

UE MAT234

Notes de cours sur les fonctions de plusieurs variables

1 Fonctions de plusieurs variables réelles

1.1 Définitions générales

1.1.1 Définition

$\mathbf{R}^n = \mathbf{R} \times \mathbf{R} \times \dots \times \mathbf{R}$ est l'ensemble des n -uplets de réels $\mathbf{x} = (x_1, \dots, x_n)$. De la même façon que l'on a défini les fonctions d'une seule variable réelle, on peut définir les fonctions de plusieurs variables réelles :

$$f : E \subset \mathbf{R}^n \longrightarrow \mathbf{R} \\ (x_1, \dots, x_n) \longrightarrow f(x_1, \dots, x_n)$$

En pratique, on travaillera le plus souvent avec $n = 2$ ou $n = 3$. On utilisera parfois dans ce cas la notation (x, y) ou (x, y, z) plutôt que (x_1, x_2) ou (x_1, x_2, x_3) .

Exemples :

- Longueur (ou norme) d'un vecteur : $f(x_1, \dots, x_n) = \sqrt{\sum_{i=1}^n x_i^2}$
- Température en un point : c'est une fonction de la position géographique (x, y, z) et du temps t : $T(x, y, z, t)$

1.1.2 Polynômes à plusieurs variables

Un **polynôme à n variables** est une fonction de la forme

$$P(x_1, \dots, x_n) = \sum_{p_1, \dots, p_n} a_{p_1, \dots, p_n} x_1^{p_1} \dots x_n^{p_n}$$

avec les exposants p_1, \dots, p_n entiers et les coefficients a_{p_1, \dots, p_n} réels.

Le **degré total** du monôme $x_1^{p_1} \dots x_n^{p_n}$ est $p_1 + \dots + p_n$.

Le **degré partiel par rapport à la variable x_i** du monôme $x_1^{p_1} \dots x_n^{p_n}$ est p_i .

Un polynôme est dit **homogène de degré p** s'il est constitué de monômes dont le degré total est toujours égal à p .

Le **degré total d'un polynôme** est le plus haut degré total des monômes qui le composent.

Le **degré partiel d'un polynôme** par rapport à une variable est le plus haut degré partiel par rapport à cette variable des monômes qui le composent.

Exemple : $P(x, y, z) = x^3y + z^4 - 3x^2z^2 + 7xz^3$ est homogène de degré 4. Ses degrés partiels par rapport à x , y et z sont respectivement 3, 1 et 4.

1.1.3 Continuité

La définition de la continuité d'une fonction de plusieurs variables est une généralisation du cas d'une fonction d'une seule variable.

Soit $f : E \subset \mathbb{R}^n \longrightarrow \mathbb{R}$. f est continue en $\mathbf{a} = (a_1, \dots, a_n)$ si et seulement si $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$. Cette définition peut aussi s'écrire de la façon suivante : f est continue en \mathbf{a} si et seulement si pour tout voisinage V de $f(\mathbf{a})$, il existe un voisinage W de \mathbf{a} tel que $\forall \mathbf{w} \in W, f(\mathbf{w}) \in V$.

Exemples :

- $f(x, y) = xy$ est continue en $(0, 0)$. En effet : $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0 = f(0, 0)$.
- Soit f définie par $f(x, y) = \frac{x^3 y}{x^2 + y^2}$ si $(x, y) \neq (0, 0)$, et $f(0, 0) = 0$. f est continue en $(0, 0)$. En effet, $0 \leq |f(x, y)| \leq |xy| \frac{x^2}{x^2 + y^2} \leq |xy|$. D'où $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0 = f(0, 0)$.
- Soit f définie par $f(x, y) = \frac{xy}{x^2 + y^2}$ si $(x, y) \neq (0, 0)$, et $f(0, 0) = 0$. f n'est pas continue en $(0, 0)$. En effet, en choisissant par exemple $y = x$, on voit que $f(x, x) = 1/2$, qui ne tend pas vers $f(0, 0) = 0$ quand x tend vers 0. Graphiquement, cela se traduit par une rupture dans le dessin de f au voisinage de $(0, 0)$.

1.1.4 Généralisation : fonctions de \mathbb{R}^n vers \mathbb{R}^p

On peut généraliser de façon naturelle les fonctions de \mathbb{R}^n vers \mathbb{R} et introduire des fonctions de \mathbb{R}^n vers \mathbb{R}^p pour $p > 1$:

$$f : E \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p \\ (x_1, \dots, x_n) \longrightarrow (f_1(x_1, \dots, x_n), \dots, f_p(x_1, \dots, x_n))$$

où chaque fonction f_i est une fonction de \mathbb{R}^n vers \mathbb{R} .

Exemple : Une station météo mobile fournit la température T , la pression P , l'humidité θ et la vitesse du vent (u, v, w) . C'est donc une fonction de \mathbb{R}^4 vers \mathbb{R}^6 :

$$M(x, y, z, t) = (T(x, y, z, t), P(x, y, z, t), \theta(x, y, z, t), u(x, y, z, t), v(x, y, z, t), w(x, y, z, t)).$$

1.2 Fonctions de deux variables réelles

On considère ici f de \mathbb{R}^2 vers \mathbb{R} .

1.2.1 Représentations graphiques

Représentation 3-D : on trace en perspective dans \mathbb{R}^3 la surface formée des points $(x, y, f(x, y))$.

Représentation par lignes de niveaux : on trace dans le plan les lignes isovaleurs (ou lignes de niveau) de la fonction. L'isoligne de niveau K est $\{(x, y) / f(x, y) = K\}$

Exemple : La ligne de niveau K ($K \geq 0$) de la fonction $f(x, y) = x^2 + y^2$ est le cercle de centre 0 de rayon \sqrt{K} .

1.2.2 Coordonnées polaires

Un point $M(x, y)$ du plan peut aussi être repéré par ses **coordonnées polaires** (r, θ) , définies par $x = r \cos \theta$, $y = r \sin \theta$, avec $r \geq 0$ et $\theta \in [0, 2\pi[$. r est appelé **rayon**, et θ **angle polaire** ou **argument**.

Ce changement de coordonnées $(x, y) \rightarrow (r, \theta)$ est bijectif de $\mathbf{R}^2 - \{(0, 0)\}$ vers $\mathbf{R}_+^* \times [0, 2\pi[$. Le changement inverse est : $r = \sqrt{x^2 + y^2}$, $\cos \theta = \frac{x}{\sqrt{x^2 + y^2}}$, $\sin \theta = \frac{y}{\sqrt{x^2 + y^2}}$.

1.3 Fonctions de trois variables réelles

On considère ici $f(x, y, z)$, de \mathbf{R}^3 vers \mathbf{R} .

1.3.1 Coordonnées cylindriques

Un point $M(x, y, z)$ de \mathbf{R}^3 peut aussi être repéré par ses **coordonnées cylindriques** (r, θ, z) , définies par $x = r \cos \theta$, $y = r \sin \theta$, avec $r \geq 0$ et $\theta \in [0, 2\pi[$. (r, θ) est donc l'expression en coordonnées polaires de (x, y) .

1.3.2 Coordonnées sphériques

Un point $M(x, y, z)$ de \mathbf{R}^3 peut aussi être repéré par ses **coordonnées sphériques** (r, θ, φ) , définies par $x = r \cos \theta \sin \varphi$, $y = r \sin \theta \sin \varphi$, $z = r \cos \varphi$, avec $r \geq 0$, $\theta \in [0, 2\pi[$ et $\varphi \in [0, \pi]$. C'est un système de repérage usuel sur une sphère : r est appelé **rayon** ou distance au centre, θ **longitude**, φ **colatitude** ($\pi/2 - \varphi$ est la latitude).

2 Dérivation d'une fonction de plusieurs variables

2.1 Dérivées premières

Soit f de \mathbf{R}^n vers \mathbf{R} , $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ la base canonique de \mathbf{R}^n , et $\mathbf{a} \in \mathbf{R}^n$.

Définition Soit \mathbf{d} un vecteur de \mathbf{R}^n . On appelle **dérivée directionnelle** de f au point \mathbf{a} dans la direction \mathbf{d} , notée $\frac{\partial f}{\partial \mathbf{d}}(\mathbf{a})$:

$$\frac{\partial f}{\partial \mathbf{d}}(\mathbf{a}) = \lim_{\alpha \rightarrow 0} \frac{f(\mathbf{a} + \alpha \mathbf{d}) - f(\mathbf{a})}{\alpha} = \lim_{\alpha \rightarrow 0} \frac{f(a_1 + \alpha d_1, \dots, a_n + \alpha d_n) - f(a_1, \dots, a_n)}{\alpha} \quad \text{si elle existe}$$

La dérivée directionnelle $\partial f / \partial \mathbf{d}$ est donc une fonction de \mathbf{R}^n vers \mathbf{R} .

Exemples :

- La pente d'un relief dans une direction donnée est la dérivée directionnelle dans cette direction.
- Soit $f(x, y) = x^3y - y$ et $\mathbf{d} = (\frac{1}{2}, \frac{\sqrt{3}}{2})$. On a : $\frac{\partial f}{\partial \mathbf{d}}(x, y) = \frac{3}{2}x^2y + \frac{\sqrt{3}}{2}(x^3 - 1)$.
- En physique, on parle souvent de **dérivée normale** et de **dérivée tangentielle** pour désigner la dérivée directionnelle en un point d'une courbe ou d'une surface donnée, dans la direction normale ou tangente à cette courbe ou surface.

Définition On appelle **dérivée partielle** de f par rapport à la i -ème variable x_i la dérivée directionnelle de f dans la direction \mathbf{e}_i . Elle est notée $\frac{\partial f}{\partial x_i}$.

En pratique on calcule $\partial f / \partial x_i$ comme une dérivée classique, en supposant les variables $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ constantes et en dérivant par rapport à x_i .

Définition Le vecteur $\left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{a}) \right)$ est appelé **gradient de f au point \mathbf{a}** , noté $\text{grad}f(\mathbf{a})$ ou $\nabla f(\mathbf{a})$.

Propriété Si $\nabla f(\mathbf{a})$ existe, alors $\frac{\partial f}{\partial \mathbf{d}}(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \mathbf{d}$

Définitions On dit que f est **de classe \mathcal{C}^1** en \mathbf{a} si chaque dérivée partielle existe au voisinage de \mathbf{a} et est continue en \mathbf{a} .

Définition On suppose que f est de classe \mathcal{C}^1 en \mathbf{a} . On appelle **différentielle de f en \mathbf{a}** l'application linéaire notée $Df[\mathbf{a}]$ définie par

$$Df[\mathbf{a}] : \mathbf{R}^n \longrightarrow \mathbf{R}$$
$$\mathbf{h} = (h_1, \dots, h_n) \longrightarrow Df[\mathbf{a}](\mathbf{h}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{a}) h_i = \nabla f(\mathbf{a}) \cdot \mathbf{h}$$

On utilise souvent la notation dx_i au lieu de h_i .

Exemple : pour $f(x, y) = x^3y - y$, on a $Df[(x, y)](dx, dy) = 3x^2y dx + (x^3 - 1) dy$.

Propriété On a alors : $f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + Df[\mathbf{a}](\mathbf{h}) + o(\|\mathbf{h}\|)$.

C'est l'équivalent de la formule des accroissements finis pour les fonctions d'une seule variable.

Définition Soit f de \mathbf{R}^n vers \mathbf{R}^p : $f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_p(x_1, \dots, x_n))$.
 f est de classe \mathcal{C}^1 en $\mathbf{a} \in \mathbf{R}^n$ si et seulement si chaque application composante f_i l'est.
La différentielle de f en $\mathbf{a} \in \mathbf{R}^n$ s'écrit alors :

$$Df[\mathbf{a}](\mathbf{h}) = (Df_1[\mathbf{a}](\mathbf{h}), \dots, Df_p[\mathbf{a}](\mathbf{h})), \text{ soit en abrégé } Df[\mathbf{a}] = (Df_1[\mathbf{a}], \dots, Df_p[\mathbf{a}]).$$

Ainsi, puisque $Df_i[\mathbf{a}](\mathbf{h}) = \nabla f_i \cdot \mathbf{h}$, on a, avec des notations en vecteurs colonnes :

$$Df[\mathbf{a}](\mathbf{h}) = \begin{pmatrix} Df_1[\mathbf{a}](\mathbf{h}) \\ \vdots \\ Df_p[\mathbf{a}](\mathbf{h}) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}[\mathbf{a}] & \cdots & \frac{\partial f_1}{\partial x_n}[\mathbf{a}] \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1}[\mathbf{a}] & \cdots & \frac{\partial f_p}{\partial x_n}[\mathbf{a}] \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix},$$

soit, $Df[\mathbf{a}](\mathbf{h}) = J_f(\mathbf{a}) \cdot \mathbf{h}$, où $J_f(\mathbf{a})$ est appelée **matrice jacobienne de f au point \mathbf{a}** .

Exemple On reprend l'exemple du changement de variables en coordonnées polaires $x = r \cos \theta$, $y = r \sin \theta$. L'application des résultats précédents donne

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix} \begin{pmatrix} dr \\ d\theta \end{pmatrix}$$

2.2 Dérivation de fonctions composées

Propriété Soit $f : E_1 \subset \mathbf{R}^n \rightarrow \mathbf{R}^p$ et $g : E_2 \subset \mathbf{R}^p \rightarrow \mathbf{R}$. On suppose, pour que $g \circ f$ soit définie, que $f(E_1) \subset E_2$. Soit $\mathbf{a} \in E_1$. On suppose que f est différentiable en \mathbf{a} et que g est différentiable en $f(\mathbf{a})$. Alors $g \circ f$ est différentiable en \mathbf{a} et $D(g \circ f)[\mathbf{a}] = Dg[f(\mathbf{a})] \circ Df[\mathbf{a}]$. Matriciellement, ceci est équivalent à $J_{g \circ f}(\mathbf{a}) = J_g(f(\mathbf{a})) \times J_f(\mathbf{a})$

Application au changement de variables Soit Φ un changement de coordonnées $(x_1, \dots, x_n) \rightarrow (y_1, \dots, y_n)$. Soit f une fonction de classe \mathcal{C}^1 et $g = f \circ \Phi : g(x_1, \dots, x_n) = f(y_1, \dots, y_n)$. La propriété précédente s'écrit

$$\forall i = 1, \dots, n \quad \frac{\partial g}{\partial x_i}(x_1, \dots, x_n) = \sum_{k=1}^n \frac{\partial f}{\partial y_k}(y_1, \dots, y_n) \frac{\partial y_k}{\partial x_i}$$

Cette formule est parfois appelée **formule des dérivées totales**.

Exemple Soit une fonction f de \mathbf{R}^2 vers \mathbf{R} . On note g son expression en coordonnées polaires $g(r, \theta) = f(r \cos \theta, r \sin \theta) = f(x, y)$. On a alors :

$$\begin{cases} \frac{\partial g}{\partial r} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} = \cos \theta \frac{\partial f}{\partial x} + \sin \theta \frac{\partial f}{\partial y} = \frac{x}{\sqrt{x^2 + y^2}} \frac{\partial f}{\partial x} + \frac{y}{\sqrt{x^2 + y^2}} \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial \theta} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \theta} = -r \sin \theta \frac{\partial f}{\partial x} + r \cos \theta \frac{\partial f}{\partial y} = -y \frac{\partial f}{\partial x} + x \frac{\partial f}{\partial y} \end{cases}$$

qui s'inverse en

$$\begin{cases} \frac{\partial f}{\partial x} = \cos \theta \frac{\partial g}{\partial r} - \frac{1}{r} \sin \theta \frac{\partial g}{\partial \theta} \\ \frac{\partial f}{\partial y} = \sin \theta \frac{\partial g}{\partial r} + \frac{1}{r} \cos \theta \frac{\partial g}{\partial \theta} \end{cases}$$

2.3 Dérivées d'ordres supérieurs

Soit une fonction f de $E \subset \mathbf{R}^n$ vers \mathbf{R} .

Définition Soient $\{i_1, \dots, i_k\} \in \{1, \dots, n\}^k$. On dit que f admet une dérivée partielle d'ordre k au point $\mathbf{a} \in \mathbf{R}^n$ par rapport aux variables x_{i_1}, \dots, x_{i_k} si et seulement si

- $\frac{\partial f}{\partial x_{i_1}}, \frac{\partial}{\partial x_{i_2}} \left(\frac{\partial f}{\partial x_{i_1}} \right), \dots, \frac{\partial}{\partial x_{i_{k-1}}} \left(\frac{\partial f}{\partial x_{i_{k-2}}} \dots \left(\frac{\partial f}{\partial x_{i_1}} \right) \dots \right)$ existent dans un voisinage de \mathbf{a}
- $\frac{\partial f}{\partial x_{i_1}}, \frac{\partial}{\partial x_{i_2}} \left(\frac{\partial f}{\partial x_{i_1}} \right), \dots, \frac{\partial}{\partial x_{i_k}} \left(\frac{\partial f}{\partial x_{i_{k-1}}} \dots \left(\frac{\partial f}{\partial x_{i_1}} \right) \dots \right) (\mathbf{a})$ existe

On note cette dérivée $\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}} (\mathbf{a})$.

Définition On dit que f est de classe \mathcal{C}^k sur E si et seulement si toutes les dérivées partielles de f jusqu'à l'ordre k existent et sont continues sur E .

Théorème de Schwarz Si f est de classe \mathcal{C}^1 sur E , si $\frac{\partial^2 f}{\partial x_i \partial x_j}$ et $\frac{\partial^2 f}{\partial x_j \partial x_i}$ existent sur E

et sont continues en \mathbf{a} , alors $\frac{\partial^2 f}{\partial x_i \partial x_j} (\mathbf{a}) = \frac{\partial^2 f}{\partial x_j \partial x_i} (\mathbf{a})$.

On en déduit que, pour f de classe \mathcal{C}^k , l'ordre de dérivation dans le calcul des dérivées partielles n'a pas d'importance.

2.4 Formule de Taylor à l'ordre 2

Soit une fonction f de \mathbf{R}^n vers \mathbf{R} , de classe \mathcal{C}^2 . Soit $\mathbf{a} = (a_1, \dots, a_n)$. On a alors pour tout incrément $\mathbf{h} = (h_1, \dots, h_n)$:

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i} (\mathbf{a}) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j} (\mathbf{a}) + o(\|\mathbf{h}\|^2)$$

Dans le cas d'une fonction de deux variables, cette formule devient :

$$\begin{aligned} f(x+h, y+k) &= f(x, y) + h \frac{\partial f}{\partial x} (x, y) + k \frac{\partial f}{\partial y} (x, y) \\ &\quad + \frac{h^2}{2} \frac{\partial^2 f}{\partial x^2} (x, y) + hk \frac{\partial^2 f}{\partial x \partial y} (x, y) + \frac{k^2}{2} \frac{\partial^2 f}{\partial y^2} (x, y) + o(\|h, k\|^2) \end{aligned}$$

On peut aussi l'exprimer sous la forme suivante : il existe $\theta \in]0, 1[$ tel que

$$\begin{aligned} f(x+h, y+k) &= f(x, y) + h \frac{\partial f}{\partial x} (x, y) + k \frac{\partial f}{\partial y} (x, y) \\ &\quad + \frac{h^2}{2} \frac{\partial^2 f}{\partial x^2} (x + \theta h, y + \theta k) + hk \frac{\partial^2 f}{\partial x \partial y} (x + \theta h, y + \theta k) + \frac{k^2}{2} \frac{\partial^2 f}{\partial y^2} (x + \theta h, y + \theta k) \end{aligned}$$

3 Optimisation d'une fonction de plusieurs variables

3.1 Extrema des fonctions de \mathbb{R}^n vers \mathbb{R}

Soit une fonction f de $E \subset \mathbb{R}^n$ vers \mathbb{R} .

Définitions

- f admet un **minimum local** en \mathbf{a} si et seulement si il existe un voisinage V de \mathbf{a} tel que $\forall \mathbf{x} \in V, f(\mathbf{x}) \geq f(\mathbf{a})$.
- f admet un **minimum global** en \mathbf{a} si et seulement si $\forall \mathbf{x} \in E, f(\mathbf{x}) \geq f(\mathbf{a})$.
- f admet un **minimum strict (local ou global)** en \mathbf{a} si et seulement si on peut remplacer les inégalités larges par des inégalités strictes dans les définitions précédentes.
- De même pour définir un **maximum, strict ou non, local ou global**.
- **Extremum** : minimum ou maximum

Définition On dit que $\mathbf{a} \in E$ est un **point critique** si et seulement si les dérivées partielles premières de f en \mathbf{a} existent et sont nulles (c'est à dire $\nabla f(\mathbf{a}) = 0$).

Théorème On suppose que les dérivées partielles de f existent en un point \mathbf{a} n'appartenant pas au bord de E . Une **condition nécessaire** pour que f admette un extremum en \mathbf{a} est que $\nabla f(\mathbf{a}) = 0$.

Théorème Si E est un domaine fermé et borné, et si f est continue sur E , alors f admet un minimum et un maximum globaux sur E .

Théorème Les extrema d'une fonction \mathcal{C}^1 sur un domaine fermé et borné sont soit des points critiques, soit des points du bord de E .

3.2 Optimisation des fonctions de \mathbb{R}^2 vers \mathbb{R}

Soit une fonction f de $E \subset \mathbb{R}^2$ vers \mathbb{R} de classe \mathcal{C}^2 au voisinage d'un point critique (x_0, y_0) . On pose (notations de Monge) :

$$r = \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \quad s = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \quad t = \frac{\partial^2 f}{\partial y^2}(x_0, y_0)$$

- Si $rt - s^2 > 0$ et $r < 0$: f admet un maximum local strict en (x_0, y_0)
- Si $rt - s^2 > 0$ et $r > 0$: f admet un minimum local strict en (x_0, y_0)
- Si $rt - s^2 < 0$: f a un point-selle (ni min, ni max) en (x_0, y_0)
- Si $rt - s^2 = 0$: on ne peut pas donner de conclusion générale. Il faut étudier localement le comportement de f au voisinage de (x_0, y_0) .

3.3 Méthode d'approximation des moindres carrés

Une question fréquemment rencontrée est celui de la modélisation du lien existant entre deux variables X et Y . On dispose en pratique d'un échantillon de n mesures $\{(x_1, y_1), \dots, (x_n, y_n)\}$ de ces deux variables. Si le nuage des points (x_i, y_i) semble suivre un certain ordre, on peut essayer de modéliser la relation entre X et Y sous la forme $Y = \sum_{k=1}^K a_k f_k(X) + \epsilon$, où les f_k sont des fonctions élémentaires (x^α , $\ln x$, $\exp x$, $\sin x$, $\cos x$, ...) et où ϵ est l'erreur entre le modèle et la réalité. On dit alors qu'"on explique Y par X ", ou encore que Y est la **variable expliquée** et X la **variable explicative**.

A partir des n relations $y_i = \sum_{k=1}^K a_k f_k(x_i) + \epsilon_i$, $i = 1, \dots, n$, on peut définir une "erreur globale" entre le modèle et la réalité:

$$E(a_1, \dots, a_K) = \sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n [y_i - a_1 f_1(x_i) - \dots - a_K f_K(x_i)]^2$$

La **méthode des moindres carrés** consiste alors à déterminer les a_k qui minimisent cette erreur. Il suffit pour cela de résoudre le système linéaire de K équations à K inconnues

$$\begin{cases} \frac{\partial E}{\partial a_1}(a_1, \dots, a_K) = -2 \sum_{i=1}^n f_1(x_i) [y_i - a_1 f_1(x_i) - \dots - a_K f_K(x_i)] = 0 \\ \vdots \\ \frac{\partial E}{\partial a_K}(a_1, \dots, a_K) = -2 \sum_{i=1}^n f_K(x_i) [y_i - a_1 f_1(x_i) - \dots - a_K f_K(x_i)] = 0 \end{cases}$$

c'est à dire

$$\begin{cases} a_1 \sum_{i=1}^n f_1^2(x_i) + a_2 \sum_{i=1}^n f_1(x_i) f_2(x_i) + \dots + a_K \sum_{i=1}^n f_1(x_i) f_K(x_i) = \sum_{i=1}^n y_i f_1(x_i) \\ \vdots \\ a_1 \sum_{i=1}^n f_K(x_i) f_1(x_i) + a_2 \sum_{i=1}^n f_K(x_i) f_2(x_i) + \dots + a_K \sum_{i=1}^n f_K^2(x_i) = \sum_{i=1}^n y_i f_K(x_i) \end{cases}$$

Lorsque ces valeurs optimales $\hat{a}_1, \dots, \hat{a}_K$ sont déterminées, l'erreur globale $E(\hat{a}_1, \dots, \hat{a}_K)$ est appelée **erreur résiduelle**, et la valeur $\sqrt{\frac{1}{n} E(\hat{a}_1, \dots, \hat{a}_K)} = \sqrt{\frac{1}{n} \sum_{i=1}^n \epsilon_i^2}$ est appelée **écart-type résiduel**.

Si l'on souhaite comparer deux modèles de régression $Y = \sum_{k=1}^K a_k f_k(X)$ et $Y = \sum_{l=1}^L b_l g_l(X)$, on comparera en général leurs écarts-types résiduels.

Cas particulier: la régression linéaire

La relation pressentie entre X et Y est de la forme $Y = a X + b$.

On détermine \hat{a} et \hat{b} en minimisant l'erreur globale $E(a, b) = \sum \epsilon_i^2 = \sum (y_i - a x_i - b)^2$.

On veut

$$\begin{cases} \frac{\partial E}{\partial a}(\hat{a}, \hat{b}) = -2 \sum_{i=1}^n x_i (y_i - \hat{a} x_i - \hat{b}) = 0 \\ \frac{\partial E}{\partial b}(\hat{a}, \hat{b}) = -\sum_{i=1}^n (y_i - \hat{a} x_i - \hat{b}) = 0 \end{cases}$$

$$\text{soit : } \begin{cases} \left(\sum_{i=1}^n x_i \right) \hat{a} + n \hat{b} = \sum_{i=1}^n y_i \\ \left(\sum_{i=1}^n x_i^2 \right) \hat{a} + \left(\sum_{i=1}^n x_i \right) \hat{b} = \sum_{i=1}^n x_i y_i \end{cases}$$

D'où

$$\hat{a} = \frac{\frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \quad \text{et} \quad \hat{b} = \bar{y} - \hat{a} \bar{x}$$

On remarque que cette droite de régression passe par le centre d'inertie (\bar{x}, \bar{y}) du nuage de points.

4 Intégrales multiples

4.1 Quelques propriétés

La notion de fonction intégrable pour les fonctions de plusieurs variables est une généralisation directe de la même notion pour les fonctions d'une seule variable.

Pour f fonction de \mathbf{R}^n vers \mathbf{R} intégrable, son intégrale sur un domaine \mathcal{D} est notée $\int \dots \int_{\mathcal{D}} f(x_1, \dots, x_n) dx_1 \dots dx_n$, ou encore en abrégé $\int_{\mathcal{D}} f(\mathbf{x}) d\mathbf{x}$.

Soient f et g des fonctions de \mathbf{R}^n vers \mathbf{R} intégrables sur un domaine \mathcal{D} de \mathbf{R}^n . On a les propriétés suivantes :

- $\int_{\mathcal{D}} (f + g) = \int_{\mathcal{D}} f + \int_{\mathcal{D}} g$ et $\int_{\mathcal{D}} \lambda f = \lambda \int_{\mathcal{D}} f \quad \forall \lambda \in \mathbf{R}$ (linéarité)
- Si $f \geq 0$ sur \mathcal{D} alors $\int_{\mathcal{D}} f \geq 0$. On en déduit que, si $f \geq g$ sur \mathcal{D} , alors $\int_{\mathcal{D}} f \geq \int_{\mathcal{D}} g$.
- Si f est intégrable sur \mathcal{D} alors $|f|$ l'est aussi, et $\int_{\mathcal{D}} |f| \geq \left| \int_{\mathcal{D}} f \right|$.
- Si $f \geq 0$ sur \mathcal{D} , si f est continue sur \mathcal{D} , et si $\int_{\mathcal{D}} f = 0$, alors $f = 0$ sur \mathcal{D} .
- $\left(\int_{\mathcal{D}} f g \right)^2 \leq \left(\int_{\mathcal{D}} f^2 \right) \left(\int_{\mathcal{D}} g^2 \right)$ (Inégalité de Schwarz)

On a aussi la **relation de Chasles** : $\int_{\mathcal{D}_1 \cup \mathcal{D}_2} f = \int_{\mathcal{D}_1} f + \int_{\mathcal{D}_2} f$ pour des domaines \mathcal{D}_1 et \mathcal{D}_2 dont l'intersection est de mesure nulle (c'est à dire de surface nulle en dimension 2, ou de volume nul en dimension 3).

Changement de variables Soit Φ un changement de variable bijectif de classe \mathcal{C}^1 : $\Phi(y_1, \dots, y_n) = (x_1, \dots, x_n)$. Alors

$$\int_{\mathcal{D}} f(\mathbf{x}) d\mathbf{x} = \int_{\Phi^{-1}(\mathcal{D})} f(\Phi(\mathbf{y})) |\det J_{\Phi}| d\mathbf{y}$$

où J_{Φ} est la matrice jacobienne de Φ (matrice formée des $\partial x_i / \partial y_j$).

4.2 Calcul d'une intégrale double

Soit f de \mathbf{R}^2 vers \mathbf{R} intégrable sur un domaine \mathcal{D} :

- Si $\mathcal{D} = [a, b] \times [c, d]$, on peut intégrer indifféremment par rapport à x puis à y , ou l'inverse : $\int_{\mathcal{D}} f(x, y) dx dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_c^d \left(\int_a^b f(x, y) dx \right) dy$

Cas particulier :

$$\text{si } f(x, y) = g(x) h(y), \text{ alors } \int_{\mathcal{D}} f(x, y) dx dy = \left(\int_a^b g(x) dx \right) \left(\int_c^d h(y) dy \right)$$

- Si $\mathcal{D} = \{(x, y) / a \leq x \leq b \text{ et } \varphi_1(x) \leq y \leq \varphi_2(x)\}$, où φ_1 et φ_2 sont continues sur $[a, b]$: $\int_{\mathcal{D}} f(x, y) dx dy = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy \right) dx$ (théorème de Fubini, sommation par tranches)
- Si $\mathcal{D} = \{(x, y) / \psi_1(y) \leq x \leq \psi_2(y) \text{ et } c \leq y \leq d\}$, où ψ_1 et ψ_2 sont continues sur $[c, d]$: $\int_{\mathcal{D}} f(x, y) dx dy = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx \right) dy$ (théorème de Fubini, sommation par tranches)

Coordonnées polaires Si Φ est le changement de variables $(r, \theta) \rightarrow (x, y) = (r \cos \theta, r \sin \theta)$, on a : $|\det J_{\Phi}| = r$

4.3 Calcul d'une intégrale triple

Les principes vus dans le cas d'une intégrale double (ordre d'intégration, théorème de Fubini) s'étendent naturellement au cas des intégrales triples.

Coordonnées cylindriques Si Φ est le changement de variables $(r, \theta, z) \rightarrow (x, y, z) = (r \cos \theta, r \sin \theta, z)$, on a : $|\det J_{\Phi}| = r$

Coordonnées sphériques Si Φ est le changement de variables $(r, \theta, \varphi) \rightarrow (x, y, z) = (r \cos \theta \sin \varphi, r \sin \theta \sin \varphi, r \cos \varphi)$, on a : $|\det J_{\Phi}| = r^2 \sin \varphi$

5 Opérateurs aux dérivées partielles

Beaucoup de phénomènes physiques peuvent être traduits par des équations aux dérivées partielles, dans lesquelles apparaissent quelques opérateurs usuels.

5.1 Opérateurs usuels

Gradient Soit une fonction f de \mathbf{R}^n vers \mathbf{R} de classe \mathcal{C}^1 . On a déjà défini le **gradient de f au point \mathbf{a}** : $\left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{a}) \right)$, noté $\text{grad}f(\mathbf{a})$ ou $\nabla f(\mathbf{a})$.

En tout point \mathbf{a} , $\nabla f(\mathbf{a})$ est orthogonal à la ligne de niveau $f(\mathbf{a})$.

Divergence Soit une fonction f de \mathbf{R}^n vers \mathbf{R}^n (on dit aussi “champ de vecteurs”) de classe \mathcal{C}^1 . On note $f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_n(x_1, \dots, x_n))$. On appelle **divergence de f** la fonction de \mathbf{R}^n vers \mathbf{R} définie par

$$\text{div } f = \frac{\partial f_1}{\partial x_1} + \dots + \frac{\partial f_n}{\partial x_n}$$

Cet opérateur caractérise la convergence ou la divergence ponctuelle d’un champ de vecteurs.

Laplacien Soit une fonction f de \mathbf{R}^n vers \mathbf{R} de classe \mathcal{C}^2 . On appelle **laplacien de f** la fonction de \mathbf{R}^n vers \mathbf{R} , notée Δf ou $\nabla^2 f$, définie par

$$\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \dots + \frac{\partial^2 f}{\partial x_n^2}$$

On a : $\Delta f = \text{div}(\nabla f)$.

Les fonctions f vérifiant $\Delta f = 0$ sont appelées **fonctions harmoniques**. Elles interviennent notamment dans la résolution de l’équation des ondes et de l’équation de la chaleur (voir plus loin).

On peut aussi étendre la définition du laplacien à des champs de vecteurs: soit f de \mathbf{R}^n vers \mathbf{R}^p de classe \mathcal{C}^2 : $f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_p(x_1, \dots, x_n))$. On définit alors $\Delta f = (\Delta f_1, \dots, \Delta f_p)$.

Rotationnel Soit f de \mathbf{R}^3 vers \mathbf{R}^3 de classe \mathcal{C}^1 : $f(x, y, z) = (f_1(x, y, z), f_2(x, y, z), f_3(x, y, z))$. On appelle **rotationnel de f** la fonction de \mathbf{R}^3 vers \mathbf{R}^3 , notée $\text{Rot } f$ et définie par :

$$\text{rot } f = \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}, \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}, \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right)$$

Cette quantité caractérise le mouvement de rotation du champ de vecteurs f autour de chaque axe. On note aussi : $\text{rot } f = \nabla \wedge f$

Par abus de langage, on parle aussi parfois du rotationnel d’un champ de vecteurs f de \mathbf{R}^2 vers \mathbf{R}^2 , en le définissant comme la fonction de \mathbf{R}^2 vers \mathbf{R} : $\text{rot } f = \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}$

Equations aux dérivées partielles usuelles La plupart des phénomènes physiques peuvent être décrits, tout au moins de façon simplifiée, par des **équations aux dérivées partielles** (EDP) qui traduisent en général des principes tels que conservation de la masse, de la chaleur, de la quantité de mouvement... Les opérateurs précédents interviennent très souvent dans ces EDP. Par exemple :

- équation de la chaleur : $\frac{\partial f}{\partial t}(\mathbf{x}, t) = \text{div} (\nu(\mathbf{x}) \text{grad } f)(\mathbf{x}, t)$ où f est la température et ν la diffusivité thermique. Cette équation devient $\frac{\partial f}{\partial t}(\mathbf{x}, t) = \nu \Delta f(\mathbf{x}, t)$ si ν est constante.
- équation des ondes : $\frac{\partial^2 f}{\partial t^2}(\mathbf{x}, t) = c^2 \Delta f(\mathbf{x}, t)$ où c est la célérité des ondes.

5.2 Quelques propriétés

Linéarité Gradient, divergence, laplacien et rotationnel sont des opérateurs linéaires. On a donc $T(f + g) = T(f) + T(g)$ et $T(\lambda f) = \lambda T(f)$, où T représente n'importe lequel de ces opérateurs.

Composition d'opérateurs Pour les propriétés qui suivent, f et g sont des fonctions de \mathbf{R}^n vers \mathbf{R} , et u est un champs de vecteurs de \mathbf{R}^n vers \mathbf{R}^n . Ces fonctions seront de classe \mathcal{C}^1 ou \mathcal{C}^2 suivant le contexte, et on supposera $n = 3$ lorsqu'on fera intervenir le rotationnel. On a alors :

$$\text{grad}(fg) = f \text{grad } g + g \text{grad } f$$

$$\text{div}(fu) = f \text{div } u + \text{grad } f \cdot u$$

$$\text{rot}(fu) = f \text{rot } u - u \wedge \text{grad } f$$

$$\text{div}(\text{rot } u) = 0$$

$$\text{rot}(\text{grad } f) = 0$$

$$\text{rot}(\text{rot } u) = \text{grad}(\text{div } u) - \Delta u$$

Expressions en coordonnées polaires

On note $(O, \mathbf{i}, \mathbf{j})$ le repère orthonormé cartésien de \mathbf{R}^2 . Pour un point M de coordonnées cartésiennes (x, y) et de coordonnées polaires (r, θ) , on définit le vecteur unitaire $\mathbf{u}_r = \cos \theta \mathbf{i} + \sin \theta \mathbf{j}$, c'est à dire $\vec{OM} = r \mathbf{u}_r$. Soit \mathbf{u}_θ le vecteur unitaire qui lui est directement orthogonal : $\mathbf{u}_\theta = -\sin \theta \mathbf{i} + \cos \theta \mathbf{j}$.

- Soit f de \mathbf{R}^2 vers \mathbf{R} de classe \mathcal{C}^1 . On note g son expression en coordonnées polaires : $f(x, y) = g(r, \theta)$. On a alors :

$$\nabla f = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} = \frac{\partial g}{\partial r} \mathbf{u}_r + \frac{1}{r} \frac{\partial g}{\partial \theta} \mathbf{u}_\theta$$

- Soit f de \mathbf{R}^2 vers \mathbf{R} de classe \mathcal{C}^2 . On note g son expression en coordonnées polaires : $f(x, y) = g(r, \theta)$. On a alors :

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 g}{\partial r^2} + \frac{1}{r} \frac{\partial g}{\partial r} + \frac{1}{r^2} \frac{\partial^2 g}{\partial \theta^2}$$

- Soit f un champ de vecteurs de \mathbf{R}^2 vers \mathbf{R}^2 de classe \mathcal{C}^1 . On note $f(M) = f_1(x, y) \mathbf{i} + f_2(x, y) \mathbf{j} = g_1(r, \theta) \mathbf{u}_r + g_2(r, \theta) \mathbf{u}_\theta$. Alors :

$$\operatorname{div} f = \frac{\partial g_1}{\partial r} + \frac{1}{r} g_1 + \frac{1}{r} \frac{\partial g_2}{\partial \theta}$$

Expressions en coordonnées cylindriques

Les expressions du laplacien et de la divergence se déduisent des expressions en coordonnées polaires. Soit f de \mathbf{R}^3 vers \mathbf{R} . On note g son expression en coordonnées cylindriques : $f(x, y, z) = g(r, \theta, z)$. On a alors :

$$\begin{aligned} \nabla f &= \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k} = \frac{\partial g}{\partial r} \mathbf{u}_r + \frac{1}{r} \frac{\partial g}{\partial \theta} \mathbf{u}_\theta + \frac{\partial g}{\partial z} \mathbf{k} \\ \Delta f &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \frac{\partial^2 g}{\partial r^2} + \frac{1}{r} \frac{\partial g}{\partial r} + \frac{1}{r^2} \frac{\partial^2 g}{\partial \theta^2} + \frac{\partial^2 g}{\partial z^2} \end{aligned}$$

Soit maintenant f de \mathbf{R}^3 vers \mathbf{R}^3 de classe \mathcal{C}^1 . On note $(O, \mathbf{i}, \mathbf{j}, \mathbf{k})$ le repère orthonormé cartésien de \mathbf{R}^3 . Pour un point M de coordonnées cartésiennes (x, y, z) et de coordonnées cylindriques (r, θ, z) , on note $f(M) = f_1(x, y, z) \mathbf{i} + f_2(x, y, z) \mathbf{j} + f_3(x, y, z) \mathbf{k} = g_1(r, \theta, z) \mathbf{u}_r + g_2(r, \theta, z) \mathbf{u}_\theta + g_3(r, \theta, z) \mathbf{k}$. Alors :

$$\begin{aligned} \operatorname{div} f &= \frac{\partial g_1}{\partial r} + \frac{1}{r} g_1 + \frac{1}{r} \frac{\partial g_2}{\partial \theta} + \frac{\partial g_3}{\partial z} \\ \operatorname{rot} f &= \left(-\frac{\partial g_2}{\partial z} + \frac{1}{r} \frac{\partial g_3}{\partial \theta} \right) \mathbf{u}_r + \left(\frac{\partial g_1}{\partial z} - \frac{\partial g_3}{\partial r} \right) \mathbf{u}_\theta + \left(\frac{1}{r} \frac{\partial(r g_2)}{\partial r} - \frac{1}{r} \frac{\partial g_1}{\partial \theta} \right) \mathbf{k} \end{aligned}$$

5.3 Formule de Green

Soit $\mathcal{D} \subset \mathbf{R}^2$ le domaine d'intégration. On note Γ le bord de \mathcal{D} . On suppose ici que Γ est suffisamment lisse, c-à-d constitué d'une réunion finie de courbes de classe \mathcal{C}^1 . On oriente la frontière Γ de sorte que \mathcal{D} soit constamment sur la gauche. Ainsi, si P et Q sont deux fonctions de classe \mathcal{C}^1 sur un ouvert contenant \mathcal{D} , on a la formule de Green-Riemann :

$$\int_{\Gamma} [P(x, y) dx + Q(x, y) dy] = \int_{\mathcal{D}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy.$$

Bibliographie

- [1] J. Lelong-Ferrand et J.-M. Arnaudiès : *Cours de mathématiques - Tome 4* (chapitres 4 et 5), Dunod.
- [2] J.-M. Monier : *Analyse MP*, Dunod, 2004.
- [3] W. Rudin et G. Auliac : *Principes d'analyse mathématique*, Ediscience 1995.