Théorème d'inversion locale

Léo Daures

Leçons 205, 214, 215, 226

Référence : le poly de cours de Karine Beauchard

1 En une dimension

En analyse réelle, pour des fonctions à valeur dans $\mathbb{R}$, on dispose du théorème suivant :

Théorème 1. *Soit I un intervalle ouvert de $\mathbb{R}$ et $a \in I$. Soit $f : I \to \mathbb{R}$ une fonction $\mathcal{C}^1$. Si $f'(a) \neq 0$, alors sur un intervalle $J \subset I$ contenant a, $f' \neq 0$ et $f_{|J} : J \to f(J)$ est bijective.*

Cette propriété semble naturelle : autour d'un point où f est strictement monotone, elle est injective, et donc inversible à condition de prendre des ensembles de départ et d'arrivée de taille raisonable. En fait, le théorème traduit la condition de stricte monotonie en une condition d'inversibilité de $f'(a)$ dans $\mathbb{R}$ (le réel $f'(a)$ doit être *non nul*). Le théorème d'inversion locale généralise ce résultat en dimensions quelconques, pourvu que les espaces de départ et d'arrivée conservent leur propriété de complétude.

2 Le théorème d'inversion locale

Théorème 2. *Soient $(E, ||\cdot||_E)$ et $(F, ||\cdot||_F)$ des Banach, Ω un ouvert de E, $a \in \Omega$, et une fonction $f \in \mathcal{C}^1(\Omega, F)$ telle que $\mathrm{d}f(a)$ est bijective. Alors, il existe :*

- *V un voisinage ouvert de a dans Ω*
- *W un voisinage ouvert de $f(a)$ dans F*

tels que $f_{|V} : V \to W$ soit un $\mathcal{C}^1$-difféomorphisme.
De plus, si f est $\mathcal{C}^k$, alors $f_{|V}^{-1}$ l'est aussi sur W.

On observe que conformément au théorème en dimension 1, c'est une condition d'inversibilité de la dérivée en un point qui a pour conséquence la bijectivité locale de f. Cela traduit que proche du point a considéré, f est proche de son linéarisé $f(a) + \mathrm{d}f(a)(\cdot)$; celui-ci étant bijectif, f doit l'être aussi à condition de ne pas trop s'en éloigner.

Corollaire 1 (dimension finie)**.** *Si $E = F = \mathbb{R}^n$, alors la condition $\det(\mathrm{d}f(a)) \neq 0$ peut remplacer la condition $\mathrm{d}f(a)$ bijective (cela revient au même).*

3 Démonstration

Commençons par monter la bijectivité locale de f. On s'occupera du caractère $\mathcal{C}^1$ de son inverse après.

3.1 Étape 1 : bijectivité de f

Comme pour la beaucoup de résultats d'existence (montrer que $f_{|V}$ est bijective revient à montrer qu'il existe un antécédent à tout point), on va utiliser le théorème du point fixe de Banach-Picard que l'on rappelle ici partiellement et sans le démontrer :

Théorème 3. *Soit (X, d_x) un espace complet et soit $\Phi : X \to X$ contractante (c'est à dire lipschitzienne, de constante de Lipschitz $K < 1$). Alors Φ admet un unique point fixe x_* sur X, et de plus x_* est la limite de la suite des itérées définie par $x_{n+1} = \Phi(x_n)$ et $x_0 \in X$ (et ce pour tout $x_0 \in X$).*

Considérons un $y \in F$ correctement choisi (nous verrons ce que signifie "correctement"). L'existence et l'unicité de l'antécédent de y sur un certain voisinage est une conséquence, jusque dans sa construction, du théorème de point fixe de Banach-Picard. En supposant que cet antécédent existe, la stratégie pour le trouver est la suivante :

- Choisir un point x_1 de E (plus particulièrement, d'un ouvert de E que l'on déterminera par des considérations plus techniques) comme antécédent potentiel de y.
- S'il s'agit vraiment de l'antécédent de y, c'est un gigantesque coup de chance et on n'a plus rien à faire.
- Dans le cas contraire, on veut savoir à quel point on est loin d'avoir trouvé l'antécédent de y. Le vecteur $y - f(x_1)$ (vecteur de F) mesure exactement de combien $f(x_1)$ est tombé loin de y. Mieux : le vecteur $(\mathrm{d}f(a))^{-1}(y - f(x_1))$ (vecteur de E) donne, à peu de choses près, comment il faudrait déplacer notre x_1 pour en faire l'antécédent de y. En effet, si f était linéaire, ce vecteur serait exactement $f^{-1}(y - f(x_1))$, c'est-à-dire $f^{-1}(y) - x_1$, soit exactement ce qu'il manque à x_1 pour atteindre $f^{-1}(y)$! Malheureusement, toutes les fonctions ne sont pas linéaires, et il est impossible dans le cas général de connaître exactement $f^{-1}(y) - x_1$. Mais c'est pour cela qu'on s'intéresse à $\mathrm{d}f(a)$: si l'on se place assez proche de a, le linéarisé $f(a) + \mathrm{d}f(a)(\cdot)$ approche raisonnablement bien f, et $(\mathrm{d}f(a))^{-1}(y - f(x_1))$ approche raisonnablement bien $f^{-1}(y) - x$. On a donc de bonnes raisons de penser que $x_1 + (\mathrm{d}f(a))^{-1}(y - f(x_1))$ est un peu plus proche de $f^{-1}(y)$ que x_1.
- Poser $x_2 = x_1 + (\mathrm{d}f(a))^{-1}(y - f(x_1))$ et recommencer, en croisant les doigts pour que le processus converge bien vers $f^{-1}(y)$.

Et, en effet, sous certaines conditions, le processus va bien converger ! Pourquoi en être aussi sûr ? Parce que la suite (x_k) ainsi construite ressemble furieusement à la suite des itérées du théorème du point fixe de Banach-Picard ! En posant $\Phi_y(x) = x + (\mathrm{d}f(a))^{-1}(y - f(x))$ sur un fermé convenable, on a $x_k = \Phi_y^k(x_1)$. Il suffira dès lors que Φ_y soit contractante (sur un ensemble complet) pour qu'elle admette un unique point fixe x_* vers lequel (x_k) convergera. Ce point fixe nous intéresse plus qu'un peu, car il vérifie $0 = \Phi_y(x_*) - x_* = \mathrm{d}f(a)^{-1}(y - f(x_*))$, et donc $f(x_*) = y$!

Tout cela n'était qu'heuristique, mais on a maintenant une meilleure idée sur la méthode à adopter pour trouver un antécédent à y. On cherche donc un fermé $W \subset F$ et un fermé $V \subset E$ sur lequel toutes les applications $(\Phi_y)_{y \in W}$ sont contractantes et laissent V stable. Commençons ici le parcours technique de la preuve.

Par le théorème d'isomorphisme de Banach, $\mathrm{d}f(a)^{-1}$ est continue donc il existe une constante $C > 0$ telle que $\forall h \in F, ||\mathrm{d}f(a)^{-1}(h)||_F \leqslant C||h||_F$. Alors,

1. il existe $r_1 > 0$ tel que $\overline{\mathcal{B}}_E(a, r) \subset \Omega$ (Ω ouvert)
2. il existe $r_2 \in]0, r_1[$ tel que $\forall x \in \overline{\mathcal{B}}_E(a, r),\ ||f(x) - f(a) - \mathrm{d}f(a)(x - a)||_F \leqslant \frac{1}{2C}||x - a||_E$ (définition de la différentielle en a)
3. il existe $r_3 \in]0, r_1[$ tel que $\forall x \in \overline{\mathcal{B}}_E(a, r),\ ||\mathrm{d}f(x) - \mathrm{d}f(a)||_{\mathcal{L}(E,F)} \leqslant \frac{1}{2C}$ (continuité en a de $x \mapsto \mathrm{d}f(x)$)

On considère donc naturellement $r := \min\{r_1, r_2, r_3\}$, qui vérifie les trois propriétés. Posons $V = \overline{\mathcal{B}}_E(a, r)$ et $W = \overline{\mathcal{B}}_F(f(a), \frac{r}{2C})$. On n'a pas choisi r par hasard : si $y \in W$,

$$\Phi_y : \begin{cases} V & \to E \\ x & \mapsto x + (\mathrm{d}f(a))^{-1}(y - f(x)) \end{cases}$$

a les propriétés qu'on attend d'elle :

Stabilité de V : $\forall x \in V, \Phi_y(x) - a = x - a + (\mathrm{d}f(a))^{-1}(y - f(x)) = (\mathrm{d}f(a))^{-1}(y) + (\mathrm{d}f(a))^{-1}(\mathrm{d}f(a)(x-a) - f(x))$ donc $||\Phi_y(x) - a||_E \leqslant C||y||_F + C||\mathrm{d}f(a)(x-a) - f(x)||_F$ par définition de la constante C, puis $||\Phi_y(x) - a||_E \leqslant C||y||_F + C\frac{1}{2C}||x-a||_E$ par la propriété (2.), et enfin comme y et x sont dans des boules bien spécifiques, $||\Phi_y(x) - a||_E \leqslant C\frac{r}{2C} + C\frac{r}{2C} = r$. Cela montre que la boule $V = \overline{\mathcal{B}}_E(a, r)$ est stable par Φ_y

Contraction : Pour tout $x \in V$, $||\mathrm{d}\Phi_y(x)||_{\mathcal{L}(E)} = ||\mathrm{d}f(a)^{-1} \circ (\mathrm{d}f(a) - \mathrm{d}f(x))||_{\mathcal{L}(E)} \leqslant C\frac{1}{2C}$ par la propriété (3.). Donc Φ_y est lipschitzienne de constante $1/2$, donc contractante.

Complétude : V est complet comme fermé d'un espace complet.

Pour tout $y \in W$, on peut appliquer le théorème du point fixe de Banach-Picard à l'application $\Phi_y : V \to V$ contractante dans V complet. Φ_y a donc un unique point fixe x qui vérifie $x = x + (\mathrm{d}f(a))^{-1}(y - f(x))$, autrement dit $0 = (\mathrm{d}f(a))^{-1}(y - f(x))$. Comme $(\mathrm{d}f(a))^{-1}$ est bijective, on a $0 = y - f(x)$!

On a maintenant, pour tout point y de W un unique antécédent (dans V) de y par f. Donc f est une bijection de $V' = V \cap f^{-1}(W)$ sur W !

3.2 Étape 2 : f est un difféomorphisme local

Il suffit de voir que f^{-1} est lipschitzienne donc continue. Alors, par la formule

$$\forall y \in W, \quad \mathrm{d}f^{-1}(y) = \mathrm{d}f(f^{-1}(y))^{-1}$$

on aura le caractère $\mathcal{C}^1$ de f^{-1}, ce qui conclura que f est bien un difféomorphisme $\mathcal{C}^1$ de V' dans W. Montrons que f^{-1} est continue. Soit y_1 et y_2 des points de W, et soient x_1 et x_2 leur antécédent respectif dans V'. On a

$$\begin{aligned}
||x_1 - x_2||_E &= ||\Phi_{y_1}(x_1) - \Phi_{y_2}(x_2)||_E \\
&= ||(x_1 - x_2) + \mathrm{d}f(a)^{-1}(y_1 - f(x_1)) + \mathrm{d}f(a)^{-1}(y_2 - f(x_2)) \\
&= ||\mathrm{d}f(a)^{-1}(\mathrm{d}f(a)(x_1 - x_2) + (y_1 - y_2) - (f(x_1) - f(x_2)))||_E \\
&\leqslant C||\mathrm{d}f(a)(x_1 - x_2) + (y_1 - y_2) - (f(x_1) - f(x_2))||_F \\
&\leqslant C||y_1 - y_2||_F + C||\mathrm{d}f(a)(x_1 - x_2) - (f(x_1) - f(x_2))||_F
\end{aligned}$$

Il reste à estimer $||\mathrm{d}f(a)(x_1 - x_2) - (f(x_1) - f(x_2))||_F$ en fonction de $||y_1 - y_2||_F$. Puisqu'apparaissent dans la même expression $f(x_1) - f(x_2)$ et la différentielle de f, on peut penser à écrire $f(x_1) - f(x_2)$ sous une forme intégrale. En posant $g(t) = f(tx_2 + (1-t)x_1)$, on a $f(x_1) - f(x_2) = \int_0^1 g'$, où par la règle de la chaîne $g'(t) = \mathrm{d}f(tx_2 + (1-t)x_1) - \mathrm{d}f(tx_2 + (1-t)x_1)$. Ceci est licite car l'intégrale de Riemann est correctement définie sur les fonction de $[0, 1]$ à valeurs dans F complet. On a alors :

$$\begin{aligned}
||\mathrm{d}f(a)(x_1 - x_2) - (f(x_1) - f(x_2))||_F &= ||\mathrm{d}f(a)(x_1 - x_2) + \int_0^1 \mathrm{d}f(tx_2 + (1-t)x_1)(x_1 - x_2)dt||_F \\
&= ||\int_0^1 (\mathrm{d}f(tx_2 + (1-t)x_1)) + \mathrm{d}f(a))(x_1 - x_2)dt||_F \\
&\leqslant \int_0^1 \underbrace{||\mathrm{d}f(tx_2 + (1-t)x_1)) + \mathrm{d}f(a)||_{\mathcal{L}(E,F)}}_{\leqslant \frac{1}{2C} \text{ par la propriété (3.)}} \times ||(x_1 - x_2)||_F dt \\
&\leqslant \frac{1}{2C}||x_1 - x_2||_E
\end{aligned}$$

Donc finalement $||x_1 - x_2||_E \leqslant \frac{1}{2}||x_1 - x_2||_E + C||y_1 - y_2||_F$ et donc $||x_1 - x_2||_E \leqslant 2C||y_1 - y_2||_F$. Autrement dit, $||f^{-1}(y_1) - f^{-1}(y_2)||_E \leqslant 2C||y_1 - y_2||_F$, donc f^{-1} est $2C-$lipschitzienne donc continue.