Calcul différentiel dans $\mathbb{R}^n$

Xavier Servot *

Mars 2019

1 Introduction

1.1 Différentiabilité

Intuitivement, une fonction va être différentiable en $\mathbf{a}$ si elle peut être bien approximée par une fonction linéaire dans un voisinage de $\mathbf{a}$. Géometriquement, cela veut dire que le graphe de f pourra s'apparenter à un plan au voisinage de $\mathbf{a}$.

Définition 1.1. Soit $f\colon E \to \mathbb{R}$, $\mathbf{a} \in E$. f est *différentiable* en $\mathbf{a}$ si il existe une application linéaire $L\colon \mathbb{R}^n \to \mathbb{R}$ tel que

$$f(\mathbf{x}) = f(\mathbf{a}) + L(\mathbf{x}-\mathbf{a}) + r(\mathbf{x})$$

où r est appelée *fonction d'erreur* et est telle que

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{r(\mathbf{x})}{||\mathbf{x}-\mathbf{a}||} = 0$$

On peut généraliser cette notion:

Définition 1.2. Soit $\mathbf{f}\colon E \to \mathbb{R}^p$. $\mathbf{f}$ est différentiable en $\mathbf{a}$ si il existe une application linéaire $\mathbf{L}\colon \mathbb{R}^n \to \mathbb{R}^p$ tel que

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + \mathbf{L}(\mathbf{x}-\mathbf{a}) + \mathbf{r}(\mathbf{x})$$

Cette fois-ci la fonction d'erreur vérifie

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{||\mathbf{r}(\mathbf{x})||}{||\mathbf{x}-\mathbf{a}||} = 0$$

*EPFL

Définition 1.3. Soit $\mathbf{f}$ différentiable. La *matrice jacobienne* de $\mathbf{f}$ est la matrice de l'application linéaire $\mathbf{L}$. On note $\mathbf{L}(\mathbf{x}) = D\mathbf{f}(\mathbf{a})\mathbf{x}$.

On va s'intéresser aux composantes de cette matrice $Df(\mathbf{a})$:

1.2 Dérivée partielle

Définition 1.4. Soit $f\colon E \to \mathbb{R}$ et $\mathbf{a} = (a_i)_{i=1}^n$. Soit $\forall k \in \{1, \dots, n\}$

$$f_k(x) := f(a_1, \dots, a_{k-1}, x, a_{k+1}, \dots, a_n)$$

Si f_k possède une dérivée en a_k, on dit que f admet une *dérivée partielle* par rapport à x_k en $\mathbf{a}$, noté

$$\frac{\partial f}{\partial x_k}(\mathbf{a}) := f_k'(a_k)$$

Alors la dérivée partielle de f par rapport à x_k en $\mathbf{a}$ peut se réecrire:

$$\lim_{x \to a_k} \frac{f(a_1, \dots, a_{k-1}, x, a_{k+1}, \dots, a_n) - f(a_1, \dots, a_{k-1}, a_k, a_{k+1}, \dots, a_n)}{x - a_k}$$

Définition 1.5. Soit $f\colon E \to \mathbb{R}$ dont la dérivée partielle par rapport à x_k existe en tout point de E, alors

$$\frac{\partial f}{\partial x_k}\colon E \to \mathbb{R}$$

est la *dérivée partielle* de f selon x_k

Définition 1.6. Soit $f\colon E \to \mathbb{R}$ dont les n dérivées partielles existent. Alors le *gradient* de f est l'application

$$\nabla f\colon \begin{array}{l} E \to \mathbb{R}^n \\ \mathbf{x} \to (\frac{\partial f}{\partial x_i}(\mathbf{x}))_{i=1}^n \end{array}$$

Pour la proposition suivant, $\mathbf{e}_k$ désigne le k-ième vecteur unitaire de $\mathbb{R}^n$.

Proposition 1.1. Soit $f\colon E \to \mathbb{R}$ différentiable en $\mathbf{a}$ et telle que $Df(\mathbf{a}) = (d_1, \dots, d_n)$. Alors les dérivées partielles de f existent en $\mathbf{a}$ et

$$d_i = \frac{\partial f}{\partial x_i}(\mathbf{a})$$

Preuve. Soit $k \in \{1, \ldots, n\}$. On va s'intéresser à la k-ième dérivée partielle. Soit

$$\mathbf{b}(y) =: (a_1, \ldots, a_{k-1}, y, a_{k+1}, \ldots, a_n)$$

La dérivée partielle se réecrit alors:

$$\frac{\partial f}{\partial x_k} = \lim_{x \to a_k} \frac{f(\mathbf{b}(x)) - f(\mathbf{b}(a_k))}{x - a_k}$$

Or

$$f(\mathbf{b}(y)) = f(\mathbf{a}) + (d_1, \ldots, d_n)(\mathbf{b}(y) - \mathbf{b}(a_k)) + r(\mathbf{b}(y))$$

qui peut se réecrire

$$f(\mathbf{b}(x)) = f(\mathbf{a}) + (d_1, \ldots, d_n)\mathbf{e}_k(x - a_k) + r(\mathbf{b}(x))$$

Et donc en particulier

$$f(\mathbf{b}(a_k)) = f(\mathbf{a}) + r(\mathbf{b}(a_k))$$

On a donc les égalités suivantes:

$$\begin{aligned}
\frac{f(\mathbf{b}(x)) - f(\mathbf{b}(a_k))}{x - a_k} &= \frac{(d_1, \ldots, d_n)\mathbf{e}_k(x - a_k) + r(\mathbf{b}(x)) - r(\mathbf{b}(a_k))}{x - a_k} \\
&= (d_1, \ldots, d_n)\mathbf{e}_k + \frac{r(\mathbf{b}(x) - r(\mathbf{b}(a_k)}{x - a_k} \\
&= d_k + \frac{r(\mathbf{b}(x)) - r(\mathbf{b}(a_k))}{x - a_k}
\end{aligned}$$

En prenant la limite, et en prenant en compte que les fonctions d'erreurs sont $o(||\mathbf{x} - \mathbf{a}||)$:

$$\begin{aligned}
\frac{\partial f}{\partial x_k} &= d_k + \lim_{x \to a_k} \frac{r(\mathbf{b}(x))}{||\mathbf{b}(x) - \mathbf{b}(a_k)||} - \frac{r(\mathbf{b}(a_k))}{||\mathbf{b}(x) - \mathbf{b}(a_k)||} \\
&= d_k
\end{aligned}$$

□

Ainsi, dans le cas d'une fonction différentiable, le gradient en un point est simplement la matrice jacobienne en ce point. Mais ce n'est pas parce que le gradient existe, qu'une fonction est différentiable. On va d'abord prouver cette proposition pour le montrer.

Proposition 1.2. Soit $f\colon E \to \mathbb{R}$ différentiable en $\mathbf{a} \in E$. Alors f est continue en $\mathbf{a}$.

Preuve. Par différentiabilité:

$$f(\mathbf{x}) = f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{x} - \mathbf{a}) + r(\mathbf{x})$$

Avec $Df(\mathbf{a}) = (d_1, \ldots, d_n)$. Soit $\varepsilon > 0$, La fonction d'erreur est en $o(\mathbf{a})$ donc

$$\exists \delta_r \text{ tel que } \forall \mathbf{x} \in B(\mathbf{a}, \delta_r), r(\mathbf{x}) \in]-\varepsilon/2; \varepsilon/2[$$

Soit $\delta_i := \frac{\varepsilon}{2nd_i}$. Alors en particulier pour $\delta_D = \min_i \delta_i$:

$$||\mathbf{x} - \mathbf{a}|| < \delta_D \Rightarrow |x_i - a_i| < \frac{\varepsilon}{2nd_i}$$

et donc

$$|Df(\mathbf{a})(\mathbf{x} - \mathbf{a})| < \frac{\varepsilon}{2}$$

Et

$$|f(\mathbf{x}) - f(\mathbf{a})| = |Df(\mathbf{a})(\mathbf{x} - \mathbf{a}) + r(\mathbf{x})| < \varepsilon$$

□

Donc pour en revenir au fait que l'existance du gradient n'implique pas la différentiabilité: on considère

$$f(x_1, x_2) = \frac{x_1 x_2}{x_1^2 + x_2^2}, \forall (x_1, x_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$$

Et $f(0,0) = 0$. On vérifie que les dérivées partielles existent et qu'elle n'est pas continue:

- Pour la continuité: on a en $f(0,0) = 0$. Or $f(x,x) = \frac{x^2}{2x^2} = \frac{1}{2}$, donc en particulier pour $x \to 0$, $f(x,x) = \frac{1}{2} \neq 0$.
- Pour l'existance des dérivées partielles, on vérifie que

  $$g_\alpha(x) = f(x, \alpha)$$

  et dérivable sur $\mathbb{R}$ et pour tout paramètre $\alpha \in \mathbb{R}$:

  $$g_\alpha(x) = \frac{x\alpha}{x^2 + \alpha^2}$$

On a que la dérivée est bien définie en 0:

$$g'_\alpha(x) = \alpha\frac{\alpha^2 - x^2}{\alpha^2 + y^2}$$

Donc

$$\frac{\partial f}{\partial x_1}(x, y) = y\frac{y^2 - x^2}{x^2 + y^2}$$

Or la proposition précédente indique que toute fonction différentiable est continue. Ici la fonction est non continue donc ne peut être différentiable et pourtant, ses dérivées partielles existent.

Définition 1.7. $f\colon E \to \mathbb{R}$ est *continument différentiable* sur un ouvert non vide E si les dérivées partielles existent en tout point de E et sont continues. Alors $f \in C^1(E)$

Proposition 1.3. Soit E un ouvert non vide. Soit $f \in C^1(E)$. Alors f est différentiable.

Preuve. D'abord, on rappelle que f est continument différentiable au voisinage de $\mathbf{a}$ si elle est dans $C^1(E \cap B(\mathbf{a}, \delta))$ pour un $\delta > 0$. Soit $\mathbf{a} \in E$. On va considérer $f(\mathbf{x}) - f(\mathbf{a})$, $\mathbf{x} \in B(\mathbf{a}, \delta)$. Or,

$$f(x_1, \ldots, x_n) - f(a_1, \ldots, a_n)$$

peut être réecrit comme une somme téléscopique

$$\sum_{i=1}^{n}(f(a_1, \ldots, a_{i-1}, x_i, \ldots, x_n) - f(a_1, \ldots, a_i, x_{i+1}, \ldots, x_n))$$

et appliquer le théorème des accroissements finis à chaque terme de la somme,

$$\begin{aligned} &f(a_1, \ldots, a_{i-1}, x_i, \ldots, x_n) - f(a_1, \ldots, a_i, x_{i+1}, \ldots, x_n) \\ &= \frac{\partial}{\partial x_i} f(a_1, \ldots, a_{i-1}, c_i, x_{i+1}, \ldots, x_n) \cdot (x_i - a_i) \end{aligned}$$

On définit

$$r_k(\mathbf{x}, \mathbf{a}) = \frac{\partial f}{\partial x_k}(a_1, \ldots, a_{i-1}, c_i, x_{i+1}, \ldots, x_n) - \frac{\partial f}{\partial x_k}(a_1, \ldots, a_n)$$

Par continuité des dérivées partielles on a

$$\lim_{\mathbf{x}\to\mathbf{a}} r_k(\mathbf{x},\mathbf{a}) = 0$$

Donc

$$f(\mathbf{x}) - f(\mathbf{a}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})(x_i - a_i) + \sum_{i=1}^{n} r_k(\mathbf{x},\mathbf{a})(x_i - a_i)$$

où

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{||\sum_{i=1}^{n} r_k(\mathbf{x},\mathbf{a})(x_i - a_i)||}{||\mathbf{x}-\mathbf{a}||} = 0$$

Donc f est différentiable. $\square$

La proposition suivante va être important pour certains cas par la suite.

Proposition 1.4. Soit $f \in C^1(E)$ et $g_1, \ldots, g_n \colon I \to \mathbb{R}$ des fonctions réelles dérivables, telles que $\{(g_1(I), \ldots, g_n(I)) \subset E\}$. Soit

$$h(t) := f(g_1(t), \ldots, g_n(t))$$

Alors h est dérivable sur I et

$$h'(t) = \sum_{i=1}^{n} g_i'(t) \frac{\partial f}{\partial x_i}(g_1(t), \ldots, g_n(t))$$

Proposition 1.5. Soit $\mathbf{g}(t) = (g_1(t), \ldots, g_n(t))$. Soit $t_0 \in I$. On veut calculer

$$h'(t_0) = \lim_{t\to t_0} \frac{f(\mathbf{g}(t+t_0) - \mathbf{g}(t_0))}{t - t_0}$$

Or f est différentiable en $\mathbf{g}(t_0)$ donc en particulier

$$f(\mathbf{g}(t)) = f(\mathbf{g}(t_0)) + Df(\mathbf{g}(t_0))(\mathbf{g}(t) - \mathbf{g}(t_0)) + r(\mathbf{g}(t), \mathbf{g}(t_0))$$

et:

$$h'(t_0) = \lim_{t\to t_0} \frac{Df(\mathbf{g}(t_0))(\mathbf{g}(t+t_0) - \mathbf{g}(t_0)) + r(\mathbf{g}(t+t_0), \mathbf{g}(t_0))}{t - t_0}$$

Par continuité des g_k, $\lim_{t\to t_0} r(\mathbf{g}(t+t_0), \mathbf{g}(t_0)) = 0$, et avec $Df(\mathbf{g}(t_0)) =: (d_1, \ldots, d_n)$:

$$h'(t) = \lim_{t\to t_0} \frac{\sum_{i=1}^{n} d_i(g_i(t+t_0) - g_i(t_0))}{t - t_0}$$

Donc par dérivabilité des g_i:

$$h'(t) = \sum_{i=1}^{n} d_i g_i'(t_0)$$

Donc par différentiabilité de f: $d_i = \frac{\partial f}{\partial x_i}(\mathbf{g}(t_0))$ et

$$h'(t) = \sum_{i=1}^{n} g_i'(t_0) \frac{\partial f}{\partial x_i}(\mathbf{g}(t_0))$$

2 Exemples

2.1 Plan tangent

Dans cet exemple, soit $E \subset \mathbb{R}^2$ un ouvert et $f \in C^1(E)$. On va s'intéresser au graphe de la fonction $\Sigma = \{x, y, f(x, y) | (x, y) \in E\}$. Soit $(a, b) \in E$. Soit I un intervalle ouvert et $t_0 \in I$. Soit $\phi, \psi \in C^1(I)$ telles que $\phi(t_0) = a, \psi(t_0) = b$. Soit $\alpha(t) = f(\phi(t), \psi(t)$. On définit la courbe paramétrique

$$\boldsymbol{\rho}(t) = (\phi(t), \psi(t), \alpha(t))$$

qui est sur le graphe de f. Le vecteur tangent au points t_0 est $\boldsymbol{\rho}'(t_0) = (\phi'(t_0), \psi'(t_0), \alpha'(t_0))$. En particulier, comme $\alpha(t) = f(\phi(t), \psi(t))$ on a

$$\alpha'(t_0) = \frac{\partial f}{\partial x}(a, b)\phi'(t_0) + \frac{\partial f}{\partial y}(a, b)\psi'(t_0)$$

Donc en définissant $\mathbf{n} = (\frac{\partial f}{\partial x}(a, b), \frac{\partial f}{\partial y}(a, b), -1)$, cette équation se réecrit

$$\langle \mathbf{n}, \boldsymbol{\rho}'(t_0) \rangle = 0$$

i.e. ces deux vecteurs sont orthogonaux. Cela veut dire que le vecteur tangent à la courbe en (a, b) se trouve dans le plan orthogonal au vecteur $\mathbf{n}$. Ce plan est appelé *plan tangent* à Σ au point (a, b), et a pour équation

$$\frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b) - (z - f(a, b)) = 0$$

2.2 Derivée directionnelle

Soit $f \in C^1(\mathbb{R}^n)$. Soit $\mathbf{a}, \mathbf{v} \in \mathbb{R}^n$ où $||\mathbf{v}|| = 1$. On définit une nouvelle fonction

$$f_{\mathbf{v}}(t) = f(\mathbf{a} + t\mathbf{v}) = f(a_1 + tv_1, \dots, a_n + tv_n)$$

f étant de classe C^1, on a:

$$f'_{\mathbf{v}}(t) = \sum_{k=1}^{n} v_k \frac{\partial f}{\partial x_k}(\mathbf{a} + t\mathbf{v})$$

et donc

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a}) = f'_{\mathbf{v}}(0) = \sum_{k=1}^{n} v_k \frac{\partial f}{\partial x_k}(\mathbf{a}) = \langle \nabla f(\mathbf{a}), \mathbf{v} \rangle$$

est appelée *dérivée directionelle* de f en $\mathbf{a}$ vers $\mathbf{v}$. Intuitivement, $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a})$ est la pente de la tangente à la courbe $C_{\mathbf{v}}$ de la fonction $f_{\mathbf{v}}$ au point $(0, f(\mathbf{a}))$.

3 Dérivées partielles de degré supérieur

Définition 3.1. Soit $f \colon E \to \mathbb{R}$. On définit la dérivée partielle du degré 2 comme la dérivée partielle de la dérivée partielle de f:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial f}{\partial x_j}(\frac{\partial f}{\partial x_i})$$

Plus généralement:

$$\frac{\partial^k f}{\partial x_{i_1}, \dots, \partial x_{i_k}} = \frac{\partial f}{\partial x_{i_k}}(\dots(\frac{\partial f}{\partial x_1}))$$

Définition 3.2. Soit E un ouvert. Alors f est dans C^m si toutes les dérivées partielles d'ordre m existent et sont continues.

Théorème 3.1 (Théorème de Shwarz)**.** Soit E ouvert. Soit $f \in C^2(E)$, i.e. les dérivées partielles de degré 2 existent: $\frac{\partial^2 f}{\partial x_i \partial x_j}$ et $\frac{\partial^2 f}{\partial x_j \partial x_i}$, et sont continues.

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

Preuve. On va considérer le cas $n = 2$ et soit $(a, b) \in E$. On doit determiner que

$$\frac{\partial^2 f}{\partial x \partial y}(a, b) = \frac{\partial^2 f}{\partial y \partial x}(a, b)$$

Soit $\overline{a} > a$ et $\overline{b} > b$, bien définis puisque E est ouvert. Intuitivement, on peut définir un rectanlge aux sommets $(a, b), (a, \overline{b}), (\overline{a}, \overline{b}), (\overline{a}, b)$. On va considérer les différences de f entre chaque paire de sommets consécutifs:

$$\begin{aligned} \square f &:= f(\overline{a}, \overline{b}) - f(a, \overline{b}) - (f(\overline{a}, b) - f(a, b)) \\ &= f(\overline{a}, \overline{b}) - f(\overline{a}, b) - (f(a, \overline{b}) - f(a, b)) \end{aligned}$$

Pour faciliter les notations, soit

$$g(x) := f(x, \overline{b}) - f(x, b), \ h(x) := f(\overline{a}, x) - f(a, x)$$

des fonctions continument dérivables deux fois. On a

$$\square f = h(\overline{b}) - h(b) = g(\overline{a}) - g(a)$$

Or par le théorème des accroissements finis appliqué à g, $\exists c_1 \in (a, \overline{a})$ tel que

$$g'(c_1)(\overline{a} - a) = g(\overline{a}) - g(a))$$

et $\exists d_1 \in (b, \overline{b})$ tel que

$$h'(d_1)(\overline{b} - b) = h(\overline{b}) - h(b)$$

Or,

$$g'(x) = \frac{\partial f}{\partial x}(x, \overline{b}) - \frac{\partial f}{\partial x}(x, b)$$

et

$$h'(x) = \frac{\partial f}{\partial y}(\overline{a}, x) - \frac{\partial f}{\partial y}(a, x)$$

On peut appliquer le théorème des accroissements finis encore une fois, mais à la fonctions à droite de l'égalité et non pas à g', h' directement: $\exists c_2 \in (\overline{b}, b), d_2 \in (\overline{a}, a)$ tel que

$$g'(c_1) = (\overline{b} - b)\frac{\partial^2 f}{\partial y \partial x}(c_1, c_2), h'(d_1) = (\overline{a} - a)\frac{\partial^2 f}{\partial x \partial y}(d_1, d_2)$$

Donc

$$\begin{aligned}\square f &= (\overline{b}-b)(\overline{a}-a)\frac{\partial^2 f}{\partial y \partial x}(c_1, c_2)\\ &= (\overline{b}-b)(\overline{a}-a)\frac{\partial^2 f}{\partial x \partial y}(d_1, d_2)\end{aligned}$$

Donc en considérant que $\overline{b} \neq b$ et $\overline{a} \neq a$, on a

$$\frac{\partial^2 f}{\partial y \partial x}(c_1, c_2) = \frac{\partial^2 f}{\partial x \partial y}(d_1, d_2)$$

Or, par continutié des dérivées partielles:

$$\begin{aligned}\frac{\partial^2 f}{\partial y \partial x}(a, b) &= \lim_{\overline{a}\to a, \overline{b}\to b} \frac{\partial^2 f}{\partial y \partial x}(c_1, c_2)\\ \frac{\partial^2 f}{\partial x \partial y}(a, b) &= \lim_{\overline{a}\to a, \overline{b}\to b} \frac{\partial^2 f}{\partial x \partial y}(d1, d_2)\end{aligned}$$

□

Définition 3.3 (Hessienne)**.** Soit $\mathbf{a} \in E$. Soit $f \colon E \to \mathbb{R}$ dont toutes les dérivées partielles d'ordre 2 existent. La *matrice Hessienne* de f en $\mathbf{a}$: $\mathrm{Hess}_f(\mathbf{a}) \in \mathbb{R}^{n\times n}$ est définie par

$$[\mathrm{Hess}_f(\mathbf{a})]_{i,j} := \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})$$

Le théorème précédent implique que si f est dans $C^2(E)$, la matrice hessienne est symétrique.

4 Fonctions de $\mathbb{R}^n \to \mathbb{R}^m$

4.1 Motivations

Peut-on élargir nos définitions pour inclures les fonctions à valeurs dans $\mathbb{R}^m$ et non seulement $\mathbb{R}$? On avait déjà défini la différentiabilité d'une fonction à valeurs dans $\mathbb{R}^m$, et on va voir que beaucoup de concepts conséquents vont aussi pouvoir être définis élégamment.

On considère

$$\mathbf{f} \colon \begin{aligned}&\mathbb{R}^n \to \mathbb{R}^m\\ &\mathbf{x} \mapsto \mathbf{f}(\mathbf{x}) := (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))\end{aligned}$$

4.2 Définitions

Définition 4.1. La k-ième dérivée partielle de $\mathbf{f}$ est définie par

$$\frac{\partial \mathbf{f}}{\partial x_k}(\mathbf{a}) := (\frac{\partial f_1}{\partial x_k}(\mathbf{a}), \ldots, \frac{\partial f_m}{\partial x_k}(\mathbf{a}))$$

Elle est définie si chacune des k-ième dérivées partielles des composantes de $\mathbf{f}$ existe.

Définition 4.2. La dérivée directionnelle de $\mathbf{f}$ est définie par

$$D\mathbf{f}(\mathbf{x}, \mathbf{v}) := (Df_1(\mathbf{x}, \mathbf{v}), \ldots, Df_m(\mathbf{x}, \mathbf{v}))$$

Elle est définie si chacune des dérivées directionnelles des composantes existe.

Définition 4.3. $\mathbf{f}$ a pour limite $\mathbf{l}$ en $\mathbf{a}$ si $\forall \varepsilon > 0$, $\exists \delta_\varepsilon > 0$ tel que

$$\forall \mathbf{x} \in E \cap B(\mathbf{a}, \delta_\varepsilon), ||\mathbf{f}(\mathbf{x}) - \mathbf{l}|| < \varepsilon$$

On voit alors en quoi cela faisait du sens d'avoir une "matrice" jacobienne comme matrice de l'application linéaire qui caractérise la différentiabilité, et non pas un "vecteur".

4.3 Application à la différentielle de compositions de fonctions

L'intuition est que calculer la différentielle de compositions de fonctions revient à multiplier la différentielle de chaque fonction. On s'intéresse au cas suivant

$$\mathbb{R}^n \xrightarrow{\mathbf{g}} \mathbb{R}^p \xrightarrow{\mathbf{f}} \mathbb{R}^q$$

Soit $\mathbf{a} \in \mathbb{R}^n$. On considère la différentielle de g en $\mathbf{a}$:

$$\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{a}) = D\mathbf{g}(\mathbf{a})(\mathbf{x} - \mathbf{a}) + \mathbf{r}_g(\mathbf{x}, \mathbf{a})$$

et celle de f en $\mathbf{g}(\mathbf{a})$:

$$\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{g}(\mathbf{a})) = D\mathbf{f}(\mathbf{g}(\mathbf{a}))(\mathbf{a} - \mathbf{g}(\mathbf{a})) + \mathbf{r}_f(\mathbf{x}, \mathbf{g}(\mathbf{a}))$$

en particulier donc pour $\mathbf{x} \in \mathbb{R}^n$:

$$\mathbf{f}(\mathbf{g}(\mathbf{x})) - \mathbf{f}(\mathbf{g}(\mathbf{a})) = D\mathbf{f}(\mathbf{g}(\mathbf{a}))(\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{a})) + \mathbf{r}_f(\mathbf{g}(\mathbf{x}), \mathbf{g}(\mathbf{a}))$$

et donc en introduisant l'expression de la différentielle de $\mathbf{g}$ en $\mathbf{a}$:

$$\mathbf{f}(\mathbf{g}(\mathbf{x})) - \mathbf{f}(\mathbf{g}(\mathbf{a})) = D\mathbf{f}(\mathbf{g}(\mathbf{a}))(D\mathbf{g}(\mathbf{a})(\mathbf{x}-\mathbf{a}) + \mathbf{r}_g(\mathbf{x},\mathbf{a})) + \mathbf{r}_f(\mathbf{g}(\mathbf{x}),\mathbf{g}(\mathbf{a}))$$

donc soit $\mathbf{r}_{\mathbf{f}\circ\mathbf{g}}(\mathbf{x},\mathbf{a}) := D\mathbf{f}(\mathbf{g}(a))\mathbf{r}_g(\mathbf{x},\mathbf{a}) + \mathbf{r}_f(\mathbf{g}(\mathbf{x}),\mathbf{g}(\mathbf{a}))$, on a bien que cette fonction est en $o(||\mathbf{x}-\mathbf{a}||)$ et l'expression se réecrit

$$\mathbf{f}\circ\mathbf{g}(\mathbf{x}) - \mathbf{f}\circ\mathbf{g}(\mathbf{a}) = D\mathbf{f}(\mathbf{g}(\mathbf{a}))D\mathbf{g}(\mathbf{a})(\mathbf{x}-\mathbf{a}) + \mathbf{r}_{\mathbf{f}\circ\mathbf{g}}(\mathbf{x}-\mathbf{a})$$

i.e.

$$D(\mathbf{f}\circ\mathbf{g})(\mathbf{a}) = D\mathbf{f}(\mathbf{g}(\mathbf{a}))D\mathbf{g}(\mathbf{a})$$

donc calculer la différentielle d'une composition de fonction revient à mutliplier les jacobiennes des fonctions intermédiaires aux points intermédiaires.

Une conséquence est que la jacobienne d'une fonction inverse est l'inverse de la jacobienne d'une fonction. En particulier avec $Id_{\mathbb{R}^n} = \mathbf{f}^{-1}\circ\mathbf{f}$ on a que

$$D(\mathbf{f}^{-1}\circ\mathbf{f})(\mathbf{a}) = D\mathbf{f}^{-1}(\mathbf{f}(\mathbf{a}))D\mathbf{f}(\mathbf{a})$$

donc

$$I_n = D\mathbf{f}^{-1}(\mathbf{f}(\mathbf{a}))D\mathbf{f}(\mathbf{a}) \Leftrightarrow D\mathbf{f}^{-1}(\mathbf{f}(\mathbf{a})) = (D\mathbf{f}(\mathbf{a}))^{-1}$$

4.4 Dérivée d'une intégrale

Théorème 4.1. *Soit $a, b \in \mathbb{R}$ tel que $a < b$. Soit $I \subset \mathbb{R}$ ouvert. Soit $f\colon [a;b]\times I$ tel que $\frac{\partial f}{\partial y}$ est continue sur $[a;b]\times I$. Soit $g(y) = \int_a^b f(x,y)dx$. Alors $g \in C^1$ et $\forall y \in I$*

$$g'(y) = \int_a^b \frac{\partial f}{\partial y}(x,y)dx$$

Preuve. Soit $y_0 \in I$. Alors on a $\forall y \neq y_0$:

$$\frac{g(y)-g(y_0)}{y-y_0} = \frac{1}{y-y_0}\int_a^b f(x,y) - f(x,y_0)dx$$

or $\exists c \in (y, y_0)$ tel que (par th. des accroissements finis)

$$\frac{\partial f}{\partial y}(x_0, c) = \frac{f(x_0,y)-f(x_0,y_0)}{y-y_0}$$

donc étant donné que $y \to y_0 \Rightarrow c \to y_0$

$$g'(y_0) = \lim_{y\to y_0}\frac{g(y)-g(y_0)}{y-y_0} = \int_a^b \frac{\partial f}{\partial y}(x,y_0)dx$$

$\square$

On peut alors utiliser ce théorème avec le théorème fondamental du calcul intégral: soit f continue,

$$F(t) := \int_a^t f(x)dx \Rightarrow F'(t) = f(t)$$

4.5 Exemple: coordonnées polaires

Définition 4.4 (Laplacien)**.** Soit $f \in C^2(E)$. On définit l'application $\Delta\colon C^2(E) \to (E \to \mathbb{R})$ par:

$$\Delta f(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}(\mathbf{x})$$

Proposition 4.1. Soit $f \in C^2(\mathbb{R}^2)$. Soit $g\colon \mathbb{R}^2 \to \mathbb{R}^2$, $g(r,\varphi) = r(\cos\varphi, \sin\varphi)$. Soit $\overline{f} = f \circ g$. Alors

$$\nabla f(x,y) = (\cos\varphi \frac{\partial \overline{f}}{\partial r} - \frac{1}{r}\sin\varphi \frac{\partial \overline{f}}{\partial \varphi}, \sin\varphi \frac{\partial \overline{f}}{\partial r} + \frac{1}{r}\cos\varphi \frac{\partial \overline{f}}{\partial \varphi})$$
$$\Delta f(x,y) = \frac{\partial^2 \overline{f}}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 \overline{f}}{\partial \varphi^2} + \frac{1}{r}\frac{\partial \overline{f}}{\partial r}$$

Preuve. On a que

$$Dg(r,\varphi) = \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix}$$

est inversible d'inverse

$$Dg(r,\varphi)^{-1} = \frac{1}{r}\begin{bmatrix} r\cos\varphi & r\sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}$$

Or

$$D\overline{f}(r,\varphi) = Df(g(r,\varphi))Dg(r,\varphi) \Leftrightarrow Df(g(r,\varphi)) = D\overline{f}(r,\varphi)Dg(r,\varphi)^{-1}$$

i.e.

$$Df(g(r,\varphi)) = \begin{bmatrix} \frac{\partial \overline{f}}{\partial r}(r,\varphi) & \frac{\partial \overline{f}}{\partial \varphi}(r,\varphi) \end{bmatrix} \cdot \frac{1}{r}\begin{bmatrix} r\cos\varphi & r\sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}$$

Et pour obtenir une expression pour le laplacien, il faut prendre en compte que

$$\frac{\partial^2 f}{\partial x^2}(g(r,\varphi)) = \cos(\varphi)\frac{\partial}{\partial r}(\frac{\partial f}{\partial x}(g(r,\varphi))) - \frac{1}{r}\cos(\varphi)\frac{\partial}{\partial \varphi}(\frac{\partial f}{\partial x}(g(r,\varphi)))$$

$\square$

Le problème avec cette expression est qu'elle est de la forme

$$Df(g(r,\varphi)) = D\overline{f}(r,\varphi) \cdot Dg(r,\varphi)^{-1}$$

donc pour obtenir le gradient pour un x, y quelconque, il faut calculer $g^{-1}(x, y)$ pour obtenir les r, φ qu'on utilise pour le calcul des jacobiennes dans la partie droite de l'équation.

5 Formule de Taylor

Soit $f \in C^{p+1}(E)$, où E est un ouvert connexe. L'intuition est de se ramener au cas réel. Soit $\mathbf{y} \in E$. Soit $h(t) = f(t\mathbf{y}), \forall t \in F = [0; 1]$.

$$h'(t) = \sum_{i=1}^{n} y_i \frac{\partial f}{\partial x_i}(t\mathbf{y})$$

or, les dérivées partielles sont dans $C^p(E)$, donc par récurrence $h \in C^{p+1}(F)$, et la formule de Taylor réelle s'applique: $\exists c \in (0; 1)$ tel que

$$h(t) = \sum_{i=1}^{p} \frac{h^{(i)}(0)}{i!} t^i + \frac{h^{p+1}(c)}{(p+1)!} t^{p+1}$$

Il faudrait maintenant calculer explicitement les $h^{(i)}(x)$. Pour rappel, n a déjà calculé $h' = h^{(1)}$:

$$h'(t) = \sum_{i=1}^{n} y_i \frac{\partial f}{\partial x_i}(t\mathbf{y})$$

Pour la dérivée seconde, avec $h_i'(t) := y_i \frac{\partial f}{\partial x_i}(t\mathbf{y})$, on a

$$(h_i')'(t) = \sum_{j=1}^{n} y_j y_i \frac{\partial^2 f}{\partial x_j \partial x_i}(t\mathbf{y})$$

donc

$$h''(t) = \sum_{j=1}^{n} \sum_{i=1}^{n} y_j y_i \frac{\partial^2 f}{\partial x_j \partial x_i}(t\mathbf{y})$$

Plus généralement, on a

$$h^{(k)}(t) = \sum_{1 \le i_1, \ldots, i_k \le n} y_{i_1} \ldots y_{i_k} \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}(t\mathbf{y})$$

On peut exprimer cette somme d'une manière plus élégante en se rappelant du fait que $f \in C^{p+1}(E)$ et donc par le théorème de Schwarz on peut échanger l'ordre dans lequel on prend les dérivées partielles de f. Da manière très informelle, on a

$$(y_1 \frac{\partial}{\partial x_1} + \ldots + y_n \frac{\partial}{\partial x_n})^k = \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} y_{i_1} \ldots y_{i_k} \frac{\partial^{i_1 + \ldots + i_k}}{\partial x_{i_1} \cdots \partial x_{i_k}}$$

qui décrit une fonction qui à f associe une dérivée partielle d'ordre supérieur multipliée par des coefficients y_{i_j}. Or l'expression de droite peut se réecrire à l'aide de coefficients multinomiaux

Proposition 5.1.

$$(x_1 + \ldots + x_n)^k = \sum_{i_1+\ldots+i_n=k} \binom{k}{n_1, \ldots, n_k} x_1^{i_1} \ldots x_n^{i_n}$$

Preuve. On va donner une preuve combinatoire. Pour développer l'expression de gauche, on peut choisir n'importe lequel des termes dans $(x_1 + \ldots + x_n)$, k fois d'affilé. Or, pour il existe $\binom{k}{n_1,\ldots,n_k}$ moyens de séparer k objects en des ensembles de $n_1, \ldots, n_k$ objects.

Donc on sait que quand on choisit x_1 i_1 fois etc. jusqu'a x_n i_n fois, il exist $\binom{k}{n_1,\ldots,n_k}$ moyens de le faire (par exemple, pour $(x_1+x_2)^2$, il existe deux manière de choisir x_1 et x_2 une fois: la premier en choisissant x_1 en premier et x_2 ensuite, la deuxième en faisant l'inverse.) $\square$

Ainsi, en introduisant les notations suivantes: $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$, $|\alpha| = \sum_{i=1}^n \alpha_i$, $\boldsymbol{\alpha}! = \alpha_1! \ldots \alpha_n!$, $\mathbf{y}^{\boldsymbol{\alpha}} = y_1^{\alpha_1} \ldots y_k^{\alpha_k}$, on a

$$\sum_{1 \leq i_1, \ldots, i_k \leq n} y_{i_1} \ldots y_{i_k} \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}(t\mathbf{y}) = \sum_{|\boldsymbol{\alpha}|=k} \frac{k!}{\boldsymbol{\alpha}!} \mathbf{y}^{\boldsymbol{\alpha}} \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(t\mathbf{y})$$

Donc pour en revenir à l'expression de départ on a

$$f(\mathbf{y}) = f(\mathbf{0}) + \sum_{1 \leq |\boldsymbol{\alpha}| \leq p} \frac{1}{\boldsymbol{\alpha}!} \mathbf{y}^{\boldsymbol{\alpha}} \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{0}) + \sum_{|\boldsymbol{\alpha}|=p+1} \frac{1}{\boldsymbol{\alpha}!} \mathbf{y}^{\boldsymbol{\alpha}} \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(c\mathbf{y})$$

6 Théorème des fonctions implicites

6.1 Le cas $n = 2$

Soit $U \subset \mathbb{R}^2$ ouvert et $f \in C^1(U, \mathbb{R})$. Soit $(a, b) \in U$ avec $f(a, b) = 0$ et $\frac{\partial f}{\partial x_2}(a, b) \neq 0$. Alors il existe un voisinage V de a et une fonction $g : V \to \mathbb{R}$ tel que $(V, g(V)) \subset U$, $g(a) = b$ et $f(x, g(x)) = 0 \forall x \in V$.

Preuve. □

6.2 Difféomorphismes locaux

6.2.1 Motivations

On sait que pour une application linéaire $F \colon \mathbb{R}^n \to \mathbb{R}^n$, avec $F(\mathbf{x}) = A\mathbf{x}$, il suffit que A soit inversible pour que l'application linéaire le soit. On aimerait elargir cette observation aux applications non linéaires mais différentiables i.e. de la forme

$$F(\mathbf{x}) = F(\mathbf{a}) + L(\mathbf{x} - \mathbf{a}) + o(||\mathbf{x} - \mathbf{a}||)$$

En particulier, on veut déterminer que F est inversible localement autours de $\mathbf{a}$ si L est inversible.

6.2.2 Théorème du point fixe de Banach

Dans cette section, on utilise des arguments de topologie dans des espaces métriques. La preuve est essentiellement la même que dans un espace vectoriel normé, ou même dans $\mathbb{R}$.

Théorème 6.1. Soit (E, d) une espace métrique complet, et $F \colon E \to E$ une application k-contractante: $\exists k \in [0; 1)$ tel que

$$d(F(x), F(y)) \leq kd(x, y)$$

alors F admet un points fixe.

Preuve. Soit $x_0 \in E$, on définit (x_k) par : $x_{i+1} = F(x_i)$. Alors, en particulier,

$$d(x_{n+1}, x_n) = d(F(x_n), F(x_{n-1})) \leq kd(x_n, x_{n-1})$$

et par récurrence, $d(x_{n+1}, x_n) < k^n d(x_1, x_0)$. Et

$$d(x_{n+p}, x_n) \leq \sum_{i=0}^{p} d(x_{n+i}, d(x_{n+i-1})) \leq (\sum_{i=0}^{p}) k^n d(x_1, x_0)$$

donc

$$d(x_{n+p}, x_n) \leq (1 - k^{p+1}) \frac{k^n}{1-k} d(x_1, x_0)$$

or l'espace métrique est compacte donc comme (x_k) est de Cauchy elle admet une limite x_*. En particulier, $F(x_*) = x_*$, i.e. c'est un point fixe □

Proposition 6.1. Soit F_1, F_2 deux contractions de (E, d) un espace métrique compacte, telles que $\exists \delta > 0$ tel que $\forall x \in E$,

$$d(F_1(x), F_2(x)) \leq \delta$$

Alors si F_1, F_2 admettent x_1, x_2 comme points fixes, respectivement,

$$d(F(x_1), F(x_2)) \leq \frac{\delta}{1-k}$$

Preuve.

$$d(x_1, x_2) \leq d(F_1(x_1), F_1(x_2)) + d(F_1(x_2, F_2(x_2))) \leq kd(x_1, x_2) + \delta$$

□

6.2.3 Conditions d'inversion locale

Soit $E \subset \mathbb{R}^n$ et $f \colon E \to \mathbb{R}^n \in C^1(E)$. On s'intéresse aux conditions pour lesquelles f est inversible localement autours de $\mathbf{a} \in E$, i.e. $\exists U \subset E$ ouvert contenant $\mathbf{a}$, et une fonction $g \colon f(U) \to U$ avec

$$g \circ f = \mathrm{Id}_U$$

On suppose que $f(U)$ est ouvert. Ceci implique que pour $\mathbf{x} \in U$, on ait

$$D(g \circ f)(f(\mathbf{x}))D(f(\mathbf{x})) = I_n$$

i.e. si f est localement inversible, sa jacobienne aussi. La question est de savoir si l'inversibilité de la jacobienne suffit pour construire U, g.

Théorème 6.2 (Conditions d'inversion locale)**.** Soit $f \in C^1(E, \mathbb{R}^n)$ et $\mathbf{a} \in E$. On suppose que $A := Df(\mathbf{a})$ est inversible. Alors il existe un voisinage $U \subset E$ de $\mathbf{a}$ tel que $f(U)$ est ouvert et $f|_U$ est une bijection.

Preuve. D'abord on va essayer de se restreindre au case où $A = I_n$. Si ce n'est pas le cas, on peut considérer la fonction $h = A^{-1} \cdot f$: si h est localement inversible, $A \cdot h$ l'est aussi.

Le premier problème va être de transformer le problème de construction de U en un problème du point fixe. On va en fait chercher une $\delta > 0$ tel que $f(U) = B(f(\mathbf{a}), \delta) := V$ et U va être de la forme

$$U = B(\mathbf{a}, 3\delta) \bigcap f^{-1}(B(f(\mathbf{a}), \delta))$$

i.e. U est l'intersection entre la préimage de $f(U)$ et la boule centrée en $\mathbf{a}$ de rayon 3δ. Soit $\mathbf{y} \in V$. On voudrait donc que l'équation:

$$f(\mathbf{x}) = \mathbf{y}, \mathbf{x} \in B(\mathbf{a}, 3\delta)$$

ait une solution. Ce qui est équivalent à avoir

$$\mathbf{x} = -f(\mathbf{x}) + \mathbf{x} + \mathbf{y}$$

Soit $F(\mathbf{x}) = -f(\mathbf{x}) + \mathbf{x} + \mathbf{y}$. Alors si F a un point fixe, on aurait une solution à l'équation cherchée. On veut montrer que F est une contraction de $B(\mathbf{a}, 3\delta)$. On a $||\mathbf{x} - \mathbf{a}|| < 3\delta$ donc

$$\begin{aligned} F(\mathbf{x}) - \mathbf{a} &= -f(\mathbf{x}) + \mathbf{x} + \mathbf{y} - \mathbf{a} \\ &= -f(\mathbf{x}) + f(\mathbf{a}) + \mathbf{x} - \mathbf{a} + \mathbf{y} - f(\mathbf{a}) \end{aligned}$$

or, $Df(\mathbf{a}) = I_n$ donc

$$f(\mathbf{x}) = f(\mathbf{a}) + \mathbf{x} - \mathbf{a} + r(\mathbf{x}, \mathbf{a})$$

et en particulier,

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{||f(\mathbf{a}) - f(\mathbf{x}) + \mathbf{x} - \mathbf{a}||}{||\mathbf{x} - \mathbf{a}||} = 0$$

donc pour $\varepsilon = \frac{1}{2}$, on peut trouver $\delta > 0$ tel que

$$||f(\mathbf{a}) - f(\mathbf{x}) + \mathbf{x} - \mathbf{a}|| < \frac{1}{2} ||\mathbf{x} - \mathbf{a}||$$

Or, $y \in B(f(\mathbf{a}), \delta)$ donc:

$$\begin{aligned} ||F(\mathbf{x}) - \mathbf{a}|| &= || - f(\mathbf{x}) + f(\mathbf{a}) + \mathbf{x} - \mathbf{a} + \mathbf{y} - f(\mathbf{a})|| \\ &\leq || - f(\mathbf{x}) + f(\mathbf{a}) + \mathbf{x} - \mathbf{a}|| + ||\mathbf{y} - f(\mathbf{a})|| \\ &\leq \frac{1}{2}||\mathbf{x} - \mathbf{a}|| + \delta < 3\delta \end{aligned}$$

Donc si c'est une contraction, c'est une contraction de $B(\mathbf{a}, \delta)$. Ensuite, on vérifie que c'est une contraction. On a

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{a}) + \mathbf{x} - \mathbf{a} + r(\mathbf{x}, \mathbf{a}) \\ f(\mathbf{z}) &= f(\mathbf{a}) + \mathbf{z} - \mathbf{a} + r(\mathbf{z}, \mathbf{a}) \end{aligned}$$

et

$$\begin{aligned} F(\mathbf{x}) &= -f(\mathbf{x}) + \mathbf{x} + \mathbf{y} \\ F(\mathbf{z}) &= -f(\mathbf{z}) + \mathbf{z} + \mathbf{y} \end{aligned}$$

donc

$$\begin{aligned} F(\mathbf{x}) &= \mathbf{a} - f(\mathbf{a}) + r(\mathbf{x}, \mathbf{a}) \\ F(\mathbf{z}) &= \mathbf{a} - f(\mathbf{a}) + r(\mathbf{z}, \mathbf{a}) \end{aligned}$$

donc

$$||F(\mathbf{x}) - F(\mathbf{z})|| = ||r(\mathbf{x}, \mathbf{a}) - r(\mathbf{z}, \mathbf{a})|| \leq ||r(\mathbf{x}, \mathbf{a})|| + ||r(\mathbf{z}, \mathbf{a})||$$

et soit $k \in [0; \frac{1}{2})$, $\exists \delta_0 \in (0; \delta]$ tel que $\forall \mathbf{x}, \mathbf{z} \in B(\mathbf{a}, \delta_0)$

$$||r(\mathbf{x}, \mathbf{a})|| \leq k||\mathbf{x} - \mathbf{a}|| \quad \text{et} \quad ||r(\mathbf{z}, \mathbf{a})|| \leq k||\mathbf{z} - \mathbf{a}||$$

donc F est un $2k$-contracture sur $B(\mathbf{a}, 3\delta_0)$.

Ceci implique que $\forall \mathbf{y} \in B(f(\mathbf{a}), \delta)$, $\exists! \mathbf{x} \in B(\mathbf{a}, 3\delta)$ tel que

$$f(\mathbf{x}) = \mathbf{y}$$

donc si on a $U = B(\mathbf{a}, 3\delta) \bigcap f^{-1}(B(f(\mathbf{a}), \delta))$, $f|_U$ est bijective.

Il resterait à prouver que l'inverse est différentiable et que sa jacobienne a $Df(\mathbf{a})$ pour inverse, démonstration qui utiliserait le résultat à propos de paires de contractures δ-proches, qu'on a prouvé dans la dernière section. On ne va pas le faire ici. $\square$

6.3 Théorème des fonctions implicites

6.3.1 Motivations

Soit $E \subset \mathbb{R}^{n+m}$ ouvert, et $f \in C^1(E, \mathbb{R}^n)$. On est face au problème

$$f(\mathbf{x}) = \mathbf{y}$$

i.e. $n+m$ inconnues pour n équations. On s'attend à ce que cette équations ait une infinité de solutions, mais le théorème d'inversion locale suggère qu'il y a une unique solutions, étant donné qu'on fixe m coordonnées de $\mathbf{x}$. Plus particulièrement, étant donné $f(\mathbf{x}) = \mathbf{y}$ et m composantes de $\mathbf{x}$, les n coordonnées restantes pourraient s'exprimer en fonction (différentiable) des m coordonnées déterminées, et de $\mathbf{y}$.

6.3.2 Le cas linéaire

Soit $A \in \mathcal{M}_{n\times(m+n)}(\mathbb{R})$. On a $f\colon \mathbf{x} \mapsto A\mathbf{x}$. On va d'abord réecrire A et $\mathbf{x}$ comme:

$$A = (A_1\ A_2), A_1 \in \mathcal{M}_{n\times n}(\mathbb{R}), A_2 \in \mathcal{M}_{n\times m}(\mathbb{R})$$

et

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \mathbf{x}_1 \in \mathbb{R}^n, \mathbf{x}_2 \in \mathbb{R}^m$$

Ce qui donne $A\mathbf{x} = A_1\mathbf{x}_1 + A_2\mathbf{x}_2$. Si A_1 est inversible, et qu'on fixe $\mathbf{x}_2$, on a un unique $\mathbf{x}_1$ qui satsfait l'équation $f(\mathbf{x}) = \mathbf{y}$:

$$\mathbf{x}_1 = A_1^{-1}(\mathbf{y} - A_2\mathbf{x}_2)$$

6.3.3 Généralisation

Dans ce cas f n'est pas forcément linéaire. Soit $\mathbf{a} \in E$. On va en fait essayer de traduire l'équation $f(\mathbf{x}) = \mathbf{y}$ par un problème d'inversion locale. Soit $F\colon E \to \mathbb{R}^{n+m}$, $F(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) & \dots & f_n(\mathbf{x}) & x_{n+1} & \dots & x_{n+m} \end{pmatrix}$. La jacobienne de F en $\mathbf{a}$ est de la forme

$$DF(\mathbf{a}) = \begin{bmatrix} G & * \\ 0 & I_m \end{bmatrix}$$

où

$$G = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{a}) & \dots & \frac{\partial f_n}{\partial x_n}(\mathbf{a}) \end{bmatrix}$$

est par hypothèse inversible. Alors $\det DF(\mathbf{a}) = \det G \neq 0$, i.e. la jacobienne de F est inversible. Donc le théorème d'inversion locale nous donne un voisinage $\overline{V}$ de $F(\mathbf{a})$, plus particulièrement de la forme $\overline{V} = V \times U_2$, où V est un voisinage de $f(\mathbf{a})$ et U_2 est un voisinage de $(\mathbf{a}_{n+1}, \ldots, \mathbf{a}_{n+m})$.