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

Rappel : si I est intervalle de $\mathbb{R}$, $f : I \to F$ est dérivable en $a \in I$ si et seulement si f admet un développement limité à l'ordre 1 en a si et seulement si on a un vecteur $b \in F$ tel que

$$f(a+h) = f(a) + hb + \underset{h\to 0_{\mathbb{R}}}{\mathrm{o}}(h)$$

(et, dans ce cas, $b = f'(a)$).

Nous allons généraliser cette idée aux fonctions définies sur une espace vectoriel normé de dimension finie E.

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,

$$f(a+h) =$$

ou encore, au voisinage de a,

$$f(x) =$$

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

$$f(a+h) =$$

Remarque

R 1 – Lorsque cela a du sens, on définit donc une application $\mathrm{d}f : \left|\begin{array}{ccc} E & \longrightarrow & \mathscr{L}(E,F) \\ a & \longmapsto & \mathrm{d}f(a) \end{array}\right.$ appelée **différentielle** de f.

R 2 – On note parfois $\mathrm{d}f(a)\cdot h$ au lieu de $\mathrm{d}f(a)(h)$.

R 3 – L'unicité vient du

Lemme 1

Si $\varphi \in \mathscr{L}(E,F)$ est telle que $\varphi(h) = \underset{h\to 0}{\mathrm{o}}(h)$, alors φ est l'application nulle.

Démonstration

$\varphi(h) = \|h\|\,\varepsilon(h)$ avec $\varepsilon(h) \xrightarrow[h\to 0_E]{} 0_F$.

Alors pour tout $x \in E$, $t \in \mathbb{R}_*^+$, $t\varphi(x) = \varphi(tx) = t\|x\|\,\varepsilon(tx)$ donc $\varphi(x) = \|x\|\,\varepsilon(tx) \xrightarrow[t\to 0]{} 0_F$.

Donc pour tout $x \in E$, $\varphi(x) = 0_F$. ■

Propriété 1 : différentiable ⇒ continue

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

Démonstration

Par continuité de l'application linéaire en dimension finie (celle de E suffit) $\mathrm{d}f(a)$, en passant à la limite dans le développement limité, $f(a+h) - f(a) \xrightarrow[h\to 0_E]{} 0_F$. ■

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,

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 $a \in \mathscr{U}$,

Démonstration

$f(a+h) = f(a) + 0_F + 0_F$ avec $h \mapsto 0_F$ linéaire et $0_F = o(h)$.
Donc f est différentiable en a et $\mathrm{d}f(a) = 0_{\mathscr{L}(E,F)}$.

Propriété 4 : Cas d'une fonction linéaire

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

Démonstration

$f(a+h) = \varphi(a+h) = \varphi(a) + \varphi(h) + 0_F$ avec φ linéaire et $0_F = o(h)$.
Donc f est différentiable en a et $\mathrm{d}f(a) = \varphi$ (donc $\mathrm{d}f$ est constante.

Exercice 1 : Montrer que $f : \left| \begin{array}{ccc} \mathscr{M}_n(\mathbb{R}) & \longrightarrow & \mathscr{M}_n(\mathbb{R}) \\ M & \longmapsto & M^2 \end{array} \right.$ est différentiable en toute $A \in \mathscr{M}_n(\mathbb{R})$ et calculer $\mathrm{d}f(A)$.

Exercice 2 : Montrer que, si E est un espace euclidien $f : \left| \begin{array}{ccc} E & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \|x\|^2 \end{array} \right.$ est différentiable en toute $a \in E$ et calculer $\mathrm{d}f(a)$.

Exercice 3 : Montrer que, si E est un espace euclidien, $u \in \mathscr{L}(E)$, $f : \left| \begin{array}{ccc} E & \longrightarrow & \mathbb{R} \\ x & \longmapsto & (x|u(x)) \end{array} \right.$ est différentiable en toute $a \in E$ et calculer $\mathrm{d}f(a)$. Que se passe-t-il si, de plus, u est symétrique ?

3 Dérivée selon un vecteur, dérivées partielles

Le fait de travailler sur un ouvert $\mathscr{U}$ assure, pour a point de $\mathscr{U}$ et v vecteur de E, l'existence d'un $\delta > 0$ (« distance de sécurité ») tel que pour tout $t \in]-\delta, \delta[$, $a + tv \in \mathscr{U}$ (on s'éloigne de a dans la direction de v), cela permet de définir l'application $\phi : t \mapsto f(a+tv)$ d'une variable réelle au voisinage de 0.

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

On note alors

Exemple

E 1 – La fonction $f : (x,y) \in \mathbb{R}^2 \mapsto \dfrac{xy^2}{x^2+y^4}$ si $(x,y) \neq (0,0)$ et $(0,0)$ sinon admet des dérivées selon tout vecteur en $(0,0)$ et si $v = (\alpha, \beta)$: $\mathrm{D}_v f((0,0)) =$

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 :

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

Remarque

R 4 – Les dérivées partielles vues pour les fonctions définies sur un ouvert de $\mathbb{R}^p$ sont les dérivées partielles dans la base canonique de $\mathbb{R}^p$.
Bien sûr la notion générale de dérivée partielle dépend de la base choisie. En ce sens, la notation $\mathrm{D}_{e_j} f(a)$ est la plus précise.

R 5 – Comme on en a déjà l'habitude, calculer la j^e dérivée partielle revient à fixer les coordonnées selon tous les autres vecteurs de bases et dériver par rapport à la seule variable x_j. : en effet, en dérive

$$\phi : t \mapsto f(a + te_j) = f(a_1 e_1 + \cdots + (a_j + t)e_j + \cdots + a_n e_n)$$

en 0, ce qui revient aussi à dériver

$$t \mapsto f(a_1 e_1 + \cdots + t e_j + \cdots + a_n e_n)$$

en a_j.

Exemple

E 2 – Avec $f : (x,y) \in \mathbb{R}^2 \mapsto \dfrac{xy}{x^2+y^2}$ si $(x,y) \neq (0,0)$ et $(0,0)$ sinon, on voit que ⚠ on peut avoir des dérivées partielles en $(0,0)$ sans avoir de dérivée selon certains vecteurs ($v = (\alpha,\beta) \in \mathbb{R}^2_*$, ici).

4 Lien entre différentielle et dérivées partielles

Propriété 5 : Lien entre différentielle et dérivées partielles

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

Démonstration

$f(a+tv) = f(a) + \mathrm{d}f(a)(tv) + \underset{t\to 0}{\mathrm{o}}(t)$ donc $\dfrac{f(a+tv)-f(a)}{t} \xrightarrow[t\to 0]{} \mathrm{d}f(a)(v)$ par linéarité de $\mathrm{d}f(a)$. ■

Cas particulier 1 : Dérivée selon un vecteur de base

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

$$\mathrm{d}f(a)(e_j) =$$

Propriété 6 : Expression de la différentielle avec les dérivées partielles

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

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

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

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 $\mathcal{B}$. Alors on a

$$\mathrm{d}f(a) =$$

Démonstration

C'est la linéarité de $\mathrm{d}f(a)$:

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

■

Remarque

R 6 – Le DL_1 de f différentiable en a s'écrit alors

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

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, $\mathscr{B} = (e_1, \ldots, e_p)$ une base de E et $\mathscr{C} = (\varepsilon_1, \ldots, \varepsilon_n)$ 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

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

Démonstration

$$\mathrm{d}f(a)(e_j) = \frac{\partial f}{\partial x_j}(a) = \sum_{i=1}^{p} \frac{\partial f_i}{\partial x_j}(a)\varepsilon_i$$

permet bien de remplir la j^{e} colonne de $J_f(a)$. ■

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 $h \in E$,

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

Démonstration

C'est le théorème de représentation de Riesz : la **forme** linéaire $\mathrm{d}f(a)$ s'écrit $x \mapsto (b|x)$ pour un certain vecteur b... ■

Remarque

R 7 – On comprend mieux la notation alternative $\mathrm{d}f(a)\cdot h$ faisant penser à un produit scalaire. Le DL$_1$ de f différentiable en a dans E euclidien se récrit

$$f(a+h) = f(a) + \underbrace{(\nabla f(a)|h)}_{\text{produit scalaire}} + o(h)$$

R 8 – Dans $\mathbb{R}^2$ muni du produit scalaire canonique,

$$\mathrm{d}f(a)(h) = h_1\frac{\partial f}{\partial x}(a) + h_2\frac{\partial f}{\partial y}(a) = \begin{pmatrix}\frac{\partial f}{\partial x}(a)\\ \frac{\partial f}{\partial y}(a)\end{pmatrix}\cdot\begin{pmatrix}h_1\\ h_2\end{pmatrix}$$

On retrouve notre gradient habituel.
Exactement le même raisonnement (expression du produit scalaire en base orthonormée) donne la propriété suivante.

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*

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

Remarque : Interprétation géométrique

R 9 – On suppose f différentiable en a et $\nabla f(a) \neq 0_E$.
Notons S la sphère unité de E pour la norme euclidienne $\|\cdot\|$. D'après l'inégalité de Cauchy-Schwarz

$$\forall h \in S,\quad \mathrm{d}f(a)(h) = (\nabla f(a)|h) \leqslant \left|(\nabla f(a)|h)\right| \leqslant \|\nabla f(a)\|\,\|h\| = \|\nabla f(a)\|.$$

Or ce majorant est atteint pour $h_0 = \dfrac{\nabla f(a)}{\|\nabla f(a)\|}$.
Ainsi le vecteur $\nabla f(a)$ donne la direction et le sens de plus forte variation de la fonction f.
C'est à la base d'algorithmes d'optimisation (descente de gradient), utilisé par exemple en Machine Learning (apprentissage automatique).
Ci-après, on représente, au cœur de Mafate, des « lignes de niveau » de la fonction $(x,y) \mapsto f(x,y)$ qui au point de coordonnées (x,y) associe son altitude (on parle d'isoplèthes d'altitude).
Ces lignes de niveau sont des courbes d'équation $f(x,y) =$ constante.
Tracer la direction et le sens de ∇f en quelques points.
On peut montrer que le gradient est normal aux lignes de niveaux en tous points.

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

Démonstration

Il suffit de remarquer, par propriétés de la dérivation que $J_{\lambda f+\mu g}(a) = \lambda J_f(a) + \mu J_g(a)$ et de repasser aux applications linéaires.

Autre possibilité : on ajoute les DL$_1$:

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

avec $\lambda \,\mathrm{d} f(a) + \mu \,\mathrm{d} g(a)$ linéaire, donc par unicité, $\lambda f + \mu g$ est différentiable et

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

Remarque

R 10 – $f \mapsto \mathrm{d} f$ est donc linéaire.

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

$$\mathrm{d}\phi(a) : h \mapsto$$

Démonstration

On traite le cas bilinéaire. Le cas général est analogue.

1re tentative On peut voir le faire avec les dérivées selon les vecteurs. On sait, sous réserve d'existence, que $\mathrm{d}\phi(a)(h) = D_h \phi(a)$. Or

$$\psi : t \mapsto \phi(a+th) = B(f(a+th), g(a+th))$$

est dérivable en 0 (car f et g sont différentiables en a) de dérivée

$$\begin{aligned}\psi'(0) = D_h\phi(a) &= B(D_h f(a), g(a)) + B(f(a), D_h g(a))\\ &= B(d f(a)(h), g(a)) + B(f(a), d g(a)(h))\end{aligned}$$

Le problème est que l'existence de dérivée selon tout vecteur ne garantit pas la différentiabilité...

On peut par exemple montrer que la fonction $f : (x, y) \mapsto \begin{cases} \dfrac{y^2}{x} & \text{si } x \neq 0 \\ y & \text{sinon} \end{cases}$ admet des dérivées selon tout vecteur en $(0,0)$ mais n'est pas continue , et encore moins différentiable.

Bonne méthode On utilise les DL$_1$: avec $\varepsilon_1(h) \xrightarrow[h\to 0]{} 0$ et $\varepsilon_2(h) \xrightarrow[h\to 0]{} 0$,

$$\begin{aligned}\phi(a+h) &= B(f(a+h), g(a+h))\\ &= B\big(f(a) + \mathrm{d} f(a)(h) + \|h\|\,\varepsilon_1(h),\ \ g(a) + \mathrm{d} g(a)(h) + \|h\|\,\varepsilon_2(h)\big)\\ &= B\big(f(a),\ g(a)\big) + B\big(\mathrm{d} f(a)(h),\ g(a)\big) + B\big(f(a),\ \mathrm{d} g(a)(h)\big) + \psi(h)\end{aligned}$$

avec $h \mapsto B\big(\mathrm{d} f(a)(h), g(a)\big) + B\big(f(a),\ \mathrm{d} g(a)(h)\big)$ linéaire et

$$\begin{aligned}\psi(h) = \|h\|\, B\big(f(a) + \mathrm{d} f(a)(h),\ \varepsilon_2(h)\big) + \|h\|\, B\big(\varepsilon_1(h),\ g(a) + \mathrm{d} g(a)(h)\big)\\ + \|h\|^2 B\,(\varepsilon_1(h),\ \varepsilon_2(h)) + B\big(\mathrm{d} f(a)(h),\ \mathrm{d} g(a)(h)\big)\end{aligned}$$

donc

$$\frac{\psi(h)}{\|h\|} = B\big(f(a)+\mathrm{d}f(a)(h),\ \varepsilon_2(h)\big) + B\big(\varepsilon_1(h),\ g(a)+\mathrm{d}g(a)(h)\big)$$
$$+\|h\|\,B(\varepsilon_1(h),\ \varepsilon_2(h)) + \frac{B\big(\mathrm{d}f(a)(h),\ \mathrm{d}g(a)(h)\big)}{\|h\|}.$$

Les trois premiers termes tendent vers 0 lorsque $h \to 0$ par continuité de l'application bilinéaire en dimension finie.
Pour le dernier, on remarque que

$$(x,y) \mapsto B(\mathrm{d}f(a)(x), \mathrm{d}g(a)(y))$$

est bilinéaire, donc on a, par continuité, un réel C tel que

$$\left\|\frac{B\big(\mathrm{d}f(a)(h),\ \mathrm{d}g(a)(h)\big)}{\|h\|}\right\| = \frac{\big\|B\big(\mathrm{d}f(a)(h),\ \mathrm{d}g(a)(h)\big)\big\|}{\|h\|}$$
$$\leqslant \frac{C\|h\|^2}{\|h\|} = C\|h\| \xrightarrow[h\to 0]{} 0.$$

Donc $\psi(h) = \underset{h\to 0}{\mathrm{o}}(h)$, ce qui permet d'obtenir toute la conclusion.

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

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

Propriété 12 : Matrice jacobienne d'une composée

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

$$J_{g\circ f}(a) =$$

Remarque

R 11 – D'où la règle de la chaîne.

Démonstration

Il suffit de composer les DL$_1$: on écrit, avec $\varepsilon_1(h) \xrightarrow[h\to 0]{} 0$ et $\varepsilon_2(k) \xrightarrow[k\to 0]{} 0$, et pour simplifier la lecture , les applications linéaires $\varphi = \mathrm{d}f(a)$ et $\psi = \mathrm{d}g(f(a))$,

$$f(a+h) = f(a) + \varphi(h) + \|h\|\,\varepsilon_1(h)$$

et

$$g(f(a)+k) = g(f(a)) + \psi(k) + \|k\|\,\varepsilon_2(k)$$

puis

$$\begin{aligned} g(f(a+h)) &= g\Big[f(a) + \underbrace{\big(\varphi(h)+\|h\|\,\varepsilon_1(h)\big)}_{k}\Big] \\ &= g(f(a)) + \psi\big(\varphi(h)+\|h\|\,\varepsilon_1(h)\big) \\ &\quad + \big\|\varphi(h)+\|h\|\,\varepsilon_1(h)\big\|\,\varepsilon_2\big(\varphi(h)+\|h\|\,\varepsilon_1(h)\big) \\ &= g\circ f(a) + [\psi\circ\varphi](h) + \zeta(h) \end{aligned}$$

avec

$$\frac{\zeta(h)}{\|h\|} = \psi(\varepsilon_1(h)) + \frac{\big\|\varphi(h)+\|h\|\,\varepsilon_1(h)\big\|}{\|h\|}\,\varepsilon_2\big(\varphi(h)+\|h\|\,\varepsilon_1(h)\big).$$

φ et ψ étant linéaires en dimension finie, elles sont continues, donc $\psi(\varepsilon_1(h)) \xrightarrow[h\to 0]{} 0$, $\varphi(h) \xrightarrow[h\to 0]{} 0$ puis $\varepsilon_2\big(\varphi(h)+\|h\|\,\varepsilon_1(h)\big) \xrightarrow[h\to 0]{} 0$.
Enfin,

$$\frac{\big\|\varphi(h)+\|h\|\,\varepsilon_1(h)\big\|}{\|h\|} \leqslant \frac{\|\varphi(h)\|}{\|h\|} + \varepsilon_1(h)$$

et par continuité de l'application linéaire φ, on a $C \in \mathbb{R}$ tel que pour tout h, $\|\varphi(h)\| \leqslant C\|h\|$ donc $\frac{\|\varphi(h)+\|h\|\,\varepsilon_1(h)\|}{\|h\|}$ est borné et finalement $\frac{\zeta(h)}{\|h\|} \xrightarrow[h\to 0]{} 0$, c'est-à-dire $\zeta(h) = \mathrm{o}(h)$.

Par unicité, on en déduit que $g\circ f$ est différentiable en a et

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

Corollaire 1 : 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

$$(f\circ\gamma)'(t) =$$

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

$$(f\circ\gamma)'(t) =$$

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

$$(f\circ\gamma)'(t) =$$

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.

Démonstration

Non exigible. ■

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

$$f(b) - f(a) =$$

Démonstration

On a vu que $\mathrm{d}f(\gamma(t))(\gamma'(t)) = (f\circ\gamma)'(t)$, c'est donc une simple application du théorème fondamental de l'analyse. ■

Corollaire 2 : Cas particulier $\gamma(t) = a + tv$

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

$$f(a+v) - f(a) =$$

Corollaire 3 : 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 $\mathrm{d}f = 0$.

Démonstration

Le sens direct a déjà été vu.

Pour le sens réciproque, seul le cas convexe est au programme. On suppose donc $\mathscr{U}$ convexe et $\mathrm{d}f = 0$.

Soient $a, b \in \mathscr{U}$, on veut montrer que $f(a) = f(b)$.

Mais on peut relier a et b par un segment inclus dans $\mathscr{U}$ par convexité : $\gamma : t \in [0,1] \mapsto (1-t)a + tb$ de classe $\mathscr{C}^1$.

Alors, la propriété précédente donne $f(b) - f(a) = 0$. ■

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 $T_a X$ l'ensemble des vecteurs tangents à X en a.

Exemple

E 3 – **Si $\mathscr{G}$ est un sous-espace affine de E de direction G (sous-espace vectoriel de E), pour tout $a \in \mathscr{G}$,**

$$\boxed{T_a \mathscr{G} = G}$$

En effet, on peut écrire $\mathscr{G} = a + G$.

Si $v \in G$, soit $\gamma : \begin{vmatrix}]-1,1[& \longrightarrow & \mathscr{G} \\ t & \longmapsto & a + tv \end{vmatrix}$ est dérivable en 0, $\gamma(0) = a$ et $\gamma'(0) = v$.

Réciproquement, si $v \in T_a\mathscr{G}$, on a $\varepsilon > 0$ et $\gamma :]-\varepsilon, \varepsilon[\to \mathscr{G}$ tel que $\gamma(0) = a$ et $\gamma'(0) = v$.

Alors v est la limite lorsque $t \to 0$ de $\dfrac{\gamma(t) - a}{t} \in G$ et comme G est un sous-espace de E qui est de dimension finie, il l'est aussi, donc il est fermé et donc $v \in G$.

E 4 – **Sphère d'un espace euclidien : si $a \in S(0_E, 1)$,**

$$\boxed{T_a S(0_E, 1) = a^{\perp}}$$

En effet, si $v \in a^{\perp}$, soit $\gamma : \begin{vmatrix}]-1,1[& \longrightarrow & S(0_E,1) \\ t & \longmapsto & \dfrac{a+tv}{\|a+tv\|} \end{vmatrix}$ est dérivable en 0, $\gamma(0) = a$ et $\gamma'(0) = v$

car, avec $\|a\| = 1$,

$$\gamma : t \mapsto \frac{a+tv}{\sqrt{\|a+tv\|^2}} = \frac{a+tv}{\sqrt{1 + 2t(a|v) + t^2\|v\|^2}}$$

donc

$$\gamma'(0) = \frac{\|a + 0\cdot v\|\, v - \frac{2(a|v) + 2\cdot 0 \cdot \|v\|^2}{2\|a+0\cdot v\|}(a + 0 \cdot v)}{\|a + 0\cdot v\|^2}$$

Réciproquement, si $v \in T_a S(0_E, 1)$, on a $\varepsilon > 0$ et $\gamma :]-\varepsilon, \varepsilon[\to S(0_E, 1)$ tel que $\gamma(0) = a$ et $\gamma'(0) = v$.

Alors $\dfrac{\gamma(t) - a}{t} \xrightarrow[t \to 0]{} v$ donc

$$\left(\frac{\gamma(t)-a}{t}\middle| a\right) = \frac{(\gamma(t)|a) - 1}{t} = \frac{(\gamma(t)|a - \gamma(t))}{t}$$
$$= -\left(\gamma(t)\middle|\frac{\gamma(t)-a}{t}\right) \xrightarrow[t\to 0]{} (v|a) = -(a|v)$$

Donc $(v|a) = 0$: $v \in a^{\perp}$.

Comme dernier exemple, on va obtenir que le plan tangent à une surface d'équation $z = f(x,y)$ admet comme vecteur normal le gradient (non nul) de $g : (x,y,z) \mapsto f(x,y) - z$.

Cela revient à transformer l'équation explicite $z = f(x,y)$ en équation implicite $g(x,y,z) = 0$.

Propriété 16 : 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

$$S = \left\{(x,y,z) \in \mathbb{R}^3,\ (x,y) \in \mathscr{U} \text{ et } z = f(x,y)\right\}.$$

Soit $g : \begin{vmatrix} \mathscr{U} \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ (x,y,z) & \longmapsto & f(x,y) - z \end{vmatrix}$

L'ensemble $T_a S$ 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 $a + P$ passant par a et de direction l'ensemble P des vecteurs tangents à X en a, d'équation*

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

Remarque

R 12 – On retrouve, pour le plan tangent, une équation type « équation de la tangente » :

Démonstration

On raisonne toujours par double inclusion.

- Si $v = (x,y,z) \in T_a S$, on a $\varepsilon > 0$ et $\gamma :]-\varepsilon, \varepsilon[\to S$ tel que $\gamma(0) = a$ et $\gamma'(0) = v$. Alors, pour tout $t \in]-\varepsilon, \varepsilon[$, $\gamma(t) = (x(t), y(t), z(t))$ avec $z(t) = f(x(t), y(t))$ donc

$$z = z'(0) = x'(0)\frac{\partial f}{\partial x}(x(0), y(0)) + y'(0)\frac{\partial f}{\partial y}(x(0), y(0))$$
$$= x\frac{\partial f}{\partial x}(x_0, y_0) + y\frac{\partial f}{\partial y}(x_0, y_0)$$

donc

$$0 = x\frac{\partial f}{\partial x}(x_0,y_0)y\frac{\partial f}{\partial y}(x_0,y_0) - z = \left(\nabla g(a)\middle|(x,y,z)\right)$$

car $\nabla g(a) = \left(\frac{\partial f}{\partial x}(x_0,y_0), \frac{\partial f}{\partial y}(x_0,y_0), -1\right)$.

- Si $v = (x,y,z)$ est orthogonal à $\nabla g(a)$, On définit

$$\gamma : \left| \begin{array}{ccc}]-1,1[& \longrightarrow & S \\ t & \longmapsto & (x_0+tx, y_0+ty, f(x_0+tx, y_0+ty)) \end{array} \right.$$

Alors γ est dérivable en 0, $\gamma(0) = a$ et $\gamma'(0) = \left(x, y, x\frac{\partial f}{\partial x}(x_0,y_0) + y\frac{\partial f}{\partial y}(x_0,y_0)\right) = (x,y,z)$ car $(\nabla g(a)|v) = 0$.

Mais toutes les surfaces n'ont pas une équation explicite $z = f(x,y)$.

Le programme propose un dernier énoncé, à la démonstration hors programme, permettant de traiter le cas général des surfaces décrites par une équation implicite $g(x,y,z) = 0$ et généralisant l'énoncé précédent.

Propriété 17 : Hyperplan tangent à une surface implicite

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 $\mathrm{d}\,g(a) \neq 0_{\mathscr{L}(E,\mathbb{R})}$, alors

$$T_a X = \mathrm{Ker}\left(\mathrm{d}\,g(a)\right) = \nabla g(a)^{\perp}$$

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

Autrement dit, X admet un hyperplan affine tangent en a qui est, $a + T_a X = a + \nabla g(a)^{\perp}$ d'équation

$$\left(\nabla g(a)\middle| x - a\right) = 0.$$

Remarque

R 13 – Bien sûr le **vecteur** x de cet énoncé ne doit pas être confondu avec le **scalaire** x d'un triplet (x,y,z) du cas particulier $\mathbb{R}^3$.

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 du-dit 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 $T_a S = \nabla g(a)^{\perp}$ 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 $g(a) = g(x_0,y_0,z_0) = 0$.

Exemple

E 5 – Équation du plan tangent en un point d'une sphère de $\mathbb{R}^3$

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 $\mathscr{U}$ **ouvert** (très important) et $f : \mathscr{U} \to \mathbb{R}$ différentiable en $a \in \mathscr{U}$.*

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

Démonstration

Il suffit de revenir aux dérivées partielles.

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

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

Autrement dit, dans le cas où E est euclidien,

$$T_a X \subset \operatorname{Ker}(\mathrm{d}\, f(a)) = \left(\nabla f(a)\right)^{\perp}$$

Démonstration

Soit $v \in T_a X$. On a $\gamma :]-\varepsilon, \varepsilon[\to X$ tel que $\gamma(0) = a$ et $\gamma'(0) = v$.
Alors, $f \circ \gamma$ admet un extremum local en $0 \in]-\varepsilon, \varepsilon[$ qui est ouvert, donc

$$0 = (f \circ \gamma)'(0) = \mathrm{d}\, f(\gamma(0))\left(\gamma'(0)\right) = \mathrm{d}\, f(a)(v).$$

■

Remarque

R 14 – Autrement dit, les vecteurs tangents à X en a extremum local de $f_{|X}$ sont dans l'hyperplan vectoriel de vecteur normal $\nabla f(a)$.

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 $\mathrm{d}\, g(a)$.

Démonstration

- Soit $\mathrm{d}\, f(a)$ est nulle et $\mathrm{d}\, f(a) = 0 \cdot \mathrm{d}\, g(a)$.
- Soit $\mathrm{d}\, f(a)$ n'est pas nulle et on prouve que $\operatorname{Ker}(\mathrm{d}\, f(a)) = \operatorname{Ker}(\mathrm{d}\, g(a))$. En effet, un résultat du programme dit que si deux formes linéaires non nulles ont même noyau, elles sont colinéaires.
  Comme $\mathrm{d}\, g(a) \neq 0$, et avec la propriété précédente, $\operatorname{Ker}(\mathrm{d}\, g(a)) = T_a X \subset \operatorname{Ker}(\mathrm{d}\, f(a))$.
  Par égalité des dimensions, $\operatorname{Ker} \mathrm{d}\, f(a) = \operatorname{Ker} \mathrm{d}\, g(a)$, ce qui conclut.

■

Terminons avec une extension hors programme, pour la culture.

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*

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

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**.*