Wykłady 1,...,10

1 Funkcje wielu zmiennych

1.1 Pojęcia wstępne

Pod nazwą „funkcje wielu zmiennych" kryją się funkcje jednej zmiennej, należącej do $\mathbb{R}^n$. Ogólniej, będziemy rozważali n-wymiarową przestrzeń wektorową X nad $\mathbb{R}$, wyposażoną w iloczyn skalarny, to znaczy w odwzorowanie $X \times X \ni (x, y) \mapsto (x|y) \in \mathbb{R}$ o własnościach:

- $(\alpha x + \beta y|z) = \alpha(x|z) + \beta(y|z)$ gdzie $x, y, z \in X$, $\alpha, \beta \in \mathbb{R}$,
- $(x|y) = (y|x)$,
- $(x|x) \geq 0$, przy czym $(x|x) = 0 \Rightarrow x = 0$.

Przestrzeń taką nazywa się n-wymiarową przestrzenią euklidesową. Będziemy także korzystali z następującego twierdzenia:

Każdy funkcjonał liniowy ϕ na X ma postać: $\phi(x) = (v|x)$, przy czym wektor v jest jednoznacznie wyznaczony przez ϕ.

Przykład. $X = \mathbb{R}^n$, $(x|y) = \sum_{k=1}^n x_k y_k$, gdzie $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$. Funkcjonały liniowe mają postać $\phi(x) = \sum_{k=1}^n c_k x_k$ czyli $v = (c_1, \ldots, c_n)$.

Podstawowe własności Wprowadzamy następującą wielkość, którą będziemy nazywali normą wektora: $||x|| := \sqrt{(x|x)}$. Własności normy:

1. $(||x|| = 0) \Rightarrow (x = 0)$,
2. $||\alpha x|| = |\alpha| \cdot ||x||$,
3. $|(x|y)| \leq ||x|| \cdot ||y||$ (nierówność Schwarza),
4. $||x + y|| \leq ||x|| + ||y||$ (nierówność Minkowskiego).

Metrykę w X definiujemy wzorem $d(x, y) = ||x - y||$. Łatwo sprawdzić, że jest to metryka, to znaczy, że ma własności, jakich oczekujemy od metryki. W podanym przykładzie mamy $||x|| = \sqrt{\sum_{k=1}^n x_k^2}$ oraz $d(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2}$.

Wprowadza się także normę operatorów liniowych. Niech $A : Y \to Y$, gdzie X, Y są przestrzeniami euklidesowymi. Definiujemy: $||A|| = \inf\{c : \forall_{x \in X} ||Ax|| \leq c||x||\}$. Łatwo zobaczyć, że $||A|| = \sup_{||x||=1} ||Ax||$. Równie łatwo jest sprawdzić, że norma operatora ma własności 1., 2. i 4.

Przypomnienie pewnych pojęć i ich własności z poprzedniego semestru

♠ $K(x_0, R) = \{x \in \mathbb{R}^n : ||x - x_0|| < R\}$ – kula (otwarta) o środku x_0 i promieniu R.

♠ $\overline{K(x_0, R)} = \{x \in \mathbb{R}^n : ||x - x_0|| \leq R\}$ – kula (domknięta) o środku x_0 i promieniu R.

♠ $\mathbb{R}^n(x_n \to y) \Leftrightarrow (d(x_n, y) = ||x_n - y|| \to 0)$. Wiemy, że zbieżność ciągu punktów $\mathbb{R}^n$ jest równoważna zbieżności ich współrzędnych.

♠ Zbiór $A \subset \mathbb{R}^n$ nazywamy otoczeniem $x \in \mathbb{R}^n$, gdy $\exists_{\epsilon>0} K(x, \epsilon) \subset A$.

♠ Zbiór $A \subset \mathbb{R}^n$ nazywamy otwartym, jeśli jest otoczeniem każdego z punktów, które do niego należą.

♠ $y \in \mathbb{R}^n$ nazywamy punktem skupienia zbioru $A \subset \mathbb{R}^n$, jeśli każde otoczenie y zawiera punkt zbioru A. Jest to równoważne temu, że istnieje ciąg $x_n \in A$ zbieżny do y.

♠ Zbiór A jest domknięty wtedy i tylko wtedy, gdy jego dopełnienie $\mathbb{R}^n \backslash A$ jest zbiorem otwartym.

♠ Funkcja $f : \mathbb{R}^n \to \mathbb{R}^m$ jest ciągła w punkcie $x \in \mathbb{R}^n$, jeśli dla każdego ciągu $x_n \to x$ zachodzi $f(x_n) \to f(x)$. Warunek równoważny: dla każdego otoczenia V punktu $f(x)$ istnieje takie otoczenie U punktu x, że $f(U) \subset V$.

1.2 Pochodna funkcji

Niech X, Y oraz Z będą skończenie wymiarowymi przestrzeniami wektorowymi wyposażonymi w iloczyny skalarne.

Definicja 1 *Niech $f : X \to Y$. Mówimy, że f jest różniczkowalna w punkcie $x \in X$, jeśli istnieje odwzorowanie $A \in L(X, Y)$ (czyli liniowe) takie, że reszta $R(x, h)$ zdefiniowana równaniem*

$$f(x + h) = f(x) + Ah + R(x, h)$$

spełnia warunek $\frac{R(x,h)}{h} \stackrel{h \to 0}{\longrightarrow} 0$. Inaczej: $R(x, h) = o(h)$.[1]

Podobnie jak dla funkcji jednej zmiennej dowodzi się, że A występujące w definicji, jeśli istnieje, to jest tylko jedno. Nazywamy je wtedy pochodną f w punkcie x i oznaczamy $f'(x)$.

Twierdzenie 1 *Niech $f : X \to Y$ i $g : Y \to Z$. Jeśli f jest różniczkowalna w punkcie $x \in X$, a g jest różniczkowalna w $f(x)$, to $g \circ f$ jest różniczkowalna w x i zachodzi wzór $(g \circ f)'(x) = g'(f(x))f'(x)$.*

[1] Tutaj i w dalszym ciągu stosujemy zapis skrócony. Chodzi o to, że funkcja, o której mowa jest określona w pewnym otoczeniu punktu, o którym mowa i ma wartości w Y.

1.2.1 Pochodna kierunkowa

Niech $f : X \to Y$ i $v \in X$. Definiujemy funkcję $g : \mathbb{R} \to Y$ wzorem $g(t) = f(x + tv)$.

Definicja 2 *Pochodną kierunkową funkcji f w punkcie x w kierunku v nazywamy $\frac{dg}{dt}|_{t=0}$ (o ile taka pochodna istnieje) i oznaczamy ją $\nabla_v f(x)$.*

Z twierdzenia o różniczkowaniu złożenia dwóch funkcji mamy natychmiastowy wniosek: jeśli f jest funkcją różniczkowalną w punkcie x, to

$$\nabla_v f(x) = f'(x)v.$$

1.2.2 Pochodne cząstkowe, macierz Jacobiego

Niech $f : \mathbb{R}^n \to \mathbb{R}^m$. W $\mathbb{R}^n$ wyróżniamy bazę $e_1, \dots, e_n$, gdzie $e_j = (0, \dots, 1, \dots, 0)$, zwaną bazą kanoniczną, gdzie jedynka występuje na j-tym miejscu. Pochodną cząstkową f względem j-tej zmiennej będziemy nazywali $\nabla_{e_j} f$. Najczęściej stosuje się oznaczenia $\frac{\partial f}{\partial x_j}$, $f_{,x_j}$ lub $f_{,j}$.

Przyjrzyjmy się bliżej.

$$\frac{\partial f}{\partial x_j}(x) = \nabla_{e_j} f(x) = \lim_{t \to 0} \frac{f(x + te_j) - f(x)}{t} =$$

$$= \lim_{t \to 0} \frac{f(x_1, \dots, x_j + t, \dots, x_n) - f(x_1, \dots, x_j, \dots, x_n)}{t}.$$

Jak wiemy z algebry, kolumny macierzy odwzorowanie liniowego $A \in L(X, Y)$ w zadanych bazach e i e' są współczynnikami rozkładu wektorów Ae_j w bazie e". Wynika stąd, że macierz pochodnej funkcji f, o ktorej tutaj mówimy, ma postać

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

Nazywamy ją macierzą Jacobiego. Podstawiając współrzędne wektorów dostajemy wzór

$$f'(x)h = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} h_j \ .$$

Ponieważ składaniu odwzorowań liniowych odpowiada mnożenie ich macierzy (w odpowiednich bazach), mamy następujący wzór dla pochodnej odwzorowania złożonego:

$$\frac{\partial (f \circ g)_i}{\partial x_j} = \sum_k \frac{\partial f_i}{\partial y_k} \frac{\partial g_k}{\partial x_j}.$$

Z różniczkowalności funkcji wynika istnienie pochodnych cząstkowych. Odwrotne wynikanie nie jest prawdziwe. Jeśli jednak pochodne cząstkowe są ciągłe, to wynika stąd różniczkowalność funkcji i ciągłość pochodnej. Dowód, wcale nie prosty, pomijamy.

1.2.3 Gradient

Jeśli $f : X \to \mathbb{R}$, a $x \in X$, to $f'(x) \in L(X, \mathbb{R})$, czyli jest funkcjonałem liniowym. Istnieje więc dokładnie jeden wektor, v taki, że $f'(x)h = (v|h)$ dla wszystkich $u \in X$. Wektor ten nazywamy gradientem f w punkcie x i oznaczamy $\operatorname{grad} f(x)$ lub $\nabla f(x)$. Interpretacja geometryczna gradientu wynika z wzoru

$$f(x+h)-f(x) = f'(x)h+r(x,h) = (\nabla f(x)|h)+r(x,h) \approx ||\nabla f(x)|| \cdot ||h|| \cos(\alpha),$$

gdzie α jest kątem między ∇f i h. Wynika stąd, że gradient ma kierunek najszybszego wzrostu funkcji, a jego długość jest miarą szybkości jej wzrostu.

1.3 Twierdzenia o wartości średniej

Zbiory wypukłe Zajmujemy się podzbiorami skończenie wymiarowej rzeczywistej przestrzeni wektorowej X wyposażonej w iloczyn skalarny.

Definicja 3 *A jest zbiorem wypukłym, jeżeli* $\forall_{x,y \in A} \forall_{\lambda \in [0,1]} (\lambda x + (1-\lambda)y) \in A$.

Łatwo jest stwierdzić, że jeśli zbiory A i B są wypukłe, to $A \cap B$ i $\overline{A}$ są wypukłe. Najmniejszy domknięty zbiór wypukły zawierający dany zbiór A nazywamy domkniętą powłoką wypukłą A i oznaczamy $\overline{\operatorname{conv}}(A)$

Twierdzenie 2 *Jeśli A jest zbiorem wypukłym domkniętym, to*

1. *Dla każdego $x \notin A$ istnieje $v \in X$ takie, że* $(v|x) > \sup_{y \in A}(v|y)$,
2. *Jeśli dla każdego $v \in X$ zachodzi* $\inf_{y \in A}(y|v) \leq (x|v) \leq sup_{z \in A}(z|v)$, *to* $x \in A$.

Przechodzimy do twierdzeń o wartości średniej.

Twierdzenie 3 *Niech f będzie funkcją <u>o wartościach w $\mathbb{R}$</u> określoną i różniczkowalną na zbiorze wypukłym A. Wtedy dla dowolnych $x, y \in A$ istnieje takie $\lambda \in]0,1[$, że*

$$f(x) - f(y) = f'(\lambda x + (1-\lambda)y)\ (x-y).$$

Dowód. Bierzemy funkcję pomocniczą $[0,1] \ni t \mapsto g(t) := f(tx + (1-t)y)$ i stosujemy do niej twierdzenie Lagrange'a o wartości średniej.

Twierdzenie 4 *Niech $f : X \to Y$ będzie funkcją określoną na zbiorze wypukłym A i różniczkowalną na tym zbiorze. Wtedy dla dowolnych $x, x' \in A$*

- $(f(x) - f(y)) \in \overline{\operatorname{conv}}(\{f'(tx + (1-t)y)(x-y) : t \in [0,1]\})$,

- $||f(x) - f(y)|| \leq \sup_{t\in[0,1]} ||f'(tx + (1-t)y)|| \cdot ||x - y||$.

Dowód. Bierzemy dowolne $v \in Y$ i stosujemy poprzednie twierdzenie do funkcji $x \mapsto (v|f(x))$, a następnie korzystamy z własności zbiorów wypukłych. Niech $g(t) := (v|f(t))$.

$$(v|f(y)-f(x)) = g(1)-g(0) = g'(\theta) = (v|(f'(\theta x + (1-\theta)y)(x-y)) \leq \sup_{s\in[0,1]} (v|f'(s)(y-x))$$

Z Twierdzenia 2. wynika pierwszy punkt tezy. Punkt drugi wynika z własności normy operatora liniowego.

W dalszym ciągu będziemy stosowali oznaczenie $[x, y] := \{\lambda x + (1-\lambda)y : \lambda \in [0,1]\}$.

1.4 Pochodne wyższych rzędów. Wzór Taylora

Niech $f : X \to Y$, gdzie X i Y są przestrzeniami euklidesowymi. Jeśli f jest różniczkowalna, to $f'(x) \in L(X, Y)$. $L(X, Y)$ jest skończenie wymiarową przestrzenią wektorową wyposażoną w normę[2], można więc pytać o pochodną funkcji $X \ni x \mapsto f'(x) \in L(X, Y)$. Pochodna tej funkcji należy do $L(X, L(X, Y))$. oznaczamy ją naturalnie $f''(x)$. Podobnie określamy pochodne wyższego rzędu. Zajmiemy się najpierw drugą pochodną. Ponieważ $f''(x) \in L(X, L(X, Y))$, dla $h \in X$ $f''(x)h \in L(X, Y)$, więc dla $k \in X$ ma sens wyrażenie $(f''(x)h)k \in Y$. We współrzędnych wygląda to tak:

$$(f''(x)h)k = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} h_j k_i \text{ , gdzie } \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right).$$

Bardzo ważne konsekwencje ma następujące twierdzenie mówiące o symetrii drugich pochodnych cząstkowych:

Twierdzenie 5 *Jeśli drugie pochodne cząstkowe są ciągłe w pewnym otoczeniu punktu x, to $(f''(x)h)k = (f''(x)k)h$ a we współrzędnych:*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

Dowód. Zauważmy najpierw, że wystarczy rozważyć przypadek funkcji dwóch zmiennych o wartościach w $\mathbb{R}$. Dla większej przejrzystości zapisu współrzędne punktu, o którym mowa w twierdzeniu, oznaczymy literami x i y. Weźmy dwie dostatecznie małe liczby rzeczywiste h i k i rozważmy wyrażenie

$$\begin{aligned}
A(h,k) &= \frac{1}{hk}[f(x+h, y+k) - f(x, y+k) - f(x+h, y) + f(x, y)] \\
&= \frac{1}{hk}\{[f(x+h, y+k) - f(x, y+k)] - [f(x+h, y) - f(x, y)]\} \\
&= \frac{1}{hk}\{[f(x+h, y+k) - f(x+h, y)] - [f(x, y+k) - f(x, y)]\}.
\end{aligned}$$

[2]Norma w $L(X, Y)$ nie została określona za pomocą iloczynu skalarnego, ale w przestrzeni z normą można bez żadnych zmian określić pojęcie różniczkowalności funkcji. Poza tym w skończenie wymiarowej przestrzeni wektorowej wszystkie normy są równoważne (to znaczy dają ten sam zbiór ciągów zbieżnych).

Wprowadzamy oznaczenia

$$F(y) = f(x+h, y) - f(x, y),$$
$$G(x) = f(x, y+k) - f(x, y).$$

Mamy

$$A(h,k) = \frac{1}{hk}[F(y+k) - F(y)] = \frac{1}{hk}[G(x+h) - G(x)].$$

Na mocy twierdzenia Lagrange'a o wartości średniej mamy

$$F(y+k) - F(y) = kF'(y+\theta k),\ 0 < \theta < 1$$

oraz

$$G(x+h) - G(x) = hG'(x+\eta h),\ 0 < \eta < 1.$$

Podstawiając mamy

$$A(h,k) = \frac{1}{h}F'(y+\theta k) = \frac{1}{h}\left[\frac{\partial f}{\partial y}(x+h, y+\theta k) - \frac{\partial f}{\partial y}(x, y+\theta k)\right].$$

Jeszcze raz stosujemy twierdzenie Lagrange'a, tym razem do funkcji $\Phi(x) = \frac{\partial f}{\partial y}(x, y+\theta k)$.

$$A(h,k) = \frac{\partial \frac{\partial f}{\partial y}}{\partial x}(x+\theta' h, y+\theta k).$$

Na mocy ciągłości pochodnych cząstkowych

$$\lim_{(h,k)\to(0,0)} A(h,k) = \frac{\partial^2 f}{\partial x \partial y}(x,y).$$

Biorąc w rachunkach rozkład $A(h,k)$ uwzględniający funkcję G stwierdzimy, że

$$\lim_{(h,k)\to(0,0)} A(h,k) = \frac{\partial^2 f}{\partial y \partial x}(x,y).$$

To kończy dowód. Przy okazji znaleźliśmy praktyczny przybliżony wzór na pochodne cząstkowe drugiego rzędu.

1.4.1 Wzór Taylora

Zajmiemy się tutaj tylko wzorem Taylora z pochodnymi do drugiego rzędu włącznie.

Twierdzenie 6 *Niech $f : X \to Y$ będzie funkcją trzykrotnie[3] różniczkowalną w pewnym wypukłym otoczeniu U punktu $x \in X$. Wtedy dla $y \in U$*

$$f(y) = f(x) + f'(x)(y-x) + \frac{1}{2!}f''(x)h^{(2)} + o(||h||^2),\ \textit{gdzie}\ f''(x)h^{(2)} = (f''(x)h)h.$$

Dowód wynika z zastosowania wzoru Taylora d0 funkcji $t \mapsto (v|f(tx+(1-t)y))$.

[3] Wystarczyłoby założyć dwukrotną różniczkowalność, ale taki prosty dowód, jak tutaj, wymaga mocniejszego założenia.

1.5 Ekstrema funkcji wielu zmiennych

Definicja terminu „lokalne ekstremum” jest taka sama, jak w przypadku funkcji jednej zmiennej i nie będziemy jej tutaj powtarzali. Oczywiście sensowne jest mówienie o ekstremach wyłącznie w przypadku funkcji $X \to \mathbb{R}$.

Twierdzenie 7 (warunek konieczny ekstremum) *Jeśli $f : X \to \mathbb{R}$ ma w punkcie x_0 lokalne ekstremum i jest w tym punkcie różniczkowalna, to $f'(x_0) = 0$.*

Dowód. Bierzemy dowolne $v \in X$ i zajmujemy się funkcją $g(t) = f(x_0) + tv)$. Funkcja g ma lokalne ekstremum dla $t = 0$ i jest w tym punkcie różniczkowalna. Z warunku koniecznego dla funkcji jednej zmiennej mamy $g'(0) = \nabla_v f(x_0) = f'(x_0)v = 0$. Ponieważ v jest dowolne, mamy $f'(x_0) = 0$.

Aby sformułować warunek wystarczający, w którym występuje druga pochodna, potrzebne będą pewne pojęcia z algebry.

Forma kwadratowa, to funkcja $Q : X \to \mathbb{R}$, która wyraża się za pomocą współrzędnych (w jakiejkolwiek, a naprawdę w każdej) bazie wzorem[4]

$$Q(h) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} h_i h_j .$$

Formę kwadratową Q nazywamy

- **dodatnio określoną**, jeśli $Q(h) > 0$ dla wszystkich $h \neq 0$,
- **ujemnie określoną**, jeśli $Q(h) < 0$ dla wszystkich $h \neq 0$.

Praktyczny sposób rozstrzygania, czy dana forma kwadratowa jest dodatnio bądź ujemnie określona albo nie należy do żadnej z tych dwóch klas, daje następujące twierdzenie:

Twierdzenie 8 *Forma kwadratowa Q określona wzorem $Q(h) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} h_i h_j$ jest*

- *dodatnio określona wtedy i tylko wtedy, gdy $\Delta_k > 0$ dla wszystkich $k = 1, \ldots, n$,*
- *ujemnie określona wtedy i tylko wtedy„ gdy $(-1)^k \Delta_k > 0$ dla wszystkich $k = 1, \ldots, n$,*
- *nieokreślona, jeśli $Q(h)$ ma wartości zarówno dodatnie, jak ujemne.*

gdzie

$$\Delta_k = \begin{vmatrix} a_{11}, & \ldots, & a_{1k} \\ \ldots, & \ldots, & \ldots \\ a_{k1}, & \ldots, & a_{kk} \end{vmatrix}$$

[4]Bardziej elegancka definicja – na wykładzie z Algebry.

Twierdzenie 9 (warunek wystarczający ekstremum) *Jeśli funkcja $f : X \to \mathbb{R}$ jest dwukrotnie różniczkowalna w punkcie $x \in X$, $f'(x) = 0$ i forma kwadratowa $f''(x)h^{(2)}$ jest dodatnio (ujemnie) określona, to f ma w punkcie x lokalne minimum (maksimum).*

Jeśli forma kwadratowa $f''(x)h^{(2)}$ jest nieokreślona, to f nie ma w punkcie x lokalnego ekstremum.

Dowód. Przeprowadzimy dowód dla minimum. Najpierw zauważmy, że jeśli $Q(h) > 0$ dla wszystkich $h \neq 0$, to istnieje takie $c > 0$, że $Q(h) \leq c||h||^2$. Rzeczywiście, wystarczy wziąć $c = \inf_{\{h:||h||=1\}} Q(h)$. Ponieważ sfera jednostkowa jest zbiorem zwartym, kres dolny jest osiągany, musi więc być większy od zera. Nierówność, o którą chodzi wynika z własności form kwadratowych: $Q(\alpha h) = \alpha^2 Q(h)$. Z wzoru Taylora mamy

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2}f''(x)h^{(2)} + o(||h||^2) = f(x) + \frac{1}{2}f''(x)h^{(2)} + o(||h||^2)$$

Wystarczy teraz wziąć $||h||$ tak małe, żeby ostatni składnik był mniejszy niż $\frac{1}{2}c||h||^2$. Będzie wtedy $f(x+h) > f(x)$ dla dostatecznie małych $||h|| \neq 0$.

1.6 Lokalna odwracalność odwzorowań

Mamy dane odwzorowanie $f : X \to Y$. Niech $f(x_0) = y_0$. Lokalna odwracalność f w otoczeniu x_0 oznacza to, że dla każdego y dostatecznie bliskiego y_0 istnieje dokładnie jedno x bliskie x_0 takie, że $f(x) = y$. Określenie „bliskie" nie było zdefiniowane, bardziej precyzyjnie rzecz formułuje następujące twierdzenie:

Twierdzenie 10 (o lokalnej odwracalności) *Niech $f : X \to Y$ będzie funkcją różniczkowalną w pewnym otoczeniu punktu $x_0 \in X$, a jej pochodna f' niech będzie ciągła w tym otoczeniu. Zakładamy ponadto, że odwzorowanie liniowe $f'(x_0)$ ma odwrotne (może to mieć miejsce tylko wtedy, gdy wymiary przestrzeni X i Y są równe). Niech $y_0 = f(x_0)$.*

Wtedy istnieją: U – otoczenie x_0 i V – otoczenie y_0 takie, że $f|_U$ jest bijekcją U na V. Ponadto, odwzorowanie $(f|_U)^{-1}$ jest różniczkowalne w y_0 i jego pochodna w tym punkcie równa się $(f'(x_0))^{-1}$.

Dowód. Chodzi o to, żeby pokazać, że równanie $f(x) = y$ ma dla $y \in \overline{K(y_0, r)}$ dokładnie jedno rozwiązanie $x \in \overline{K(x_0, s)}$, gdzie r i s są odpowiednio dobrane. Do rozwiązania zastosujemy metodę kolejnych przybliżeń opartą na zasadzie Banacha. Niech

$$F(x) := x - (f'(x_0))^{-1}(f(x) - y).$$

Sprawdzamy najpierw, że $F(x) = x$ jest równoważne temu, że $f(x) = y$. Oczywiście tak jest, bo $f'(x_0)$ jest bijekcją.

Szukamy teraz warunków, jakie mają spełniać r i s, żeby F odwzorowywało $\overline{K(x_0, s)}$ w ten sam zbiór. Niech $||x - x_0|| \leq s$ wtedy

$$||F(x) - x_0|| = ||x - x_0 - (f'(x_0))^{-1}(f(x) - y)|| =$$

$$
\begin{aligned}
&||(f'(x_0))^{-1}[f'(x_0)(x-x_0)-f(x)+y]|| = \\
&||(f'(x_0))^{-1}[f'(x_0)(x-x_0)-f(x_0)-f'(x_0)(x-x_0)-R(x_0,x-x_0)+y|| = \\
&||(f'(x_0))^{-1}[-y_0+y-R]|| \leq ||(f'(x_0))^{-1}|| \cdot ||y-y_0+R||
\end{aligned}
$$

Niech $C = ||(f'(x_0))^{-1}||$. Jeśli weźmiemy takie małe s, że $R < ||y-y_0||$, to będziemy mieli $||F(x)-x_0|| \leq 2Cr$. Wystarczy więc wziąć takie r, żeby $2Cr < s$ i sprawa jest załatwiona.

Pytamy teraz, kiedy F jest odwzorowaniem zbliżającym.

$$
\begin{aligned}
&||F(x)-F(x')|| \leq \sup_{c\in[x,x']} ||F'(c)|| \cdot ||x-x'|| = \\
&\sup_{c\in[x,x']} ||I-(f'(x_0))^{-1}f'(c)|| \cdot ||x-x'|| = \\
&\sup_{c\in[x,x']} ||I-(f'(x_0))^{-1}(f'(x_0)+f'(c)-f'(x_0))|| \cdot ||x-x'|| = \\
&\sup_{c\in[x,x']} ||(f'(x_0))^{-1}(f'(c)-f'(x_0)|| \cdot ||x-x'|| \leq \\
&\sup_{c\in[x,x']} ||(f'(x_0))^{-1}|| \cdot ||(f'(c)-f'(x_0)|| \cdot ||x-x'||.
\end{aligned}
$$

Ponieważ f' jest z założenia ciągłe, możemy dobrać tak małe s, żeby we wszystkich punktach $c \in \overline{K(x_0,s)}$ spełniona była nierówność $||(f'(x_0))^{-1}|| \cdot ||(f'(c)-f'(x_0)|| < q < 1$ dla pewnego q. Jeśli teraz wybierzemy s spełniające wszystkie nałożone do tej pory warunki, to możemy skorzystać z zasady Banacha ($\overline{K(x_0,s)}$ jest zbiorem zwartym), co kończy dowód istnienia odwzorowanie odwrotnego. Różniczkowalność i wzór na pochodną odwzorowania odwrotnego dowodzi się dość podobnie, jak w przypadku funkcji jednej zmiennej.

Teraz trochę elementarnego wyjaśnienia. Znajdowanie odwzorowanie odwrotnego jest rozwiązywaniem układu n równań z n niewiadomymi:

$$
\begin{aligned}
&f_1(x_1,\ldots,x_n) = y_1, \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
&f_n(x_1,\ldots,x_n) = y_n.
\end{aligned}
$$

Gdzie $y_1,\ldots,y_n$ są dane, a $x_1,\ldots,x_n$ szukane.Twierdzenie mówi o tym, że jeśli $f'(x_0)$ jest odwracalne, a to ma miejsce wtedy, gdy

$$
\det f'(x_0) = \begin{vmatrix} \frac{\partial f_1}{\partial x_1}, & \ldots, & \frac{\partial f_1}{\partial x_n} \\ \ldots, & \ldots, & \ldots \\ \frac{\partial f_n}{\partial x_1}, & \ldots, & \frac{\partial f_n}{\partial x_n} \end{vmatrix} \neq 0.
$$

a także, gdy interesują nas tylko rozwiązania dostatecznie bliskie jakiegoś znanego rozwiązania przy nieznacznie zmienionych danych. Wyznacznik, który tutaj występuje, nazywa się **jakobianem**.

1.7 Funkcje uwikłane

Zajmiemy się tutaj następującym zagadnieniem: Dany jest układ m równań

$$f_1(x_1,\ldots,x_n,y_1,\ldots,y_m)=0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$f_m(x_1,\ldots,x_n,y_1,\ldots,y_m)=0.$$

Czy można $y_1,\ldots,y_m$ uzyskać stąd jako funkcje $x_1,\ldots,x_n$?

Sformułujemy to pytanie w sposób podobny problemu z poprzedniego podrozdziału. Dane jest odwzorowanie $f: X\times Y\to Y$. Przy jakich założeniach można twierdzić, że istnieje odwzorowanie $\phi: X\to Y$ takie, że $(y=\phi(x))\Leftrightarrow(f(x,y)=0)$? Odpowiedź (oczywiście lokalną) daje następujące twierdzenie (w sformułowaniu twierdzenia użyjemy symbolu f'_Y dla oznaczenia pochodnej odwzorowania $y\mapsto f(x,y)$):

Twierdzenie 11 (o funkcjach uwikłanych) *Zakładamy, że odwzorowanie $f: X\times Y\to Y$ ma ciągłą pochodną w pewnym otoczeniu punktu (x_0,y_0). Poza tym niech $f(x_0,y_0)=0$. Jeśli $f'_Y(x_0,y_0)$ ma odwzorowanie odwrotne (czyli $\det f'_Y(x_0,y_0)\neq 0$), to istnieją takie otoczenia U – otoczenie x_0 i V – otoczenie y_0 oraz odwzorowanie $\phi: U\to V$, że równanie $f(x,y)=0$ z warunkami $x\in U, y\in V$ jest równoważne temu, że $y=\phi(x)$. Ponadto odwzorowanie ϕ jest różniczkowalne w x_0 i*

$$\phi'(x_0)=-(f'_Y(x_0,y_0))^{-1}\circ f'_X(x_0,y_0).$$

Dowód. Stosujemy twierdzenie o lokalnej odwracalności do odwzorowania $G: X\times Y\to X\times Y$ określonego wzorem $G(x,y)=(x,f(x,y))$ i otoczenia punktu (x_0,y_0). Z założeń o różniczkowalności i ciągłości pochodnej f wynika taka sama różniczkowalność G. Sprawdzamy odwracalność $G'(x_0,y_0)$:

$$\det G'(x_0,y_0)=\det\begin{bmatrix} I & 0\\ f'_X(x_0,y_0) & f'_Y(x_0,y_0)\end{bmatrix}=\det f'_Y(x_0,y_0)\neq 0.$$

Odwzorowanie G jest więc lokalnie odwracalne w pewnym otoczeniu punktu (x_0,y_0). Definiujemy $\phi(x):=G^{-1}(x,0)$, gdzie G^{-1} oznacza odwzorowanie odwrotne do G ograniczonego do odpowiednio małego otoczenia (x_0,y_0). Sprawdzamy, że $f(x,y)=0$ jest równoważne (dla (x,y) dostatecznie bliskiego (x_0,y_0)) temu, że $y=\phi(x)$. Rzeczywiście, $f(x,y)=0$, to to samo, co $G(x,y)=(x,0)$. Różniczkowalność ϕ wynika z różniczkowalności G^{-1}, a wzór na pochodną otrzymuje się różniczkując stronami tożsamość $f(x,\phi(x))=0$:

$$f'_X+f'_Y\phi'=0,\ \text{skąd}\ \phi'=-(f'_Y)^{-1}\circ f'_X.$$

1.8 Powierzchnie w $\mathbb{R}^n$. Ekstrema związane

1.8.1 Powierzchnia. Styczna do powierzchni

Powierzchnią w $\mathbb{R}^n$ będziemy nazywali podzbiór, który lokalnie opisuję się za pomocą układu równań $G(x)=0$, gdzie $G:\mathbb{R}^n\to\mathbb{R}^m$. Zamiast $\mathbb{R}^n$ i $\mathbb{R}^m$

będziemy pisali X i Y i możemy myśleć ogólniej o skończenie wymiarowych przestrzeniach wektorowych euklidesowych.

Definicja 4 *Bardziej precyzyjnie, powierzchnią $n-m$ wymiarową w X nazywamy taki podzbiór $M \subset X$, dla którego istnieje pokrycie (skończone) zbiorami otwartymi $U_i, i \in I$ w X oraz odwzorowania $\varphi_i : U_i \to \mathbb{R}^n$ odpowiednio wiele razy różniczkowalne[5] wraz z odwrotnymi do nich, takie, że każdy ze zbiorów $M \cap U_i$ jest odwzorowywany na podzbiór otwarty w $\mathbb{R}^{n-m} \subset \mathbb{R}^n$, przy czym to zawieranie rozumie się jako odwzorowanie*

$$\mathbb{R}^{n-m} \ni (x_1, \dots, x_{n-m}) \mapsto (x_1, \dots, x_{n-m}, 0, \dots, 0) \in \mathbb{R}^n.$$

Odwzorowania φ_i nazywa się lokalnymi mapami powierzchni M.

Pytamy teraz, jakie warunki wystarczają na to, żeby równanie $G(x) = 0$ wyznaczało powierzchnię.

Twierdzenie 12 *Niech $G : X \to Y$, gdzie X, Y są przestrzeniami wektorowymi nad $\mathbb{R}$ odpowiednio n i m – wymiarowymi, przy czym $n > m$. Załóżmy, że G jest odwzorowaniem różniczkowalnym o pochodnych ciągłych. Niech*

$$M := \{x \in X : G(x) = 0\}$$

Punkt $x_0 \in M$ nazywamy regularnym, jeśli rząd $G'(x_0)$ ma największą możliwą wartość, czyli m. Jeśli x_0 jest punktem regularnym M, to istnieje taki zbiór otwarty $U \subset X$, że $M \cap U$ jest powierzchnią.

Idea dowodu. Jeśli rząd $G'(x_0) = m$, to przestrzeń X można przedstawić w postaci sumy prostej $X = X_1 \oplus X_2$, gdzie $X_1 = \mathrm{Ker} G'(x_0)$. Wtedy $G'(x_0)|_{X_2}$ jest odwracalne i można zastosować twierdzenie o funkcjach uwikłanych do równania $G(x) = 0$; otrzymuje się funkcję $\phi : X_1 \to X_2$ taką, że $G(x_1 + \phi(x_1)) = 0$ dla wszystkich $x_1 \in X_1$. Odwzorowanie φ, o którym mowa w definicji powierzchni określamy wzorem $\varphi(x) = (x_1, G(x))$.

Styczna Niech M będzie podzbiorem przestrzeni wektorowej X. Krzywą w M nazywamy odwzorowanie $]a, b[\ni t \mapsto \gamma(t) \in M$. Zakładamy, że γ jest różniczkowalne oraz $0 \in]a, b[$. Interesują nas krzywe takie, że $\gamma(0) = x_0$.

Definicja 5 *Wektorem stycznym do M w punkcie x_0 nazywamy $\gamma'(0)$.*

Twierdzenie 13 *Niech $G : X \to Y$, gdzie X, Y są przestrzeniami wektorowymi nad $\mathbb{R}$ odpowiednio n i m – wymiarowymi, przy czym $n > m$. Załóżmy, że G jest odwzorowaniem różniczkowalnym o pochodnych ciągłych. Niech*

$$M := \{x \in X : G(x) = 0\}$$

[5] Odpowiednio do tego powierzchnię M nazywamy raz, dwa itd. razy różniczkowalną albo gładką, gdy φ_i są dowolnie wiele razy różniczkowalne.

Załóżmy, że $x_0 \in M$ jest punktem regularnym M. Wtedy

$$T_{x_0}(M) := \{v \in X : G'(x_0)v = 0\}$$

jest zbiorem (wszystkich) wektorów stycznych do M.

$T_{x_0}(M)$ nazywamy przestrzenią styczną do M w punkcie x_0.

Dowód. Najpierw pokażemy, że każdy wektor styczny w punkcie x_0 jest elementem $T_{x_0}(M)$. Rzeczywiście, jeśli γ jest krzywą w M taką, że $\gamma(t) \in M$ dla wszystkich t oraz $\gamma(0) = x_0$. Wtedy $G(\gamma(t)) \equiv 0$; różniczkując to w punkcie $t = 0$ mamy $G'(x_0)v = 0$, czyli $v \in T_{x_0}(M)$.

Jeśli $v \in T_{x_0}(M)$, to bierzemy krzywą γ opisaną równaniem $]-\epsilon, \epsilon[\ni t \mapsto \gamma(t) := tv + \phi(tv)$, gdzie ϕ jest odwzorowaniem występującym w dowodzie poprzedniego twierdzenia. Z jego definicji wynika, że $\gamma(t) \in M$, a z twierdzenia o funkcjach uwikłanych mamy $\phi'(0) = 0$, wobec czego $\gamma'(0) = v$.

1.8.2 Ekstrema związane

Problem jest następujący: mamy daną funkcję $f : X \to \mathbb{R}$ oraz powierzchnię M zadaną równaniem $G(x) = 0$, gdzie $G : X \to Y$, $\dim X > \dim Y$. Szukamy lokalnych ekstremów funkcji f na M, czyli takich punktów $x_0 \in M$, w których f ma największą bądź najmniejszą wartość spośród wartości przyjmowanych w punktach M leżących w jakimś otoczeniu x_0.

Zakładamy w tej części, że f i G są odwzorowaniami różniczkowalnymi tyle razy ile będzie potrzeba i pochodne ich są ciągłe.

Twierdzenie 14 (warunek konieczny) *Jeśli f ma lokalne ekstremum związane w $x_0 \in M$, to dla każdego $v \in T_{x_0}(M)$ zachodzi*

$$\nabla_v f(x_0) = 0.$$

Dowód. Jeśli $v = \gamma'(0)$, gdzie γ jest krzywą leżącą na M taką, że $\gamma(0) = x_0$, to funkcja $t \mapsto f(\gamma(t))$ ma w punkcie $t = 0$ lokalne ekstremum, skąd wynika, że jej pochodna zeruje się w zerze, a to właśnie jest napisane jako teza twierdzenia.

Twierdzenie 15 (warunek konieczny jeszcze raz) *Jeśli f ma lokalne ekstremum związane w $x_0 \in M$, a x_0 jest punktem regularnym M, to istnieje funkcjonał liniowy $\Lambda \in Y'$ taki, że*

$$f'(x_0) = \Lambda \circ G'(x_0).$$

Λ nazywa się funkcjonałem Lagrange'a, a jego współrzędne mnożnikami Lagrange'a.

Dowód. Ponieważ x_0 jest punktem regularnym M, Przestrzeń X jest sumą prostą podprzestrzeni $X = X_1 \oplus X_2$, gdzie $X_1 = T_{x_0}(M)$, a $G'(x_0)$ odwzorowuje wzajemnie jednoznacznie X_2 na Y. Wprowadźmy oznaczenie $A :=$

$(G'(x_0)|_{X_2})^{-1}$. Bierzemy dowolne $x \in X$. $x = v + v_2$, gdzie $v \in X_1, v_2 \in X_2$. $v_2 = Ay$ dla pewnego $y \in Y$.

$$\begin{aligned} 0 &= \nabla_v f(x_0) = f'(x_0)v, \text{więc } f'(x_0)x = f'(x_0)v_2 = f'(x_0)Ay = \\ &= f'(x_0)AG'(x_0)v_2 = f'(x_0)AG'(x_0)x \end{aligned}$$

Mamy więc tezę dla $\Lambda = f'(x_0)A$.

Twierdzenie 16 (warunek wystarczający) *Jeśli spełniony jest warunek konieczny z poprzedniego twierdzenia oraz forma kwadratowa*

$$f''(x_0) - \Lambda \circ G''(x_0)$$

jest dodatnio (ujemnie określona) ***na podprzestrzeni*** *$T_{x_0}(M)$, to f ma w x_0 lokalne minimum (maksimum) związane.*

Dowód. Najpierw pewne nowe pojecie i lemat z nim związany.

Definicja 6 *Płaszczyzną styczną do M w punkcie regularnym x_0 nazywamy zbiór $\mathcal{T}_{x_0}(M) := x_0 + T_{x_0}(M)$.*

Lemat 1 *Istnieje różniczkowalne wzajemnie jednoznaczne odwzorowanie π pewnego otoczenia x_0 w $\mathcal{T}_{x_0}(M)$ na otoczenie x_0 w M takie, że $||x_0+v-\pi(x_0+v)|| = o(||v||)$.*

Dowód lematu. Równanie $G(x) = 0$ zapisujemy używając rozkładu na sumę prostą $X = T_{x_0}(M) \oplus X_2$: $G(x_0 + v + z) = 0$, gdzie $v \in T_{x_0}(M), z \in X_2$ i traktujemy jako równanie wyznaczające y jako uwikłaną funkcję v. Założenia twierdzenia o funkcjach uwikłanych są spełnione w otoczeniu punktu $v = 0 = y$, ponieważ $G'(x_0)|_{x_2}$ jest odwracalne. Wobec tego w pewnym otoczeniu punktu $v = 0$ równanie to wyznacza odwzorowanie $v \mapsto \phi(v)$ takie, że $G(x_0 + v + \phi(v)) = 0$. Obliczamy $\phi'(0)$: $G'(x_0)|_{T_{x_0}(M)} + G'(x_0)|_{X_2}\phi'(0) = 0$. Ponieważ $G'(x_0)|_{T_{x_0}(M)} = 0$ mamy $\phi'(0) = 0$. Odwzorowanie π określamy wzorem $\pi(x_0 + v) := x_0 + v + \phi(v)$. Ponieważ $\phi'(0) = 0$ mamy $\phi(v) = o(||v||)$, co kończy dowód lematu.

Powracamy do dowodu twierdzenia. Bierzemy punkt $x \in M$ leżący w otoczeniu x_0 w M, o którym mówi lemat. Wtedy $x = \pi(x_0 + v)$ dla pewnego $v \in T_{x_0}(M)$. Wzór Taylora w otoczeniu x_0 dla funkcji $f - \Lambda \circ G$:

$$\begin{aligned} & f(x) - \Lambda \circ G(x) = f(x_0) - \Lambda G(x_0) + [f'(x_0) - \Lambda G'(x_0)](x - x_0) + \\ & + \frac{1}{2}[f''(x_0) - \Lambda G''(x_0)](x - x_0)^{(2)} + o(||x - x_0||) = \\ & = f(x_0) + \frac{1}{2}[f''(x_0) - \Lambda G''(x_0)](x - x_0)^{(2)} + o(||x - x_0||^2). \end{aligned}$$

Na podstawie lematu wiemy, że $x = \pi(x_0 + v)$, gdzie $v \in T_{x_0}(M)$ oraz $x - x_0 = x_0 + v + o(||v||)$. Wobec tego

$$[f''(x_0) - \Lambda G''(x_0)](x - x_0)^{(2)} = [f''(x_0) - \Lambda G''(x_0)]v^{(2)} + o(||v||^2),$$

a $o(||v||^2)$ jest także $o(||x - x_0||^2)$, więc postępując dalej tak, jak w przypadku zwykłych ekstremów otrzymujemy tezę twierdzenia.