Диференцијални рачун функција више променљивих

Настављајући као у једнодимензионалном случају, након увођења појма непрекидности, следеће је увођење појма непрекидности. Надаље разматрамо само функције дефинисане на подскуповима A од $\mathbb{R}^m$, са вредностима у $\mathbb{R}$, које називамо реалне (скаларне) функције више променљивих, као и са вредностима у $\mathbb{R}^n$, које називамо векторске функције. Даље, настављамо са нотацијом од раније, тачке x скупа A записујемо и координатно као $x = (x_1, x_2, \ldots, x_m)$ и аналогно $y = f(x) = f(x_1, x_2, \ldots, x_m)$. Сам скуп $\mathbb{R}^m$ посматрамо са метриком d_2, односно нормом $\|\cdot\|_2$, ако се не нагласи другачије. Даље, као и раније, уколико је кодомен $\mathbb{R}^n$, тада је $y = (y_1, y_2, \ldots, y_n)$ и имамо дефинисане координатне функције $f_i\colon A \to \mathbb{R}$, са $f_i(x_1, x_2, \ldots, x_m) = y_i$.

Даље, приметимо да ако хоћемо да нам је израз $f(a+h)$ дефинисан за h довољно мало, неопходно је да тачка a припада скупу A са неком околином. Отуда, надаље по питањима диференцијабилности, претпостављаћемо да је скуп A отворен. Као и у једнодимензионалном случају, израз $f(a+h)$ зависи само од вредности функције на околини тачке a, а отуда ћемо претпостављати и да је скуп A повезан. Овакве скупове (отворене и повезане) називамо још и област у простору $\mathbb{R}^m$.

Парцијални изводи и диференцијабилност скаларних функција

Нека је A област у $\mathbb{R}^m, f\colon A \to \mathbb{R}$ и $a \in A$. Посматрајмо вектор $h \in \mathbb{R}^m$ чије су све координате осим $i-$те једнаке нули, а $i-$та износи h_i, тј $h = (0, \ldots, 0, h_i, 0, \ldots, 0)$. Приметимо да за h_i довољно мало, из отворености скупа A, имамо да је $a + h \in A$. Отуда има смисла разматрати лимес $\frac{f(a+h)-f(a)}{h_i}$. Прецизније

Дефиниција Ако, под претходним претпоставкама, постоји гранична вредност

$$\lim_{h_i \to 0} \frac{f(a_1, a_2, \ldots, a_{i-1}, a_i + h_i, a_{i+1} \ldots, a_m) - f(a_1, a_2, \ldots, a_m)}{h_i},$$

називамо је парцијалним изводом функције f по $i-$тој променљивој у тачки a. Означава се углавном са $\dfrac{\partial f}{\partial x_i}(a)$, као и са $f'_{x_i}(a), \partial_i f(a)$...

Парцијални извод суштински није нов појам. Он је заправо обичан извод функције коју добијамо до функције f када фиксирамо све променљиве осим $i-$те. Самим тим, његово рачунање се своди на уобичајена правила диференцирања функција једне променљиве.

Из тога имамо и да вредност парцијалног извода (као и његово постојање), зависи само од вредности које функција узима на дуж $i-$те координатне осе, те не узима у обзир вредности у целој околини. Отуда, не чуди да постоје функције које су довољно лепе по неким правцима одакле имају парцијалне изводе, али да са друге стране нису непрекидна. Нпр.

Пример Ако посматрамо функцију $f\colon \mathbb{R}^2 \to \mathbb{R}$, дату са

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2}, (x,y) \neq (0,0) \\ 0, (x,y) = (0,0), \end{cases}$$

она није непрекидна у нули, јер је нпр. $f(x,x) = \frac{1}{2}$, што не тежи нули када $x \to 0$. Са друге стране је

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h \to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h \to 0} \frac{\frac{h0}{h^2+0^2} - 0}{h} = 0,$$

а слично се показује и да је $\frac{\partial f}{\partial y}(0,0) = 0$.

Дакле, за аналгију појма диференцијабилности нам је потребно више од постојања парацијалних извода. Ако посматрамо сада да је h произвољан вектор, такав да је $a + h \in A$, и ако посматрамо прираштај функције $f(a+h) - f(a)$, сетимо се да је један од еквивалената диференцијабилности у једнодимензионалном случају било да је главни део прираштаја линеаран. Преносећи ту идеју, имали би да је $f(a+h) - f(a) = Lh + o(h)$, где је $L\colon \mathbb{R}^m \to \mathbb{R}$ сада линеарна функција, а $o(h)$ брже тежи нули од $\|h\|$. Прецизније

Дефиниција Нека је $A \subset \mathbb{R}^m$ област, $f\colon A \to \mathbb{R}$ и $a = (a_1, a_2, \ldots, a_m) \in A$ фиксирана тачка. Нека је $h = (h_1, \ldots, h_m) \in \mathbb{R}^m$ такво да је $\|h\|$ довољно мала да би $a + h \in A$. Уколико се прираштај може записати у облику

$$f(a_1 + h_1, a_2 + h_2, \ldots, a_m + h_m) - f(a_1, a_2, \ldots, a_m) = L_1 h_1 + L_2 h_2 + \ldots + L_m h_m + o(h),$$

где су $L_1, \ldots, L_m$ константе, а са $o(h)$ је означена функција од h за коју важи $\lim_{h \to 0} \frac{o(h)}{\|h\|} = 0$, кажемо да је функција f диференцијабилна у тачки a, а линеарну функију $h \to L_1 h_1 + L_2 h_2 + \ldots + L_m h_m$ називамо диференцијалом функције f у тачки a и пишемо

$$df(a)h = L_1 h_1 + L_2 h_2 + \ldots + L_m h_m.$$

Став Нека је $f\colon A \to \mathbb{R}$ диференцијабилна функција у тачки $a \in A$. Тада

1. Функција f је непрекидна у тачки a.

2. Функција f има парцијалне изводе по свим променљивим у тачки a, а важи и $\dfrac{\partial f}{\partial x_i} = L_i$, у нотацији претходне дефиниције.

Доказ погледати у књизи.

Напоменимо да не важи обрат овог става, наиме ако је функција непрекидна и има све парцијалне изводе она не мора бити диференцијабилна (пример се може видети у књизи).

Са друге стране, слично као у једнодимензионалном случају, непрекидност парцијалних извода је јачи услов од диференцијабилности, прецизније

Став Нека у некој околини U тачке a функција f има дефинисане све парцијалне изводе (дакле у свакој тачки околине има све парцијалне изводе) и нека су они непрекидни у тачки a. Тада је f диференцијабилна у тачки a.

Ако посматрамо разлику

$$\begin{aligned}&f(a_1+h_1,a_2+h_2,\dots,a_m+h_m)-f(a_1,a_2,\dots,a_m)\\ =&f(a_1+h_1,a_2+h_2,\dots,a_m+h_m)-f(a_1,a_2+h_2,\dots,a_m+h_m)+f(a_1,a_2+h_2,\dots,a_m+h_m)-f(a_1,a_2,\dots,a_m)\\ =&\frac{\partial f}{\partial x_1}(a_1+\theta h_1,a_2+h_2,\dots,a_m+h_m)h_1+f(a_1,a_2+h_2,\dots,a_m+h_m)-f(a_1,a_2,\dots,a_m),\end{aligned}$$

где је $\theta\in(0,1)$, а последњи корак је последица Лагранжеве Теореме. Њена примена је оправдана тиме да за h довољно мало, цела дуж $a+th$, за $t\in[0,1]$, припада околини U. Одатле, ако посматрамо f само као функцију прве променљиве, она је и непрекидна и диференцијабилна на поменутој дужи те можемо да применимо Лагранжеву Теорему. Даље, из непрекидности парцијалног извода имамо да

$$\frac{\partial f}{\partial x_1}(a_1+\theta h_1,a_2+h_2,\dots,a_m+h_m)\to\frac{\partial f}{\partial x_1}(a_1,a_2,\dots,a_m),$$

а самим тим разлика

$$\Big(\frac{\partial f}{\partial x_1}(a_1+\theta h_1,a_2+h_2,\dots,a_m+h_m)-\frac{\partial f}{\partial x_1}(a_1,a_2,\dots,a_m)\Big)h_1,$$

је $o(h)$. Поступак настављамо даље, примењујући претходно резоновање на променљиву x_2, итд. Детаље погледати у књизи.

Као и у једнодимензионалном случају, претходни услов није неопходан, а функције чији су изводи непрекидни називамо непрекидно диференцијабилне. Често се користи ознака $C^1(A)=\{f\colon A\to\mathbb{R}|f-$ непрекидно диференцијабилна$\}$.

Извод у правцу, градијент

Као што је речено у претходном, парцијални извод по некој променљивој је само обичан извод функције коју добијамо ако фиксирамо све остале променљиве. Другачије га можемо интерпретирати као извод дуж правца паралелног одговарајућој координатној оси. Одатле видимо да је њихов избор донекле вештачки, јер би аналогно могли да разматрамо и било који правац (то што су "фаворизовани" правци паралелни координатним осама је само ствар записа). Из тога имамо следећу дефиницију

Дефиниција Нека је A- област, $f\colon A\to\mathbb{R}^m$ и нека је $l\in\mathbb{R}^m$ вектор. Коначна гранична вредност (ако постоји)

$$\lim_{t\to0}\frac{f(a+tl)-f(a)}{t},$$

назива се извод функције f у тачки a у правцу вектора l. Означава се са $f'_l(a)$, $\dfrac{\partial f}{\partial l}(a)$.

Претходна дефиниција има смисла, јер за t довољно мало, због отворености скупа A и тачка $a+tl$ припада скупу A.

Уколико онзачимо са $e_k\in\mathbb{R}^m$ вектор који на k-том месту има јединицу а на свим осталим нуле, видимо да је

$$\frac{\partial f}{\partial e_k}(a)=\frac{\partial f}{\partial x_k}(a).$$

Наредни став нам даје везу између извода у правцу и диференцијала.

Став Нека је $A\subset\mathbb{R}^m$ област, $f\colon A\to\mathbb{R}$ диференцијабилна функција у тачки $a\in A$ и нека је $l=(l_1,l_2,\dots,l_m)$ вектор. Тада функција f има извод у правцу вектора l и важи

$$\frac{\partial f}{\partial l}(a)=\sum_{k=1}^m\frac{\partial f}{\partial x_k}(a)l_k=df(a)l.$$

Доказ погледати у књизи.

Приметимо прво да из претходног става имамо инваријантну дефиницију диференцијала, прецизније дефиницију која не зависи од координата. Дакле, диференцијал можемо и да уведемо са

$$df(a)l=\lim_{t\to0}\frac{f(a+tl)-f(a)}{t},$$

у коме не фигурише координатни запис $(x_1,\dots,x_m)$. Приметимо да је диференцијал (у фиксираној тачки a) функционал, тј. он је линеарно пресликавање вектора у скаларе.

Даље, вектор $(\dfrac{\partial f}{\partial x_1}(a),\dfrac{\partial f}{\partial x_2}(a),\dots,\dfrac{\partial f}{\partial x_m}(a))$ називамо градијентом функције f у тачки a и означава се са

$\operatorname{grad} f(a)$, као и са $\nabla f(a)$. Даље, уколико означимо са $\langle , \rangle$ уобичајени скаларни производ у $\mathbb{R}^m$, дат са $\langle x, y\rangle \sum_{k=1}^m x_k y_k$, имамо да нам претходни став даје важну везу (дуалност) градијента и диференцијала

$$df(a)l = \langle l, \operatorname{grad} f\rangle.$$

Диференцијабилност векторских функција

За почетак, укратко прођимо мало формалније кроз линеарна пресликавања у случају којим се бавимо. За пресликавање $L\colon \mathbb{R}^m \to \mathbb{R}^n$ кажемо да је линеарно ако важи

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y),$$

где су $x, y \in \mathbb{R}^m$ а $\alpha, \beta \in \mathbb{R}$. Приметимо да свака матрица $A \in \mathcal{M}(n \times m, \mathbb{R})$ задаје пример оваквог пресликавања са

$$Lx = Ax = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix},$$

јер множење матрица има одговарајућа својства. Са друге стране, ако опет означимо са $e_k \in \mathbb{R}^m$ векторе који на k-ој координати имају јединицу а иначе нуле, приметимо прво да вектор $x = (x_1, \dots, x_m)$ можемо да запишемо са $x = \sum_{k=1}^m x_k e_k$, а одатле, ако је L произвољно линеарно пресликавање, имамо да је

$$L(x) = L\left(\sum_{k=1}^m x_k e_k\right) = \sum_{k=1}^m x_k L(e_k).$$

Одавде, видимо да је свако линеарно пресликавање одређено са сликама вектора e_k. Даље, $L(e_k) \in \mathbb{R}^n$ по претпоставци, односно $L(e_k)$ је вектор одређен са n координата. Ако формирамо матрицу $A_L \in \mathcal{M}(n \times m, \mathbb{R})$ чија се k-та колона састоји од координата тог вектора, имамо да је $Lx = A_L x$ (заправо вектори e_k се пресликавају у исте векторе при оба пресликавања, а тиме су ова два пресликавања иста јер су оба линеарна). Тиме имамо да су заправо сва линеарна пресликавања облика Ax где је $A \in \mathcal{M}(n \times m, \mathbb{R})$. Из тог разлога је између осталог уобичајено да се не пишу заграде код линеарних преслиакврања.

За линеарно пресликавање, односно матрицу A, уводимо норму са

$$\|A\| = \left(\sum_{j=1}^n \sum_{k=1}^m a_{jk}^2\right)^{1/2}.$$

Тада, из Коши - Шварцове неједнакости имамо да је (у 2 нормама на $\mathbb{R}^n$ односно $\mathbb{R}^m$)

$$\|Ax\|_2^2 = \sum_{j=1}^n \left(\sum_{k=1}^m a_{jk} x_k\right)^2 \leqslant \sum_{j=1}^n \left(\sum_{k=1}^m a_{jk}^2 \sum_{k=1}^m x_k^2\right) = \sum_{j=1}^n \sum_{k=1}^m a_{jk}^2 \|x\|_2^2,$$

односно, узимањем корена

$$\|Ax\|_2 \leqslant \|A\| \|x\|_2.$$

Вратимо се сада на произвољне функције

Дефиниција Нека је $A \subset \mathbb{R}^m$ област, $f\colon A \to \mathbb{R}^n$ и $a \in A$. За функцију f кажемо да је диференцијабилна у тачки a ако је главни део њеног прираштаја $f(a+h) - f(a)$ линеаран, тј. ако важи

$$f(a+h) - f(a) = Lh + o(h),$$

где је L линеарно пресликавање, а $o(h)$ је функција за коју важи $\lim_{h\to 0} \frac{\|o(h)\|}{\|h\|} = 0$. Сама линеарна функција L назива се диференцијалом функције f у тачки a.

Уколико уобичајено означимо диференцијал са $df(a)h$, другачије речено, функција је диференцијабилна у тачки a ако важи

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

где је норма у бројиоцу норма простора $\mathbb{R}^n$ а у имениоцу је норма простора $\mathbb{R}^m$.

Пређимо сада на координатни запис претходне релације. За почетак, $df(a)h$ је линеарно пресликавање, а по претходном је дакле у питању матрица, тј. $df(a) \in \mathcal{M}(n \times m, \mathbb{R})$. Даље, функцију $o(h)$ означимо са $\alpha(h)$. Како су f и α векторске функције, оне су облика $f = (f_1, f_2, \dots, f_n)$ односно $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, где су $f_i\colon A \to \mathbb{R}$ и $\alpha_i\colon U \to \mathbb{R}$, где је U нека околина тачке a. Такође знамо да $\alpha_i(h) \to 0$, када $h \to a$, (следи из тога што је конвергенција у $\mathbb{R}^n$ еквивалентна конвергенцији по координатама). Најзад, нека је матрица диференцијала дата са

$$df(a) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}.$$

Посматрајмо сада $k-$ту координату израза $f(a+h)-f(a)-df(a)h$, која је једнака $\alpha_k(h)$. Дакле, што се функције f тиче имамо њену координатну функцију f_k, а што се $df(a)h$ тиче, на $k-$том месту поменутог производа матрица имамо $k-$ту врсту у производу са вектором h, што износи $\sum_{j=1}^m a_{kj}h_j$. Прецизније

$$f_k(a+h)-f_k(a)=\sum_{j=1}^m a_{kj}h_j+\alpha_j(h),$$

а како $|\alpha_j(h)|\leqslant(\sum_{k=1}^n\|\alpha_k\|^2)^{1/2}=\|\alpha\|$, то и $\frac{|\alpha_j(h)|}{\|h\|}\to 0$. Прецизније, функција f_k је диференцијабилна и њен диференцијал је дат са $df_k(a)h=\sum_{j=1}^m a_{kj}h_j$. Са друге стране, од раније знамо да је $df_k(a)h=\sum_{j=1}^m \frac{\partial f_k}{\partial x_j}hj$, односно $a_{kj}=\frac{\partial f_k}{\partial x_j}$. Дакле, матрица линеарног пресликавања је дата са

$$\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \frac{\partial f_1}{\partial x_m}(a)\\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \dots & \frac{\partial f_2}{\partial x_m}(a)\\ \vdots & \vdots & \dots & \vdots\\ \frac{\partial f_n}{\partial x_1}(a) & \frac{\partial f_n}{\partial x_2}(a) & \dots & \frac{\partial f_n}{\partial x_m}(a)\end{pmatrix}=\begin{pmatrix}\text{grad } f_1\\ \text{grad } f_2\\ \vdots\\ \text{grad } f_n\end{pmatrix}.$$

Последњу матрицу називамо Јакобијевом матрицом, а и првим изводом функције f у тачки a и означавамо је са $f'(a)$. Такође, обрнутим поступком, имамо и да ако су све функције f_k диференцијабилне, тада је и f диференцијабилна. Наиме, у истој нотацији, тада имамо да све функције $\alpha_k\to 0$, а одатле би и одговарајућа векторска функција тежила нули, опет због поменуте еквивалентности конвергенције по координатама и у норми.

Приметимо да је претходна дефиниција извода сагласна са претходним, у једнодимензионалном случају, као и у случају скаларних функција, наиме, извод је увек био линеарна апроксимација пресликавања у датој тачки.

У складу са тиме, приметимо да уколико имамо пресликавање $L\mathbb{R}^m\to\mathbb{R}^n$ које је већ линеарно, тј. матрица, тада је Јакобијева матрица пресликавања L у свакој тачки једнака L. Ово се наравно може проверити директним рачуном по координатама, али уместо тога ћемо само закључити да је линеарна апроксимација dL линеарног пресликавања L опет L, тј. у овом случају је грешка $\alpha(h)$ идентички једнака нули.

У случају $m=n$ дефинисана је и детерминанта Јакобијеве матрице, која се назива Јакобијан. Згодна ознака за њу је

$$\frac{D(f_1,f_2,\dots,f_m)}{D(x_1,x_2,\dots,x_m)}=\det\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \frac{\partial f_1}{\partial x_m}(a)\\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \dots & \frac{\partial f_2}{\partial x_m}(a)\\ \vdots & \vdots & \dots & \vdots\\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \dots & \frac{\partial f_m}{\partial x_m}(a)\end{pmatrix}.$$

Даље, из претходне везе диференцијабилности и непрекидности координатних функција и векторске функције имамо и

Став Диференцијабилна функција $f\colon A\to\mathbb{R}^n$, где је $A\subset\mathbb{R}^m$, у тачки $a\in A$ је такође непрекидна у тој тачки.

Правила диференцирања

Даље пролазимо кроз основна својства извода, односно његово слагање са уобичајеним алгебарским операцијама. Углавном су тврђења аналогна једнодимензионалном случају, а сами доказе погледати у књизи.

Теорема Нека су функције $f,g\colon A\to\mathbb{R}^n(A\subset\mathbb{R}^m)$ диференцијабилне у тачки $a\in A$ и нека су $\lambda,\mu\in\mathbb{R}$. Тада је функција $\lambda f+\mu g$ диференцијабилна у a и важи

$$d(\lambda f+\mu g)(a)=\lambda df(a)+\mu dg(a).$$

Теорема Нека су $f,g\colon A\to\mathbb{R}(A\subset\mathbb{R}^m)$ диференцијабилне у тачки a. Тада су и функције fg и f/g диференцијабилне у тачки a (друга ако је $g(a)\neq 0$) и важи

$$d(fg)(a)=g(a)df(a)+f(a)dg(a),$$
$$d\Big(\frac{f}{g}\Big)(a)=\frac{g(a)df(a)-f(a)dg(a)}{g(a)^2}.$$

Наравно претходне две Теореме могу се формулисати и у терминима Јакобијевих матрица, где важе аналогна правила рачунања.

Теорема Нека су $A\subset\mathbb{R}^m$ и $B\subset\mathbb{R}^n$ отворени скупови, $f\colon A\to B$ и $g\colon B\to\mathbb{R}^p$. Ако је функција f диференцијабилна у тачки a, а g диференцијабилна у тачки $b=f(a)$, тада је композиција $g\circ f$ диференцијабилна у тачки a и важи

$$d(g\circ f)(a)=dg(b)\circ df(a),$$

где је са десне стране једнакости композиција одговарајућих линеарних пресликавања. У терминима Јакобијевих маатрица, са десне стране је производ матрица $g'(b)$ и $f'(a)$. Запишимо последњу формулу експлицитно преко Јакобијевих матрица. У том контексту, означимо координате у $\mathbb{R}^m$ са x, а у $\mathbb{R}^n$ са y (или ти означимо функцију f као функцију од x а g коа функцију од y). Тада је Јакобијева матрица функције f у тачки a дат са $[\frac{\partial f_j}{\partial x_i}]$, где је $j = 1, \ldots, n$ и $i = 1, \ldots m$, дока је матрица $g'(a)$ дата са $[\frac{\partial g_k}{\partial y_j}]$, где је $j = 1, \ldots, n$ и $k = 1, \ldots p$. Извод сложене функције $g \circ f$ записујемо тада са

$$\begin{pmatrix} \frac{\partial(g_1\circ f)}{\partial x_1}(a) & \ldots & \frac{\partial(g_1\circ f)}{\partial x_m}(a) \\ \vdots & \vdots & \vdots \\ \frac{\partial(g_p\circ f)}{\partial x_1}(a) & \ldots & \frac{\partial(g_p\circ f)}{\partial x_m}(a) \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(b) & \ldots & \frac{\partial g_1}{\partial y_n}(b) \\ \vdots & \vdots & \vdots \\ \frac{\partial g_p}{\partial y_1}(b) & \ldots & \frac{\partial g_p}{\partial y_n}(b) \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \ldots & \frac{\partial f_1}{\partial x_m}(a) \\ \vdots & \vdots & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \ldots & \frac{\partial f_n}{\partial x_m}(a) \end{pmatrix}.$$

Изједначавањем одговарајућих елемената имамо да је

$$\frac{\partial(g_k\circ f)}{\partial x_i}(a) = \sum_{j=1}^{n} \frac{\partial g_k}{\partial y_j}(b)\frac{\partial f_j}{\partial x_i}(a).$$

Напоменимо још да у специјалном случају $p = 1$ формула гласи

$$\frac{\partial g(f_1(x_1,\ldots,x_m), f_2(x_1,\ldots,x_m),\ldots,f_n(x_1,\ldots,x_m))}{\partial x_i}(a) = \frac{\partial g}{\partial y_1}(b)\frac{\partial f_1}{\partial x_i}(a) + \frac{\partial g}{\partial y_2}(b)\frac{\partial f_2}{\partial x_i}(a) + \ldots + \frac{\partial g}{\partial y_n}(b)\frac{\partial f_n}{\partial x_i}(a).$$

У претходном запису је нотација $\frac{\partial g}{\partial y_i}$ за парцијални извод g по i-тој променљивој мало вештачка, јер се променљиве y и не појављују нигде. Отуда може да буде згоднија ознака $\partial_i g$.

Из претходне формуле имамо и да извод комутира са линеарним пресликавањима. Прецизније, уколико је пресликавање g линеарно, како смо раније приметили то је $dg(a) = g$, а самим тим за извод композиције имамо

$$d(g\circ f)(a) = dg(b)\circ df(a) = g\circ df(a),$$

односно линеарно пресликавање g и диференцијал d су само заменили места.

Применом претходне теореме и Бине - Кошијеве формуле, за јакобијан сложене функције имамо да важи

Последица Нека су испуњени услови претходне теореме и нека је још $m = n = p$. Тада важи

$$\det(d(g\circ f)(a)) = \det dg(b)\det df(a).$$

Теореме о средњој вредности

За почетак, покажимо да за скаларне функције важи исти реазултат као у једнодимензионалном случају

Теорема Нека је $A \subset \mathbb{R}^m$ отворен скуп, $a = (a_1,\ldots,a_m) \in A$ и $h = (h_1,\ldots,h_m) \in \mathbb{R}^m$ такви да је дуж (сегмент)

$$[a, a+h] = \{a + th | t \in [0,1]\}$$

садржана у A (приметимо да је из отворености скупа A ово увек случај за h мало). Ако је $f\colon A \to \mathbb{R}$ непрекидна у свим тачкама $[a, a+h]$ и диференцијабилна у свим тачкама $(a, a+h)$ тада постоји тачка $c \in (a, a+h)$ таква да је

$$f(a+h) - f(a) = df(c)h = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(c).$$

Доказ се своди на посматрање функције $g\colon \mathbb{R} \to \mathbb{R}^m$ дате са $g(t) = a + th$. Посматрајмо даље композицију $\varphi = f\circ g\colon \mathbb{R} \to \mathbb{R}$. За њу по Лагранжевој Теореми постоји $\theta \in (0,1)$ такво да је

$$\varphi(1) - \varphi(0) = \varphi'(\theta).$$

Са друге стране је, по теореми о изводу сложене функције $\varphi'(\theta) = f'(g(\theta))g'(\theta)$. Даље, $g' = h$, у свим тачкама, док је f' задан уобичајено преко парцијалних извода. Одатле имамо да је $f'(g(\theta))g'(\theta) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(a + \theta h)h_i$. Детаље погледати у књизи.

Специјално из претходне теореме имамо често коришћену процену за диференцијабилне функције, дату са

$$\left|f(a+h) - f(a)\right| = \left|df(c)h\right| \overset{H(2,2)}{\leqslant} \left(\sum_{i=1}^{m}\left|\frac{\partial f}{\partial x_i}(c)\right|^2\right)^{1/2}\left(\sum_{i=1}^{m} h_i^2\right)^{1/2} \leqslant \sup_{t\in[0,1]}\left\|df(a+th)\right\|\left\|h\right\|,$$

где је у наведеном кораку коришћена Коши - Шварцова неједнакост. За векторске функције не постоји аналогија претходне теореме. Ипак, имамо алтернативу у виду претходне неједнакости. Прецизније

Теорема Нека је $A \subset \mathbb{R}^m$ отворен скуп, $f\colon A \to \mathbb{R}^n$ непрекидно диференцијабилна функција и нека су $a, h \in \mathbb{R}^m$ такве да $[a, a+h] \in A$. Тада важи

$$\left\|f(a+h) - f(a)\right\| \leqslant \left\|h\right\| \cdot \sup_{t\in[0,1]}\left\|df(a+th)\right\|.$$

Доказ погледати у књизи.

Дакле, у векторском случају не мора постојати тачка у којој се достиже одговарајућа једнакост.

Једна директна последица претходне две теорем је, слично као у једнодимензионалном случају, да једино константне функције на повезаном скупу имају извод нула.

Став Нека је $A \subset \mathbb{R}^m$ област и $f\colon A \to \mathbb{R}^m$ диференцијабилна функција. Ако за свако $a \in A$ важи $df(a) = 0$, тада је f константна функција.

Проблем у претходном је што се не може директно применити претходна теорема, јер она захтева да је функција дефинисана на дужи, док скуп A иако повезан, за две тачке $x, y \in A$ не мора да садржи дуж која спаја x и y. Са тим у виду погледати доказ у књизи.

Парцијални изводи вишег реда. Тејлорова формула

Нека је функција $f\colon A \to \mathbb{R}$, где је $A \subset \mathbb{R}^m$ област, таква да у свакој тачки постоји парцијални извод $\frac{\partial f}{\partial x_i}$. Као и у једнодимензионалном случају, тада је $\frac{\partial f}{\partial x_i}$ опет функција из A у $\mathbb{R}$ те има смисла рачунати даље парцијалне изводе те функције, уколико постоје. Самим тим $\frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i})$ називамо другим парцијалним изводом и записујемо га са $\frac{\partial^2 f}{\partial x_j \partial x_i}$. У случају $x_i = x_j$ користимо запис $\frac{\partial^2 f}{\partial x_i^2}$, по аналогији са једнодимензионалним случајем. Слично се даље дефинишу парцијални изводи произвољног реда.
Природно се поставља питање колико се разликују парцијални изводи када се замени редослед диференцирања, $\frac{\partial^2 f}{\partial x_j \partial x_i}$ и $\frac{\partial^2 f}{\partial x_i \partial x_j}$. Иако у општем случају нису исти, за шта се може погледати пример у књизи, у већини разумних случајева редослед диференцирања није битан. Прецизније важи
Став Нека $f\colon A \to \mathbb{R}$ има друге мешовите парцијалне изводе

$$\frac{\partial^2 f}{\partial x_j \partial x_i} \quad \text{и} \quad \frac{\partial^2 f}{\partial x_i \partial x_j},$$

за све $x \in A$, и нека су они још и непрекидни у тачки $a \in A$. Тада су они у тој тачки и једнаки.
Доказ погледати у књизи.
Директна последица претходног става је да ако су сви други парцијални изводи непрекидни у свим тачкама (што је углавном случај којим ћемо се бавити), тада је редослед диференцирања небитан. Применом индукције исто важи и за изводе вишег реда. Нека је $f\colon A \to \mathbb{R}$ таква да је сваки n-ти парцијални извод

$$\frac{\partial^n f}{\partial x_{i_n} \cdots \partial x_{i_1}},$$

где је $i_1, \ldots, i_n \in \{1, \ldots, m\}$ и наравно може бити једнаких, непрекидан. Тада је претходни извод исти при било ком распореду индекса $i_1, \ldots, i_n$. Скуп свих таквих функција, по аналогији са претходним, означавамо са $C^n(A)$, а саме такве функције називамо и n- пута непрекидно диференцијабилним.
Тејлорова формула
Ако се сетимо, Тејлорова формула са остатком у Лагранжевом облику, у једнодимензионалном случају, је гласила

$$f(a+h) - f(a) = \sum_{k=1}^{n} \frac{1}{k!} h^k \frac{d^k}{dx^k} f(a) + \frac{1}{(n+1)!} h^{n+1} \frac{d^{n+1}}{dx^{n+1}} f(a + \theta h).$$

Заправо, слично као теорема о средњој вредности за скаларне функције, која се сводила на Лагранжеву Теорему и рачунање извода сложене функције, и претходна формула се директно преноси на случај функција више промењљивих. Остаје дакле само идентификовати ознаке (као уосталом већина ствари код функција више промењљивих).
Нека је дакле $f \in C^{n+1}(A)$, где је A област у $\mathbb{R}^m$, и $a, h \in \mathbb{R}^m$ такви да цела дуж $[a, a+h] \subset A$. Означимо са $g\colon \mathbb{R} \to \mathbb{R}^m$ функцију дату са $g(t) = a + th$, и са $\varphi = f \circ g\colon \mathbb{R} \to \mathbb{R}$. Како је φ реална функција, по Тејлоровој формули у Лагранжевом облику имамо да је

$$\varphi(1) - \varphi(0) = \varphi'(0) + \frac{1}{2!}\varphi''(0) \cdot 1 + \ldots + \frac{1}{n!}\varphi^(n)(0) \cdot 1^n + \frac{1}{(n+1)!}\varphi^{(n+1)}(\theta) \cdot 1^{(n+1)},$$

где је $\theta \in (0, 1)$. Претходна примена је оправдана, јер је φ $n+1$ пута диференцијабилна на интервалу $[0, 1]$. Заиста, функција g има парцијалне изводе сваког реда (који су сви нула, осим првог реда), док је f по претпоставци $n+1$ пута диференцијабилна. Од свега дакле, остаје да се израчуна шта је k-ти извод од φ.
Прво за $k = 1$, претходно имамо из извода сложене функције, дакле

$$\varphi'(t) = f'(g(t)) \cdot g'(t) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a+th) & \ldots & \frac{\partial f}{\partial x_m}(a+th) \end{pmatrix} \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix} = \frac{\partial f}{\partial x_1}(a+th)h_1 + \ldots + \frac{\partial f}{\partial x_m}(a+th)h_m.$$

Дакле, g је векторска функција, и њене координатне функције g_i су редом $a_i + th_i$, а извод претходног по t износи h_i. Одатле, њена Јакобијева матрица се састоји од само једне колоне (јер је функција само једне промењљиве). Са друге стране, Јакобијева матрица f се састоји само од једне врсте, јер је f скаларна функција, одакле свезаједно имамо претходну формулу. Такође, одавде видимо директно и да је $g''(t) = 0$.
Израчунајмо сада други извод, који је први извод првог извода. Тада ће бити

$$\varphi''(t) = (f'(g(t)) \cdot g'(t))' = f''(g(t)) \cdot g'(t) + f'(g(t)) \cdot g''(t) = f''(g(t)) \cdot g'(t) + 0.$$

У претходном смо прошли са изводом кроз производ матрица као кроз уобичајени производ, прецизније искористили смо да ако су $A(t)$ и $B(t)$ функције матрица које могу да се множе да е тада $(A(t)B(t))' =$

$A'(t)B(t)+A(t)B'(t)$, што се аналогно показује као и формула за обично множење. Дакле, сада остаје да се израчуна $f''(g(t))$.
Један начин на који можемо претходно да израчунамо јесте да ако посматрамо претходни производ матрица који диференцирамо $f'(g(t))\cdot g'(t)$, ту је $f'(g(t))$ градијент, али је записан као врста а не као колона како уобичајено записујемо векторе. Отуда да би тај производ имао смисла, претходно можемо да запишемо са $(\operatorname{grad} f(g(t)))^T\cdot g'(t)$. Отуда, треба да израчунамо $((\operatorname{grad} f(g(t)))^T)'$. Како је транспоновање линеарно, то је прво $((\operatorname{grad} f(g(t)))^T)' = ((\operatorname{grad} f(g(t)))')^T$, а даље, како је $\operatorname{grad} f\colon A\to\mathbb{R}^m$, векторска функција, њен извод је задат уобичајено Јакобијевом матрицом. Прецизније, координатне функције градијента су $\frac{\partial f}{\partial x_i}$, тако да је Јакобијева матрица $\operatorname{grad} f$ дата са

$$\begin{pmatrix}\operatorname{grad}\frac{\partial f}{\partial x_1}\\ \vdots\\ \frac{\partial f}{\partial x_m}\end{pmatrix} = \begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2\partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_m\partial x_1}\\ \vdots & \vdots & \vdots & \vdots\\ \frac{\partial^2 f}{\partial x_1\partial x_m} & \frac{\partial^2 f}{\partial x_2\partial x_m} & \cdots & \frac{\partial^2 f}{\partial x_m^2}\end{pmatrix}.$$

Како је даље у птиању сложена функција $\operatorname{grad} f(g(t))$, то је њен извод

$$\big(\operatorname{grad} f(a+th)\big)' = \begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2}(a+th) & \frac{\partial^2 f}{\partial x_2\partial x_1}(a+th) & \cdots & \frac{\partial^2 f}{\partial x_m\partial x_1}(a+th)\\ \vdots & \vdots & \vdots & \vdots\\ \frac{\partial^2 f}{\partial x_1\partial x_m}(a+th) & \frac{\partial^2 f}{\partial x_2\partial x_m}(a+th) & \cdots & \frac{\partial^2 f}{\partial x_m^2}(a+th)\end{pmatrix}\cdot\begin{pmatrix}h_1\\ \vdots\\ h_m\end{pmatrix},$$

аналогно као у рачуну за први извод. Најзад, ако још транспонујемо одговарајуће матрице, други извод функције φ је дат са

$$\varphi''(t) = h^T\cdot\begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2}(a+th) & \frac{\partial^2 f}{\partial x_2\partial x_1}(a+th) & \cdots & \frac{\partial^2 f}{\partial x_m\partial x_1}(a+th)\\ \vdots & \vdots & \vdots & \vdots\\ \frac{\partial^2 f}{\partial x_1\partial x_m}(a+th) & \frac{\partial^2 f}{\partial x_2\partial x_m}(a+th) & \cdots & \frac{\partial^2 f}{\partial x_m^2}(a+th)\end{pmatrix}\cdot h = \sum_{i=1}^m\sum_{j=1}^m\frac{\partial^2 f}{\partial x_i\partial x_j}h_ih_j,$$

где се последња једнакост добија директним множењем матрица. Квадратну матрицу из претходне једнакости $[\frac{\partial^2 f}{\partial x_i\partial x_j}]_{i,j=1}^m$ називамо још и матрица другог извода.
Други начин да се претходно израчуна је директно. Дакле, из претходног знамо да је

$$f'(g(t))g'(t) = \sum_{i=1}^m\frac{\partial f}{\partial x_i}(a+th)h_i.$$

Диференцирањем по t, имамо да је

$$\varphi''(t) = \Big(\sum_{i=1}^m\frac{\partial f}{\partial x_i}(a+th)h_i\Big)' = \sum_{i=1}^m\Big(\frac{\partial f}{\partial x_i}(a+th)h_i\Big)' = \sum_{i=1}^m\sum_{j=1}^m\frac{\partial^2 f}{\partial x_j\partial x_i}h_jh_i,$$

где је само искоришћено да је

$$\frac{d}{dt}\frac{\partial f}{\partial x_i}(a+th)h_i = \sum_{j=1}^m\frac{\partial^2 f}{\partial x_j\partial x_i}h_jh_i,$$

у суштини потпуно исто као што смо рачунали први извод φ. Приметимо да у претходном не морамо да водимо рачуна о редоследу парцијалних извода.
Настављајући поступак (други начин конкретно се лако показује за све изводе), добија се да је

$$\varphi^{(k)}(t) = \sum_{i_1,i_2,\ldots,i_k=1}^m\frac{\partial^k f}{\partial x_{i_1}\cdots\partial x_{i_k}}(a+th)h_{i_1}h_{i_2}\cdot\ldots\cdot h_{i_k}.$$

Последњи израз се још означава са

$$\Big(h_1\frac{\partial}{\partial x_1}+h_2\frac{\partial}{\partial x_2}+\ldots+h_m\frac{\partial}{\partial x_m}\Big)^k f(a+th),$$

где се користи уобичајено степеновање у загради (тј. полиномијална формула), али када се појаве парцијални изводи не множимо их, већ диференцирамо функцију f. Опет приметимо редослед није битан под датим претпоставкама.
На крају сам запис Тејлорове формуле испада превише рогобатан у општем случају па га нећемо наводити у једној формули. Преко последњег записа може се записати са

$$f(a+h)-f(a) = \sum_{k=1}^n\frac{1}{k!}\Big(h_1\frac{\partial}{\partial x_1}+h_2\frac{\partial}{\partial x_2}+\ldots+h_m\frac{\partial}{\partial x_m}\Big)^k f(a)+\frac{1}{(n+1)!}\Big(h_1\frac{\partial}{\partial x_1}+h_2\frac{\partial}{\partial x_2}+\ldots+h_m\frac{\partial}{\partial x_m}\Big)^{n+1}f(a+th).$$

Локални екстремуми

Када имамо Тејлорову формулу, слично као у једнодимензионалном случају можемо дати неопходне а и довољне услове за локалне екстремуме. За почетак

Дефиниција Нека је $f\colon A \to \mathbb{R}$ и $a \in A$. Кажемо да је a локални максимум (минимум) ако постоји околина U тачке a таква да је $f(x) \leqslant f(a)(f(x) \geqslant f(a))$ за све $x \in U$. Ако је тачка било шта од претходна два кажемо да је a локални екстремум.

Неопходни услови локалног екстремума су аналогни

Став Нека функција $f\colon A \to \mathbb{R}$, где је $A-$област, у тачки $a \in A \subset \mathbb{R}^m$ има све парцијалне изводе првог реда и нека у тој тачки има локални екстремум. Тада су сви парцијални изводи у тој тачки једнаки нули.

Ако фиксирамо све кординате осим једне и посматрамо функцију $g(x) = f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_m)$, она ће бити дефинисана на околини $(a_i - r, a_i + r)$ за r довољно мало, јер је A отворен скуп. Такође, a ће бити и локални екстремум функције g. Одатле је, по Фермаовом ставу $g'(a_i) = 0$, али извод функције g је заправо $\frac{\partial f}{\partial x_i}(a)$. Отуда следи закључак става.

Као и раније, тачке у којима је извод нула називају се стационарне тачке функције f.

Дакле, претходни став нам даје све кандидате за локалне екстремуме, уз напомену да се ради о унутрашњим тачкама (претпоставили смо са отворенсохћу скупа A да је a унутрашања тачка, а и неопходно нам је у доказу). У једнодимензионалном случају је даље углавном било једноставно закључити које су од тих тачака заиста екстремуми, тј. довољно је било само посматрати знак извода лево и десно од те тачке. Како у вишедимензионалном случају тачки можемо прићи по много смерова (не само два као раније), претходно питање постаје компликованије.

Развијмо f у Тејлоров полином другог степена у околини a. Како су изводи нула, то је

$$f(a+h) = f(a) + \frac{1}{2!}\sum_{i,j=1}^{m} \frac{\partial^2 f}{\partial x_j \partial x_i} f(a) h_i h_j + o(\|h\|^2).$$

Очекујемо даље, да $o(\|h\|^2)$ не утиче на знак, тј. да ли је разлика $f(a+h) - f(a)$ позитивна или негативна да зависи само од знака $\sum_{i,j=1}^{m} \frac{\partial^2 f}{\partial x_j \partial x_i} f(a) h_i h_j$. Пре него што прецизирамо претходно резоновање, треба нам нешто линеарне алгебре, односно начин на који испитујемо знак израза $\sum_{i,j=1}^{m} a_{ij} h_i h_j$.

Дефиниција Реална функција m променљивих дата са

$$\Phi(h_1, \dots, h_m) = \sum_{i,j=1}^{m} a_{ij} h_i h_j,$$

назива се квадратна форма. Матрица $A(\Phi) = [a_{ij}]_{i,j=1}^{m}$ назива се матрица квадратне форме. Даље, претпостављамо да је матрица A увек симетрична, тј. да је $a_{ij} = a_{ji}$, односно $A^T = A$ (ово није општи случај, али матрица другог извода је таква, што је последица једнакости мешовитих извода). За квадратну форму кажемо да је

1. Позитивно (негативно) полудефинитна, ако је $\Phi(h_1, \dots, h_m) \geqslant 0$, односно $(\leqslant 0)$, за све $(h_1, \dots, h_m) \in \mathbb{R}^m$.
2. Позитивно (негативно) дефинитна, ако је $\Phi(h_1, \dots, h_m) > 0$, односно (< 0), за све $(h_1, \dots, h_m) \in \mathbb{R}^m$.
3. Променљивог знака, ако постоје $h_1, h_2 \in \mathbb{R}^m$ такви да је $\Phi(h_1) < 0$ а $\Phi(h_2) > 0$.

Приметимо и да форму можемо да запишемо матрично

$$\Phi(h_1, \dots, h_m) = \sum_{i,j=1}^{m} a_{ij} h_i h_j = h^T A h,$$

где је $h = (h_1, \dots, h_m)$.

За испитивање дефинитности, користимо следећи Силвестеров критеријум

Став Нека је

$$A(\Phi) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{pmatrix}$$

симетрична матрица квадратне форме $\Phi\colon \mathbb{R}^m \to \mathbb{R}$ и нека су

$$A_1 = a_{11}, \quad A_2 = \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \dots \quad A_m = \det A(\Phi),$$

детерминанте њених главних минора. Тада важи

1. Φ је позитивно дефинитна ако и само ако су сви A_i позитивни, тј.

$$A_1 > 0, \quad A_2 > 0, \quad \dots, \quad A_m > 0.$$

2. Φ је негативно дефинитна ако и само ако A_i наизменично мењају знак (или ти да су $(-1)^n A_n$ позитивни), тј. ако је

$$A_1 < 0, \quad A_2 > 0, \quad A_3 < 0, \dots$$

Пре формулисања довољних услова локалног екстремума неопходно је још једно тврђење

Став Нека је $A \subset \mathbb{R}^m$ и $a_{ij}\colon A \to \mathbb{R}$ непрекидне функције, такве да је још $a_{ij}(x) = a_{ji}(x)$. Означимо са Φ_x квадратну форму са матрицом $[a_{ij}(x)]$ за свако $x \in A$. Ако је за неко $a \in A$, квадратна форма Φ_x позитивно дефинитна, тада постоји околина U тачке a, таква да је Φ_x позитивно дефинитна за свако $x \in U$.

Како је детерминанта (јер је само производ и збир уноса матрице) непрекидна функција, то су све функције

$$A_1(x) = a_{11}(x), \quad ,A_2(x) = \det\begin{pmatrix} a_{11}(x) & a_{12}(x) \\ a_{21}(x) & a_{22}(x) \end{pmatrix}, \quad \ldots, \quad A_m(x) = \det A(\Phi_x).$$

непрекидне. Како су све оне позитивне у тачки a, по Силвестеровом критеријуму, то за сваку од функција $A_i(x)$ посотји околина на којој је она сталног знака, тј. позитивна. На пресеку свих тих околина, што је опет околина тачке a, је свака од њих позитивна, а самим тим, опет по Силвестеровом критеријуму, Φ је позитивно дефинитна у свакој тачки из те околине.

Формулишимо сада довољне услове локалног екстремума

Теорема Нека је $A \subset \mathbb{R}^m$ отворен, $a \in A$ и $f \in C^2(A)$, при чему је a стационарна тачка функције f. Нека је Φ квадратна форма која је задата другим изводом f, тј.

$$\Phi(h_1, \ldots, h_m) = \frac{\partial^2 f}{\partial x_j \partial x_i} f(a) h_i h_j.$$

Тада:

1. Ако је квадратна форма Φ позитивно дефинитна, функција f има строги локални минимум у тачки a.
2. Ако је квадратна форма Φ негативно дефинитна, функција f има строги локални максимум у тачки a.
3. Ако форма Φ мења знак, тада f нема у a локални екстремум.

Доказ погледати у књизи, али није неопходно.

Претходна теорема, иако технички није једноставна за доказ, нема неочекиван закључак. Према резоновању са почетка, део 1. одговара томе да је $f(a+h) - f(a) \geqslant 0$, тј. да је $f(a+h) \geqslant f(a)$ и аналогно за део 2. Део 3. нема аналогију у једнодимензионалном случају, али ни он није неочекиван, тј. одговара томе да $f(a+h) - f(a)$ мења знак, дакле, a није екстремум.

Наравно, остаје проблем уколико је форма другог извода полудефинитна, дакле уколико има вектора h, на којима је $\Phi(h) = 0$, где мора експлицитно да се проверава да ли је a екстремум. Такође, напоменимо опет да се претходно односи само на унутрашње тачке домена функције (као и у једнодимензионалном случају, између осталог). Испитивањем руба домена бавићемо се касније.