unique racine.

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

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

si bien qu'il y a deux racines

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

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 $x = r$ est racine, alors $(x - r)$ se met en facteur si bien que l'on peut écrire le trinôme sous la forme

$$f(x) = (x - r)(\alpha x - \beta)$$

ce qui permet, par identification, de déterminer immédiatement α et β .

Proposition (somme et produit des racines d'un trinôme)

Etant 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 $p = c/a$.

Proposition (sur l'extremum d'un trinôme du second degré)

Le trinôme $ax^2 + bx + c$ admet un extremum au point $\bar{x} = -b/2a$. Si $a > 0$, il s'agit d'un minimum global et si $a < 0$ il s'agit d'un maximum global. Si le trinôme a deux racines distinctes, l'extremum est situé exactement au milieu de celles-ci.

Le graphe d'un trinôme du second degré s'appelle une parabole. Celle-ci peut, suivant les valeurs prises par les coefficients a , b ou c , rencontrer ou non l'axe des abscisses. Les points de rencontre, s'ils existent, correspondent aux racines du trinôme. Dans le cas d'un seul point de rencontre on dit que la racine est double (on peut interpréter cela comme deux points de rencontre confondus). On peut aussi remarquer que l'orientation des branches paraboliques ne dépend que du signe du coefficient a du terme de plus haut degré.

1.2 Sur la factorisation des polynômes

Définition (d'un polynôme)

Un polynôme est une fonction du type

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n,$$

où les a_i sont des réels et $a_n \neq 0$. L'entier n s'appelle le degré du polynôme et les a_i sont les coefficients.

Factoriser un polynôme consiste à le mettre sous la forme d'un produit de plusieurs facteurs, ces facteurs étant des polynômes de degré strictement plus petit que n . Factoriser facilite la résolution des équations et la détermination du signe d'une expression. Par exemple, factoriser le polynôme $p(x) := x^2 + x - 2$ sous 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.

Tout polynôme n'est pas factorisable en produit de polynômes de degré 1. Par exemple le polynôme $p(x) = 1 + x + x^2$ ne l'est pas.

On dit que $r \in \mathbb{R}$ est racine du polynôme p si $p(r) = 0$.

Proposition (Lien entre racine et factorisation)

Si $r \in \mathbb{R}$ est racine du polynôme p , on peut le factoriser sous la forme

$$p(x) = (x - r)q(x)$$

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

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

$$p(x) = (x - 1)(\alpha x^2 + \beta x + \gamma).$$

En identifiant les termes de plus haut et de plus bas degré, on obtient immédiatement $\alpha = 5$ et $\gamma = 2$ si bien qu'on peut écrire directement

$$p(x) = (x - 1)(5x^2 + \beta x + 2).$$

Il ne reste alors plus qu'à déterminer β . En identifiant les termes de degré 1, (ou ceux de degré 2) on obtient $\beta = -3$.

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.

Proposition

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

En effet, si les coefficients respectifs de deux polynômes sont égaux il va de soi que les deux fonctions polynômes sont identiques. Réciproquement si deux fonctions polynômes $a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ et $b_0 + b_1x + b_2x^2 + \dots + b_nx^n$ sont égales on a

$$a_0 + a_1x + a_2x^2 + \dots + a_nx^n = b_0 + a_1x + b_2x^2 + \dots + b_nx^n \quad \text{pour tout } x.$$

En prenant $x = 0$ on obtient $a_0 = b_0$. Tenant alors compte de l'égalité $a_0 = b_0$ l'égalité des deux fonctions polynômes devient

$$a_1x + a_2x^2 + \dots + a_nx^n = b_1x + b_2x^2 + \dots + b_nx^n \quad \text{pour tout } x$$

ce qui implique (en simplifiant par x)

$$a_1 + a_2x + \dots + a_nx^{n-1} = b_1 + b_2x + \dots + b_nx^{n-1} \quad \text{pour tout } x \neq 0.$$

Cette dernière égalité, bien que valable uniquement pour des x non nuls, reste vraie pour des x infiniment petits, ce qui conduit, par *passage à la limite*, à l'égalité $a_1 = b_1$. Tenant compte du fait que $a_0 = b_0$ et $a_1 = b_1$ l'égalité des deux fonctions polynômes devient

$$a_2x^2 + \dots + a_nx^n = b_2x^2 + \dots + b_nx^n \quad \text{pour tout } x$$

ce qui implique (en simplifiant par x^2)

$$a_2 + \dots + a_nx^{n-2} = b_2 + \dots + b_nx^{n-2} \quad \text{pour tout } x \neq 0.$$

De nouveau cette dernière égalité n'est valable que pour des x non nuls mais reste vraie pour des x infiniment petits, ce qui conduit, par *passage à la limite*, à l'égalité $a_2 = b_2$. De proche en proche on montre ainsi que tous les coefficients respectifs sont égaux.

1.3 Opérateur Σ

Une expression du type

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

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 $S := 2 + 2^2 + 2^3 + 2^4 + 2^5 + \dots + 2^{15}$ 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

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

1.4 Valeur absolue

Définition(de la valeur absolue)

Etant 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

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

1.5 Système d'équations linéaires

On appelle système de deux équations linéaires à deux inconnues un système d'équations du type

$$(\mathcal{S}) : \begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2. \end{cases}$$

Il y a deux façons d'aborder un tel système, l'une purement analytique et l'autre géométrique. D'un point de vue géométrique, l'ensemble des solutions d'une équation du type

$$\alpha_1x_1 + \alpha_2x_2 = \beta$$

est constitué des points d'une droite. Si $\alpha_2 \neq 0$, cette droite a pour pente $p = -(\alpha_1/\alpha_2)$ et elle coupe l'axe des ordonnées en $x_2 = \beta/\alpha_2$. Si $\alpha_2 = 0$ la droite est verticale.

Il résulte de cette observation que résoudre le système $(\mathcal{S})$ revient à trouver l'intersection de deux droites du plan. Ces deux droites se coupent en un point unique dans la mesure où elles ne sont pas parallèles, i.e. si leurs pentes $p_1 = -(a_{11}/a_{12})$ et $p_2 = -(a_{21}/a_{22})$ sont différentes.

Ecrire que les deux pentes sont différentes, i.e. que

$$-\frac{a_{11}}{a_{12}} \neq -\frac{a_{21}}{a_{22}}$$

revient à dire que la quantité $a_{11}a_{22} - a_{12}a_{21}$ n'est pas nulle.

Définition(du déterminant d'un système linéaire de deux équations à deux inconnues)

Etant donné un système

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2. \end{cases}$$

de deux équations linéaires à deux inconnues, on appelle déterminant du système la quantité

$$\Delta := a_{11}a_{22} - a_{12}a_{21}.$$

Ainsi, le système $(\mathcal{S})$ a une solution unique si son déterminant est non nul. Si ce déterminant est nul, les deux droites ont même pente. En général, elles ne se coupent pas et le système $(\mathcal{S})$ n'a donc pas de solution mais les deux droites peuvent aussi être confondues, auquel cas $(\mathcal{S})$ a une infinité de solutions.

D'un point de vue pratique on peut facilement résoudre le système en exprimant x_2 en fonction de x_1 (ou le contraire) dans l'une des équations et en remplaçant x_2 par son expression en fonction de x_1 dans l'autre

équation. On se ramène ainsi à une seule équation et une seule inconnue. Cette façon de procéder n'est pas très élégante et ne peut malheureusement pas se généraliser aux systèmes de 3 équations à 3 inconnues et de façon plus générale aux systèmes de n équations à n inconnues. Prenons un exemple et résolvons le système

$$\begin{cases} 2x_1 + x_2 = 10 \\ 3x_1 + 2x_2 = 1. \end{cases}$$

De la première équation on tire $x_2 = 10 - 2x_1$ et en remplaçant x_2 par sa valeur en fonction de x_1 dans la seconde équation on obtient $x_1 = 19$. Il suffit alors de reprendre la première équation pour obtenir $x_2 = -28$.

Pour résoudre le système on peut aussi le transformer en un système triangulaire équivalent. Cela consiste à conserver la première équation et à remplacer la seconde par une combinaison des deux, de façon à éliminer la variable x_1 .

Avec l'exemple précédent, en gardant la première équation et en remplaçant la seconde par la combinaison *2 fois la seconde équation moins 3 fois la première* on obtient le système triangulaire équivalent

$$\begin{cases} 2x_1 + x_2 = 10 \\ x_2 = -28. \end{cases}$$

qui permet d'obtenir immédiatement la valeur de x_2 , puis celle de x_1 en remplaçant x_2 par sa valeur dans la première équation.

Cette seconde façon de procéder est à préférer à la première car elle est à la base de la méthode de résolution du pivot de Gauss qui, en l'absence de structure particulière du système à résoudre, est l'une des meilleures méthodes directes de résolution.

2 L'espace $\mathbb{R}^n$

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 $x_1, x_2, \dots, x_n$ 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 $(x_1, x_2, \dots, x_n)$. 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

$$\mathbb{R}^n := \{(x_1, x_2, \dots, x_n); x_i \in \mathbb{R}, i = 1, 2, \dots, n\}.$$

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 $(x_1, x_2, \dots, x_n)$ et $(y_1, y_2, \dots, y_n)$ sont égaux si, et seulement si, $x_i = y_i$ pour tout $i = 1, 2, \dots, n$.

Pour un élément $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, 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 peut ainsi représenter les $x \in \mathbb{R}^2$ par les points d'un plan et les $x \in \mathbb{R}^3$ par les points de l'espace à trois dimensions. Cela permet alors d'utiliser le langage classique de la géométrie et de parler de droite, de plan, de triangle, de rectangle, etc. On peut aussi voir un élément $x \in \mathbb{R}^n$ comme une flèche issue de l'origine et dont l'extrémité est le point x . Dans ce cas, on parle de vecteur. Pour éviter toute confusion dans les notations, on peut mettre une flèche sur la lettre qui désigne le vecteur.

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

$$x + y := (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

et le produit de x par le scalaire λ est défini par

$$\lambda x := (\lambda x_1, \lambda x_2, \dots, \lambda x_n).$$

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

Définition(de la colinéarité entre deux vecteurs de $\mathbb{R}^n$)

On dit que deux vecteurs x et y de $\mathbb{R}^n$ sont colinéaires s'il existe un scalaire $t \neq 0$ tel que $x = ty$.

Définition(d'une droite de $\mathbb{R}^n$)

Soit $x \in \mathbb{R}^n$ et $d \in \mathbb{R}^n$. La droite (affine) passant par le point x et dirigée par le vecteur d est l'ensemble

$$D := \{x + td; t \in \mathbb{R}\}.$$

La droite passant par les points $x, y \in \mathbb{R}^n$ est dirigée par le vecteur $y - x$. Par conséquent, elle est donnée par

$$D := \{x + t(y - x); t \in \mathbb{R}\}.$$

Définition(d'une demi-droite de $\mathbb{R}^n$)

Soit $x \in \mathbb{R}^n$ et $d \in \mathbb{R}^n$. La demi-droite issue du point x et dirigée par le vecteur d est l'ensemble

$$D_+ := \{x + td; t \geq 0\}.$$

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 x et y sont des points de $\mathbb{R}^n$, on peut considérer la droite qui passe par ces deux points et définir le segment $[x, y]$ comme l'ensemble des points de cette droite situés entre x et y .

Définition(d'un segment de $\mathbb{R}^n$)

Soit $x, y \in \mathbb{R}^n$. On appelle segment d'extrémités x et y l'ensemble

$$\begin{aligned}[x, y] &:= \{x + t(y - x); t \in [0, 1]\} \\ &= \{ty + (1 - t)x; t \in [0, 1]\}.\end{aligned}$$

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

Définition(du produit scalaire)

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

Proposition (propriétés du produit scalaire de deux vecteurs)

Si x, y et z sont des vecteurs de $\mathbb{R}^n$ et λ est un scalaire, on a

- . $\langle x, y \rangle = \langle y, x \rangle$ symétrie
- . $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$, et $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$ linéarité par rapport au premier membre
- . $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$, et $\langle x, \lambda y \rangle = \lambda \langle x, y \rangle$ linéarité par rapport au second membre.

Définition(vecteurs orthogonaux)

On dit de deux vecteurs x et y non nuls qu'ils sont orthogonaux ou encore qu'ils forment un angle droit si $\langle x, y \rangle = 0$.

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

Définition(de la norme euclidienne d'un vecteur)

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

On observera que par définition on a $\|x\| = \sqrt{\langle x, x \rangle}$. 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

- . $\|x\| \geq 0$ non négativité
- . $\|x\| = 0 \Leftrightarrow x = 0$ non dégénérescence
- . $\|\lambda x\| = |\lambda| \|x\|$ pour $\lambda \in \mathbb{R}$ homogénéité de degré 1
- . $\|x + y\| \leq \|x\| + \|y\|$ inégalité triangulaire

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

Proposition (identités remarquables vectorielles)

Si x et y sont des vecteurs de $\mathbb{R}^n$ on a les identités remarquables

- . $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle$,
- . $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x, y \rangle$,
- . $\|x\|^2 - \|y\|^2 = \langle x + y, x - y \rangle$.

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

$$d(x, y) := \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} = \|x - y\|.$$

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

Définition(de la distance euclidienne entre deux points de $\mathbb{R}^n$)

On appelle distance euclidienne entre deux points $x, y \in \mathbb{R}^n$ la quantité définie par

$$d(x, y) := \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

3 Généralités sur les fonctions

3.1 Définitions

Définition(d'une fonction)

Etant donné un ensemble de départ A et un ensemble d'arrivée B , une fonction f de A dans B fait correspondre à tout $x \in A$ au plus un élément de B , que l'on note $f(x)$.

Si l'élément $f(x)$ est bien défini, il s'appelle l'image de x .

Définition(domaine de définition)

Le domaine de définition d'une fonction $f : A \rightarrow B$ est l'ensemble

$$D_f := \{x \in A, f(x) \text{ est bien défini.}\}$$

Si $D_f = A$, on dit que $f : A \rightarrow B$ est une application de A dans B . Dans ce cas, chaque élément de A possède une et une seule image.

Exemple

Considérons la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par

$$f(x) := \sqrt{1 - x^2}.$$

Pour donner un sens à la racine, il faut imposer $1 - x^2 \geq 0$, i.e. $x \in [-1, 1]$. Finalement, on a $D_f = [-1, 1]$.

Définition(du graphe d'une fonction)

Si $f : A \rightarrow \mathbb{R}$ est une fonction définie sur A , on appelle graphe de f l'ensemble

$$\text{Gr}(f) := \{(x, y) \in A \times \mathbb{R}; \quad f(x) = y\}.$$

Le graphe d'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est un sous-ensemble de $\mathbb{R}^2$, c'est une *courbe*. Le graphe d'une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ est un sous-ensemble de $\mathbb{R}^3$, c'est une *surface*.

Définition(d'un ensemble de niveau d'une fonction)

Si $f : A \rightarrow \mathbb{R}$ est une fonction définie sur A , on appelle ensemble de niveau de f un ensemble du type

$$L_\lambda := \{x \in A; \quad f(x) = \lambda\} \quad \text{où } \lambda \in \mathbb{R}.$$

Lorsque $A \subset \mathbb{R}^2$, on parle également de courbe 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 ».

3.2 Opérations sur les fonctions

Définition(de la composée de deux applications)

Si f est une application de A dans un ensemble B et si g est une application de B dans C , on peut définir la composée $g \circ f$ qui est une application de A dans C , en posant

$$(g \circ f)(x) := g[f(x)].$$

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.

Par exemple la fonction

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

se décompose sous forme

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

De même, la fonction de deux variables

$$\varphi(x_1, x_2) := (x_1 - x_2)^4$$

se décompose sous forme

$$\varphi = g \circ f \quad \text{avec} \quad f(x_1, x_2) := x_1 - x_2 \quad \text{et} \quad g(t) := t^4.$$

On peut additionner, multiplier, diviser deux fonctions à valeurs réelles, multiplier une fonction à valeurs réelles par un scalaire, ou encore prendre l'enveloppe supérieure ou l'enveloppe inférieure de deux fonctions à valeurs réelles. On définit ainsi de nouvelles fonctions notées $f + g$, fg , f/g , λf , $\max\{f, g\}$ et $\min\{f, g\}$ en posant

- . $(f + g)(x) := f(x) + g(x)$
- . $(fg)(x) := f(x)g(x)$
- . $(f/g)(x) := f(x)/g(x)$ sous réserve que $g(x) \neq 0$
- . $(\lambda f)(x) := \lambda f(x)$
- . $\max\{f, g\}(x) := \max\{f(x), g(x)\}$
- . $\min\{f, g\}(x) := \min\{f(x), g(x)\}$.

En comparant les valeurs d'une fonction on est conduit à la notion de minimum et de maximum.

Définition(d'un minimum global et d'un maximum global d'une fonction)

Soit A un ensemble quelconque, $f : A \rightarrow \mathbb{R}$ et $\Omega \subset A$. On dit qu'un point $\bar{x} \in \Omega$ est un minimum global (resp. maximum global) de f sur Ω si

$$\forall x \in \Omega, \quad f(\bar{x}) \leq f(x) \quad (\text{resp. } f(\bar{x}) \geq f(x)).$$

Remarque

On prendra garde, pour les fonctions, de ne pas confondre minimum et valeur minimale. Un minimum d'une fonction est un point du domaine de définition sur lequel la fonction prend sa valeur minimale. Par exemple le point $\bar{x} = 1$ est minimum de la fonction $f(x) := x^2 - 2x + 5$ tandis que la valeur minimale de la fonction vaut $\bar{f} = 4$.

3.3 Fonctions d'une variable réelle

Travailler avec des fonctions d'une variable réelle permet de définir les notions classiques suivantes.

Définition(d'une fonction croissante et d'une fonction décroissante)

On dit qu'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est croissante (resp. décroissante) si pour tout $x_1 \leq x_2$ on a $f(x_1) \leq f(x_2)$ (resp. $f(x_1) \geq f(x_2)$). Lorsque l'inégalité précédente est stricte pour $x_1 \neq x_2$, on dit que f est strictement croissante (resp. strictement décroissante).

De façon équivalente, on peut caractériser la croissance (resp. la décroissance) d'une fonction par la propriété

$$\forall x_1, x_2 \in \mathbb{R}, \quad (x_1 - x_2)(f(x_1) - f(x_2)) \geq 0 \quad (\text{resp. } (x_1 - x_2)(f(x_1) - f(x_2)) \leq 0).$$

Définition(d'une fonction monotone)

On dit d'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ qu'elle est monotone (resp. strictement monotone) si elle est croissante ou décroissante (resp. strictement croissante ou strictement décroissante).

Il arrive que le graphe d'une fonction possède des propriétés de symétrie qui permettent d'en réduire le domaine d'étude. Cela justifie la définition suivante.

Définition(d'une fonction paire et d'une fonction impaire)

On dit qu'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est paire (resp. impaire) si pour tout x on a $f(x) = f(-x)$ (resp. $f(x) = -f(-x)$).

3.4 Fonctions usuelles

3.4.1 Logarithme

On définit la fonction logarithme népérien $\ln$ au moyen d'une intégrale.

Définition(de la fonction logarithme)

Pour tout $x > 0$, on pose

$$\ln x := \int_1^x \frac{1}{t} dt.$$

Le nombre $\ln x$ correspond à l'aire algébrique de la surface comprise entre l'axe des abscisses, le graphe de la fonction $t \mapsto 1/t$ et les deux droites verticales d'équation $t = 1$ et $t = x$. Cette aire doit être comptée positivement si $x > 1$ et négativement si $x < 1$. Pour $x = 1$, la surface étant réduite à un trait vertical, l'aire est nulle. On a donc $\ln 1 = 0$.

Proposition (sur les propriétés algébriques du logarithme népérien)

- . $\forall x > 0$ et $\forall y > 0$, $\ln xy = \ln x + \ln y$,
- . $\forall x > 0$ et $\forall y > 0$, $\ln\left(\frac{x}{y}\right) = \ln x - \ln y$,
- . $\forall x > 0$, $\forall \alpha \in \mathbb{R}$, $\ln(x^\alpha) = \alpha \ln x$.

La dernière propriété ci-dessus prendra son sens lorsque nous aurons défini la fonction puissance $x \mapsto x^\alpha$ à partir des fonctions logarithme et exponentielle.

Proposition (sur les propriétés analytiques du logarithme népérien)

La fonction logarithme

- . a pour domaine de définition l'ensemble $]0, +\infty[$,
- . est une fonction strictement croissante,
- . admet une racine en $x = 1$ (donc $\ln 1 = 0$),
- . vérifie

$$\lim_{x \rightarrow 0^+} \ln x = -\infty \quad \text{et} \quad \lim_{x \rightarrow +\infty} \ln x = +\infty.$$

3.4.2 Exponentielle

Le logarithme népérien est une fonction strictement croissante, donc injective. Pour $\alpha \in \mathbb{R}$ donné, l'équation

$$\ln x = \alpha$$

ne peut donc admettre qu'une seule solution. L'existence d'une solution est assurée par le théorème des valeurs intermédiaires (admis). On peut ainsi définir une fonction réciproque au logarithme.

Définition(de la fonction exponentielle)

On appelle fonction exponentielle, et on note $\exp$ la fonction réciproque de la fonction logarithme $\ln$.

Pour tout $\alpha \in \mathbb{R}$, le réel $\exp \alpha$ est donc par définition l'unique solution de l'équation $\ln x = \alpha$

$$\ln x = \alpha \iff x = \exp \alpha.$$

De la définition il résulte que $\exp 0 = 1$. On retiendra aussi que $e := \exp 1 \sim 2,718$ et que e n'est pas un nombre rationnel. Les propriétés de l'exponentielle se déduisent directement des propriétés du logarithme.

Proposition (sur les propriétés algébriques de l'exponentielle)

- . $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}$, $\exp(x + y) = (\exp x)(\exp y)$ et $\exp(x - y) = \frac{\exp x}{\exp y}$,
- . $\forall x, \forall \alpha \in \mathbb{R}$, $\exp(\alpha x) = (\exp x)^\alpha$.

La deuxième propriété ci-dessus prendra son sens lorsque nous aurons défini la fonction puissance.

Proposition (sur les propriétés analytiques de l'exponentielle)

La fonction exponentielle

- . a pour domaine de définition l'ensemble $\mathbb{R}$,
- . est strictement croissante,
- . est à valeurs strictement positives, ce qui signifie que $\exp x > 0$ pour tout $x \in \mathbb{R}$,

. vérifie

$$\lim_{x \rightarrow -\infty} \exp x = 0 \quad \text{et} \quad \lim_{x \rightarrow +\infty} \exp x = +\infty.$$

3.4.3 Fonction puissance

Définition (de la puissance d'un nombre réel positif)

Etant donné $\alpha \in \mathbb{R}$, la fonction puissance est définie pour tout $x > 0$, par

$$x^\alpha := \exp(\alpha \ln x).$$

Observons que

$$e^\alpha = \exp(\alpha \ln e) = \exp \alpha,$$

car $\ln e = 1$. Ceci justifie l'emploi de la notation e^α pour l'exponentielle de α .

En utilisant la définition, on voit également que

$$\ln(x^\alpha) = \ln[\exp(\alpha \ln x)] = \alpha \ln x$$

et

$$(\exp x)^\alpha = \exp(\alpha \ln[\exp(x)]) = \exp(\alpha x),$$

qui correspondent à des propriétés déjà énoncées du logarithme et de l'exponentielle.

Proposition (sur les propriétés algébriques de la fonction puissance)

- . $\forall x > 0, \forall \alpha \in \mathbb{R}, \forall \beta \in \mathbb{R}, \quad x^\alpha x^\beta = x^{\alpha+\beta} \quad \text{et} \quad \frac{x^\alpha}{x^\beta} = x^{\alpha-\beta},$
- . $\forall x > 0, \forall y > 0, \forall \alpha \in \mathbb{R}, \quad (xy)^\alpha = x^\alpha y^\alpha \quad \text{et} \quad \left(\frac{x}{y}\right)^\alpha = \frac{x^\alpha}{y^\alpha},$
- . $\forall x > 0, \forall \alpha \in \mathbb{R}, \forall \beta \in \mathbb{R}, \quad (x^\alpha)^\beta = x^{\alpha\beta}.$

Proposition (sur les propriétés analytiques de la fonction puissance)

La fonction $x \mapsto x^\alpha$ ($\alpha \in \mathbb{R}$) a pour domaine de définition l'ensemble $]0, +\infty[$.

Si $\alpha > 0$ alors

- . la fonction $x \mapsto x^\alpha$ est strictement croissante,
- . $\lim_{x \rightarrow 0^+} x^\alpha = 0$ et $\lim_{x \rightarrow +\infty} x^\alpha = +\infty$.

Si $\alpha < 0$ alors

- . la fonction $x \mapsto x^\alpha$ est strictement décroissante,
- . $\lim_{x \rightarrow 0^+} x^\alpha = +\infty$ et $\lim_{x \rightarrow +\infty} x^\alpha = 0$.

4 Fonctions linéaires, homogènes, convexes

4.1 Fonctions linéaires

Définition(de la linéarité d'une fonction)

On dit qu'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ 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 $f : \mathbb{R}^n \rightarrow \mathbb{R}$. 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(d'une fonction affine)

On dit d'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ qu'elle est affine si elle est de la forme :

$$f(x) = \ell(x) + c$$

où ℓ est une forme linéaire et c est une constante.

4.2 Fonctions homogènes

Définition(d'un cône)

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

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

Définition(d'une fonction homogène)

Soit Ω un cône et $f : \Omega \rightarrow \mathbb{R}$. 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.

4.3 Fonctions convexes et fonctions concaves

Définition(d'un ensemble convexe)

On dit qu'un ensemble $\Omega \subset \mathbb{R}^n$ est convexe si pour tout $x, y \in \Omega$ on a $[x, y] \subset \Omega$.

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(d'une fonction convexe)

Soit $\Omega \subset \mathbb{R}^n$ un convexe et $f : \Omega \rightarrow \mathbb{R}$. On dit que f est convexe sur Ω si

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

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.

On verra dans le chapitre suivant qu'une fonction $f : I \rightarrow \mathbb{R}$ 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 (sur les 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 (d'une fonction concave)

Soit $\Omega \subset \mathbb{R}^n$ un convexe et $f : \Omega \rightarrow \mathbb{R}$. On dit que f est concave sur Ω si

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

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.

On verra dans le chapitre suivant qu'une fonction $f : I \rightarrow \mathbb{R}$ 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 (sur les 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 .

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.

5 Dérivées et dérivées partielles

5.1 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.

Définition(de la dérivée d'une fonction d'une variable en un point)

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

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

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

Interprétation

En posant $\Delta f := f(\bar{x} + h) - f(\bar{x})$ et $\Delta x := h$, la définition de la dérivabilité peut se réécrire sous la forme

$$\frac{\Delta f}{\Delta x} \rightarrow f'(\bar{x}) \quad \text{quand } \Delta x \rightarrow 0.$$

Suite à une variation Δx de la variable, la variation Δf de la fonction est donc approximativement donnée par

$$\Delta f \sim \Delta x \times f'(\bar{x}).$$

Si on regarde le graphe de la fonction f , on voit immédiatement que la quantité $\tau(h)$ s'interprète comme la pente de la droite qui interpole f aux points $\bar{x}$ et $\bar{x} + h$ (*interpoler* signifie *passer par*). Dire que le quotient différentiel admet une limite revient à dire que cette droite admet une position limite qui définit une droite appelée la tangente au graphe de f en $(\bar{x}, f(\bar{x}))$. La dérivée en $\bar{x}$ mesure donc la pente de cette tangente, qui a pour équation

$$y = f(\bar{x}) + f'(\bar{x})(x - \bar{x}).$$

Définition(de la 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 $\bar{x}$, on dit que f est deux fois dérivable en $\bar{x}$. La dérivée de f' en $\bar{x}$ s'appelle la dérivée seconde de f en $\bar{x}$ et se note $f''(\bar{x})$ ou parfois $f^{(2)}(\bar{x})$.

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 .

5.2 Dérivées des fonctions usuelles

Proposition (dérivées des fonctions usuelles)

f	D_f	f'
$x^n, n \neq 0, n \in \mathbb{Z}$	$\mathbb{R}$ si $n > 0$ et $\mathbb{R} \setminus \{0\}$ si $n < 0$	nx^{n-1}
$\frac{1}{x}$	$\mathbb{R} \setminus \{0\}$	$-\frac{1}{x^2}$
$\sqrt{x}$	$[0, +\infty[$	$\frac{1}{2\sqrt{x}}$
$\ln x$	$]0, +\infty[$	$\frac{1}{x}$
e^x	$\mathbb{R}$	e^x
x^α	$]0, +\infty[, \alpha \in \mathbb{R}$	$\alpha x^{\alpha-1}$

5.3 Propriétés des fonctions dérivables

Proposition (propriétés des fonctions dérivables)

Si $f, g : \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions dérivables en un point $\bar{x}$, alors :

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

Ces propriétés restent évidemment valables si les fonctions ne sont définies que sur une partie $\Omega \subset \mathbb{R}$ dans la mesure où Ω est un voisinage de $\bar{x}$.

Proposition (règle de dérivation en chaîne)

Soient $u : \mathbb{R} \rightarrow \mathbb{R}$ une fonction dérivable en un point $\bar{x}$ et $g : \mathbb{R} \rightarrow \mathbb{R}$ une fonction dérivable au point $u(\bar{x})$. Alors $g \circ u$ est dérivable au point $\bar{x}$ et on a

$$(g \circ u)'(\bar{x}) = g'(u(\bar{x}))u'(\bar{x}).$$

Ce résultat est essentiel pour obtenir la dérivée de fonctions dont l'expression analytique est complexe, mais qui peuvent se décomposer en fonctions élémentaires faciles à dériver. Il reste valable si les fonctions u et g sont définies uniquement au voisinage respectivement des points $\bar{x}$ et $u(\bar{x})$.

Exemple

Considérons la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par

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

La règle de dérivation en chaîne donne

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

5.4 Elasticité des fonctions d'une variable

On peut être amené à comparer les variations relatives d'une fonction par rapport aux variations relatives de la variable. Le but est de formaliser une affirmation du type : *si la variable augmente de 2%, alors les valeurs de la fonction varient approximativement de 5%*.

Cela conduit au taux de variation relative $\frac{\Delta f/f}{\Delta x/x}$ donné par

$$\frac{\frac{f(\bar{x}+h) - f(\bar{x})}{f(\bar{x})}}{\frac{(\bar{x}+h) - (\bar{x})}{\bar{x}}} = \frac{f(\bar{x}+h) - f(\bar{x})}{h} \frac{\bar{x}}{f(\bar{x})}.$$

On introduit la notion d'élasticité qui est la limite, si elle existe, du taux de variation relative. Évidemment cela ne peut avoir un sens que si les valeurs de la variable et de la fonction sont non nulles (pour ne pas diviser par zéro).

Définition (de l'élasticité d'une fonction en un point)

Soit f une fonction d'une variable définie au voisinage d'un point $\bar{x} \neq 0$ sur lequel f ne s'annule pas. On appelle élasticité de f en $\bar{x}$ la quantité

$$E(f; \bar{x}) := \bar{x} \cdot \frac{f'(\bar{x})}{f(\bar{x})}.$$

On retiendra que, suite à une variation relative $\Delta_r x := \Delta x/x$ de la variable, la variation relative $\Delta_r f$ de la fonction est approximativement donnée par

$$\Delta_r f \sim \Delta_r x \times E(f; \bar{x}).$$

Par exemple, si la demande d'un produit est de 354 unités, et si pour cette valeur l'élasticité en fonction du prix du produit vaut 1.7, une augmentation du prix de 2% aura pour effet une augmentation (approximative) de la demande de $2 \times 1.7 = 3.4\%$, c'est-à-dire d'une douzaine d'unités.

5.5 Quelques applications des dérivées

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 $f : I \rightarrow \mathbb{R}$ 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.

Proposition (caractérisation des fonctions convexes/concaves par le signe de la dérivée seconde)

Soit $I \subset \mathbb{R}$ un intervalle ouvert et $f : I \rightarrow \mathbb{R}$ 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 (condition nécessaire d'optimalité sans contrainte pour les fonctions d'une variable)

Soit $I \subset \mathbb{R}$ un intervalle ouvert et $f : I \rightarrow \mathbb{R}$ une fonction dérivable sur I . Si un point $\bar{x} \in I$ est un minimum (resp. maximum) global de f sur I , alors $f'(\bar{x}) = 0$.

La condition $f'(\bar{x}) = 0$ n'est pas en général une condition suffisante d'optimalité. Pour s'en convaincre il suffit de prendre la fonction $x \mapsto x^3$.

Définition (d'un point critique d'une fonction d'une variable)

Soit $I \subset \mathbb{R}$ un intervalle ouvert et $f : I \rightarrow \mathbb{R}$ une fonction dérivable sur I . On appelle point critique de f un point $\bar{x}$ tel que $f'(\bar{x}) = 0$.

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 ni minimum ni maximum et que l'on appelle points d'inflexion.

Proposition (caractérisation des extrema globaux des fonctions convexes/concaves dérivables)

Soit $I \subset \mathbb{R}$ un intervalle ouvert et $f : I \rightarrow \mathbb{R}$ une fonction convexe (resp. concave) dérivable. Alors un point $\bar{x}$ est minimum (resp. maximum) global de f sur I si et seulement si $f'(\bar{x}) = 0$.

5.6 Dérivées partielles

Définition (d'une dérivée partielle)

Soit f une fonction définie au voisinage d'un point $\bar{x} = (\bar{x}_1, \bar{x}_2) \in \mathbb{R}^2$. On appelle dérivée partielle de f par rapport à la première variable au point $\bar{x}$ la limite (si elle existe)

$$\frac{\partial f}{\partial x_1}(\bar{x}) := \lim_{h \rightarrow 0} \frac{f(\bar{x}_1 + h, \bar{x}_2) - f(\bar{x}_1, \bar{x}_2)}{h},$$

qui est la dérivée de la fonction d'une variable $x_1 \mapsto f(x_1, \bar{x}_2)$, la variable x_2 étant fixée (à la valeur $\bar{x}_2$). On note aussi parfois cette dérivée partielle $D_1 f(\bar{x})$ ou encore $f'_{x_1}(\bar{x})$.

De façon similaire on appelle dérivée partielle de f par rapport à la seconde variable au point $\bar{x}$ la limite (si elle existe)

$$\frac{\partial f}{\partial x_2}(\bar{x}) := \lim_{h \rightarrow 0} \frac{f(\bar{x}_1, \bar{x}_2 + h) - f(\bar{x}_1, \bar{x}_2)}{h},$$

qui est la dérivée de la fonction d'une variable $x_2 \mapsto f(\bar{x}_1, x_2)$, la variable x_1 étant fixée (à la valeur $\bar{x}_1$).

Il est d'usage de regrouper les dérivées partielles en un vecteur que l'on appelle le gradient de la fonction.

Définition(du gradient d'une fonction de 2 variables)

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction admettant des dérivées partielles en un point $\bar{x}$. On appelle gradient de f en $\bar{x}$ le vecteur

$$\nabla f(\bar{x}) := \left(\frac{\partial f}{\partial x_1}(\bar{x}), \frac{\partial f}{\partial x_2}(\bar{x}) \right).$$

On a alors le résultat immédiat suivant.

Proposition (condition nécessaire d'optimalité d'une fonction de 2 variables)

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction admettant des dérivées partielles en un point $\bar{x}$. Si $\bar{x}$ est minimum ou maximum global, alors son gradient est nul, i.e. $\nabla f(\bar{x}) = 0$.

Définition(d'un point critique d'une fonction de 2 variables)

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de 2 variables définie au voisinage d'un point $\bar{x}$ et admettant en $\bar{x}$ des dérivées partielles. Si $\nabla f(\bar{x}) = 0$, on dit du point $\bar{x}$ que c'est un point critique.

La condition $\nabla f(x) = 0$ n'est pas une condition suffisante d'optimalité. Pour s'en convaincre il suffit de prendre la fonction $f(x_1, x_2) = x_1 x_2$. On a clairement

$$\nabla f(x_1, x_2) = (x_2, x_1),$$

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