# Twierdzenie o funkcji odwrotnej

Patryk Prewendowski

12 listopada 2019

Sformułowanie i dowód twierdzenia pochodzą z książki M. Spivak - Analiza na rozmaitościach (wyd. 2005, Wydawnictwo Naukowe PWN).

## 1 Wprowadzenie - przypadek $\mathbb{R} \to \mathbb{R}$.

Niech $f : \mathbb{R} \to \mathbb{R}, f \in C^1, x_0 \in \mathbb{R}, f'(x_0) \neq 0$. Wtedy istnieją pewne otoczenie $x_0 \in U$ i funkcja $g : U \to R$, takie, że $g = f^{-1}$ na $U$.

Załóżmy dla przykładu, że $f'(x_0) > 0$. $f'(x)$ jest ciągła, więc istnieją $a, b$ takie, że $f(x) > 0$ dla $x \in (a, b)$. Funkcja jest rosnąca, zatem różnowartościowa, z tego powodu odwracalna, zatem istnienie funkcji $g$ zostało wykazane. Jeżeli założymy, że $g$ jest różniczkowalna (a jest, czego dowód łatwo znaleźć w różnych źródłach, a nawet nieco niżej - w nieco bardziej ogólnym sformułowaniu), to z reguły łańcucha: $1 = x' = (g(f(x)))' = f'(x)g'(f(x))$, zatem $f'(x) = g'(f(x))^{-1}$.

(Tutaj brakuje ładnej grafiki, np. funkcja $f(x) = x \cdot \sin(x)$ i lokalne przedziały, w których jest odwracalna.)

## 2 Przypadek ogólny

### 2.1 Sformułowanie

Niech $f : \mathbb{R}^n \to \mathbb{R}^n$ ma ciągłą pochodną na zbiorze otwartym zawierającym $a$ oraz $\det f'(a) \neq 0$. Wtedy istnieje zbiór otwarty $V$ zawierający $a$ i zbiór otwarty $W$ zawierający $f(a)$ takie, że $f : V \to W$ ma ciągłą funkcję odwrotną $f^{-1} : W \to V$, która jest różniczkowalna i dla wszystkich $y \in W$ spełnia

$$\left(f^{-1}\right)'(y) = \left[f'(f^{-1}(y))\right]^{-1} \tag{1}$$

## 2.2 Wprowadzenie

Niech $A \subset \mathbb{R}^n$ będzie kulą i niech $f : A \to \mathbb{R}^n$, $f \in C^1$. Jeżeli istnieje liczba $M$ taka, że dla każdego $x \in \text{int}A$ zachodzi $\frac{\partial f_i}{\partial x_j}(x) \leqslant M$, to dla wszystkich $x, y \in A$ zachodzi:

$$|f(x) - f(x)| \leqslant n^2 M|x - y| \tag{2}$$

Dodatkowo zauważmy, że dla liniowego odwzorowania $F : \mathbb{R}^n \to \mathbb{R}^n$ istnieje liczba $M$ taka, że $|T(x)| \leqslant M|x|$ dla każdego $x \in \mathbb{R}^n$.

W każdym punkcie pochodna odwzorowania liniowego $F : \mathbb{R}^n \to \mathbb{R}^n$ to $m(F)$, czyli macierz tego odwzorowania. Inaczej: dla macierzy $M \in M_{nxn}(\mathbb{R})$ pochodna odwzorowania zadanego tą macierzą (ozn. $F_M$ - odwzorowanie zadane przez $M$) to $M$, co możemy zapisać: $DF_M(x) = M$ dla dowolnego $x$.

Ponadto odnotujmy, że funkcja $\det N$ jest ciągła dla macierzy kwadratowych dowolnego wymiaru.

Będziemy stosować oznaczenia $x = (x_1, x_2, \ldots, x_n), y = (y_1, \ldots, y_n), f(x) = (f_1(x), f_2(x), \ldots, f_n(x))$.

Wszędzie, gdzie pojawia się $x_0$, $x_1$, $x_2$, $y_1$ lub $y_2$ (oprócz akapitu wyżej), mam na myśli punkt w przestrzeni $n$-wymiarowej, nie współrzędne! Jeżeli będę miał na myśli współrzędne, to będę pisał po prostu $x_i$ lub $y_i$ (w zależności od kontekstu będą się pojawiać indeksy $i$ lub $j$ lub inne). Gdyby w pewnych miejscach były wątpliwości co oznaczają poszczególne symbole - śmiało kontaktować się.

## 2.3 Dowód twierdzenia

Traktujemy macierz $\lambda = Df(a)$ jako macierz odwzorowania liniowego, które będziemy oznaczać $F_\lambda$.

### 2.3.1 Podstawowe ustalenia

Reguła łańcucha mówi nam, że $D(f \circ g)(a) = Df(g(a)) \cdot Dg(a)$. Dla odwzorowania zadanego macierzą $\lambda$ mamy: $D(F_\lambda \circ g)(a) = DF_\lambda(g(a)) \cdot Dg(a) = \lambda \cdot Dg(a)$.

Bez straty ogólności możemy założyć, że $\lambda$ jest identycznością. Dlaczego?

Otóż jeżeli $F_\lambda \neq \text{Id}_{\mathbb{R}^n}$, to wystarczy wtedy rozpatrzyć funkcję $f' = F_\lambda^{-1} \circ f = F_{\lambda^{-1}} \circ f$. Taka funkcja $f'$ spełnia $D(f')(a) = \text{Id}_{\mathbb{R}^n}$.

Z prawdziwość twierdzenia dla $f'$ wynikałaby prawdziwość twierdzenia $f$. Jeżeli $f'$ jest odwracalne na pewnym zbiorze, to $f$ również musi być odwracalne. Jeżeli mamy $D\left(\left(F_{\lambda^{-1}} \circ f\right)^{-1}\right)(a) = \left(D\left(F_{\lambda^{-1}} \circ f\right)(a)\right)^{-1}$, to z równości $D\left(\left(F_{\lambda^{-1}} \circ f\right)^{-1}\right)(a) = D\left(f^{-1}\right)(a) \cdot \lambda$ oraz $\left(D\left(F_{\lambda^{-1}} \circ f\right)(a)\right)^{-1} = Df(a)^{-1} \cdot \lambda$, a także z odwracalności $\lambda$ wynikałoby, że $(Df(a))^{-1} = D(f^{-1})(a)$.

Załóżmy więc, że $F_\lambda = \mathrm{Id}_{\mathbb{R}^n}$.

Niech $f(a+h) = f(a)$ dla pewnego $h$. Wtedy:

$$\frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} = 1$$

Ale jako, że $\lambda$ jest pochodną, to:

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

Dlatego istnieje taki zbiór domknięty $U$ zawierający $a$ w swoim wnętrzu, że dla $x \in U, x \neq a$:

$$f(x) \neq f(a) \tag{3}$$

Skoro $f$ ma ciągłą pochodną na zbiorze otwartym zawierającym $a$ to możemy założyć, że dla wszystkich $i, j$ i wszystkich $x \in U$:

$$\det f'(x) = \det Df(x) \neq 0 \tag{4}$$

$$\left| \frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a) \right| < \frac{1}{2n^2} \tag{5}$$

Pierwsze - z ciągłości wyznacznika, drugie - z ciągłości pochodnych cząstkowych.

Chwilowo rozpatrzmy funkcję pomocniczą, $h : U \to \mathbb{R}^n$.

Zauważmy, że dla dowolnych $i, j$ mamy $\frac{\partial h_i}{\partial x_j}(a) = \frac{\partial f_i}{\partial x_j}(a) - \mathrm{Id}(a) = 0$. Pochodne cząstkowe $h$ są zerowe w punkcie $a$, wobec tego na odpowiednio bliskim otoczeniu punktu $a$ spełniają $\left|\frac{\partial h_i}{\partial x_j}(x)\right| < \frac{1}{2n^2}$. Będziemy zatem stosować (2) do funkcji $h$, a z (2) otrzymujemy:

$$|g(x_1) - g(x_2)| \leqslant n^2 \frac{1}{2n^2} |x_1 - x_2|$$

Zatem:

$$|f(x_1) - x_1 - (f(x_2) - x_2)| \leqslant \frac{1}{2} |x_1 - x_2|$$

dla $x_1, x_2 \in U$. Ponieważ z nierówności trójkąta:

$$|x_1 - x_2| - |f(x_1) - f(x_2)| \leqslant |f(x_1) - x_1 - (f(x_2) - x_2)|$$

to dla $x_1, x_2 \in U$:

$$|x_1 - x_2| \leqslant 2|f(x_1) - f(x_2)| \tag{6}$$

Zauważmy, że obraz brzegu $U$ jest zbiorem zwartym i nie zawiera $f(a)$, dlatego istnieje taka liczba $d > 0$, że $|f(a) - f(x)| \geqslant d$ dla każdego $x$ z brzegu $U$. (Tutaj przydałby się jakiś rysunek!) Niech $W = \{y : |y - f(a)| < \frac{1}{2}d\}$. Jeżeli $y \in W$ i $x$ należy do brzegu $U$, to ponownie z nierówności trójkąta:

$$|y - f(a)| < |y - f(x)| \tag{7}$$

Pokażemy, że dla każdego $y \in W$ istnieje dokładnie jeden taki punkt $x$ z wnętrza $U$, że $f(x) = y$. Ustalmy $y \in W$.

### 2.3.2 Lokalna surjektywność

Rozważmy funkcję $g : U \to \mathbb{R}$ określoną wzorem:

$$g(x) = |y - f(x)|^2 = \sum_{i=1}^{n} (y_i - f_i(x))^2$$

Funkcja ta jest ciągła i dlatego przyjmuje minimum na $U$. Na mocy (7) dla $x$ z brzegu $U$ mamy $g(a) < g(x)$, dlatego $g$ nie przyjmuje minimum na brzegu $U$. Zatem istnieje taki punkt $x_0$ z wnętrza $U$, że $\frac{\partial g}{x_j}(x_0) = 0$ dla każdego $1 \leqslant j \leqslant n$, czyli z reguły łańcucha:

$$\sum_{i=1}^{n} 2(y_i - f_i(x_0)) \cdot \frac{\partial f_i}{\partial x_j}(x_0) = 0$$

Ale z drugiej strony warunek ten oznacza, że:

$$2Df(x_0)(y - f(x_0)) = 0$$

Z (4) i z faktu, że $x_0 \in U$ mamy, że $\det Df(x_0) \neq 0$, czyli macierz ta jest odwracalna - ma trywialne jądro, wobec tego $y - f(x_0) = \vec{0}$. Zatem $y = f(x_0)$, zatem istnienie takiego punktu zostało wykazane.

### 2.3.3 Lokalna różnowartościowość

Jedyność wynika natychmiast z (6):

$$|x_1 - x_2| \leqslant 2|f(x_1) - f(x_2)|$$

Dla $x_1 \neq x_2$ mamy: $0 < |x_1 - x_2| \leqslant 2|f(x_1) - f(x_2)|$, zatem $f(x_1) \neq f(x_2)$.

### 2.3.4 Ciągłość funkcji odwrotnej

Niech $V = \mathrm{int} U \cap f^{-1}[W]$. Pokazaliśmy, że $f : V \to W$ ma funkcję odwrotną $f^{-1} : W \to V$. Możemy zapisać (6) $|x_1 - x_2| \leqslant 2|f(x_1) - f(x_2)|$ inaczej, mianowicie dla $y_1, y_2 \in W$:

$$|f^{-1}(y_1) - f^{-1}(y_2)| \leqslant 2|y_1 - y_2| \tag{8}$$

Stąd widać, że $f^{-1}$ jest funkcją ciągłą.

### 2.3.5 Różniczkowalność funkcji odwrotnej

Niech $x_1 \in V$, użyjemy oznaczenia $\mu = Df(x_1)$ i pokażemy, że $f^{-1}$ jest różniczkowalna w $y_1 = f(x_1)$ i ma pochodną $\mu^{-1}$. $\mu$ ma wyznacznik niezerowy, zatem traktowana jako odwzorowanie liniowe jest odwracalna, stąd $\mu^{-1}$ istnieje. Mamy

$$f(x) = f(x_1) + \mu(x - x_1) + \varphi(x - x_1) \tag{9}$$

gdzie

$$\lim_{x_1 \to x} \frac{|\varphi(x - x_1)|}{|x - x_1|} = 0 \tag{10}$$

A to dlatego, że wystarczy zdefiniować:

$$\varphi(x - x_1) = f(x) - f(x_1) - \mu(x - x_1)$$

Wtedy:

$$0 \leqslant \frac{|\varphi(x - x_1)|}{|x - x_1|} = \frac{|f(x) - f(x_1) - \mu(x - x_1)|}{|x - x_1|}$$

Przy przejściu granicznym dostajemy 0 z założenia, że $\mu$ jest pochodną. (Wystarczy oznaczyć np. $x = x_1 + h$.)

Z równania (9) (przenosimy $f(x_1)$ na drugą stronę i nakładamy $\mu^{-1}$ obustronnie):

$$\mu^{-1}(f(x) - f(x_1)) = x - x_1 + \mu^{-1}(\varphi(x - x_1))$$

Inaczej:

$$x - x_1 - \mu^{-1}(f(x) - f(x_1)) = -\mu^{-1}(\varphi(x - x_1))$$

Możemy to zapisać również następująco:

$$f^{-1}(y) - f^{-1}(y_1) - \mu^{-1}(y - y_1) = -\mu^{-1}\left[\varphi\left(f^{-1}(y) - f^{-1}(y_1)\right)\right]$$

Aby $\mu^{-1}$ było pochodną $f^{-1}$ wystarczy pokazać, że

$$\lim_{y \to y_1} \frac{|\mu^{-1}\left[\varphi\left(f^{-1}(y) - f^{-1}(y_1)\right)\right]|}{|y - y_1|} = 0$$

W tym celu wystarczy pokazać (wstępny fakt 2), że:

$$\lim_{y \to y_1} \frac{|\varphi\left(f^{-1}(y) - f^{-1}(y_1)\right)|}{|y - y_1|} = 0$$

Ale lewą stronę możemy zapisać jako:

$$\frac{|\varphi\left(f^{-1}(y) - f^{-1}(y_1)\right)|}{|f^{-1}(y) - f^{-1}(y_1)|} \cdot \frac{|f^{-1}(y) - f^{-1}(y_1)|}{|y - y_1|}$$

$f^{-1}$ jest ciągła, zatem $f^{-1}(y) \to f^{-1}(y_1)$ gdy $y \to y_1$. Zatem na mocy (10) lewy czynnik dąży do 0. Na mocy (6) natomiast prawy czynnik należy do $[0, 2]$, zatem iloczyn dąży do 0.