Dowód: wystarczy ograniczyć się do przypadku $m=1$, bo f jest różniczkowalna w x wtedy i tylko wtedy, gdy wszystkie f_i są różniczkowalne w x, podobnie f ma w punktach $B(x,\varepsilon)$ pochodne cząstkowe, ciągłe w x wtedy i tylko wtedy, gdy mają je f_i dla $i=1,2,\dots,m$.

Ustalmy $h=(h_1,\dots,h_n)$ takie, że $x+h\in B(x,\varepsilon)\subset U$ (a więc $\|h\|<\varepsilon$).

Przyjmijmy $p_0=x$, $p_1=x+h_1e_1=(x_1+h_1,x_2,\dots,x_n)$

$p_2=p_1+h_2e_2=(x_1+h_1,x_2+h_2,x_3,\dots,x_n)$ itd

$p_k=p_{k-1}+h_ke_k=(x_1+h_1,\dots,x_k+h_k,x_{k+1},\dots,x_n)$

$p_n=(x_1+h_1,\dots,x_n+h_n)=x+h$.

Przyjmijmy teraz $F_k:[0,h_k]\to\mathbb{R}$ dla $k=1,2,\dots,n$

$$F_k(t)=f(p_{k-1}+te_k)=$$
$$=f(x_1+h_1,\dots,x_{k-1}+h_{k-1},x_k+t,x_{k+1},\dots,x_n)$$

$F^k:[0,h_k]\to\mathbb{R}$ jest ciągła (na $[0,h_k]$) i różniczkowalna na $(0,h_k)$, możemy więc do F^k użyć tw. Lagrange'a:

$$f(p_k)-f(p_{k-1})=F_k(h_k)-F_k(0)=F_k'(\xi_k)\cdot h_k$$
dla pewnego $\xi_k\in(0,h_k)$

Obliczmy $F_k'(t)$:

$$F_k'(t) = \lim_{s\to 0} \frac{F_k(t+s) - F_k(t)}{s} =$$

$$= \lim_{s\to 0} \frac{f(p_{k-1} + te_k + se_k) - f(p_{k-1} + te_k)}{s} =$$

$$= \frac{\partial f}{\partial x_k}(p_{k-1} + te_k), \text{ zatem}$$

$$f(p_k) - f(p_{k-1}) = \frac{\partial f}{\partial x_k}(p_{k-1} + \xi_k e_k) h_k$$

dla pewnego $\xi_k \in (0, h_k)$

Stąd

$$f(x+h) = f(p_n) = f(p_0) + (f(p_1) - f(p_0)) + (f(p_2) - f(p_1)) +$$
$$+ \dots + (f(p_n) - f(p_{n-1})) =$$
$$= f(x) + \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(p_{k-1} + \xi_k e_k) \cdot h_k$$
$$= f(x) + \underbrace{\sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x) \cdot h_k}_{\substack{= Df(x)\cdot h \\ \text{o ile } f \text{ różniczkowalna} \\ \text{w } x}} + \underbrace{\sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}(p_{k-1} + \xi_k e_k) - \frac{\partial f}{\partial x_k}(x) \right) h_k}_{\substack{\text{kandydat na resztę} \\ R(h)}}$$

Co się dzieje z $\frac{R(h)}{\|h\|}$ gdy $h \to 0$?

Zauważmy, że gdy $h \to 0$, to ~~również~~ $p_{k-1} + \xi_k e_k \to x$; z ciągłości pochodnych cząstkowych $\frac{\partial f}{\partial x_k}$ w $B(x, \epsilon)$ dostajemy, że $\left(\frac{\partial f}{\partial x_k}(p_{k-1} + \xi_k e_k) - \frac{\partial f}{\partial x_k}(x)\right) \xrightarrow[h \to 0]{} 0$.

Równocześnie $\left|\frac{h_k}{\|h\|}\right| \leq 1$, zatem $\frac{R(h)}{\|h\|} \xrightarrow[h \to 0]{} 0$,
jak powinno.

a w zasadzie i twierdzenie.

<u>Definicja</u>: Niech X będzie przestrzenią Hilberta, $\mathcal{U} \subset X$ otwarty, $f: \mathcal{U} \to \mathbb{R}$. Załóżmy, że f jest różniczkowalna w $x \in \mathcal{U}$. Wówczas istnieje dokładnie jeden element $g \in X$ taki, że $\forall_{h \in X}\ Df(x)h = \langle g, h \rangle_X$; element ten nazywamy gradientem f w punkcie x i oznaczamy $\nabla f(x)$ lub $\operatorname{grad} f(x)$.

Dowód: Musimy wykazać istnienie i jednoznaczność takiego g; dowód opiera się na tzw. Lemacie Riesza o reprezentacji.

Lemat (Frigyesa Riesza o reprezentacji)
Dla każdego ciągłego funkcjonału liniowego φ na przestrzeni Hilberta X istnieje dokładnie jeden element $y_\varphi \in X$ taki, że $\forall_{x \in X}\ \varphi(x) = \langle y_\varphi, x \rangle_X$.

Gdy mamy lemat Riesza, nic nie pozostaje do dowiedzenia, bo $Df(x): X \to \mathbb{R}$ jest ciągłym funkcjonałem liniowym na X.

Dowód lematu Riesza

- Przypomnijmy tzw. tożsamość równoległoboku: jeżeli X jest p-nią Hilberta, to $\forall_{x,y\in X}$

$$\|x+y\|^2+\|x-y\|^2 = 2\|x\|^2+2\|y\|^2.$$

- Rozważmy teraz $V=\ker\varphi=\varphi^{-1}(0)$.

To jest liniowa, domknięta podprzestrzeń X.

Mamy 2 możliwości: albo $\forall_x \ \varphi(x)=0$, wtedy $V=X$, a lemat Riesza zachodzi z $y_\varphi=0$, albo istnieje $x_0\in X$ t.ż. $\varphi(x_0)\neq 0$. Wykażemy, że w X istnieje niezerowy element prostopadły do V, a więc $0\neq g\in X$ t.ż. $\forall_{v\in V} \ \langle g,v\rangle=0$.

Rozważmy $\alpha=\inf\{\|x_0+v\|^2 : v\in V\}$.

Istnieje ciąg (v_n) taki, że $\forall_n \ v_n\in V$, $\|x_0+v_n\|^2\to\alpha$.

~~Z tożsamości równoległoboku, z x=~~

Ustalmy $\varepsilon>0$ i niech $N\in\mathbb{N}$ będzie takie, że $\forall_{n>N} \ \alpha\le\|x_0+v_n\|^2<\alpha+\varepsilon$. Wtedy, dla $n,m>N$, z tożsamości równoległoboku ($x=x_0+v_n$, $y=x_0+v_m$)

$$\|v_n-v_m\|^2=2\left(\|x_0+v_n\|^2+\|x_0+v_m\|^2\right)-\|2x_0+v_n+v_m\|^2$$

$$\le 2(\alpha+\varepsilon+\alpha+\varepsilon)-4\left\|x_0+\frac{v_n+v_m}{2}\right\|^2=(*)$$

ale $\frac{v_n+v_m}{2} \in V$, więc $\|x_0+\frac{v_n+v_m}{2}\|^2 \geq \alpha$,

stąd $\|v_n-v_m\|^2 \leq (*) \leq 4\alpha + 4\varepsilon - 4\alpha = 4\varepsilon$.

(o ile $n,m > N$)

Tym samym wykazaliśmy, że ciąg (v_n) jest ciągiem Cauchy'ego w X, ma zatem granicę:

$v_n \xrightarrow[n\to\infty]{} v$. Co więcej, $v \in V$, bo V jest domknięta.

Niech zatem $g = x_0 + v$. Dla każdego $w \in V$, $\lambda \in \mathbb{R}$

$$\underbrace{\alpha = \|g\|^2 \leq \|g+\lambda w\|^2} = \|x_0 + \underbrace{v+\lambda w}_{\in V}\|^2$$

$$\|g\|^2 \leq \|g\|^2 + 2\lambda\langle g,w\rangle + \lambda^2\|w\|^2, \text{ czyli}$$

$$\lambda^2\|w\|^2 + 2\lambda\langle g,w\rangle \geq 0 \quad \text{dla wszystkich } \lambda > 0$$

Wyróżnik Δ musi więc być ≤ 0, $4\langle g,w\rangle^2 \leq 0$

$$\Rightarrow \langle g,w\rangle = 0 \quad \forall_{w\in V}.$$

Musimy jeszcze sprawdzić, czy $g \neq 0$, ale gdyby $g=0$, to $x_0 + v = 0 \Rightarrow$ ~~$x_0 \in V$~~ $x_0 = -v \in V = \ker\varphi$ $\Rightarrow \varphi(x_0) = 0$, wbrew założeniu.

- No to mamy $g \neq 0$, $g \perp \ker\varphi$.
  Dla każdego $x \in X$ zachodzi

$$\varphi(g)\cdot x - \varphi(x)g \in \ker\varphi, \text{ bo}$$

$$\varphi(\varphi(g)x - \varphi(x)g) = \varphi(g)\varphi(x) - \varphi(x)\varphi(g) = 0$$

No to wiemy, że $\forall_{x \in X} \langle g, \varphi(g)x - \varphi(x)g \rangle = 0$,

cyli
$$\varphi(x)\|g\|^2 = \varphi(g)\langle g, x\rangle$$

$$\varphi(x) = \left\langle \frac{\varphi(g)g}{\|g\|^2}, x \right\rangle$$

i wystarczy przyjąć $y_\varphi = \frac{\varphi(g)}{\|g\|^2} g$.

- Pozostaje wykazanie, że jest dokładnie jedno takie y_φ. Załóżmy, że dla pewnego $\varphi : X \to \mathbb{R}$ mamy $\forall_{x \in X}$ $\varphi(x) = \langle y_\varphi, x\rangle = \langle z_\varphi, x\rangle$.

Wtedy, $\forall_{x \in X}$ $\langle y_\varphi - z_\varphi, x\rangle = 0$,

w szczególności dla $x = y_\varphi - z_\varphi$

$$\|y_\varphi - z_\varphi\|^2 = 0 \Rightarrow z_\varphi = y_\varphi.$$

$\square$.

Twierdzenie: Jeżeli $f: \mathcal{U} \to \mathbb{R}$ jest różniczkowalna
$\mathcal{U} \subset X$, X Hilberta

w $x_0 \in \mathcal{U}$ i $v = \nabla f(x_0) \neq 0$,

to dla każdego $w \in X$ ~~zachodzi~~ takiego, że $\|v\| = \|w\|$

zachodzi $D_w f(x_0) \leq D_v f(x_0)$.

(a więc gradient wskazuje kierunek najszybszego wzrostu f).

Dowód:

$$D_w f(x_0) = Df(x_0)w = \langle \nabla f(x_0), w \rangle = \langle v, w \rangle \overset{\text{nier. Schwarza}}{\leq} \|v\| \cdot \|w\|$$

$$\|v\| \cdot \|w\| = \|v\|^2$$

$$D_v f(x_0) = Df(x_0)v = \langle \nabla f(x_0), v \rangle = \langle v, v \rangle = \|v\|^2$$

$\square$

Uwaga: Trywialne modyfikacje tego dowodu pokażą, że

$$\forall_{w \in V} \quad D_w f(x_0) \geq D_{(-v)} f(x_0),$$

więc kierunek najwolniejszego wzrostu wskazany jest przez $-v = -\nabla f(x_0)$.

Uwaga 2: W nierówności Schwarza równość zachodzi tylko wtedy, gdy wektory w i v są równoległe, a skoro $\|w\| = \|v\|$, to tylko wtedy, gdy $w = v$. (dla $w = -v$ mamy $\langle v, w \rangle = -\|v\|^2$)

Stąd w twierdzeniu nierówność jest ostra, chyba, że $v = w$.

Def. Mówimy, że $f\colon U \to \mathbb{R}$ ma w $x_0 \in U$ ekstremum (minimum/maksimum) lokalne, jeżeli istnieje $\varepsilon > 0$ takie, że $\forall_{y \in B(x_0, \varepsilon)}$

- $f(y) \geq f(x_0)$
  ~~maksimum~~ minimum lokalne
- $f(y) \leq f(x_0)$
  maksimum lokalne
- $f(y) > f(x_0)$, o ile $y \neq x_0$
  minimum lokalne właściwe
- $f(y) < f(x_0)$, o ile $y \neq x_0$
  maksimum lokalne właściwe.

Tw. (lemat Fermata) Jeżeli f ma w x_0 ekstremum lokalne i istnieje pochodna kierunkowa

$$D_v f(x_0) = \lim_{\underline{t \to 0}} \frac{f(x_0 + tv) - f(x_0)}{t},$$

$\uparrow$
$\boxed{\text{nie } t \to 0^+}$

to jest ona równa zero.

W szczególności jeżeli f ma w x_0 pochodną (Gateaux lub silną, Frécheta), to $Df(x_0) = 0$.

Dowód: Funkcja jednej zmiennej $g(t) = f(x_0 + tv)$ ma w $t = 0$ ekstremum lokalne, więc
$0 = g'(t) = D_v f(x_0)$.

Uogólnieniem pojęcia ekstremum lokalnego na przekształcenia $f: \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}^m$ jest pojęcie <u>punktu krytycznego</u>.

<u>Def:</u> Niech $\mathcal{U} \subset \mathbb{R}^n$ będzie otwarty, $f: \mathcal{U} \to \mathbb{R}^m$ niech będzie różniczkowalna w $x_0 \in \mathcal{U}$. Mówimy, że f ma w x_0 <u>punkt krytyczny</u>, jeżeli macierz $Df(x_0) \in M_{n \times m}$ ma rząd mniejszy niż $\min(n,m)$, a więc mniejszy, niż maksymalny.

Dla $f: \mathcal{U} \to \mathbb{R}$ oznacza to, że f ma w x_0 punkt krytyczny, gdy $\operatorname{rank} Df(x_0) < 1$, a więc $Df(x_0) = 0$.

~~Punkty zbioru~~ Jeżeli f jest różniczkowalna na $\mathcal{U}$, to ~~zbiór tych $x \in$~~ punkty $x \in \mathcal{U}$, które nie są krytyczne, nazywamy <u>punktami regularnymi</u> f.

<u>Zadanie:</u> Niech $f \in C^1(\mathcal{U}, \mathbb{R}^m)$. Wykazać, że zbiór punktów regularnych f jest otwarty.

Dla funkcji określonych na nieskończenie wymiarowych przestrzeniach Banacha nie możemy mówić o rzędzie pochodnej, ale jeżeli X jest p-nią Banacha, ~~Jeżeli~~ $U \subset X$ jest otwarty, $f: U \to \mathbb{R}$ i funkcja (funkcjonał, tylko niekoniecznie liniowy) f jest w x_0 różniczkowalna w sensie Gateaux, to mówimy, że f ma w x_0 punkt krytyczny, gdy ta pochodna Gateaux jest w x_0 równa 0:

$Df(x_0) = 0$, czyli $\forall_{y \in X}$

$$\lim_{t \to 0} \left(f(x_0 + ty) - f(x_0) \right) / t = 0 .$$

Oczywiście każde ekstremum lokalne funkcji różniczkowalnej jest punktem krytycznym i (równie oczywiście) nie jest odwrotnie, wystarczy przyjrzeć się $f(x,y) = xy$ w pobliżu $(0,0)$.

Polecam Państwu staranną lekturę stron 30-35 skryptu P. Strzeleckiego i przeanalizowanie przedstawionych tam przykładów.