26

Calcul différentiel et optimisation

Dans tout le chapitre, E et F désignent des $\mathbb{R}$-espaces vectoriels normés de dimension finie nulle, $\mathscr{U}$ désigne un ouvert de E.

I Différentielle

1 Différentielle en un point

Définition 1 : Application différentiable en un point

Soit f définie sur un ouvert $\mathscr{U}$ de E, à valeurs dans F. Soit a un point de $\mathscr{U}$. On dit que f est **différentiable** en a lorsqu'il existe une application linéaire ℓ_a de E dans F telle que, au voisinage de 0_E,

$$f(a+h) = f(a) + \ell_a(h) + \underset{h\to 0_E}{\mathrm{o}}(h)$$

ou encore, au voisinage de a,

$$f(x) = f(a) + \ell_a(x-a) + \underset{x\to a}{\mathrm{o}}(x-a).$$

Lorsqu'elle existe, l'application ℓ_a est unique et appelée **différentielle** de f au point a ou encore **application linéaire tangente** à f en a, notée $\mathrm{d}f(a) \in \mathscr{L}(E,F)$. On a donc

$$f(a+h) = f(a) + \mathrm{d}f(a)(h) + \underset{h\to 0_E}{\mathrm{o}}(h)$$

Propriété 1 : différentiable ⇒ continue

Si f est différentiable en a, alors f est continue en a.

2 Cas particuliers

Propriété 2 : Cas d'une fonction d'une variable réelle

Dans le cas d'une fonction $f : I \to F$, f est dérivable en a si et seulement si elle est différentiable en a.

Dans ce cas, $\mathrm{d}f(a) : h \mapsto hf'(a)$ et en particulier $f'(a) = \mathrm{d}f(a)(1)$.

Propriété 3 : Cas d'une fonction constante

Si $f : \mathscr{U} \subset E \to F$ est constante, elle est différentiable en tout point de $\mathscr{U}$ et pour tout $a \in \mathscr{U}$, $\mathrm{d}f(a) = 0_{\mathscr{L}(E,F)}$.

Propriété 4 : Cas d'une fonction linéaire

Si $\tilde{f} : \mathscr{U} \subset E \to F$ est la restriction à $\mathscr{U}$ d'une application linéaire $f \in \mathscr{L}(E,F)$, elle est différentiable en tout point de $\mathscr{U}$ et pour tout $a \in \mathscr{U}$, $\mathrm{d}\tilde{f}(a) = f$ (et donc $\mathrm{d}\tilde{f}$ est constante).

En particulier, f est différentiable sur E et $\forall a \in E$, $\mathrm{d}f(a) = f$.

3 Dérivée selon un vecteur, dérivées partielles

Définition 2 : Dérivée selon un vecteur

On dit que $f : \mathscr{U} \subset E \to F$ est **dérivable selon le vecteur** $v \in E$ **au point** $a \in \mathscr{U}$, lorsque $\phi : t \mapsto f(a+tv)$ est dérivable en 0.

On note alors $\mathrm{D}_v f(a) = \phi'(0) = \lim\limits_{t\to 0} \dfrac{f(a+tv) - f(a)}{t} \in F$.

Définition 3 : Dérivées partielles

Soient $\mathscr{B} = (e_1, \ldots, e_p)$ une base de E, $f : \mathscr{U} \subset E \to F$, $a \in \mathscr{U}$, $j \in [\![1,p]\!]$.

On appelle j^e **dérivée partielle de** f **en** a, lorsqu'elle existe, la dérivée de f selon le vecteur e_j de base en a :

$$\partial_j f(a) = \frac{\partial f}{\partial x_j}(a) = \mathrm{D}_{e_j} f(a) \in F.$$

4 Lien entre différentielle et dérivées partielles

Propriété 5 : Lien entre différentielle et dérivées partielles

Soit $f : \mathscr{U} \to F$, $a \in \mathscr{U}$. Si f est différentiable en a, alors f est dérivable en a selon tout vecteur $v \in E$ et $D_v f(a) = \mathrm{d}f(a)(v)$.

Cas particulier 1 : Dérivée selon un vecteur de base

Soit $f : \mathscr{U} \to F$, $a \in \mathscr{U}$, $\mathscr{B} = (e_1, \ldots, e_p)$ une base de E. On note $\dfrac{\partial f}{\partial x_j}$ les dérivées partielles de f dans la base $\mathscr{B}$. Si f est différentiable en a, alors pour tout $j \in [\![1,p]\!]$,

$$\mathrm{d}f(a)(e_j) = D_{e_j} f(a) = \frac{\partial f}{\partial x_j}(a).$$

Propriété 6 : Expression de la différentielle avec les dérivées partielles

Soit $\mathscr{B} = (e_1, \ldots, e_p)$ une base de E, $f : \mathscr{U} \to F$, $a \in \mathscr{U}$ tel que f est différentiable en a.

Alors pour tout vecteur $h = \sum\limits_{j=1}^{p} h_j e_j \in E$,

$$\mathrm{d}f(a)(h) = \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a)$$

On note $\mathrm{d}x_j : \left| \begin{array}{ccc} E & \longrightarrow & \mathbb{R} \\ h & \longmapsto & h_j \end{array} \right.$ *la forme linéaire j^e coordonnée dans $\mathscr{B}$. Alors on a*

$$\mathrm{d}f(a) = \sum_{j=1}^{p} \frac{\partial f}{\partial x_j}(a)\,\mathrm{d}x_j.$$

5 Matrice jacobienne

Définition 4 : Matrice jacobienne

Soit $p = \dim E$, $n = \dim F$, $\mathscr{U}$ ouvert de E et $f : E \to F$ différentiable, $\mathscr{B} = (e_1, \dots, e_p)$ une base de E et $\mathscr{C} = (\varepsilon_1, \dots, \varepsilon_n)$ une base de F. On note $f = \sum_{i=1}^{n} f_i \varepsilon_i$.

On appelle **matrice jacobienne** de f en a dans les bases $\mathscr{B}$ et $\mathscr{C}$ la matrice $J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]}$.

Propriété 7 : Matrice jacobienne et différentielle

$$J_f(a) = \mathrm{Mat}_{\mathscr{B},\mathscr{C}}(\mathrm{d}f(a))$$

6 Gradient

Définition 5 : Gradient

Soit E un espace **euclidien**, $\mathscr{U}$ un ouvert de E, $f : \mathscr{U} \to \mathbb{R}$ une fonction différentiable. Alors pour tout $a \in \mathscr{U}$, il existe un unique vecteur noté $\nabla f(a)$ ou $\overrightarrow{\mathrm{grad}}\, f(a)$ et appelé **gradient de f en a** tel que pour tout $h \in E$,

$$\mathrm{d}f(a)(h) = (\nabla f(a) | h).$$

Propriété 8 : Coordonnées du gradient

*Soit E un espace **euclidien**, $\mathscr{U}$ un ouvert de E, $f : \mathscr{U} \to \mathbb{R}$ une fonction différentiable. Si l'on fixe une base **orthonormée** de E, le coordonnées de $\nabla f(a)$ dans cette base sont*

$$\left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_p}(a)\right).$$

II Opérations sur les différentielles

1 Combinaisons linéaires

Propriété 9 : Linéarité

Soit $\mathscr{U}$ ouvert de E, $f, g : \mathscr{U} \to F$ différentiables en $a \in \mathscr{U}$, $\lambda, \mu \in \mathbb{R}$.

Alors $\lambda f + \mu g$ est différentiable et

$$\mathrm{d}(\lambda f + \mu g)(a) = \lambda\,\mathrm{d}f(a) + \mu\,\mathrm{d}g(a).$$

2 Image par une application multilinéaire

Propriété 10 : Image par une application multilinéaire

Soit $\mathscr{U}$ ouvert d'un $\mathbb{R}$-espace vectoriel normé de dimension finie, $E_1, \dots, E_q, F$ des $\mathbb{R}$-espaces vectoriels normés de dimension finie, $f_1 : \mathscr{U} \to E_1, \dots, f_q : \mathscr{U} \to E_q$ des applications différentiables en $a \in \mathscr{U}$ et $M : E_1 \times \cdots \times E_q \to F$ une fonction q-linéaire.

Alors $\phi : \mathscr{U} \to F$ définie par $\phi : x \mapsto M(f_1(x), \dots, f_q(x))$ est différentiable en a et

$$\mathrm{d}\phi(a) : h \mapsto \sum_{k=1}^{q} M\big(f_1(a), \dots, \mathrm{d}f_k(a)(h), \dots, f_q(a)\big).$$

Corollaire 1 : Différentielle d'un produit, d'un produit scalaire

- *Si f et g sont des fonctions différentiables sur $\mathscr{U}$ à valeurs réelles, alors $f \times g$ l'est et*

$$\mathrm{d}(f \times g)(a) : h \mapsto \mathrm{d}f(a)(h) \times g(a) + f(a) \times \mathrm{d}g(a)(h).$$

- *Si f et g sont des fonctions différentiables sur $\mathscr{U}$ à valeurs dans un espace euclidien $(E, (\cdot|\cdot))$, alors $(f|g)$ l'est et*

$$\mathrm{d}\left(f|g\right)(a) : h \mapsto \big(\mathrm{d}f(a)(h) \big| g(a)\big) + \big(f(a) \big| \mathrm{d}g(a)(h)\big).$$

3 Composition

Propriété 11 : Différentielle d'une composée

Soit $\mathscr{U}$ ouvert de E, $\mathscr{V}$ ouvert de F, $f : \mathscr{U} \to \mathscr{V}$ différentiable en $a \in \mathscr{U}$ et $g : \mathscr{V} \to G$ différentiable en $b = f(a)$.

Alors $g \circ f$ est différentiable en a et

$$\mathrm{d}(g \circ f)(a) = \mathrm{d}g(f(a)) \circ \mathrm{d}f(a).$$

Propriété 12 : Matrice jacobienne d'une composée

En munissant E, F et G de bases, on obtient

$$J_{g\circ f}(a) = J_g(f(a)) \times J_f(a).$$

Corollaire 2 : Dérivée le long d'un arc

Soit $\mathscr{U}$ ouvert de E, $f : \mathscr{U} \to F$, I intervalle de $\mathbb{R}$ et $\gamma : I \to \mathscr{U}$ une application dérivable.

On suppose que f est différentiable en $\gamma(t)$ où $t \in I$. Alors $f \circ \gamma$ est dérivable en t et

$$(f \circ \gamma)'(t) = \mathrm{d}f(\gamma(t))\left(\gamma'(t)\right).$$

En munissant E d'une base, on retrouve l'expression vue avec la règle de la chaîne, si les coordonnées de $\gamma(t)$ sont $(\gamma_1(t), \dots, \gamma_p(t))$,

$$(f \circ \gamma)'(t) = \sum_{k=1}^{p} \gamma_k'(t) \frac{\partial f}{\partial x_k}(\gamma(t)).$$

En particulier, si $\gamma : t \mapsto x + tv$ où $x, v \in \mathscr{U}$ sont fixés,

$$(f \circ \gamma)'(t) = \mathrm{d}f(\gamma(t))(v) = D_v f(\gamma(t)).$$

III CLASSE $\mathscr{C}^1$

Définition 6 : Fonctions de classe $\mathscr{C}^1$

$f : \mathscr{U} \to F$ est dite **de classe** $\mathscr{C}^1$ lorsque f est différentiable sur $\mathscr{U}$ et $\mathrm{d} f : \mathscr{U} \to \mathscr{L}(E,F)$ est continue.

Propriété 13 : Opérations

Toute combinaison linéaire, toute composée d'applications de classe $\mathscr{C}^1$ l'est encore.

Si M est q-linéaire et $f_1, \dots, f_q$ sont $\mathscr{C}^1$, $M(f_1, \dots, f_q)$ l'est.

Propriété 14 : Caractérisation par les applications coordonnées

Soit $f : \mathscr{U} \to F$, d'applications coordonnées $(f_1, \dots, f_n)$ dans une base de F. f est de classe $\mathscr{C}^1$ si et seulement si toutes les f_i le sont.

Théorème 1 : IMPORTANT - Caractérisation avec les dérivées partielles

f est de classe $\mathscr{C}^1$ si et seulement si, dans une base quelconque de E, toutes les dérivées partielles de f existent et sont continues.

Propriété 15 : Expression intégrale le long d'un arc

Soit $f \in \mathscr{C}^1(\mathscr{U}, F)$, $\gamma \in \mathscr{C}^1([0,1], \mathscr{U})$, $a = \gamma(0)$ et $b = \gamma(1)$. Alors

$$f(b) - f(a) = \int_0^1 \mathrm{d} f(\gamma(t))\big(\gamma'(t)\big)\,\mathrm{d}t.$$

Corollaire 3 : Cas particulier $\gamma(t) = a + tv$

En particulier, avec $\gamma(t) = a + tv$ sur $[0,1]$,

$$f(a+v) - f(a) = \int_0^1 \mathrm{d} f(a+tv)(v)\,\mathrm{d}t = \int_0^1 D_v f(a+tv)\,\mathrm{d}t.$$

Corollaire 4 : Caractérisation des fonctions constantes sur un ouvert connexe par arcs

*Si $f \in \mathscr{C}^1(\mathscr{U}, F)$ et $\mathscr{U}$ est un ouvert **connexe par arcs**, alors*

f est constante si et seulement si $\mathrm{d} f = 0$.

IV VECTEURS TANGENTS

Définition 7 : Vecteur tangent

Si X est une partie de E, $a \in X$. Un vecteur $v \in E$ est dit **tangent** à X en a lorsqu'il existe $\varepsilon > 0$ et un arc $\gamma :]-\varepsilon, \varepsilon[\to X$, dérivable en 0 tel que $\gamma(0) = a$ et $\gamma'(0) = v$.

On note $\mathrm{T}_a X$ l'ensemble des vecteurs tangents à X en a.

Propriété 16 : Hyperplan tangent à une partie décrite implicitement

Soit g une fonction numérique de classe $\mathscr{C}^1$ sur l'ouvert $\mathscr{U}$, $X = \{x \in \mathscr{U},\ g(x) = 0_{\mathbb{R}}\}$ et $a \in X$.

*Si $\mathrm{d} g(a) \neq 0_{\mathscr{L}(E,\mathbb{R})}$ (a est dit **régulier**, dénomination hors-programme), alors*

$$\mathrm{T}_a X = \mathrm{Ker}\big(\mathrm{d} g(a)\big) = \nabla g(a)^{\perp}.$$

(L'espace étant supposé euclidien pour utiliser le gradient.)

Autrement dit, X admet un hyperplan affine tangent en a qui est, $a + \mathrm{T}_a X = a + \nabla g(a)^{\perp}$ d'équation

$$\big(\nabla g(a) \big| x - a\big) = 0.$$

On utilisera en particulier ce résultat pour décrire les vecteurs tangents en un point d'une surface.

Une surface est en général décrite par une équation explicite $z = f(x,y)$ (attention le vecteur x de l'énoncé précédent correspond ici au vecteur $(x,y,z) \in \mathbb{R}^3$ dans lequel x désigne un nombre réel.)

On se ramène alors au cas implicite en posant $g : (x,y,z) \mapsto f(x,y) - z$. La surface est alors décrite par l'équation implicite $g(x,y,z) = 0$.

Propriété 17 : Plan tangent pour une surface explicite

Soit $\mathscr{U}$ ouvert de $\mathbb{R}^2$, $f : \mathscr{U} \to \mathbb{R}$ différentiable. Soit $S \subset \mathbb{R}^3$ la surface représentative de f, c'est-à-dire $S = \{(x,y,z) \in \mathbb{R}^3,\ (x,y) \in \mathscr{U} \text{ et } z = f(x,y)\}$. On définit

$$g : \left|\begin{array}{ccc} \mathscr{U} \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ (x,y,z) & \longmapsto & f(x,y) - z \end{array}\right.$$

L'ensemble $\mathrm{T}_a S$ des vecteurs tangents à S en $a = (x_0, y_0, z_0) \in S$ est un plan vectoriel $P = \nabla g(a)^{\perp}$ de vecteur normal le vecteur (non nul) $\nabla g(a)$, donc d'équation

$$\big(\nabla g(a) \big| (x,y,z)\big) = 0.$$

*On appelle **plan tangent à** S **en** a le plan affine $\mathscr{P} = a + P$ passant par a et de direction l'ensemble P des vecteurs tangents à X en a, d'équation*

$$\big(\nabla g(a) \big| (x - x_0, y - y_0, z - z_0)\big) = 0$$

Méthode 1 : Trouver une équation d'hyperplan tangent

On retiendra que dans tous les cas, mieux vaut repasser par une équation **implicite** $g(x) = 0$ pour trouver une équation d'hyperplan tangent.

Le théorème précédent donne alors une équation dudit hyperplan tangent en un point a n'annulant pas la $\mathrm{d}g$.

Par exemple, le plan tangent à la surface S d'équation $g(x, y, z) = 0$ en $a = (x_0, y_0, z_0)$ est le plan affine $a + \nabla g(a)^{\perp}$ passant par a et de direction le plan vectoriel $\mathrm{T}_a S = \nabla g(a)^{\perp}$ de vecteur normal $\nabla g(a) \neq (0,0,0)$, d'équation

$$\frac{\partial g}{\partial x}(x_0, y_0, z_0)(x - x_0) + \frac{\partial g}{\partial y}(x_0, y_0, z_0)(y - y_0) + \frac{\partial g}{\partial z}(x_0, y_0, z_0)(z - z_0) = 0$$

où $g(a) = g(x_0, y_0, z_0) = 0$.

V Optimisation (recherche d'extremums)

Définition 8 : Point critique

Soit $f : \mathscr{U} \to \mathbb{R}$ différentiable. On dit que a est un **point critique** de f lorsque $\mathrm{d} f(a) = 0$.
Si E est euclidien, cela équivaut à $\nabla f(a) = 0$.

Propriété 18 : Condition nécessaire d'extremum local

Soit X partie de E, $f : X \to \mathbb{R}$ et $a \in X$. On suppose

H1 ***Très important*** *: $a \in \overset{\circ}{X}$ est un **point intérieur** à X (c'est le cas par exemple si X est un **ouvert**)*

H2 *f différentiable en a,*

H3 *f présente un extremum local en a,*

alors a est un point critique de f, c'est-à-dire $\mathrm{d} f(a) = 0$.

Propriété 19 : Vecteurs tangents en un extremum local

Si f est une fonction numérique définie sur l'ouvert $\mathscr{U}$, X une partie de $\mathscr{U}$, $f|_X$ admet un extremum local en $a \in X$ et f est différentiable en a, alors

$$\forall\, v \in \mathrm{T}_a X, \quad \mathrm{d} f(a)(v) = 0.$$

Autrement dit, dans le cas où E est euclidien,

$$\mathrm{T}_a X \subset \mathrm{Ker}(\mathrm{d} f(a)) = \left(\nabla f(a)\right)^{\perp}$$

Théorème 2 : d'optimisation sous contrainte

Soient $f, g \in \mathscr{C}^1(\mathscr{U}, \mathbb{R})$ où $\mathscr{U}$ est un ouvert de E, et $X = \{x \in \mathscr{U},\ g(x) = 0\}$.
Si $f|_X$ admet un extremum local en $a \in X$ et si $\mathrm{d} g(a) \neq 0_E$, alors $\mathrm{d} f(a)$ est colinéaire à $\mathrm{d} g(a)$.

Théorème 3 : HP : multiplicateurs de Lagrange

*Soit $f, g_1, \dots, g_p$ des fonctions numériques de classe $\mathscr{C}^1$ sur l'**ouvert** $\mathscr{U}$ de E et*

$$X = \{x \in \mathscr{U},\ g_1(x) = \cdots = g_p(x) = 0\}.$$

Si $f|_X$ admet un extremum local en $a \in X$ et si les formes linéaires $\mathrm{d} g_1(a), \dots, \mathrm{d} g_p(a)$ sont linéaires indépendantes, alors il existe $\lambda_1, \dots, \lambda_p \in \mathbb{R}$ tels que

$$\mathrm{d} f(a) = \lambda_1 \mathrm{d} g_1(a) + \cdots + \lambda_p \mathrm{d} g_p(a).$$

*Les réels $\lambda_1, \dots, \lambda_p$ sont appelés **multiplicateurs de Lagrange**.*