Chap I - Inversion locale et fonctions implicites

Où l'on se persuade que les systèmes d'équations
sont finalement tous linéaires

1 Introduction

Question : La Grande Idée du Calcul Différentiel, c'est que toute fonction raisonnable peut être, au voisinage d'un point, approchée par une application linéaire. Peut-on se servir de cette idée pour résoudre des équations ?

Considérons un système (pas forcément linéaire) d'équations, d'inconnue $x \in \mathbb{R}^n$

$$\begin{cases} f_1(x_1, \dots, x_n) & = u_1 \\ \vdots & \\ f_p(x_1, \dots, x_n) & = u_p \end{cases}$$

Si f est différentiable, alors localement, f est presque linéaire et il se trouve qu'on sait résoudre les systèmes linéaires:

- Il y a une unique solution si $n = p =$ rang du système
- Il y a une infinité de solutions, décrites par des variables paramètres, si $n > p$.

Les deux grands théorèmes de ce chapitre précisent ce qu'on peut obtenir, pour un système différentiable non linéaire, dans chacun de ces deux cas.

- Le Théorème d'Inversion Locale s'applique dans le cas $n = p$: à quelle condition peut-on "inverser" le système, c'est à dire résoudre $f(x) = u$ en $x = f^{-1}(u)$?
- Le Théorème des Fonctions Implicites s'applique dans le cas $n > p$: à quelle condition peut-on choisir des variables "paramètres" $z_{p+1}, \dots, z_n$ et résoudre $f(x, z) = 0$ en $x = \psi(z)$?

2 Théorème d'Inversion Locale

Premier cas: fonctions linéaires

Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire.

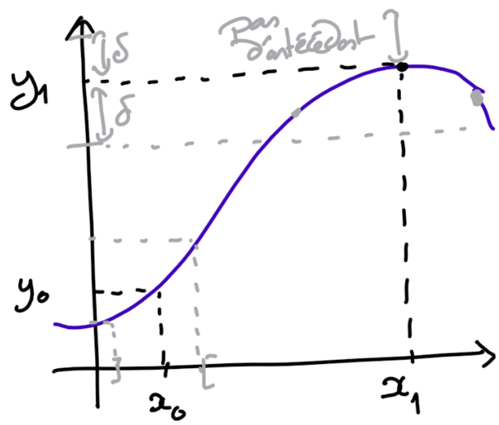
Alors l'équation $y = f(x)$ a une unique solution $x \in \mathbb{R}^n$ si, et seulement si, $n = p$ et si f est un isomorphisme; ce qui équivaut à dire que la matrice A de f dans la base canonique de $\mathbb{R}^n$ est inversible.

D'au autre côté, si $f \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ alors f est différentiable en tout point de $\mathbb{R}^n$ et, pour tout $x \in \mathbb{R}^n$, $Df(x) = f$ et $\mathrm{Jac}_f(x) = A$.

$\rightsquigarrow$ Donc f est inversible ssi $Df(x)$ est inversible.

Deuxième cas: fonctions réelles dérivables

Soit $f : I \subset \mathbb{R} \to \mathbb{R}$ une fonction dérivable sur I. Alors, pour tout $x \in I$, $Df(x) : h \in \mathbb{R} \mapsto f'(x_0)h$.



- En x_0: $f'(x_0) > 0$ donc $Df(x_0) : h \mapsto f'(x_0)h$ est inversible.

  Et du coup, au voisinage de x_0, f est strictement croissante donc bijective: il existe un voisinage V de x_0 tel que $f_{|V} : V \to f(V)$ et la réciproque f^{-1} est différentiable/dérivable avec $Df^{-1}(y_0)(k) = \frac{1}{f'(x_0)}k = \frac{1}{f'(f^{-1}(y_0))}k$.

- En x_1: $f'(x_1) = 0$, donc $Df(x_1)$ est l'application linéaire nulle, notoirement non iversible. Et d'ailleurs, $\forall \delta > 0$,

  - si $y \in]y_1, y_1 + \delta[$, $f(x) = y$ n'a pas de solution
  - si $y \in]y_1 - \delta, y_1[$, $f(x) = y$ a 2 solutions

  $\rightsquigarrow$ Pas d'inverse local !

Cas général: Heuristique

Soient $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$ $\mathscr{C}^1$, $x_0 \in U$ et $y_0 = f(x_0)$.

Le problème: On souhaite résoudre $f(x) = y$ pour (x, y) proches de (x_0, y_0).

On écrit, pour $\|x - x_0\|$ petit, $f(x) = f(x_0 + (x - x_0))$ ce qui donne, par définition de la différentielle

$$\begin{aligned} f(x) = y &\iff f(x_0 + (x - x_0)) = y \\ &\iff f(x_0) + Df(x_0)(x - x_0) + \text{chouïa} = y \end{aligned}$$

avec un chouïa d'autant plus microscopique que x est proche de x_0.

Donc, si $Df(x_0)$ est inversible, la solution est donnée par

$$x = x_0 + Df(x_0)^{-1}(y - f(x_0)) + \text{autre chouïa}$$

$\rightsquigarrow$ Comment rigorifier tout ça ? C'est le rôle du théorème d'inversion locale, qui va nous assurer que les chouïa sont effectivement inoffensifs.

Théorème 1

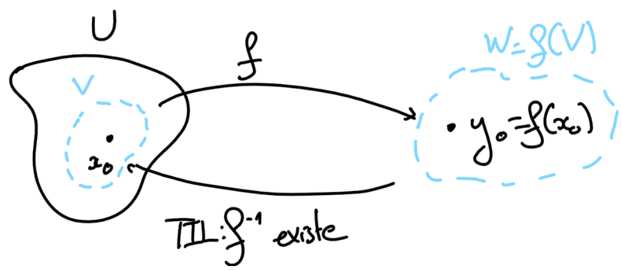
Soient E, F deux espaces de Banach, $f : U \subset E \to F$ une application $\mathscr{C}^1$ et soit $x_0 \in U$.

On suppose que $Df(x_0) \in \mathcal{L}(E, F)$ est un isomorphisme linéaire[1].

Alors il existe $V \subset U$ voisinage de x_0, W voisinage de $y_0 = f(x_0)$ tels que

$$f_{|V} : V \to W$$

soit un $\mathscr{C}^1$-difféomorphisme.



Exemple: Considérons $f : (x, y) \in \mathbb{R}^2 \mapsto (e^x - e^y, y)$. Alors $f \in \mathcal{C}^1(\mathbb{R}^2)$ et pour tout $(x_0, y_0) \in \mathbb{R}^2$,

$$\mathrm{Jac}(f)(x_0, y_0) = \begin{pmatrix} e^{x_0} & -e^{y_0} \\ 0 & 1 \end{pmatrix} \text{ est inversible.}$$

$\rightsquigarrow$ D'après le TIL, il y a donc un voisinage V de (x_0, y_0) tel que $f_{|V}$ soit un $\mathscr{C}^1$-difféomorphisme: autrement dit,

$$f_{|V} : V \to f(V)$$

est une bijection dont la réciproque est différentiable en tout point (a, b) de $f(V)$, et on a

$$D(f^{-1})(a, b) = Df(f^{-1}(a, b))$$

Autrement dit, si on note $(x, y) = f^{-1}(a, b)$,

$$\mathrm{Jac}\, f^{-1}(a, b) = (\mathrm{Jac}\, f(x, y))^{-1} = \begin{pmatrix} e^x & -e^y \\ 0 & 1 \end{pmatrix}^{-1} = \frac{1}{e^x} \begin{pmatrix} 1 & e^y \\ 0 & e^x \end{pmatrix}$$

Remarque: En fait, dans cet exemple, on peut trouver explicitement le voisinage et l'application réciproque:

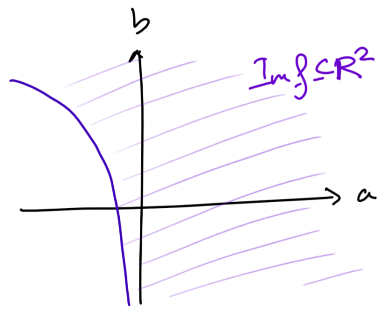
$$f(x, y) = (a, b) \iff e^x = a + e^b, y = b$$

$\rightsquigarrow$ Pour tous a, b tels que $a + e^b > 0$, ces équations ont une unique solution. Donc f est bijective de $\mathbb{R}^2$ dans $\mathcal{V} = \{(a,b) \in \mathbb{R}^2, a + e^b > 0\} = f(\mathbb{R}^2)$ et on a

$$f^{-1}(a,b) = (\ln(a+e^b), b) \in \mathscr{C}^1(\mathcal{V})$$

et on a bien

$$Df^{-1}(a,b) = \begin{pmatrix} \frac{1}{a+e^b} & \frac{e^b}{a+e^b} \\ 0 & 1 \end{pmatrix}, Df^{-1}(f(x,y)) = \begin{pmatrix} \frac{1}{e^x} & \frac{e^y}{e^x} \\ 0 & 1 \end{pmatrix} = (Df(x,y))^{-1}$$



On va maintenant s'attaquer à la démonstration du TIL. L'ingrédient principal de la preuve va être le *théorème du point fixe*: rappelons donc que

Théorème 2

Soit $B \subset E$ un fermé dans un e.v.n. E complet et soit $f : B \to B$ contractante, autrement dit

$$\exists\, k \in]0,1[, \forall\, x, y \in B, \|f(x) - f(y)\| \leq k\|x - y\|.$$

Alors f admet un unique point fixe $\overline{x} \in B$.

De plus, pour tout $x_0 \in B$, la suite définie à partir de x_0 par $x_{n+1} = f(x_n)$ converge vers $\overline{x}$, avec

$$\|x_n - \overline{x}\| \leq \frac{k^n}{1-k}\|x_1 - x_0\|$$

$\rightsquigarrow$ Si besoin, un petit complément à ce sujet se trouve par ici: `http://carolinevernier.website/memos/compl_point_fixe.pdf`

Preuve du TIL

Soit $x_0 \in U$ tel que $Df(x_0)$ est inversible, et $y_0 = f(x_0)$. On se fixe un y proche[2] de y_0. On veut montrer qu'il existe un unique x tel que $f(x) = y$.

[2] On précisera, pendant la preuve, à quel point y soit être proche de y_0 pour que tout fonctionne.

Plan de bataille:

1. On se ramène à un problème de point fixe: on trouve $\varphi_y : U \to E$ tel que
$$y = f(x) \iff \varphi_y(x) = x$$
2. On montre qu'on peut appliquer le théorème du point fixe, en trouvant $B \subset U$ fermé tel que $\varphi_y(B) \subset B$ et $\varphi_y : B \to B$ est contractante.
3. On note $x = g(y)$ l'unique point fixe de φ_y. On montre que g est continue, différentiable de différentielle $Df(x)^{-1}$, et enfin, que g est $\mathscr{C}^1$ (i.e., que $y \mapsto Dg(y)$ est continue).

Etape 1 - Transformation en problème de point fixe

Puisqu'on a $Df(x_0)^{-1} \in \mathcal{L}(F, E)$, on peut définit l'application
$$\varphi_y : x \in U \subset E \mapsto x - Df(x_0)^{-1}(f(x) - y) \in E$$

Autrement dit, $\varphi_y = \mathrm{Id}_E - Df(x_0)^{-1} \circ f + \underbrace{Df(x_0)^{-1}(y)}_{cste}$, donc φ_y est $\mathscr{C}^1$ sur U et, pour tout $x \in U$,
$$D\varphi_y(x) = \mathrm{Id}_E - Df(x_0)^{-1}Df(x).$$

De plus, on a
$$\begin{aligned} f(x) = y &\iff f(x) - y = 0 \\ &\iff Df(x_0)^{-1}(f(x) - y) = 0 \text{ (car } Df(x_0)^{-1} \text{ est un isomorphisme)} \\ &\iff x - Df(x_0)^{-1}(f(x) - y) = x \\ &\iff \varphi_y(x) = x \end{aligned}$$

Autrement dit, les solutions de notre équation $f(x) = y$ sont exactement les points fixes de φ_y.

$\rightsquigarrow$ On cherche maintenant à déterminer pour quels y l'application φ_y admet un unique point fixe. On va voir qu'il nous faut un y "pas trop loin" de y_0; et de plus, on ne pourra garantir que φ_y est contractante que sur un petit voisinage de x_0. D'où le côté "local" de ce théorème.

Etape 2 - Application du théorème du point fixe

Montrons qu'on peut appliquer le théorème du point fixe à φ_y quand y est proche de y_0 et x proche de x_0.

Dans un premier temps, on montre que φ_y est contractante sur $B = \overline{B}(x_0, r)$ pour un petit $r > 0$.

$\triangleright$ Remarquons que $D\varphi_y(x_0) = \mathrm{Id} - Df(x_0)^{-1}Df(x_0) = 0_{\mathcal{L}(\mathbb{R}^n)}$.

Or, φ_y n'est pas seulement différentiable: elle est $\mathcal{C}^1$ en x_0, autrement dit $x \mapsto D\varphi_y(x) \in \mathcal{L}(E)$ est continue en x_0. Donc, il existe $r > 0$ tel que

$$\|x - x_0\|_E \leq r \Rightarrow \|D\varphi_y(x)\|_{\mathcal{L}(E)} \leq \frac{1}{2}$$

▷ Mais alors, par l'inégalité des accroissements finis, pour tous $x_1, x_2 \in \overline{B}(x_0, r)$,

$$\|\varphi_y(x_1) - \varphi_y(x_2)\|_E \leq \frac{1}{2}\|x_1 - x_2\|_E$$

Pour pouvoir appliquer le théorème du point fixe, il faut aussi vérifier que $\varphi_y(\overline{B}(x_0, r)) \subset \overline{B}(x_0, r)$. On va voir que c'est le cas pour y assez proche de y_0.

▷ Par continuité de l'application linéaire $Df(x_0)^{-1}$, il existe $\delta > 0$ tel que

$$\|y - y_0\|_F \leq \delta \Rightarrow \|Df(x_0)^{-1}(y - y_0)\|_E \leq \frac{r}{2}$$

▷ Alors, pour $y \in B(y_0, \delta)$ et $x \in \overline{B}(x_0, r)$,

$$\begin{aligned} \|\varphi_y(x) - x_0\|_E &\leq \|\varphi_y(x) - \varphi_y(x_0)\|_E + \|\varphi_y(x_0) - x_0\|_E \\ &\leq \frac{1}{2}\|x - x_0\|_E + \| - Df(x_0)^{-1}(f(x_0) - y)\|_E \leq r \end{aligned}$$

autrement dit, $\varphi_y(x) \in \overline{B}(x_0, r)$.

Par le théorème du point fixe, pour $y \in B(y_0, \delta)$, il existe un unique $x \in \overline{B}(x_0, r)$ tel que $\varphi_y(x) = x$, autrement dit tel que $f(x) = y$.

Notons-le $x = f^{-1}(y)$: alors $f^{-1} : B(y_0, \delta) \to \overline{B}(x_0, r)$ est bien la bijection réciproque de f.

Il reste maintenant à montrer que f^{-1} est de classe $\mathscr{C}^1$.

Etape 3 - Régularité de f^{-1}

▷ **Montrons que f^{-1} est continue.** Soient $y_1, y_2 \in B(y_0, \delta)$. On note $x_1 = f^{-1}(y_1)$, $x_2 = f^{-1}(y_2)$. Alors

$$\begin{aligned} \|f^{-1}(y_1) - f^{-1}y_2)\|_E = \|x_1 - x_2\|_E &= \|\varphi_{y_1}(x_1) - \varphi_{y_2}(x_2)\|_E \\ &\leq \|\varphi_{y_1}(x_1) - \varphi_{y_1}(x_2)\|_E + \|\varphi_{y_1}(x_2) - \varphi_{y_2}(x_2)\|_E \\ &\leq \frac{1}{2}\|x_1 - x_2\|_E + \|Df(x_0)^{-1}(y_1 - y_2)\|_E \end{aligned}$$

donc, en soustrayant $\frac{1}{2}\|x_1 - x_2\|_E$ de part et d'autre de l'inégalité, puis en multipliant par 2:

$$\|f^{-1}(y_1) - f^{-1}(y_2)\|_E = \|x_1 - x_2\|_E \leq 2\|Df(x_0)^{-1}\|_{\mathcal{L}(F,E)}\|y_1 - y_2\|_F.$$

⇝ Ainsi, f^{-1} est Lipschitzienne, donc continue.

$\triangleright$ **Montrons que f^{-1} est différentiable**, de différentielle

$$Df^{-1}(y) = Df(x)^{-1} \text{ où } y = f(x)$$

Tout d'abord, il faut justifier que $Df(x)$ est inversible pour $x \in B(x_0, r)$. Remarquons que, puisque

$$D\varphi_y(x) = \mathrm{Id} - Df(x_0)^{-1} Df(x),$$

on a

$$Df(x) = Df(x_0) \circ (\mathrm{Id} - D\varphi_y(x))$$

Or, on a vu que pour $x \in B(x_0, r)$,

$$\|D\varphi_y(x)\|_{\mathcal{L}(E)} \leq \frac{1}{2} < 1.$$

Il se trouve que, dans l'espace vectoriel normé $\mathcal{L}(E)$, la boule ouverte de centre Id_E et de rayon 1 est incluse dans l'ensemble des isomorphismes linéaires $Isom(E)$. Autrement dit, si $H \in \mathcal{L}(E)$ vérifie $\|H\|_{\mathcal{L}(E)} < 1$, alors $\mathrm{Id} + H$ est une application linéaire inversible, d'inverse continue. Pour une preuve, voir par ici:

`http://carolinevernier.website/complement-appli-lin/appli-lin-inverse.html`

Donc, $\mathrm{Id} - D\varphi_y(x)$ est bien un isomorphisme linéaire. Et comme $Df(x_0)$ en est un aussi, $Df(x)$ est une composée d'isomorphismes, donc c'est un isomorphisme. Donc $Df(x)^{-1}$ existe, et c'est une application linéaire continue $F \to E$.

Il s'agit maintenant de montrer que, pour tout $y \in B(y_0, \delta)$ et $k \in F$,

$$\frac{1}{\|k\|_F} \left\| f^{-1}(y+k) - f^{-1}(y) - Df(x)^{-1}(k) \right\|_E \xrightarrow[k \to 0_F]{} 0$$

Cette douloureuse majoration est reléguée à l'annexe technique suivante: `http://carolinevernier.website/compl/TIL-preuve.pdf`

$\triangleright$ **Enfin, on veut montrer que f^{-1} est $\mathscr{C}^1$**. Autrement dit, il s'agit de montrer que

$$y \in B(y_0, \delta) \mapsto Df^{-1}(y) = Df(f^{-1}(y))^{-1} \in \mathcal{L}(F, E)$$

est continue.

Or, on peut l'écrire comme la composée $I \circ Df \circ f^{-1}$, où

$$I : A \in \mathcal{GL}(E, F) \mapsto A^{-1} \in \mathcal{L}(F, E)$$

et on a

1. f^{-1} est continue, on l'a vu;
2. $x \mapsto Df(x)$ puisque f est $\mathscr{C}^1$

3. I est continue:
   Dans le cas où $E = F = \mathbb{R}^n$, on peut le voir en mettant A sous forme matricielle. On a alors

$$I(A) = \frac{1}{\det(A)}\,{}^t\mathrm{Com}(A)$$

   donc les coefficients de $I(A)$ sont des fractions rationnelles des coefficients de A, qui ne s'annulent pas sur $\mathcal{GL}_n(\mathbb{R})$.

   Pour le cas général, voir par ici:

   `http://carolinevernier.website/complement-appli-lin/appli-lin-inverse.html`

   Et ceci termine la preuve du théorème d'inversion locale. On a bien mérité un petit carré:

$\square$

Remarque : Il est essentiel au fonctionnement de la preuve que f soit $\mathscr{C}^1$, et non simplement différentiable. Et ce n'est pas juste parce qu'on n'a pas eu l'intelligence de trouver une meilleure preuve: le théorème devient faux si f est simplement supposée différentiable.

Contre exemple:

$$f : x \in \mathbb{R} \mapsto \begin{cases} x + x^2 \sin\left(\frac{1}{x}\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases}$$

Alors f est différentiable en 0, et $f'(0) = 1 \neq 0$:

$$\frac{1}{h}(f(h) - f(0)) = 1 + h \sin\left(\frac{1}{h}\right) \to 1$$

Mais f n'est inversible sur aucun voisinage de 0. En effet, si on considère les suites

$$x_n = \frac{1}{2\pi n} \geq y_n = \frac{1}{\pi(2n + \frac{1}{2})} \geq z_n = \frac{1}{\pi(2n+1)}$$

alors on calcule que $f(z_n) < f(x_n) < f(y_n)$, donc, par le TVI, tout $u \in \left] f(x_n), f(y_n) \right[$ a deux antécédents par f.

▷ Ceci étant vrai pour tout n, on en déduit qu'il n'existe aucun voisinage de 0 où f soit injective.
  En particulier, quel que soit $\varepsilon > 0$, $f :]-\varepsilon, \varepsilon[\to \mathbb{R}$ n'est pas une bijection.

▷ Le problème, c'est que f n'est pas $\mathscr{C}^1$: pour $x \neq 0$,

$$f'(x) = 1 + 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)$$

Donc $f'(x)$ n'a pas de limite quand x tend vers 0.

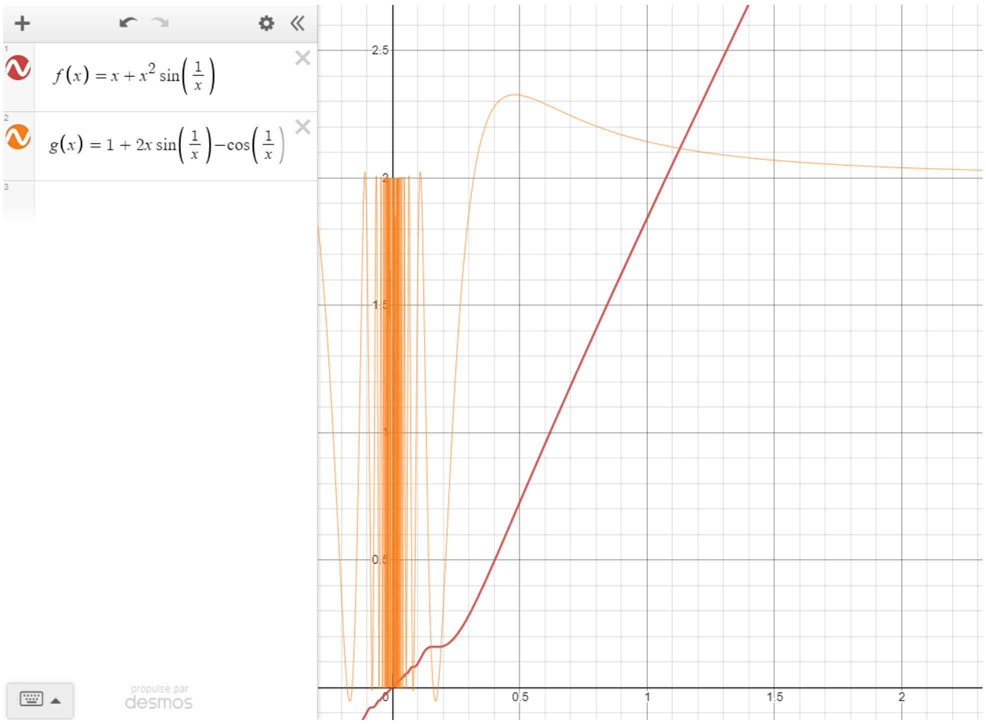


Figure 1: Source: `https://www.desmos.com/calculator/cqkmt4buei`

Sauf que ! J'ai lu récemment sur le blog de Terrence Tao le résultat suivant:

Théorème 3

Soit $U \subset \mathbb{R}^n$ un ouvert de $\mathbb{R}^n$, et $f : U \to \mathbb{R}^n$ une application telle que

- *f est différentiable sur U;*
- *Pour tout $x \in U$, $Df(x)$ est inversible.*

Alors, pour tout $x_0 \in U$, il existe un voisinage V_{x_0} de x_0 et un voisinage $W_{f(x_0)}$ de $f(x_0)$ tels que $f_{|V_{x_0}}$ soit un homéomorphisme[3]

Mais alors, et le contre-exemple ?

$\rightsquigarrow$ En fait, si on y regarde de plus près, le contre-exemple au TIL classique est aussi un contre-exemple à ce nouveau Taorème.

En effet, la contradiction venait du fait que, dans tout intervalle $]-\varepsilon, \varepsilon[$, on pouvait trouver deux points $a_\varepsilon, b_\varepsilon$ tels que $f(a_\varepsilon) = f(b_\varepsilon)$, et donc f n'est bijective sur aucun voisinage de 0.

Mais du coup, d'après le Théorème de Rolle, il doit y avoir un point $c_\varepsilon \in]a_\varepsilon, b_\varepsilon[$ tel que $f'(c_\varepsilon) = 0$.

Et donc, même si la première hypothèse du Taorème est vérifiée, la deuxième ne l'est pas: il n'y a pas d'ouvert contenant 0 sur lequel $Df(x)$ soit inversible en tout point.

Et comment "sauve-t-on" la preuve qu'on a faite ?

$\rightsquigarrow$ Au lieu d'utiliser le théorème du point fixe de Picard-Banach, on utilise un autre théorème de point fixe, le *théorème du point fixe de Brouwer*:

Théorème 4

Point fixe de Brouwer Soit $K \subset \mathbb{R}^n$ un compact convexe. Toute application continue $f : K \to K$ admet un point fixe.

Voilà. Point. Le théroème du point fixe de Brouwer n'est pas évident à démontrer, mais sa simplicité le rend facile à utiliser: par exemple une de ses versions, le théorème de Kakutani, est un ingrédient essentiel de la démonstration de l'existence de l'équilibre général dans le modèle des économies de marché d'Arrow et Debreu. Il sert aussi à démontrer d'importants théorèmes de topologie, comme le théorème de la boule chevelue [4]

Remarquons quand même que, par conséquent, contrairement au TIL, le Taorème d'Inversion Locale ne peut pas s'étendre gratuitement aux espaces de Banach: il est limité à $\mathbb{R}^n$. Ce qui est déjà pas mal, mais certaines applications, notamment pour la résolution d'équations dont les inconnues sont des fonctions, nécessitent des Banach plus massifs, comme par exemple les espaces $L^p(\mu)$ dont on reparlera.

Théorème des extrema liés - cas à une contrainte

Peut-on enlever le "L" de TIL ?

Il peut sembler très limitant d'avoir une inversion seulement locale, restreinte à de petits voisinages d'une solution $f(x_0) = y_0$.

On pourrait s'attendre à ce que si $Df(x)$ est inversible pour tout $x \in U$, alors $f : U \to f(U)$ soit un $\mathscr{C}^1$-difféo.

C'est d'ailleurs le cas dans $\mathbb{R}$: si $f'(x) \neq 0$ pour tout x et est continue, alors f' est de signe constant et f est strictement monotone, donc bijective.

Mais ce n'est pas le cas dans $\mathbb{R}^n$:

Contre-exemple: Considérons, sur $U = \mathbb{R}^2 \setminus \{(0,0)\}$, l'application

$$\begin{aligned} f : U &\to \mathbb{R}^2 \\ (x,y) &\mapsto (x^2 - y^2, 2xy) \end{aligned}$$

Alors pour tout $(x,y) \in U$, f est différentiable en (x,y) et

$$\mathrm{Jac}_f(x,y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}, \quad \det(\mathrm{Jac}_f(x,y)) = 4(x^2+y^2) \neq 0,$$

donc $Df(x,y)$ est inversible et f est un $\mathscr{C}^1$-difféomorphisme au voisinage de (x,y).

Mais f n'est pas un $\mathscr{C}^1$-difféo global: pour tout $(x,y) \in U$, alors $f(x,y) = f(-x,-y)$, donc f n'est pas injective.

Inversion globale

Cependant, en ajoutant l'hypothèse d'injectivité, on obtient une version globale:

Théorème 5

Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$ $\mathscr{C}^1$. On suppose que

1. f est injective sur U,

[4]Si, si, vraiment. Chevelue. Wikipedia nous recommande de ne pas le confondre avec le théorème de calvitie sur les propriétés des trous noirs.

2. *Pour tout $x \in U$, $Df(x)$ est inversible.*

Alors $f : U \to f(U)$ est un $\mathscr{C}^1$-difféomorphisme.

Preuve: Pour tout $a \in U$, il existe un voisinage U_a tel que

$$f : U_a \to \underbrace{f(U_a)}_{\text{ouvert}} \text{ est un } \mathscr{C}^1 - \text{difféo.}$$

Donc $f : U \to f(U)$ est bijective et, pour tout $b = f(a) \in f(U)$,

$$f^{-1} : f(U_a) \to U_a \text{ est } \mathscr{C}^1,$$

donc f^{-1} est $\mathscr{C}^1$ en b. Et donc, f est un $\mathscr{C}^1$-difféomorphisme global. □

Exemple: $f : (x, y) \in \mathbb{R}^2 \mapsto (x^3 + x, y - x^2) \in \mathbb{R}^2$ est un $\mathscr{C}^1$-difféomorphisme.

▷ C'est un C^1-difféo local: pour $(x, y) \in \mathbb{R}^2$,

$$Jac_f(x, y) = \begin{pmatrix} 3x^2 + 1 & 0 \\ -2x & 1 \end{pmatrix}, \quad \det(\text{Jac}_f(x, y)) = 3x^2 + 1 \neq 0.$$

▷ $f(\mathbb{R}^2) = \mathbb{R}^2$ et f injective.
Soit $(a, b) \in \mathbb{R}^2$, montrons qu'il existe un unique $(x, y) \in \mathbb{R}^2$ t.q. $f(x, y) = (a, b)$. On a

$$f(x, y) = (a, b) \iff \begin{cases} x^3 + x = a \\ y - x^2 = b \end{cases} \iff \begin{cases} x^3 + x - a = 0 \\ y = b + x^2 \end{cases}$$

Or $x \mapsto x^3 + x - a$ est strictement croissante sur $\mathbb{R}$, il existe un unique $x_0 \in \mathbb{R}$ t.q. $x_0^3 + x_0 - a = 0$.

Alors, en prenant $y_0 = b + x_0^2$, (x_0, y_0) est l'unique solution de $f(x, y) = (a, b)$.

⇝ $f : \mathbb{R}^2 \to \mathbb{R}^2$ est un $\mathscr{C}^1$-difféomorphisme global.

3 Théorème des fonctions implicites

Question: Qu'appelle-t-on résoudre un système d'équations

$$\begin{cases} f_1(u_1 \dots, u_m) & = 0 \\ \vdots \\ f_p(u_1, \dots, u_m) & = 0 \end{cases}$$

quand il y a plus d'inconnues que d'équations (i.e. $m > p$) ?

Dans le cas où $f_1, \dots, f_p$ sont linéaires, on choisit parmi les inconnues $u_1, \dots, u_m$ des *paramètres* $x_1, \dots, x_n$, et on exprime les autres variables (qu'on appelle *inconnues principales*) $y_1, \dots, y_{m-n}$ en fonction des paramètres. La solution est donc décrite par

$$\begin{cases} y_1 & = \phi_1(x_1 \dots, x_n) \\ \vdots \\ y_{m-n} & = \phi_{m-n}(x_1, \dots, x_n) \end{cases}$$

Exemple: Considérons le système

$$(\mathcal{S}) \begin{cases} x+y+z &= 0 \\ 2x+2y+z &= 1 \end{cases} \iff \begin{cases} x+y+z &= 0 \\ \quad x+y &= 1 \end{cases} \iff \begin{cases} z &= -1 \\ y &= 1-x \end{cases}$$

Ainsi, l'ensemble des solutions de $(\mathcal{S})$ est le graphe $\{(x,\varphi(x)), x \in \mathbb{R}\}$ de la fonction $\varphi : x \in \mathbb{R} \mapsto (1-x,-1) \in \mathbb{R}^2$.

Pour savoir combien d'inconnues mettre en paramètre et lesquelles, on dispose de l'algorithme du pivot de Gauss, qui nous permet d'échelonner le système.

$\rightsquigarrow$ Quel est le bon critère pour faire ça quand les équations ne sont pas linéaires ?

Revenons à notre système linéaire, maintenant qu'on sait qu'on va exprimer (y,z) en fonction de x. On a

$$\begin{cases} x+y+z &= 0 \\ 2x+2y+z &= 1 \end{cases} \iff \begin{cases} y+z &= -x \\ 2y+z &= 1-2x \end{cases} \iff \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} -x \\ 1-2x \end{pmatrix}$$

$\rightsquigarrow$ Ce qui permet de conclure, c'est que la matrice $\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ est inversible (d'inverse $\begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}$) ce qui donne

$$\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} -x \\ 1-2x \end{pmatrix} \iff \begin{cases} y = 1-x \\ z = -1 \end{cases}$$

$\rightsquigarrow$ On a donc choisi les inconnues principales en trouvant une sous-matrice carrée inversible dans la matrice qui représente le système.

Quelques notations

Soient U un ouvert de $\mathbb{R}^n \times \mathbb{R}^p$ et f une application $\mathscr{C}^1$ sur U:

$$\begin{aligned} f : U &\to \mathbb{R}^p \\ (x,y) &\mapsto f(x,y) \end{aligned}$$

Pour $(x_0, y_0) \in U$, on note:

- $D_x f(x_0,y_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ la différentielle de l'application $x \mapsto f(x, y_0)$,
- $D_y f(x_0,y_0) \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^p)$ la différentielle de l'application $y \mapsto f(x_0, y)$.

En d'autres termes, on décompose la jacobienne de f en deux parties:

$$\left(\begin{array}{ccc|ccc} \frac{\partial f_1}{\partial x_1}(x_0,y_0) & \dots & \frac{\partial f_1}{\partial x_n}(x_0,y_0) & \frac{\partial f_1}{\partial y_1}(x_0,y_0) & \dots & \frac{\partial f_1}{\partial y_p}(x_0,y_0) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(x_0,y_0) & \dots & \frac{\partial f_p}{\partial x_n}(x_0,y_0) & \frac{\partial f_p}{\partial y_1}(x_0,y_0) & \dots & \frac{\partial f_p}{\partial y_p}(x_0,y_0) \end{array}\right)$$

$$\underbrace{\qquad\qquad\qquad\qquad}_{\text{Matrice de } D_x f(x_0,y_0)} \quad \underbrace{\qquad\qquad\qquad\qquad}_{\text{Matrice de } D_y f(x_0,y_0)}$$

On a donc, pour $(h,k) \in \mathbb{R}^n \times \mathbb{R}^p$,

$$\begin{aligned} Df(x_0,y_0)(h,k) &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0,y_0) h_i + \sum_{i=1}^{p} \frac{\partial f}{\partial y_i}(x_0,y_0) k_i \\ &= D_x f(x_0,y_0)(h) + D_y f(x_0,y_0)(k). \end{aligned}$$

Exemple: Considérons $f : (x,y) \in \mathbb{R}^2 = \mathbb{R} \times \mathbb{R} \mapsto x^2 + y^2 - 1 \in \mathbb{R}$. Alors:

- $D_x f(x_0,y_0)$ est la différentielle en x_0 de l'application $x \mapsto x^2 + y_0^2 - 1$. Or cette application est une fonction dérivable $\mathbb{R} \to \mathbb{R}$, donc, d'après le lien dérivée-différentielle, c'est l'application linéaire

$$D_x f(x_0,y_0) : h \in \mathbb{R} \mapsto \frac{\partial f}{\partial x}(x_0,y_0) h = 2x_0 h$$

- $D_y f(x_0,y_0)$ est la différentielle en y_0 de l'application $y \mapsto x_0^2 + y^2 - 1$. Donc

$$D_y f(x_0,y_0) : k \in \mathbb{R} \mapsto \frac{\partial f}{\partial y}(x_0,y_0) k$$

- Au total, on a bien

$$Df(x_0,y_0)(h,k) = \frac{\partial f}{\partial x}(x_0,y_0) h + \frac{\partial f}{\partial y}(x_0,y_0) k = D_x f(x_0,y_0) h + D_y f(x_0,y_0) k$$

$\rightsquigarrow$ Dans le cas des fonctions $\mathbb{R}^2 \to \mathbb{R}$, $D_x f$ et $D_y f$ sont données par les dérivées partielles.

Heuristique

Soit $f : U \subset \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^p$, et $(x_0, y_0) \in U$ tel que $f(x_0, y_0) = 0_{\mathbb{R}^p}$.

On veut résoudre l'équation $f(x, y) = 0$ près de (x_0, y_0) en fonction de x, autrement dit on cherche $\varphi : V \subset \mathbb{R}^n \to \mathbb{R}^p$ telle que

$$f(x, y) = 0 \iff y = \varphi(x)$$

Or, au voisinage de (x_0, y_0), on a

$$\begin{aligned} f(x, y) &= f(x_0, y_0) + Df(x_0, y_0)(x - x_0, y - y_0) + \text{chouïa} \\ &= D_x f(x_0, y_0)(x - x_0) + D_y f(x_0, y_0)(y - y_0) + \text{chouïa} \end{aligned}$$

Donc l'équation $f(x, y) = 0$ se réécrit

$$D_y f(x_0, y_0)(y - y_0) = -D_x f(x_0, y_0)(x - x_0) + \text{chouïa}$$

$\rightsquigarrow$ Si $D_y f(x_0, y_0)$ est inversible, il y a une unique solution

$$y = \varphi(x) = y_0 - (D_y f(x_0, y_0)^{-1} D_x f(x_0, y_0))(x - x_0) + ...$$

$\rightsquigarrow$ C'est le théorème des fonctions implicites qui va nous permettre de rigorifier tout ça:

Théorème 6

Soit $U \subset \mathbb{R}^n \times \mathbb{R}^p$ un ouvert, et $f : U \to \mathbb{R}^p$ une application $\mathscr{C}^1$. Soit $(x_0, y_0) \in U$. On suppose:

- $f(x_0, y_0) = 0$
- $D_y f(x_0, y_0) : \mathbb{R}^p \to \mathbb{R}^p$ *est inversible*

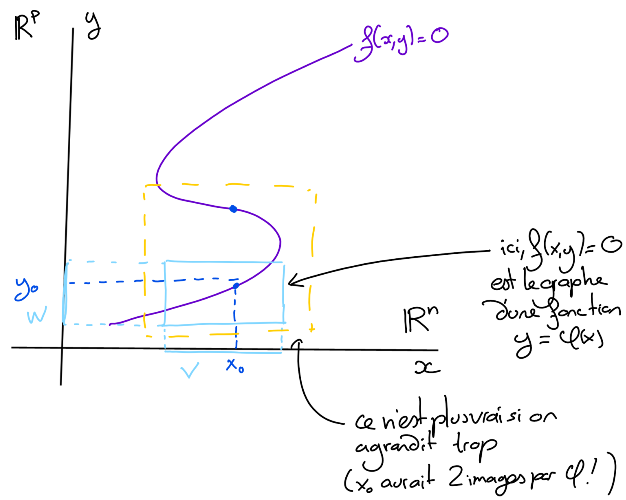
Alors il existe un voisinage $V \subset \mathbb{R}^n$ de x_0, un voisinage $W \subset \mathbb{R}^p$ de y_0 et une application $\mathscr{C}^1$ $\varphi : V \to W$ tels que

- $V \times W \subset U$
- *Pour tout* $(x, y) \in V \times W$, $D_y f(x, y)$ *est inversible*
- $((x, y) \in V \times W \text{ et } f(x, y) = 0) \iff (x \in V, y = \varphi(x))$.

De plus, pour $x \in V$, $D\varphi(x) = -(D_y f(x, \varphi(x))^{-1} \circ D_x f(x, \varphi(x)))$.

Remarque: Plus généralement, comme le TIL, ce théorème s'applique aussi à une application $\mathscr{C}^1$ $f : U \subset E \times F \to G$ où E, F, G sont des espaces de Banach.

Dans le cas, le théorème s'applique si $f(x_0, y_0) = 0_G$ et $D_y f(x_0, y_0)$ est un isomorphisme linéaire entre F et G.



Exemple: Soit $f : (x, y) \in \mathbb{R} \times \mathbb{R} \mapsto x^2 + y^2 - 1 \in \mathbb{R}$. On a

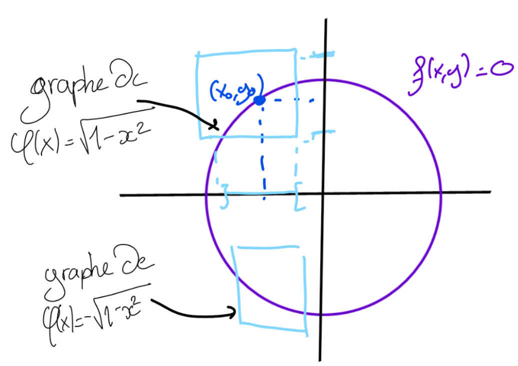
$$\frac{\partial f}{\partial y}(x_0, y_0) = 2y_0$$

On a vu que

$$D_x f(x_0, y_0)(h) = \frac{\partial f}{\partial x}(x_0, y_0)h = 2x_0 h \text{ et } D_y f(x_0, y_0)(k)\frac{\partial f}{\partial y}(x_0, y_0)k = 2y_0 k$$

donc $D_y f(x_0, y_0)$ est inversible ssi $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$, c'est-à-dire ici $y_0 \neq 0$. Du coup:

- Soit (x_0, y_0) tel que $f(x_0, y_0) = 0$, avec $y_0 > 0$ et $x_0 \neq \pm 1$. Alors le TFI s'applique, et d'un autre côté on voit directement qu'il existe un intervalle ouvert V autour de x_0 et un intervalle W sur lequel $y > 0$, tels que sur $V \times W$, $f(x, y) = 0 \iff y = \sqrt{1 - x^2}$.
- On a alors aussi $f(x_0, -y_0) = 0$ et $D_y f(x_0, -y_0) = -2y_0$ inversible. Et on trouve de même qu'il existe aussi un voisinage V' de x_0 et W' de $-y_0$ tels que sur $V' \times W'$, $f(x, y) = 0 \Leftrightarrow y = -\sqrt{1 - x^2}$.
- Si $(x_0, y_0) = (\pm 1, 0)$, il n'y a aucun voisinage de x_0 sur lequel $f(x, y) = 0$ soit le graphe d'une fonction de x.



En général, on ne connaît pas l'expression de la fonction φ: on sait seulement qui est $D\varphi(x)$.

Parfois, on peut en déduire des informations sur φ. Par exemple, dans l'exemple précédent, on a via le TFI

$$\begin{aligned} D\varphi(x) &= -D_y f(x,\varphi(x)) \circ D_x f(x,\varphi(x)) = -\frac{2x}{2\varphi(x)} \\ &\text{donc } \varphi'(x)\varphi(x) = -x \\ &\text{donc } \frac{1}{2}(\varphi(x)^2)' = -x \\ &\text{donc } \varphi(x)^2 = c - x^2 \end{aligned}$$

Or $c = \phi(x)^2 + x^2 = y^2 + x^2 = 1$, donc, si $\varphi(x) > 0$, on retombe sur $\varphi(x) = \sqrt{1-x^2}$.

Preuve du TFI

On va voir qu'on peut se ramener au théorème d'inversion locale. Pour cela, on pose

$$\begin{aligned} g : U \subset \mathbb{R}^n \times \mathbb{R}^p &\to \mathbb{R}^n \times \mathbb{R}^p \\ (x,y) &\mapsto (x, f(x,y)) \end{aligned}$$

Du coup, $f(x,y) = 0 \iff g(x,y) = (x,0)$. On va voir qu'on peut appliquer le TIL à g. En effet, avec $u_0 = (x_0, y_0)$:

$$Dg(u_0)(h,k) = (h, Df(u_0)(h,k)) = (h, D_x f(u_0)h + D_y f(u_0)k)$$

Montrons que $Dg(u_0)$ est inversible. Soient $(a,b) \in \mathbb{R}^n \times \mathbb{R}^p$, on a

$$\begin{cases} h & = a \\ D_x f(u_0)h + D_y f(u_0)k & = b \end{cases} \iff \begin{cases} h & = a \\ k & = D_y f(u_0)^{-1}(b - D_x f(u_0)(a)) \end{cases}$$

autrement dit, $Dg(u_0)^{-1}(a,b) = (a, D_y f(u_0)^{-1}(b - D_x f(u_0)(a)))$.

On peut aussi mettre l'inversibilité de $Dg(u_0)$ en évidence en calculant la jacobienne de g:

$$Jac_g(u_0) = \left(\begin{array}{c|c} I_n & 0_{\mathcal{M}_{n,p}(\mathbb{R})} \\ \hline D_x f(x_0,y_0) & D_y f(x_0,y_0) \end{array}\right)$$

donc $\det(Jac_g(u_0)) = \det(D_y f(u_0)) \neq 0$[5].

$\rightsquigarrow$ Du coup, par le TIL, il existe un voisinage $\tilde{U}$ de (x_0,y_0) tel que $g : \tilde{U} \to g(\tilde{U})$ soit un $\mathscr{C}^1$-difféomorphisme.

Remarque: Sur $\tilde{U}$, $\det(Jac_g(u)) = \det(D_y f(u)) \neq 0$, donc $D_y f(x,y)$ reste inversible.

Soient $\tilde{V}, W$ des voisinages de x_0, y_0 respectivement tels que $\tilde{V} \times W \subset \tilde{U}$, alors $g(\tilde{V} \times W)$ est un ouvert contenant $(x_0, 0)$. Il existe donc un voisinage V de x_0 tel que $V \times \{0\} \subset g(\tilde{V} \times W)$.

[5] On peut calculer ça en développant par rapport à la première ligne, puis la deuxième, etc...jusqu'à avoir "fait disparaître" le bloc I_n

Alors, pour tout $x \in V$, il existe un unique $y \in W$ tel que

$$g(x,y) = (x,0) \text{ i.e. } (x,y) = g^{-1}(x,0)$$

Ainsi, la fonction $\mathscr{C}^1$ $\varphi : V \to W$ donnée par la deuxième composante de g^{-1} répond à la question.

De plus, pour tout $x \in V$, on a $f(x,\varphi(x)) = 0_{\mathbb{R}^p}$. Posons $\alpha : x \in \mathbb{R}^n \mapsto f(x,\varphi(x))$. Alors α est différentiable et on a

- d'un côté, $\alpha(x) = 0_{\mathbb{R}^p}$ pour tout $x \in \mathbb{R}^n$ donc $D\alpha(x) = 0_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)}$
- d'un autre côté, via la formule de composition,

$$D_x f(x,\varphi(x))(h) + D_y f(x,\varphi(x)) \circ D\varphi(x)(h) = 0$$

Ce qui, au total, donne

$$D\varphi(x)(h) = -D_y f(x,\varphi(x))^{-1} \circ D_x f(x,\varphi(x))(h).$$

comme attendu. Petit carré. $\square$