18 Wektor normalny, funkcje uwikłane

18.1 Odwzorowania klasy C^1

Zacznijmy od wprowadzenia pojęcia funkcji (i odwzorowań) klasy C^1 określonych na zbiorze otwartym $D \subset \mathbb{R}^d$.

Definicja. Odwzorowanie $F : D \to \mathbb{R}^k$ o współrzędnych $(f_1, f_2, \dots f_k)$ jest klasy C^1, co oznaczamy symbolem $F \in C^1(D, \mathbb{R}^k)$, gdy szystkie pochodne cząstkowe $\frac{\partial f_m}{\partial x_j}$ są ciągłe (w każdym punkcie dziedziny). Gdy $k = 1$, mówimy o funkcjach klasy C^1 i zbiór wszystkich funkcji klasy C^1 na takim zbiorze D oznaczamy $C^1(D)$.

Innymi słowy, klatki macierzy Jacobiego odwzorowania F mają być funkcjami ciągłymi. Oczywiście, suma odwzorowań klasy C^1 oraz iloczyn funkcji o wartościach rzeczywistych, klasy C^1 przez odwzorowanie $F \in C^1(D, \mathbb{R}^k)$ jest klasy C^1. Z reguły łańcucha zapisanej w postaci macierzowej wynika, że złożenie odwzorowań klasy C^1 jest odwzorowaniem klasy C^1. Mamy więc do dyspozycji sporą kolekcję odwzorowań klasy C^1, nie tylko wielomiany, czy zestawienia funkcji 1 zmiennej z $C^1[a, b]$ (krzywe klasy C^1).

Nawiasem mówiąc- gdy $D = (a, b) \subset \mathbb{R}$, to często potrzebna jest ciągłość pochodnych w przedziale domkniętym -nawet jeśli na końcach przedziału są to pochodne jednostronne. Dla wielu zmiennych odpowiednikiem końców przedziału będzie cały brzeg obszaru D. Zamiast mówić o pochodnych "jednostronnych" warto skorzystać z twierdzenia o przełużalności pewnych funkcji ciągłych ze zbioru D (gęstego w swoim domknięciu $\bar{D}$) -na to domknięcie. Warunkiem równoważnym w przypadku ograniczonych zbiorów otwartych $D \subset \mathbb{R}^d$ jest jednostajna ciągłość danych funkcji. Poprzez analogię z przypadkiem gdy $D = (a, b) \subset \mathbb{R}$ możemy więc przyjąć, że $f \in C^1(\bar{D})$, gdy wszystkie pochodne cząstkowe f są jednostajnie ciągłe na zbiorze D, traktując ich wartości w punktach brzegu jako wartości tego ich (jedynego) przedłużenia ciągłego na zbiór $\bar{D}$. Tak jest przypadku zbiorów ograniczonych i wówczas dzięki zwartości $\bar{D}$ uzyskujemy ograniczoność tych pochodnych, co pozwala zdefiniować normą w taki sam sposób, jak w przypadku funkcji jednej zmiennej -jako sumy norm supremowych po zbiorze D tych pochodnych cząstkowych z dodanym modułem $|f(P_0)|$ dla pewnego $P_0 \in D$ - o ile D jest zbiorem spójnym.

Przestrzeń wektorowa macierzy o k wierszach i d kolumnach ma wymiar skończony (równy kd), więc jak już wiemy, wszystkie normy są tu równoważne, zbieżność (np. ciągów) takich macierzy jest równoważna zbieżności każdej z klatek do klatki macierzy granicznej. Możemy więc mówić o ciągłości odwzorowania przypisującego punktowi $P \in D$ różniczkę $d_P F \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^k)$ i taka ciągłość zachodzi wtedy i tylko wtedy, gdy $F \in C^1(D, \mathbb{R}^k)$.

18.2 Wektory normalne, orientacja powierzchni

Równanie płaszczyzny w $\mathbb{R}^3$ przechodzącej przez punkt (x_0, y_0, z_0) ma postać

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0.$$

Jeśli Q oznacza punkt o współrzędnych (x, y, z), zaś $Q_0 = (x_0, y_0, z_0)$, to $(x - x_0, y - y_0, z - z_0) = Q - Q_0$ jest wektorem o początku Q i końcu Q_0, a równanie opisuje zbiór $\{Q \in \mathbb{R}^3 : Q - Q_0 \perp \vec{n}\}$, gdzie $\vec{n}$ jest tzw. wektorem normalnym o współrzędnych (A, B, C). Widzieliśmy to przy opisie płaszczyzny stycznej do wykresu funkcji $f : D \subset \mathbb{R}^2 \to \mathbb{R}$. W przypadku f różniczkowalnej w punkcie $P_0 = (x_0, y_0)$ w równaniu płaszczyzny stycznej mieliśmy $z_0 = f(x_0, y_0), A = \frac{\partial f}{\partial x}(P_0), B = \frac{\partial f}{\partial y}(P_0)$, zaś $C = -1$. W przypadku funkcji różniczkowalnej $F : D \subset \mathbb{R}^3 \to \mathbb{R}$ w punkcie $Q_0 = (x_0, y_0, z_0)$ mogliśmy opisać płaszczyznę styczną do powierzchni

$$S = \{(x, y, z) \in D : F(x, y, z) = 0\} \qquad (1)$$

i wówczas $(A, B, C) = \nabla F(Q_0)$, czyli wektor normalny jest gradientem F. Na przykad, jeśli $F(x, y, z) = f(x, y) - z$, to równanie $F(x, y, z) = 0$ opisuje dokładnie wykres f i faktycznie, $\nabla F(x_0, y_0, z_0) = (\frac{\partial f}{\partial x}(P_0), \frac{\partial f}{\partial y}(P_0), -1)$.

Pewnych powierzchni zakrzywionych nie da się opisać jednym równaniem, za to w otoczniu U każdego z jej punktów mamy jej opis typu $F_U(x, y, z) = 0$ dla pewnej funkcji $F_U : U \to \mathbb{R}$ trzech zmiennych, mającej ciągłe pochodne cząstkowe. Takie powierzchnie nazywamy rozmaitościami różniczkowalnymi, jednak ich teorii nie będziemy opisywać w tym kursie. Dziedziną matematyki badającą rozmaitosci jest geometria różniczkowa.

Jeśli Państwo słyszeli o wstędze Möbiusa, to można wykazać, że nie istnieje tam funkcja ciągła, która każdemu punktowi z tej powierzchni przypisuje niezerowy wektor normalny. Wskazanie takiego "ciągłego, nieznikającego wektora normalnego" determinuje stronę powierzchni. Np. na sferze wektory w kierunku promienia, zwrócone ku środkowi kuli, wyróżniają wewnętrzną stronę płata sfery. Ale przy użyciu kredki, farby lub flamastra warto przekonać się, że powierzchnia wstęgi Möbiusa ma tylko jedną stronę. Jej model uzyskujemy biorąc pasek papieru i przed sklejeniem -przekręcamy jeden z jego końców o 180 stopni. Wektor normalny służyć nam będzie w teorii całek powierzchniowych do określania tzw. orientacji powierzchni.

Powróćmy do powierzchni S opisanej w (1). Czy da się wyrazić ją jako wykres pewnej funkcji f? Niekoniecznie. Np. równanie sfery jednostkowej S^2 ma postać $x^2 + y^2 + z^2 - 1 = 0$, rozwiązania dla zmiennej z są na ogół dwa różne; $z = \pm\sqrt{1 - x^2 - y^2}$ - są to funkcje opisujące dwie półkule: dolną i górną. Jednak w otoczeniu dowolnego punktu na S leżącego poza równikiem lokalnie S jest wykresem funkcji klasy C^1. A co "jest nie tak" z punktami równika? W dowolnym ich otoczeniu nad punktem $(x, y, 0)$ są dwa rozwiązania różniące się znakiem. Czym jeszcze różnią się te punkty z równika od pozostałych? Mamy tam $\frac{\partial F}{\partial z}$ równą zero (bo w każdym punkcie ta pochodna, cząstkowa, to $2z$, a na równiku $z = 0$).

18.3 Funkcje uwikłane

Dla uproszczenia rozważmy najpierw funkcje dwu zmiennych $F : D \to \mathbb{R}$ klasy C^1 w pewnym zbiorze otwartym $D \subset \mathbb{R}^2$. Zamiast $\frac{\partial F}{\partial x}$ będziemy tu używać wygodniejszego oznaczenia F'_x dla pochodnych cząstkowych.

Definicja. Mówimy, że funkcja $g : U \subset \mathbb{R} \to \mathbb{R}$ jest uwikłana w równaniu $F(x, y) = 0$, jeśli $\forall_{x \in U}\ (x, g(x)) \in D$ `oraz` $F(x, g(x)) = 0$.

Na przykład, niech $U = (-1, 1), F : \mathbb{R}^2 \ni (x, y) \to x^2 + y^2 - 1 \in \mathbb{R}$. Funkcjami uwikłanymi są tu $y = -\sqrt{1 - x^2}$ oraz $y = \sqrt{1 - x^2}$, zbiór $\mathbb{T} := F^{-1}(\{0\})$ jest tu okręgiem i poza dwoma punktami $(-1, 0)$ oraz $(1, 0)$, gdzie pochodne względem y zerują się, półokręgi: dolny i górny -są wykresami tych funkcji. Natomiast w żadnym z otoczeń punków $(\pm 1, 0)$ zbiór T nie jest wykresem żadnej funkcji. Jest to ogólna prawidłowość, mamy bowiem następujący rezultat.

Twierdzenie o funkcjach uwikłanych. Jeżeli funkcja $F : D \to \mathbb{R}$ jest klasy C^1 w pewnym zbiorze otwartym $D \subset \mathbb{R}^2$, $P_0 = (x_0, y_0) \in D$ jest takim punktem, że $F(P_0) = 0$ oraz $\frac{\partial F}{\partial y}(P_0) \neq 0$, to istnieją: otoczenia U punktu x_0 oraz W -punktu y_0 w $\mathbb{R}$ takie, że równanie $F(x, y) = 0$ ma w zbiorze U dokładnie jedno takie rozwiązanie $y = g(x)$, że $g(x) \in W$. (Zamiast warunku dotyczącego zbioru W można dla uzyskania jednoznaczności -postulować w tym miejscu ciągłość funkcji g). Ponadto g jest funkcją klasy C^1 oraz

$$g'(x) = -\frac{F'_x(x, g(x))}{F'_y(x, g(x))}. \tag{2}$$

Dowód. Z ciągłości F'_y wynika, że ta funkcja (różna od zera w punkcie P_0) przyjmuje wartości niezerowe i stałego znaku w pewnym otoczeniu tego punktu -np. postaci otwartego prostokąta $U \times W$. Zamieniając, w razie potrzeby F na $-F$ -co nie zmienia zbioru $F^{-1}(\{0\})$- możemy bez straty ogólności przyjąć, że $F'_y > 0$ w zbiorze $U \times W$. Funkcja $\phi(y) := F(x_0, y)$ ma więc dodatnią pochodną w przedziale W, jest równa zero w punkcie $y = y_0$, bo $F(P_0) = 0$. Istnieją więc punkty $y_-, y_+ \in W$ takie, że $F(x_0, y_-) < 0$, $F(x_0, y_+) > 0$. Ale z ciągłości F wzgl. pierwszej zmiennej, takie znaki $F(x, y_-), F(x, y_+)$ utrzymają się również dla $|x - x_0|$ dostatecznie małych (powiedzmy, dla $|x - x_0| < \delta$. Zamieniając U na przedział $(x_0 - \delta, x_0 + \delta)$, zastępując W przez przedział (y_-, y_+) otrzymamy pierwszą część tezy, czyli istnienie i jednoznaczność rozwiązania równania $F(x, y) = 0$ gdy $x \in U, y \in W$.

Faktycznie, istnienie wynika z własności Bolzano-Darboux dla zmieniającej znak funkcji $W \ni y \to F(x, y) \in \mathbb{R}$ gdy $x \in U$ pozostaje ustalone. Gdyby było dla jednego takiego x więcej rozwiązań: np. $y_1, y_2 \in W$, to z twierdzenia Rolle'a -pochodna z tej funkcji zmiennej y przyjęła by wartość zero dla pewnego y leżącego pomiędzy y_1 a y_2, co wykluczyliśmy na początku dowodu. (Dla alternatywnej wersji tezy- nie zakładając warunku $g(x) \in W$ -możemy postępować w dowodzie tak jak poprzednio, uzyskując pewne otoczenie W punktu y_0. Ale wtedy ciągłość pozwoli nam, po ewentualnym zmniejszeniu otoczenia U punktu x_0 uzyskać zawieranie zbioru $g(U) \subset W$.) Jeśli wykażemy różniczkowalność funkcji uwiklanej g, to ostatnia równość z tezy wyniknie z reguły łańcucha. Faktycznie, na zbiorze U funkcja $x \mapsto F(x, g(x))$ stale równa zero, ma zerową pochodną. Ale z Reguły łańcucha w jej najprostszej wersji, ta pochodna, to $F'_x \cdot 1 + F'_y \cdot g'(x)$. Skoro to jest zerem, wyliczając $g'(x)$ otrzymujemy równość z tezy.

Pozostaje tylko sprawdzenie, że g' istnieje w zbiorze U. Skorzystamy z lematu o przyrostach całkowitych z wykładu 14. Zamiast x_0, x_1 wstawmy tam $x, x + h$, zamiast y_0 -wartość $y = g(x)$, zamiast y_1- wartość $y + \Delta_g$, gdzie $\Delta_g = g(x + h) - g(x)$. Wzór na

$$F(x+h, y+\Delta_g) - F(x, y) = F(x+h, y+\Delta_g) - F(x, y+\Delta_g) + F(x, y+\Delta_g) - F(x, y)$$

(który też bezpośrednio uzyskamy stosując Tw. Lagrange'a) przyjmie postać

$$\exists_{\alpha,\beta\in(0,1)}\ F(x+h, y+\Delta_g) - F(x, y) = hF'_x(x+\alpha h, y+\Delta_g) + \Delta_g F'_y(x, y+\beta\Delta_g). \tag{3}$$

Lewa strona tej równości, to zero, bo $y = g(x), y + \Delta_g = g(x + h)$. Stąd wyliczamy

$$\Delta_g = -\frac{hF'_x(x + \alpha h, y + \Delta_g)}{F'_y(x, y + \beta\Delta_g)}.$$

Zmniejszając otoczenia U, W w razie potrzeby -możemy otrzymać ograniczoność pochodnych cząstkowych F'_x oraz ograniczenie przez pewną stałą $c > 0$ od dołu F'_y w zbiorze $U \times W$. Wówczas z ostatniego wzoru na Δ_g widzimy, że przyrost funkcji g zmierza do zera gdy $h \to 0$, co oznacza ciągłość tej funkcji. Dzieląc stronami przez $h = \Delta_x$ otrzymujemy wzór na iloraz różnicowy dla funkcji g. Przy $h \to 0$ zarówno $\alpha h \to 0$, jak i $\beta\Delta_g \to 0$, więc z ciągłości pochodnych cząstkowych funkcji F otrzymujemy istnienie granicy ilorazów różnicowych $\frac{\Delta_g}{h}$ przy $h \to 0$, czyli różniczkowalność g oraz wzór na jej pochodną, jak w tezie. $\square$

Należy zwrócić uwagę na fakt, że czasami bardzo trudno jest znaleźć konkretny wzór na funkcję uwikłaną g, współrzędne $(x, g(x))$ punktu z danej powierzchni możemy znać *a priori* i wtedy wzór (2) na pochodną g bywa bardzo przydatny. Przy szukaniu ekstremów warunk konieczny: $g' = 0$ oznacza wtedy zerowanie się licznika. Okaże się, że w punktach, gdzie $g' = 0$ wzór na drugą pochodną jest całkiem podobny do tego na pierwszą -tylko licznik zawiera drugą, zamiast pierwszej pochodną cząstkową względem zmiennej x. Ze względu na znak "minus" -np. przy dodatniej F'_y będziemy mieli odwrotną zależność typu ekstremum od znaku $F''_x(x_0, g(x_0))$ -będzie np. maksimum lokalne funkcji g w punkxie x_0, jeśli ta pochodna F''_x jest dodatnia, zaś $F'_x(x_0, g(x_0)) = 0$, $F'_y(x_0, g(x_0)) > 0$.

Stosując podobne metody możemy wykazać analogiczne twierdzenie dla F zależnej od trzech, lub nawet od d zmiennych. Rozwikłujemy zazwyczaj przedstawiając ostatnią zmienną jako funkcję poprzednich zmiennych.

Na przykład, mając równanie $F(x, y, z) = 0$, gdzie $F'_z(x_0, y_0, z_0) \neq 0$ znajdujemy rozwiązanie, w którym zamiast $g(x, y)$ oznaczę funkcję rozwikłującą przez $z(x, y)$.
Istnieje wówczas dokładnie jedno takie ciągłe rozwiązanie $z(x, y)$ równania

$$F(x, y, z(x, y)) = 0$$

w pewnym otoczeniu badanego punktu, a jego pochodne cząstkowe wyrażają się podobnym wzorem: np.

$$z'_x(x, y) = -\frac{F'_x(x, y, z(x, y))}{F'_z(x, y, z(x, y))}.$$

Inny dowód twierdzenia o funkcjach uwikłanych -pozwalający na rozwiązywanie nie jednego, a układów paru równań -otrzymamy z Twierdzenia o Lokalnej Odwracalności. Jego trudność polega między innymi na niedostępności w sytuacji odwzorowań o wartościach wektorowych metod rozwiązywania równań opartych na twierdzeniu Bolzano-Darboux. Dla $d > 1$ brak jakiejś naturalnej struktury liniowego porządku w przestrzeni $\mathbb{R}^d$. Zamiennikiem dla tych metod, świetnie działającym nawet w przestrzeniach Banacha nieskończenie wymiarowych będzie Twierdzenie o Punkcie Stałym. Jest to jedno z najczęściej stosowanych twierdzeń Banacha. Ale to już w następnym wykładzie.