20 Dowód twierdzenia o lokalnej odwracalności

Dowodzimy teraz następującego twierdzenia:

Twierdzenie o lokalnej odwracalności. Załóżmy, że $D \subset \mathbb{R}^n$ jest zbiorem otwartym, a różniczka $d_P F$ odwzorowania $F : D \to \mathbb{R}^n$ klasy C^1 jest odwzorowaniem odwracalnym. Wówczas istnieją otoczenia **otwarte:** U -punktu P oraz W-punktu $Q = F(P)$ takie, że restrykcja $F|_U : U \to W$ jest bijekcją, zaś bijekcja odwrotna $(F|_U)^{-1} : W \to U$ jest również klasy C^1. Jej różniczka $d_Q((F|_U)^{-1})$ jest odwrotnością $d_P F$.

Na początku zauważmy, że dowód można sprowadzić do przypadku, gdy $d_p F = I$ (gdzie I oznacza operator identyczności na $\mathbb{R}^n$). Faktycznie, w razie potrzeby zastąpmy F przez złożenie $F_1 := (d_p F)^{-1} \circ F$. Z reguły łańcucha wynika, że różniczką odwzorowania F_1 w punkcie P_1 jest $(d_p F)^{-1} \circ d_{P_1} F$, gdyż różniczka odwzorowania zewnętrznego w tym złożeniu, czyli $d_{Q_1}(d_p F)^{-1}$ w punkcie $Q_1 = F(P_1)$ (i w każdym innym punkcie) jest równa $(d_p F)^{-1}$. Dla $P_1 = P$ otrzymamy więc $d_P F_1 = I$. Obraz zbioru otwartego przez bijekcję liniową ciągłą T jest zbiorem otwartym. Faktycznie, wiemy, że przeciwobrazy zbiorów otwartych przez odwzorowania ciągłe są otwarte. Obraz $T(U)$ jest w naszym przypadku równy przeciwobrazowi przez T^{-1} zbioru U, zaś T^{-1} -jako odwzorowanie liniowe w przestrzeni skończenie wymiarowej - jest ciągłe. Gdy więc znajdziemy otwarte otoczenie U punktu P, które F_1 odwzorowuje bijektywnie na pewien zbiór otwarty W będący otoczeniem punktu $F_1(P)$, to restrykcja $F|_U = d_P F \circ F_1|_U$ -jako złożenie bijekcji -będzie bijekcją zbioru U na zbiór $d_P F(W)$ -który, jak właśnie zauważyliśmy, będzie też otwarty. Klasa C^1 dla odwzorowania $(F_1|_U)^{-1}$ oznacza jego różniczkowalność oraz ciągłą zależność od punktu $y \in W$ różniczki $d_y(F_1|_U)^{-1}$. Ale $(F_1|_U)^{-1} = (F|_U)^{-1} \circ d_P F$, $(F|_U)^{-1} = (F_1|_U)^{-1} \circ (d_P F)^{-1}$ i z reguły łańcucha, $d_w(F|_U)^{-1} = d_y(F_1|_U)^{-1} \circ (d_P F)^{-1}$, gdzie $w = (d_P F)^{-1}(y)$. Zależność od $y \in W$ (jak również od $w = (d_P F)^{-1}(y)$ różniczki z $(F|_U)^{-1}$ jest więc ciągła.

Bez straty ogólności możemy więc założyć, że $d_P F = I$. Podobnie, jak w przypadku odwzorowań liniowych (poprzedni wykład), decydujące dla odwracalności F w otoczeniu punktu P będzie zachowanie się odwzorowania $\Phi := I - F$. Oczywiście, $d_P \Phi = d_P I - d_P F = 0$, więc klasa C^1 dla F implikuje, że w pewnym otoczeniu otwartym U punktu P dla będzie

$$x \in U \Rightarrow \|d_x \Phi\| < \frac{1}{2}. \tag{1}$$

Możemy dodatkowo założyć, że zbiór U jest wypukły (np. biorąc U jako małą kulę o środku P). Wtedy z twierdzenia o przyrostach wynika, że

$$\|\Phi(x_1) - \Phi(x_2)\| \leqslant \frac{1}{2}\|x_1 - x_2\|, \tag{2}$$

ponieważ $\frac{1}{2}$ jest majorantą dla norm operatorowych różniczek Φ na odcinku łączącym punkty $x_1.x_2$. Z oszacowania (2) wynika, że

$$\|F(x_1) - F(x_2)\| \geqslant \frac{1}{2}\|x_1 - x_2\|. \tag{3}$$

Rzeczywiście, oznaczmy na chwilę $A := F(x_1) - F(x_2)$, $B := x_1 - x_2$. Wówczas mamy wykazać, że $\|A\| \geqslant \frac{1}{2}\|B\|$. Ale (2) mówi, że $\|B - A\| \leqslant \frac{1}{2}\|B\|$. Żądaną nierówność otrzymamy więc z ostatniej z nierówności:

$$\|B\| = \|A + B - A\| \leqslant \|A\| + \|B - A\| \Rightarrow \|A\| \geqslant \|B\| - \|B - A\|.$$

Nierówność (3) gwarantuje różnowartościowość f na zbiorze U.

Pozostaje wykazać otwartość zbioru $W := F(U)$ oraz różniczkowalność $(F|_U)^{-1} : W \to U$. Weźmy więc dowolny punkt $y_0 = F(x_0) \in W$. Wtedy pewna kula $K(x_0, r)$ o środku x_0 i promieniu $r > 0$ zawiera się wraz ze swoim domknięciem: $\overline{K(x_0, r)}$ w zbiorze U.

Wykażemy, że wówczas $F(\overline{K(x_0,r)})$ zawiera kulę $K(y_0,\frac{1}{2}r)$, używając w tym celu twierdzenia Banacha o punkcie stałym. Najpierw wybierzmy i ustalmy dowolny punkt y_1 z kuli $K(y_0,r)$ i niech

$$G(x) := y_1 + \Phi(x) = y_1 + x - F(x).$$

Zauważmy, że G, różniąc się od Φ o wektor stały y_1, spełnia również warunek Lipshitza ze stałą $\frac{1}{2}$ w zbiorze U, czyli jest zwężające, zaś równanie na punkt stały: $G(x_1) = x_1$ oznacza, że $y_1 + x_1 - F(x+1) = x_1$, czyli dokładnie równość $F(x_1) = y_1$. W ten sposób widzimy, jak twierdzenie o punkcie stałym daje rozwiązanie równania. W naszej sytuacji chcemy zastosować twierdzenie Banacha o kontrakcji dla odwzorowania zwężającego $G : U \to \mathbb{R}^n$ na zbiorze domkniętym $K = \overline{K(x_0,r)} = \{x \in \mathbb{R}^n : \|x - x_0\| \leqslant r\}$. W tym celu wystarczy sprawdzić, że $G(K) \subset K$. Jeśli więc $x \in K$, czyli $\|x - x_0\| \leqslant r$, to $\|G(x) - x_0\| = \|y_1 + x - F(x) - x_0\|$ możemy oszacować przez

$$\|y_1 - y_0\| + \|y_0 - x_0 + \Phi(x)\| = \|y_1 - y_0\| + \|\Phi(x) - \Phi(x_0)\| \leqslant r.$$

Faktycznie, każdy ze składników po prawej stronie nie przekracza $\frac{r}{2}$: pierwszy z definicji kuli, drugi -z warunku Lipshitza dla Φ i z nierówności $\|x - x_0\| \leqslant r$.

Pozostaje wykazać różniczkowalność i klasę C^1 dla $H := (F|_U)^{-1}$. Zbadajmy istnienie różniczki w punkcie y_0. Przy $y \to y_0$ możemy przyjąć, że $x := H(y)$ oraz $x_0 = H(y_0)$ Dzięki nierówności (3), $\|y - y_0\| \geqslant \frac{1}{2}\|x - x_0\|$, więc również $x \to x_0$, o ile tylko $y \to y_0$. Z różniczkowalności F mamy

$$y - y_0 = d_{x_0}F(x - x_0) + \alpha(x)\|x - x_0\|, \tag{4}$$

gdzie $\|\alpha(x)\| \to 0$ przy $x \to x_0$. Obkładając ostatnią równość stronami przez odwzorowanie liniowe $d_{x_0}F^{-1}$ i przenosząc na drugą stronę ze znakiem "minus" składnik zawierający $\alpha(x)$ otrzymamy równość

$$H(y) - H(y_0) = (d_{x_0}F)^{-1}(y - y_0) + (d_{x_0}F)^{-1}(-\alpha(x))\|y - y_0\|\frac{\|x - x_0\|}{\|y - y_0\|},$$

która świadczy o różniczkowalności odwzorowania H i o równosci $d_{y_0}H = (d_{x_0}F)^{-1}$. Faktycznie, ułamek występujący na prawej stronie ostatniej równości jest ograniczony przez 2 (dzięki (3)) , więc pomnożony przez $(d_{x_0}F)^{-1}(-\alpha(x))$ zmierza do zera.

Na koniec, ciągła zależność d_yH od y wynika z ciągłej zależności (d_xF) od x i z ciągłości operacji odwracania w zbiorze GL_n wykazanej na poprzednim wykładzie. I jeszcze trzeba tu skorzystać z ciągłosci odwzorowania H: gdy $x \to x_0$, to $y \to y_0$, będącej konsekwencją nierównosci (3) $\square$.

Bijekcje klasy C^1 przekształcają więc zbiory otwarte U w $\mathbb{R}^n$ na zbiory otwarte W i bijekcje do nich odwrotne też są klasy C^1. Odwzorowania takie nazywamy dyfeomorfizmami zbioru U na zbiór W. Powstaje pytanie, czy gdy $U = \mathbb{R}^n$ i różniczka odwzorowania $F : \mathbb{R}^n \to \mathbb{R}^n$ klasy C^1 jest w każdym punkcie jest odwracalna, to czy już F musi być dyfeomorfizmem? Odpowiedź brzmi "NIE", ale są pewne założenia gwarantujące w takiej sytuacji injektywność F. Założenia te można zmaleźć w Twierdzeniu Hadamarda-Levy'ego (założenie wystarczające, to $\sup\{\|(d_xF)^{-1}\| : x \in \mathbb{R}^n\} < \infty$), jednak jego dowód wykracza poza ramy obecnego kursu.

Najważniejsze są dla nas pewne konsekwencje twierdzenia o lokalnej odwracalności. Jedną z nich jest następujące uogólnienie twierdzenia o funkcjach uwikłanych:

Tym razem niech dziedziną odwzorowania $F = (f_1, \dots, f_m)$ klasy C^1 będzie zbiór otwarty $D \subset \mathbb{R}^{k+m}$, zaś $F : D \to \mathbb{R}^m$.

Zmienne P ze zbioru D pogrupujmy w dwa bloki: $\mathbf{x} = (x_1, \ldots, x_k) \in \mathbb{R}^k$ oraz $\mathbf{y} = (y_1, \ldots, y_m)$, tak, że $P = (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{m+k}$. Analogicznie $P_0 = (\mathbf{x}_0, \mathbf{y}_0)$. Niech $d_P^{\mathbf{y}} F$ oznacza macierz kwadratową

$$\begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_m} \end{pmatrix}$$

pochodnych cząstkowych względem zmiennych z bloku $\mathbf{y}$ (oraz wyznaczone przez nią odwzorowanie liniowe $d_P^{\mathbf{y}} F \in \mathcal{L}(\mathbb{R}^m)$). Jest to w pewnym sensie $k-$wymiarowa "różniczka cząstkowa". Odpowiednikiem niezerowania się pochodnej cząstkowej w punkcie P będzie nieosobliwość tej macierzy, czyli niezerowanie się jej wyznacznika.

Twierdzenie o odwzorowaniach uwikłanych. Jeżeli w punkcie P_0 o współrzędnych blokowych $(\mathbf{x_0}, \mathbf{y_0})$ macierz $d_P^{\mathbf{y}} F$ jest nieosobliwa dla naszego odwzorowania klasy C^1 oraz $F(P_0) = 0$, to istnieją otwarte otoczenia: U punktu $\mathbf{x_0}$ w $\mathbb{R}^k$ oraz W -dla punktu $\mathbf{y_0}$ w $\mathbb{R}^m$ i odwzorowanie klasy C^1 -oznaczmy je $G : U \to W$ takie, że dla $\mathbf{x} \in U$ oraz $\mathbf{y} \in W$ warunek $F(\mathbf{x}, \mathbf{y}) = 0$ jest równoważny warunkowi $\mathbf{y} = G(\mathbf{x})$. Innymi słowy, w otoczeniu punktu P przeciwobraz zera przez F jest lokalnie wykresem odwzorowania G klasy C^1 -czyli można rozwikłać równanie $F(\mathbf{x}, G(\mathbf{x})) = 0$.

Dowód: Niech $P_{\mathbf{x}}$ oznacza odwzorowanie $\mathbb{R}^{m+k} \ni (\mathbf{x}, \mathbf{y}) = \mathbf{x} \in \mathbb{R}^k$, zaś $P_{\mathbf{y}}(\mathbf{x}, \mathbf{y}) = \mathbf{y}$. Niech Ψ oznacza zestawienie

$$\Psi = (P_{\mathbf{x}}, F) : \mathbb{R}^{m+k} \ni (\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x}, F(\mathbf{x}, \mathbf{y})) \in \mathbb{R}^{m+k}.$$

Nietrudno zauważyć, że macierz odwzorowania $d_P \Psi$ ma postać macierzy blokowej

$$\begin{pmatrix} I_k & * \\ 0 & d_P^{\mathbf{y}} F \end{pmatrix},$$

gdzie I_k oznacza macierz identyczności rozmiaru $k \times k$, 0 oznacza macierz prostokątną złożoną z samych zer, symbol * oznacza pewne wyrazy dla nas nieistotne. Rozwinięcie kolumnowe Laplace'a dla wyznacznika tej macierzy pokazuje, że jest on taki sam, jak wyznacznik bloku w prawym dolnym rogu, czyli macierzy $d_P^{\mathbf{y}} F$. Jest to więc macierz nieosobliwa dla $P = P_0$ i odwzorowanie Ψ jest w pewnym otoczeniu $U \times W$ punktu P_0 odwracalne. $\Psi^{-1}(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \Lambda(\mathbf{x}, \mathbf{y}))$ dla pewnego odwzorowania klasy C^1 , powiedzmy $\Lambda : \mathbb{R}^{m+k} \to \mathbb{R}^m$. Jak nietrudno sprawdzić, funkcja rozwikłująca G jest teraz dana wzorem $G(\mathbf{x}) = \Lambda(\mathbf{x}, 0)$, bo $F = P_{\mathbf{y}} \circ \Psi$, więc $F(\mathbf{x}, \Lambda(\mathbf{x}, \mathbf{0}) = P_{\mathbf{y}}(\Psi \circ \Psi^{-1})(\mathbf{x}, \mathbf{0}) = \mathbf{0}$. Regularność klasy C^1 dla takiego odwzorowania G jest oczywista. $\square$.

Rozpisując na współrzędne, $G = (g_1, \ldots, g_m)$ jest więc w otoczeniu U rozwiązaniem klasy C^1 dla układu k równań nieliniowych:

$$\left\{ \begin{array}{ccc} f_1\big(x_1, \ldots, x_k, g_1(x_1, \ldots, x_k), \ldots, g_m(x_1, \ldots, x_k)\big) & = & 0 \\ \cdots\cdots\cdots & \cdots & \cdots \\ \cdots\cdots\cdots & \cdots & \cdots \\ f_m\big(x_1, \ldots, x_k, g_1(x_1, \ldots, x_k), \ldots, g_m(x_1, \ldots, x_k)\big) & = & 0 \end{array} \right.$$

Dalszą konsekwencją twierdzenia o funkcjach uwikłanych jest tak zwane Twierdzenie o Rzędzie. Z grubsza- mówi ono, że gdy rząd różniczki odwzorowania F klasy C^1 jest stały, równy k, to w pewnym otoczeniu można złożyć F z pewnym dyfeomorfizmem H_1 od wewnątrz i z drugim dyfeonorfizmem H_2 od zewnątrz tak, by powstałe złożenie było rzutem na k pierwszych zmiennych:

$$H_2 \circ F \circ H_1(\mathbf{x}, \mathbf{y}) = \mathbf{x}.$$

Zainteresowanych odsyłam np. do książki W. Rudina "*Podstawy analizy matematycznej*"